# Microfinance and agricultural productivity

## A study exploring the relationship between access to credit and agricultural productivity

Simon Andersson & Albert Wallgren

**Abstract:**

This thesis establishes a macro-level analysis of how credit and microfinance affect agricultural productivity. We use fixed effects regression for panel data with countries as the cross-sectional variation from 2000 to 2018. We also use income categorization of countries from the UN to better understand for which income categories microfinance has an effect. The model consists of cereal yield as the dependent variable and disbursement to agriculture, fertilizer consumption, secondary school enrollment, and population growth as the independent variables. The research question of the study is: **In what respect does access to credit affect agricultural productivity on a macro-level?**

Our results show a significant effect of disbursement to agriculture on cereal yield, indicating that credit and microfinance have a positive effect on agricultural productivity. Secondary school enrollment has a positive effect on cereal yield, higher enrollment in secondary school leads to higher cereal yield. Population growth is also significant; higher age-dependency ratios lead to higher productivity. The results indicate no significant relationship between fertilizer and cereal yield.

The thesis adds value to the existing body of literature regarding development economics, especially credit and microfinance's effect on the well-being of the developing countries. Based on our results, access to credit leads to higher agricultural productivity. However, for low-income countries, no relationship is established. It supports earlier studies indicating a mismatch between microfinance and agriculture. Therefore, efforts to construct microfinance to better match the needs of farmers are required to promote higher agricultural productivity in low-income countries.

Bachelor's thesis in Economics, 15 credits
Spring Semester 2022
Supervisor: Christer Ljungwall

Department of Economics
School of Business, Economics and Law
University of Gothenburg

# Table of contents

# 1. Introduction

Microfinance is often cited as one of the most promoted methods of reducing world poverty. However, it has given mixed results in reaching and helping people out of poverty (Köhn, 2013). Agriculture is a large share of what the poorest people in the world have as their livelihood. Data from 2013 shows that 84 percent of the world's total population lives in developing countries, a majority of which are involved in agriculture. (de Janvry & Sadoulet, 2021) Therefore, growth in agriculture is an essential factor for aggregate economic growth and poverty reduction in many developing countries (World Bank, 2008). Essential for increasing agricultural productivity is farmers' adoption of new technology. However, several obstacles impede the rate of adoption, one of these is financial constraints. Microfinance is a method to reduce financial constraints (de Janvry & Sadoulet, 2021).

Microfinance loans usually consist of loans with repayments that start shortly after the loan is issued. It means that microfinance loans usually only reach poor people with businesses that can generate fast income returns (Weber, 2013). There is a mismatch between one of the most promoted poverty reduction methods, microfinance, and the primary source of livelihood for the poorest people (de Janvry & Sadoulet, 2021, Weber, 2013). It takes time from planting a seed to harvesting the crop; there is a time lag in agriculture before investments can generate higher income streams (de Janvry & Sadoulet, 2021). It leads to farmers easily being left out of microfinance support (Weber, 2013).

The mismatch leads to an interesting question about the relationship between better access to credit and increased productivity in agriculture. Since technology adoption is an essential factor in increasing productivity in agriculture, it is interesting to examine how better access to credit can increase the adoption of new technology and if this can lead to higher productivity in agriculture. It leads to the purpose and research question of the study:
**In what respect does access to credit affect agricultural productivity on a macro-level?**

The thesis investigates the relationship between access to credit and productivity in agriculture, with a focus on how technology adoption affects productivity. Using data from the World Bank and the Food and Agriculture Organization of the United Nations, spanning the time range from 2000 to 2018. The model and variables build on existing literature investigating the problem on a micro-level.

Between 2002 and 2019, the world spent $ 40.2 billion on Agricultural development (Atteridge et al., 2019). Therefore, it is essential to analyze if and how the funds affect agricultural development. An increasing body of literature has drawn attention to the fact that previous studies on microfinance mainly focus on its impact at a micro-level where the focus is on individuals, families, villages, or regions. However, a low proportion of previous literature focus on the aggregate impact of microfinance at a macro level (Islam & O'Gorman, 2019, Alimukhamedova, 2017). This study contributes to the literature by investigating the impact of microfinance on a macro-level. More studies that examine the impact of microfinance on an aggregate level are needed. To our knowledge, no previous studies have focused on the problem of agricultural productivity on a macro level.

The organization for the rest of the thesis is as follows. Section 2 literature review highlights previous studies related to the thesis topic and focuses on microfinance, agriculture, technology, information, and learning. Section 3 focuses on the theoretical framework of agricultural productivity. Section 4 methodology describes the advantages of panel data and the fixed effects regression model we use for causal analysis. Section 5 data gives detailed information about the variables and presents assumptions for linear regression and various tests for the regression analysis. Section 6 presents the results from our model. Finally, sections 7 and 8 contain the analysis and conclusion, respectively.

# 2. Literature review

A vast literature in development economics examines the relationship between microfinance and poverty reduction in developing countries. Studies in the area take different focal points, such as poverty reduction, women empowerment, business startups, and the prosperity of poor people. In addition, some previous studies have examined microfinance and technological innovations in agriculture. The overall conclusion is that technological innovation positively affects productivity growth in agriculture. However, there are mixed results for microfinance's impact on agriculture and technology adoption for farmers.

## 2.1 Barriers for the adoption of new technologies

In low-income countries, much of the agricultural production is done by smallholder farmers. Smallholder farmers face barriers when it comes to the adoption of new technology. Examples of potential barriers for farmers to adopt new technologies are profitability, information and learning, credit and insurance constraints, and lack of secure property rights. The profitability of new technologies may be too low for farmers to adopt them. Farmers who lack property rights over their land have lower incentives to invest in technologies to increase productivity. Due to the risk of losing the land and the potential higher yield. (de Janvry & Sadoulet, 2021)

Microfinance is a substitute for the lack of access to commercial loans for poorer people due to the credit and insurance market failures often appearing in low-income countries. The credit and insurance market failures in many low developed countries are an obstacle for poor households and farmers to access credit and insurance. It is common for microfinance providers to require the repayment to start shortly after issuing standard loans. (de Janvry & Sadoulet, 2021) Although the number of microfinance programs has increased significantly, it has not benefited farmers and agriculture (World Bank, 2008). Many smallholder farmers do not get access to credit, and microfinance loans are too costly for many agricultural activities. In addition, the difficulty of gaining access to credit creates high barriers for smallholder farmers to access many new technologies since they have a high upfront cost that requires credit. (de Janvry & Sadoulet, 2021)

There is a mismatch between short-term microfinance loans and long-term investments in agriculture. Investments in agriculture are often tied up in production for a long time since

there typically are long lags from planting to harvesting. Lags in agriculture can be from four to six months up to four to five years. Ordinary microfinance loans have a short time before the repayment starts. (de Janvry & Sadoulet, 2021) Weber (2013) conducted a study examining the impact of flexible microfinance loans adapted for agriculture with a mixed-methods approach with observations from field visits and in-depth portfolio analyses. The results show that flexible loans are an important factor for farmers with seasonal income flows not to default on their loans and to have access to credit. Flexible loans increase the possibility of seasonal agriculture farmers getting access to microfinance loans. An essential aspect of flexible loans is grace periods, leading to a significantly lower loan default rate than standard loans. Standard loans are not eligible for seasonal agriculture and are better suited for businesses with faster turnover in income-generating activities. (Weber, 2013)

There is thus a difference in when farmers receive their income compared to other sectors. Farmers gain most of their income after their harvest. The main part of the microfinance should instead be paid back after the harvest to better suit the farmers. Compared with the frequent repayment directly after the start of the lending implemented by the usual microfinance programs. (Nakano & Magezi, 2020) One consequence of this may be that investments that had been profitable in agriculture are not realized due to the mismatch between income flows from harvest and repayment obligations from short-term microfinance loans (Weber, 2013).

## 2.2 Positive impacts of microfinance on technology adoption

Two of the studies that find a positive impact of microfinance on technology adoption and productivity growth in agriculture are written by Mariyono (2018a & 2018b). The first article by Mariyono (2018a) examines the impact of microfinance and agriculture technology and factors essential for farmers' access to microfinance, using a structural equation model (SEM). The results in the first article show that microfinance positively impacts farm households both directly and indirectly. There is a significant positive impact on farmers' prosperity through access to credit and the adoption of new technology in agriculture. The new technology in agriculture enabled farmers to achieve a higher profit. (Mariyono, 2018a)

de Janvry and Sadoulet (2002) finds a similar conclusion about the direct and indirect effects of technological change in agriculture and how it affects poverty. The technological change

can increase production for the farmers who adopt the new technology through the direct effect. The increased production due to the new technology can lead to higher revenue from sales, lower costs in production, and higher consumption for the own household. The indirect effect of new technology in agriculture can positively affect poverty through, among other things, lower food prices, higher employment, and wages in agriculture and through other sectors of the country's economy. (de Janvry & Sadoulet, 2002)

Microfinance was used to finance technology acquisition and acted as a helping hand for farmers to adopt the technology (Mariyono, 2018a). Since credit constraint is a hindrance that stops farmers from adopting new technology (de Janvry & Sadoulet, 2021). However, just having increased access to microfinance is not enough to increase the prosperity of farmers. Microfinance needs to be used to acquire new technology for agriculture to increase farmers' profits (Mariyono, 2018a). Mariyono (2018b) shows that microfinance has an essential role in improving farmers' productivity and prosperity through the access and adoption of new technology.

Osabohien et al. (2022) is another study finding a positive impact of credit on agriculture. They examine the long-run effect of access to credit and agricultural performance in Nigeria using the Autoregressive Distribution Lag (ARDL) method. The results show a significant relationship between access to credit and increased agricultural performance in Nigeria. They highlight the importance of low interest rates for farmers to access credit. High interest rates lead to a lower possibility for farmers to access credit, which lowers investments in agriculture and leads to lower outputs. Their conclusion is thus that farmers need to be allowed to access sufficient credit to acquire better inputs such as fertilizers or machinery that can lead to higher output. (Osabohien et al., 2022)

Mariyono (2018b) uses an SEM method to analyze the role of microfinance in improving the standard of living in rural areas. This study further includes an investigation of which factors influence access to microfinance and adoption of new technologies for farmers. The results show that microfinance positively impacts living standards in rural areas. Moreover, microfinance once again acts as an intermediary for farmers to adopt new technology (Mariyono, 2018b). New technologies are often expensive to acquire, which requires farmers to have access to credit to acquire them. The reasoning is that new technology will increase

productivity for farmers to increase their profitability, which leads to a higher standard of living for farmers. (Mariyono, 2018b)

Other studies find positive impacts of credit to agriculture on productivity; for example, Peng et al. (2020) that use a random-effects model combined with a generalized method of moments method. They take a broader scope and examine the impact of credit to agriculture on regional agricultural growth and regional economic development in Jiangsu Province in China. Their results show a significant positive impact of credit to agriculture on both the growth in agriculture and the economic development in the region. Through credit to agriculture, growth is created in agriculture, which leads to economic growth in the region. (Peng et al., 2020)

There is usually a lack of physical capital in rural areas, and access to credit is an essential factor in reducing this lack of physical capital. In addition, investments are often an important factor for economic development. Therefore, access to credit is a significant factor for farmers in rural areas to make investments that can increase their production and productivity, leading to economic development. (Peng et al., 2020)

Twumasi et al. (2021) have a slightly different focus but still find the importance of credit for rural farm households. Using an endogenous switching regression (ESR) model and a problem confronting index (PCI) model, the study examines if access to credit affects farmers from abandoning their farmland in Ghana. The study shows that access to credit is an important factor for farmers not to abandon their land. Furthermore, it shows the importance of creating policies that increase the opportunities for farmers to access credit. Access to credit will improve participation in agriculture and promote rural financial development, for example, by reducing production costs. They also examine the factors that affect farmers' access and find that age, education, household size, off-farm work, belonging to an association, and land registration increase the possibility of getting access to credit. (Twumasi et al., 2021) These studies by Mariyono (2018a & 2018b), de Janvry and Sadoulet (2002), Osabohien et al. (2022), Peng et al. (2020) and Twumasi et al. (2021) all provide a clear understanding of why microfinance is an essential factor for farmers to acquire and adopt new technologies for agriculture.

## 2.3 Information and human capital for adoption of new technologies

Insufficient information and learning can be a hindrance to the adoption of new technology. If information and learning about new technologies, their expected benefits, and how to use them are missing or undersupplied, farmers' adoption of new technologies will be lower (de Janvry & Sadoulet, 2021). Foster and Rosenzweig (1996) also highlight the importance of education for better and more effective adoption of technological innovations. They examine the connection between education and technical change. Using panel data, they conclude that more educated individuals are better at taking advantage of technological innovations. Technological innovations are also likely to have a more significant effect in a country with a higher proportion of educated people than in a country with a lower level of education. (Foster and Rosenzweig, 1996)

Schultz (1964) (referred to in Kaldor, 1964) reach a similar conclusion about the importance of human capital. He emphasizes the importance of investing in agriculture, focusing on modern technology inputs and human resources. By formulating a conceptual model of traditional agriculture, he examines several agricultural issues in low-income countries. He argues that investments in traditional agriculture will give low returns and have a negligible effect on income growth. Instead, investments to generate a transformation of agriculture must improve human capital, such as farmers' knowledge, skills, and health. Investments also need to be directed at modern technologies such as fertilizers and seeds to increase agricultural productivity. (Schultz, 1964, referred to in Kaldor, 1964).

Mariyono (2018b) also shows that information is essential for technology adoption. Therefore, more information sources lead to more extensive technology adoption by farmers. Furthermore, as the number of information sources increases, the asymmetric information decreases, increasing the opportunities for farmers' productivity. (Mariyono, 2018b)

Mariyono (2018b & 2018a), de Janvry and Sadoulet (2021), Foster and Rosenzweig (1996), and Schultz (1964) (referred to in Kaldor, 1964) all point to the importance of human capital for farmers to have access to microfinance and to adoption of new technology. Higher human capital means, among other things, a higher level of education and experience (Mariyono, 2018b).

Adoption of new technology requires greater access to credit but also a higher level of human capital. Just giving farmers access to new technology will not suffice either; it requires support for how to use them. New technology is often more advanced than previous technology, requiring either a higher level of human capital or information and training support is needed to help farmers adopt the new technology. (Mariyono, 2018b)

## 2.4 No impact of microfinance on agriculture

Despite the positive results of the previously mentioned studies, several studies offer different evidence and do not find any impact of microfinance on agriculture. These are especially Shahidur et al. (2014) and Nakano and Magezi (2020). Shahidur et al. (2014) examine rural credit's effect on agricultural households in their article. They define rural credit as both microfinance and loans from formal banks. Their results find no effect on increased crop incomes, productivity gains, or similar agriculture activities from microfinance (Shahidur et al., 2014).

This result is in line with findings that Weber (2013) finds when crop incomes require longer investments to give a positive impact from microfinance. Furthermore, Shahidur et al. (2014) show that microfinance is not used to acquire technology that can be used to increase productivity in agriculture. It suggests that microfinance is not a factor used to acquire new technology that can be used to increase productivity in agriculture (Shahidur et al., 2014). It gives the opposite result to what Joko Maryiano (2018a & 2018b) finds, where microfinance is used to acquire new technology used to increase productivity in agriculture.

Shahidur et al. (2014) results show, however, that microfinance leads to significantly increased income for agriculture activities such as livestock rearing. They find the most significant effect in the households that do not or only have limited access to land. They are consistent with the effect of livestock rearing, which does not require as much land (Shahidur et al., 2014).

The fact that increased agricultural productivity requires the adoption of technological innovations and that lack of access to microfinance is a constraint for farmers to adopt technology is also the starting point for Nakano and Magezi (2020) in their study. They examine the impact of microfinance on the adoption of technology and productivity of rice

cultivation in Tanzania, conducted using randomized controlled trials (Nakano & Magezi, 2020).

In general, their results show no increase in the use of fertilizers due to the availability of microfinance. Credit seems to have no impact on the adoption of fertilizers and yield. The farmers who already have better access to irrigation systems already use a recommended level of fertilizers and therefore do not increase the use of fertilizers. Their results do not find any general increase in yield in the farmers' rice production, profit, or income. (Nakano & Magezi, 2020)

In their study, they use microfinance loans that are adapted to suit agricultural production better. For example, they use loans that extend one cultivating season because farmers gain most of their income after harvest. Ordinary standard microfinance loans are not suitable for investments in agriculture as these require that the repayment of the loan starts early after the loan has been issued. (Nakano & Magezi, 2020) The same mismatch discussed by Weber (2013), Nakano and Magezi (2020) and de Janvry and Sadoulet (2021).

On the other hand, Nakano and Magezi (2020) find that farmers that have limited access to irrigation water and use a small proportion of fertilizers increase the use of fertilizers due to increased access to credit. However, the increased use of fertilizers for this group of farmers does not increase production yield. It suggests that increasing credit supply alone is not enough to achieve higher agricultural productivity. (Nakano & Magezi, 2020) Adopting new technology requires additional support, which is in line with Mariyono (2018b), who says that new technology requires support.

Nakano and Magezi (2020) show that microfinance to agriculture may not be an effective method to increase the adoption of new technologies and productivity in agriculture. Their results indicate that more complementary methods together with microfinance are needed to successfully adopt new technology and productivity growth in agriculture. However, their results need to be taken with caution as their statistical power is low due to a small sample size. (Nakano & Magezi, 2020)

# 3. Theoretical framework

## 3.1 Determinants of Agricultural Growth

There are mainly three determinants that affect the production and growth of agriculture. The first is price and market incentives. Farmers in low-income countries allocate their resources in response to price signals. Additionally, food prices are suppressed in the world market because of the protection by high-income countries for their domestic producers. Therefore, the lower world food prices make it difficult for low-income country farmers to export to the world market. The second is sustainability in using agricultural resources, as agriculture has a major impact on the environment. (de Janvry & Sadoulet, 2021) These two are out of focus for this thesis and will not be considered.

The third on which this thesis builds its focus is the availability of sources that can promote growth in agriculture, such as technology, inputs, institutions, and public goods. For agriculture to produce more output, there are two alternatives. The first alternative is to expand the land used for agriculture so that there is more land to cultivate. However, there is a limit to exploiting new agricultural land; land is a finite resource and cannot be cultivated endlessly. It emphasizes the second alternative for continued growth in agriculture, which is to increase the productivity on already cultivated agricultural land through gains in total factor productivity (TFP). Also called vertical expansion or growth on the intensive margin. (de Janvry & Sadoulet, 2021) The required growth in productivity per cultivated land, yield gains per hectare, give the output variable and dependent variable of interest for this study.

In order to generate growth in TFP in agriculture, technological and institutional innovations are highlighted as two essential factors to accomplish this (de Janvry & Sadoulet, 2021). Institutional innovations are outside the scope of this study and will therefore not be covered.

Technological innovations can increase the TFP by increasing the output for a fixed combination of land and labor inputs. Furthermore, technological innovations can decrease land and labor input costs for a given output level. Various categories of technological innovations exist that promote productivity increases in agriculture. This study concentrates on the agriculture innovation of fertilizer as a variable of technology adoption and increased productivity in agriculture. Fertilizers function as a yield-increasing technology. Other

agricultural technology innovations can be labor- and cost-saving, risk-reducing, and quality improvement (de Janvry & Sadoulet, 2021). Out of the scope of the study, these other technologies are not considered.

## 3.2 Induced Innovation Theory

The various technological innovations developed and adapted by agriculture can be explained by the Induced Innovation Theory, developed by Hayami and Ruttan (referred to in Hazell & Herdt, 1987) in their book Agricultural Development: An International Perspective. According to the theory, what drives technological innovation for agriculture and farmers is the relative factor scarcity (Hayami & Ruttan referred to in Hazell & Herdt, 1987).

Agricultural factors of production are labor and land (de Janvry & Sadoulet, 2021), and technology innovation can substitute for the factor with the highest relative price. For example, fertilizers work as land saving technology where the relative price of land is higher than for labor. Fertilizers increase the yields in the output per unit of cultivated land. If the land is relatively more expensive than labor, fertilizers and other land-saving technologies will lower the relative price of land compared to labor. If the opposite holds and labor is the factor with the highest relative price, labor-saving technologies like harvesters and tractors will lower the relative price for labor. Therefore, innovations and research in agriculture occur for the highest relative price between labor and land, the scarce resource. (Hayami & Ruttan referred to in Hazell & Herdt, 1987)

## 3.3 Demographic transition model

The demographic transition model refers to the different stages of development that countries go through. It explains how death rates, birth rates, and population growth relates to economic and social development. As death rates drop due to better healthcare and income, people eventually decrease the number of children they have. Quality of children is preferred instead of quantity, increasing human capital and labor productivity. The growth rate of the population is higher for less developed countries and lower for more developed countries (Galor, 2011).

Low death and birth rates lead to lower population growth, increasing the resources available per person. The distribution of the population demographic changes when countries' death

and birth rates decrease over time. Death rates decrease first and are followed, with a delay, by birth rates. When death rates are low and birth rates are still high, there is an increase in the ratio of working-age people to the total population, increasing the total productivity. (Galor, 2011) Which can also lead to higher labor productivity in agriculture (de Janvry & Sadoulet, 2021).

## 3.4 Cognitive load

Due to psychological and behavioral consequences, being poor can have a negative impact on one's cognitive function. If one’s whole focus is on survival, it becomes more difficult to concentrate on anything else, such as long-term planning, self-control, or developing skills. It can lead to a poverty trap where poor people do not have the capability to escape poverty. (Fehr & Haushofer, 2014)

Mullainathan et al. (2010) have a similar argument, describing that extremely poor people have limited mental bandwidth. Suppose a significant part of one focus is directed to one's own or family's survival. In that case, there is limited capacity to deal with less pressing matters, even if those matters would have a long-term positive effect on the quality of life.

Mani et al. (2013) organized a study in India investigating cognitive load by examining farmers' cognitive function pre-harvest, when they were at their poorest, compared to the same farmer post-harvest. They found that farmers had significantly worse cognitive functioning pre-harvest compared to post-harvest (Mani et al., 2013). It can affect their ability to adopt new technology. (Bruns et al., 2022)

## 3.5 Theory and Literature Review

Based on both theory and literature review, the importance of technological innovations to increase productivity in agriculture is evident. The theory shows that technological innovations can increase TFP by, among other things, increasing productivity on already cultivated land. The technological innovation for different areas is driven by the relative price between land and labor. There is a difference between the type of technological innovation that works as an effective productivity booster in agriculture. It may explain why, for example, fertilizers do not increase productivity on certain occasions. The need for labor-saving technological innovations may be needed to increase productivity. Based on previous

studies, a further explanation for this is the importance of information and human capital in agriculture. Human capital and increased information are essential factors in increasing the adoption of technological innovation, which can lead to increased productivity in agriculture.

However, several obstacles can prevent farmers from adopting new technological innovations, as previous studies show. Credit constraint is a common hindrance that stops farmers from adopting new technologies. Previous studies show that farmers need to access sufficient amounts of credit to be able to acquire better and newer technologies. Microfinance acts as a helping hand for farmers to access new technology. Technological innovations do not increase productivity if it does not reach farmers due to various obstacles such as credit constraints. Therefore, it increases the importance of policies that enable technological innovations to reach farmers and that it is used in the right way. Previous studies show that it is not only enough for farmers to have access to technological innovations; they must know how to use them in the right way for increased productivity.

# 4. Methodology

## 4.1 Panel data

Panel data regression refers to data that varies by two factors, one is time, and the other is cross-sectional. Using a generic model for panel data to exemplify:

$$Y_{it} = \delta_i + \beta_1 X_{1,it} + \beta_2 X_{2,it} + \varepsilon_{it} \quad (1:1)$$

This thesis observes 102 countries constituting the *i* in the model above. There are 19 years of observation for every $i^{th}$ country, representing the *t* in the model, from 2000 to 2018. The power of panel data allows the observation of several time series and how they vary cross-sectionally between countries which allows for more observations to be used. (Stock & Watson, 2020)

### 4.1.1 Fixed effects

Panel data is either analyzed using random or fixed-effects models. For random-effects the cross-sectional variation is random and for fixed-effects the variation is systemic. Choosing either random or fixed effects model is therefore essential for an unbiased, causal analysis. Therefore, we perform the Hausman test to find proof of random or systematic variation. Thus, determining whether our data is best suited for random or fixed-effects models.

**Table 1: Hausman test**

| Model | P-value |
|---|---|
| $H_0$: Difference in coefficients not systematic | 0.002 |

**Source**: StataCorp. 2021
**Note**: The probability that the variation in the model is due to randomness is 0.002. The null hypothesis is therefore rejected, and a fixed effects model is preferred. Interpretation of the null and alternative hypothesis is: $H_0: Random\ effects\ models\ are\ appropriate.\quad H_A: Fixed\ effects\ models\ are\ appropriate.$

The test shows that the probability that the variation in the model is due to randomness is 0.002. This suggests that we reject the null hypothesis and say that fixed-effects models are appropriate for our analysis.

Fixed-effects regression allows for using the rate of change to control for omitted variables bias. The theory is that a variable which only varies between countries but not within country can be excluded from the analysis. Taking the rate of change excludes variables that do not

change over time. We use equation (1:1) above as an example. By changing $X_{2,it}$ to $X_{2,i}$ it is independent of time, only varying between countries.

$$Y_{it} = \delta_i + \beta_1 X_{1,it} + \beta_2 X_{2,i} + \varepsilon_{it} \rightarrow$$
$$\Delta Y_{it} = \Delta\delta_i + \Delta\beta_1 X_{1,it} + \Delta\beta_2 X_{2,i} + \Delta\varepsilon_{it}$$
$$= 0 + \Delta\beta_1 X_{1,it} + \Delta\beta_2 0 + \Delta\varepsilon_{it}$$
$$= \Delta\beta_1 X_{1,it} + \Delta\varepsilon_{it} \qquad (1:2)$$

Looking at the last equality in (1:2), the rate of change for variable $X_{2,i}$ is equal to zero since it is time-invariant. We can therefore control for unobserved omitted variables that do not change over time using the fixed effects method. Variables that change over time can still cause bias if not included in the model. (Stock & Watson, 2020)

### 4.1.2 Linear regression assumptions

All parametric tests assume that the variable distribution is a standard normal distribution. When the distribution is too dissimilar, it creates a problem for drawing accurate conclusions from performed tests. (Gujarati & Porter, 2009) Therefore tests and graphs for normal distributions will be performed.

We base the methodology of the thesis on investigating that our model satisfies the assumptions of linear regression, more specifically:

1. Linear relationship between dependent and independent variables.
2. Independent variables and the unobserved variable are uncorrelated.
3. The unobserved variable has a mean value of zero.
4. Homoscedastic residuals.
5. No autocorrelation.
6. More observations than coefficients
7. Variation in the independent variables.
8. No perfect linear relationship between independent variables, multicollinearity.
9. No specification bias.

Additionally, there is a discussion about the assumption of stationarity for panel data and time series. [1]

[1] Further explanation and reading, see Gujarati and Porter (2009)

## 4.2 Model

We build the variable selection from Mariyono (2018a), that presents a model for analyzing the causal relationship between access to microfinance and technology adoption on a micro-level. We expand upon this by investigating if there is a significant relationship on a macro level. Mariyono (2018a) argues using a structural equation model that allows for simultaneous analysis of multiple endogenous regression equations. We do not possess the ability for such an approach because it is beyond the scope of a bachelor's education level. However, we base the choice of variables on what Mariyono argues are the most important factors to explain the relationship between agricultural productivity and microfinance. This thesis uses panel data fixed effects regression to minimize endogeneity.

Our model is:

$$ceryiel_{it} = \delta_i + \beta_1 disagri_{it} + \beta_2 fertikg_{it} + \beta_3 enrlsec_{it} + \beta_4 popgrow_{it} + \varepsilon_{it} \quad (1:3)$$

Cereal yield (ceryiel) is our dependent variable which is measured in kg per hectare. Disbursement to agriculture (disagri) is our main independent variable of interest, measured in constant million, 2019 USD. Fertilizer consumption (fertikg) is an independent variable measured in kg per hectare of arable land. Another independent variable is secondary school enrollment (enrlsec), which is measured in gross percentage enrollment. The last independent variable is population growth (popgrow), measured in annual percentage growth.

Since disbursement to agriculture is our variable of interest, the null and alternative hypothesis for this thesis is:

$$H_0: \beta_1 = 0 \ (Disbursement\ to\ agriculture\ has\ no\ effect\ on\ agricultural\ productivity)$$
$$H_A: \beta_1 \neq 0 \ (Disbursement\ to\ agriculture\ has\ an\ effect\ on\ agricultural\ productivity)$$

# 5. Data

## 5.1 Data collection

We assemble our dataset from the World Bank's indicator database (World Bank, n.d.) and the UN's sub-organ Food and Agriculture Organization (FAOSTAT, n.d.). The dependent variable, disbursement to agriculture, is gathered from FAO since it was unavailable in the World Bank database. The other variables are gathered from the World Bank. The aim is to use one database to the extent it is possible. To reduce the risk differences in how actors gather information, having random effects in the dataset.

The availability of the variables forms the time range we use in the thesis. Many countries have limited records of the variables of interest pre-2000. For that reason, we use observations from 2000 to 2018 for a more balanced dataset.

Countries have been categorized by the UN's Gross National Income (GNI) per capita for 2021, which we use for the analysis to enhance the understanding at what income level disbursement to agriculture has an effect. The categories are high-income, upper-middle-income, lower-middle-income, and low-income. (UN, 2022) Full list of the countries and their respective income categories can be found in Appendix 3.

**Table 2: Descriptive statistics**

| Variable | Obs | Mean | Std. dev. | Min | Max |
|---|---|---|---|---|---|
| Log Cereal yield | 1,938 | 7.864 | 0.693 | 5.028 | 9.545 |
| Log Disbursement | 1,890 | 2.344 | 2.039 | -7.094 | 7.268 |
| Log Fertlizer | 1,933 | 3.933 | 1.831 | -8.694 | 7.821 |
| Log Secondary school | 1,833 | 4.245 | 0.552 | 1.824 | 5.099 |
| Population growth | 1,938 | 1.356 | 1.215 | -3.848 | 7.350 |

Source: StataCorp. 2021
Note: Descriptive statistics for the variables of this thesis.

## 5.2 Dependent variable

This study uses cereal yield, kg per hectare, as the dependent variable, downloaded from the World Bank indicator database. The argument for using yield instead of the total production of cereals is that when we use a relative term, we capture the productivity of the hectares used. Additionally, it is easier to compare a relative measure to another country than total

production. Yield is, therefore, better for the analysis of how microfinance affects technological adoption by farmers. Cereals and grains are some of the most produced crops globally (Statista, 2019), especially in developing countries (Grote et al., 2021), Making it suitable for this study's dependent variable of interest. The variable is transformed to a logarithmic scale to better emulate a normal distribution.

## 5.3 Independent variables

Our main independent variable of interest is disbursement to agriculture, forestry, and fishing in millions, constant 2019 US dollar prices. Disbursement, as opposed to commitment, is the paid amount countries receive from all donors, multi and bilateral. The variable is a proxy for credit received by farmers from all donors worldwide from the UN's Food and Agriculture Organization database. The variable is transformed to a logarithmic scale to better emulate a normal distribution.

Fertilizer consumption measures in kilograms per hectare of arable land. The argument for using the "per hectare" measure instead of total fertilizer is similar to yield. By having a relative measure instead of absolute consumption, comparisons are easier to make of fertilizer utilization by country or category. Fertilizer consumption is a proxy for the level of technology used by farmers. Fertilizer is collected from the World bank database and transformed into a logarithmic scale to better emulate a normal distribution.

Secondary school enrollment is the percentage of gross school enrollment. That is the number of students, regardless of age, that are undertaking a secondary school level of education, divided by the official number of students that should be undertaking that level of education. (UNESCO, 2020) Secondary school enrollment is collected from the World Bank and transformed into a logarithmic scale to de-trend and de-season the time series for the assumption of stationarity to hold. (Wooldridge, 2012)

Population growth, the annual percentage change in population, is collected from the World Bank database. It is a proxy for human capital, and as discussed in theory, different levels of population growth are associated with certain levels of economic and social development. Mariyono (2018a), among others, argues that human capital is crucial to explaining technological adoption by farmers in developing countries.

## 5.4 Data limitations

The credibility of the variables may differ between countries. A totalitarian state might be more inclined to showcase better numbers than the actual ones. We use two, credible databases as data sources to minimize the probability of this creating bias.

Documentation for the variables of choice differs between countries. An unbalanced dataset can be problematic if the reason why observations are missing is not understood. (Wooldridge, 2012) Countries with many missing values are excluded from our analysis because it reduces the test's statistical power, it can also create bias in the results (Young & Johnson, 2015). Stock and Watson (2020) explain that the basis for fixed effects regression models is balanced panel data. An unbalanced dataset can be used if the statistical software of choice can handle it, which Stata can (Torres-Reyna, 2007).

To further investigate reducing missing values, we use a linear interpolation function to fill the missing values with the weighted moving average. It is a method to use information from known values to predict a value between two observations, filling in a missing value. On the assumption that there is an ongoing trend. (The Education Life, 2019)

Mariyono (2018a) argues that other variables have a causal relationship to agricultural productivity. However, we exclude them in this thesis due to multicollinearity, poor distribution, or lack of observations.

## 5.5 Regression model analysis

From the methodology stated in 4.1.2 we test for normal distribution, the linear regression assumption and stationarity.  Assumption nr 2 is not tested since it is impossible to test the unobserved variable. Assumption nr 3 is not tested either since the unobserved variable has a mean value of zero per construction in a linear regression.

### 5.5.1 Normal distribution

We use The Shapiro-Wilks test for normality to test the overall model and the different income categories. We reject the null hypothesis; the data does not display a normal distribution.

**Table 3: Shapiro-Wilks test for normality**

| | All countries | Low income | Lower middle income | Higher middle income | High income |
|---|---|---|---|---|---|
| **Variable** | **P-value** | **P-value** | **P-value** | **P-value** | **P-value** |
| Log Cereal yield | < 0.001 | < 0.001 | 0.001 | < 0.001 | < 0.001 |
| Log Disbursement | < 0.001 | < 0.001 | < 0.001 | 0.014 | < 0.001 |
| Log Fertlizer | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |
| Log Secondary school | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |
| Population growth | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |
| $H_0$: The data is normally distributed | | | | | |

**Source**: StataCorp. 2021
**Note**: The test shows that the probability that the respective variables are normally distributed is less than 5 %. We, therefore, reject the null hypothesis that the data is normally distributed.

Based on information from Stata help pages, Shapiro-Wilks may be inaccurate at larger sample sizes. (StataCorp, 2021a) We perform a skewness and kurtosis test and reject the null hypothesis of normally distributed data for the overall model.

**Table 4: Skewness test**

| | All countries | Low income | Lower middle income | Higher middle income | High income |
|---|---|---|---|---|---|
| **Variable** | **P-value** | **P-value** | **P-value** | **P-value** | **P-value** |
| Log Cereal yield | < 0.001 | 0.001 | 0.0096 | < 0.001 | < 0.001 |
| Log Disbursement | < 0.001 | 0.006 | < 0.001 | 0.153 | < 0.001 |
| Log Fertlizer | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |
| Log Secondary school | < 0.001 | 0.001 | < 0.001 | < 0.001 | < 0.001 |
| Population growth | < 0.001 | < 0.001 | 0.1579 | < 0.001 | < 0.001 |
| $H_0$: The data is normally distributed | | | | | |

**Source**: StataCorp. 2021
**Note**: The test shows that variables are significant at a 5% level for all categories. We, therefore, reject the null hypothesis that data is normally distributed. However, the stratified results show some better behavior. We fail to reject population growth and disbursement for lower and higher middle income, respectively.

Furthermore, the skewness test shows that disbursement and population growth are non-significant for higher middle income and lower middle income, respectively, indicating that they display a normal distribution.

**Table 5: Kurtosis test**

| | All countries | Low income | Lower middle income | Higher middle income | High income |
|---|---|---|---|---|---|
| Variable | P-value | P-value | P-value | P-value | P-value |
| Log Cereal yield | 0.025 | 0.296 | 0.899 | < 0.001 | < 0.001 |
| Log Disbursement | < 0.001 | < 0.001 | 0.117 | 0.041 | < 0.001 |
| Log Fertlizer | < 0.001 | < 0.001 | 0.134 | < 0.001 | < 0.001 |
| Log Secondary school | < 0.001 | 0.015 | 0.174 | < 0.001 | < 0.001 |
| Population growth | < 0.001 | < 0.001 | 0.031 | < 0.001 | < 0.001 |
| $H_0$: The data is normally distributed | | | | | |

**Source**: StataCorp. 2021
**Note**: The test shows that variables are significant at a 5% level for all categories. We, therefore, reject the null hypothesis that data is normally distributed. The stratified results do, however, once again show some better behavior. We fail to reject cereal yield for low-income countries and all variables, except population growth, for lower-middle-income countries.

The kurtosis test shows that all the variables are significant for the overall model. On closer inspection, the lower-middle-income countries are non-significant except for population growth. Furthermore, cereal yield is non-significant for low-income levels.

We conduct a Kolmogorov-Smirnov test for additional investigation, which is more robust to large sample sizes. (StataCorp, 2021b)

**Table 6: Kolmogorov-Smirnov test for normality**

| | All countries | Low income | Lower middle income | Higher middle income | High income |
|---|---|---|---|---|---|
| Variable | P-value | P-value | P-value | P-value | P-value |
| Log Cereal yield | 0.048 | 0.025 | 0.221 | < 0.001 | 0.003 |
| Log Disbursement | 0.141 | 0.179 | 0.328 | 0.551 | 0.040 |
| Log Fertlizer | < 0.001 | 0.015 | 0.334 | 0.036 | 0.015 |
| Log Secondary school | < 0.001 | < 0.001 | < 0.001 | 0.011 | < 0.001 |
| Population growth | 0.012 | 0.001 | 0.010 | < 0.001 | < 0.001 |
| $H_0$: The data is normally distributed | | | | | |

**Source**: StataCorp. 2021
**Note**: A more robust test for larger sample sizes shows some improvement when testing for normality. There are, however, still problems regarding that all variables should be normally distributed as stated in 4.1 Panel data.

It shows that there are still problems regarding the normality assumption. There is evidence to support that the distribution of the variables is non-normal overall, but some variables behave a little better when divided into income categories.

Schmidt and Finan (2018) argue that linear regression models are robust to violations of normal distribution for large sample sizes. An ocular examination explains the tests in Appendix 2. The distributions of the variables are sufficiently similar to a normal distribution

to continue the analysis. However, we maintain some reservations about the robustness of the result.

### 5.5.2 Linear relationship assumption nr 1

One assumption of linear regression is that there should be a linear relationship between the dependent and independent variables. In Appendix 1, we plot the average of the dependent and independent variables, and there seems to be linear relationships.

### 5.5.3 Homoscedasticity assumption nr 4

Homoscedasticity is the assumption for linear regression that the variability about the mean for residuals is constant. Heteroscedasticity is the problem created by a systemic change in the variance of the residuals. (Wooldridge, 2012) The latter creates a problem for causal effects due to endogeneity problems. There is systematic variation in the unobserved variable. We conducted a Breusch–Pagan and a White test for constant variance and homoscedasticity.

**Table 7: Breuch-Pagan test**

| | All countries | Low income | Lower middle income | Higher middle income | High income |
|---|---|---|---|---|---|
| **Variable** | **P-value** | **P-value** | **P-value** | **P-value** | **P-value** |
| $H_0$: Constant variance | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |

**Source**: StataCorp. 2021
**Note**: The probability that the model has constant variance is less than 0.1%. We, therefore, reject the null hypothesis of constant variance of the residuals.

We reject the null hypothesis for the overall and stratified models for the Breusch–Pagan test; there is evidence for heteroscedastic residuals.

**Table 8: White's test**

| | All countries | Low income | Lower middle income | Higher middle income | High income |
|---|---|---|---|---|---|
| **Variable** | **P-value** | **P-value** | **P-value** | **P-value** | **P-value** |
| $H_0$: Homoscedasticity | < 0.001 | 0.001 | < 0.001 | < 0.001 | 0.244 |

**Source**: StataCorp. 2021
**Note**: The probability that the model has constant variance is less than 0.1%. We, therefore, reject the null hypothesis of homoscedastic residuals. We fail to reject the stratified result of high-income countries.

White's test is in line with the Breusch-Pagan test; the null hypothesis for the overall level is rejected, indicating heteroscedastic residuals. The test shows that the lower-income countries also display heteroscedasticity. High income countries are non-significant and interpreted as exhibiting homoscedastic residuals.

### 5.5.4 Autocorrelation assumption nr 5

Serial or auto-correlation is a phenomenon in time series, where there is a systemic correlation in the error terms over time, from year to year. (Wooldridge, 2012) Creating an endogeneity problem within each but not between countries for panel data.

**Table 9: Wooldridge test for autocorrelation**

| | All countries | Low income | Lower middle income | Higher middle income | High income |
|---|---|---|---|---|---|
| Variable | P-value | P-value | P-value | P-value | P-value |
| $H_0$: no first-order autocorrelation | 0.871 | < 0.001 | 0.604 | 0.724 | 0.320 |

**Source**: StataCorp. 2021
**Note**: The test shows no significant autocorrelation for the overall model. We fail to reject the null hypothesis. There is significant autocorrelation for the stratified low-income model. We reject the null hypothesis of no first-order autocorrelation.

We fail to reject the null hypothesis for the overall model using the Wooldridge test for autocorrelation. We can, with some certainty, say that the model does not display autocorrelation. Upon further investigation, low-income countries are significant, and we reject the null hypothesis. There is significant autocorrelation for low-income countries.

We handle the heteroscedasticity and the autocorrelation problems using clustered standard errors by country. Allowing for autocorrelated and heteroscedastic residuals within a country over time, but not between countries. This thesis applies clustered standard errors by country to all regression models.

### 5.5.5 More observations than coefficients assumption nr 6

Observing the result tables in results 6.1 and 6.2 there is evidence that the assumption of there being more observations than coefficients hold.

### 5.5.6 Variation in the independent variables assumption nr 7

Observing Appendix 4 shows that there is variation of the independent variables over time, the assumption holds.

### 5.5.7 Multicollinearity assumption nr 8

Multicollinearity may arise when a significant linear relationship exists between the independent variables. The covariance between two regressors explains too much of the same

variation in the regressand. We, therefore, present a correlation table to analyze the potential problem.

**Table 10: Correlation matrix**

| Variable | Log Cereal yield | Log Disbursement | Log Fertlizer | Log Secondary school | Population growth |
|---|---|---|---|---|---|
| Log Cereal yield | **1.000** | | | | |
| Log Disbursement | -0.016 | **1.000** | | | |
| Log Fertlizer | 0.668 | -0.056 | **1.000** | | |
| Log Secondary school | 0.591 | <0.001 | 0.643 | **1.000** | |
| Population growth | -0.433 | 0.053 | -0.359 | -0.585 | **1.000** |

**Source**: StataCorp. 2021
**Note**: Correlation matrix using all countries, the full sample. The correlation table shows that no variables defy the rule of thumb for correlation higher than $\pm 0.8$ display multicollinearity.

Gujarati and Porter (2009) suggest a rule of thumb is that a correlation higher than $\pm 0.8$ is considered to display multicollinearity. All the variables in Table 10 are within the accepted range. There is some concern for secondary school that correlates with fertilizer and population growth by ca 0.64. Therefore, further investigation is necessary, using a Variance Inflation Factor (VIF) test.

**Table 11: Variance inflation factor**

| | All countries | Low income | Lower middle income | Higher middle income | High income |
|---|---|---|---|---|---|
| **Variable** | **VIF** | **VIF** | **VIF** | **VIF** | **VIF** |
| Log Disbursement | 1.01 | 1.40 | 1.01 | 1.05 | 1.06 |
| Log Fertlizer | 1.69 | 1.22 | 1.17 | 1.05 | 1.28 |
| Log Secondary school | 2.26 | 1.46 | 1.48 | 1.20 | 1.06 |
| Population growth | 1.52 | 1.12 | 1.43 | 1.25 | 1.22 |

**Source**: StataCorp. 2021
**Note**: The Variance inflation factor table shows that no variables defy the rule of thumb for a score higher than ten to display multicollinearity.

As a rule of thumb, a score above 10 for a variable display multicollinearity. (UCLA, 2021) None of the variables score a number higher than ten. The analysis will therefore include all variables.

### 5.5.8 Specification assumption nr 9

We conduct a specification test to examine if there are any omitted variables potentially causing bias in the model.

**Table 12: Ramsey RESET test**

| Model | P-value |
|---|---|
| $H_0$: Model has no omitted variables | 0.002 |

**Source**: StataCorp. 2021
**Note**: The probability that the model has no omitted variables is less than 1%. We reject the null hypothesis and conclude that there are omitted variables.

Suggesting we reject the null hypothesis, concluding that there are omitted variables. Potentially creating bias and affecting the robustness of our result.

### 5.5.9 Stationarity

Stationarity is a property in time series data where the probability that the fluctuations of the series are stable over time (Wooldridge, 2012). Hyndman and Athanasopoulos (2018) explain that the property of a stationary time series is that it displays time invariance. In other words, if the mean and variance are constant over time. Series that exhibit trends and seasonality are non-stationary. The test used is the Dickey-Fuller: Fisher unit-root test.

**Table 13: Fisher unit-root test: Dickey-Fuller**

| | All countries | Low income | Lower middle income | Higher middle income | High income |
|---|---|---|---|---|---|
| Variable | P-value | P-value | P-value | P-value | P-value |
| Log Cereal yield | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |
| Log Disbursement | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |
| Log Fertlizer | < 0.001 | < 0.001 | 0.004 | < 0.001 | < 0.001 |
| Log Secondary school | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |
| Population growth | < 0.001 | < 0.001 | < 0.001 | < 0.001 | < 0.001 |
| $H_0$: All panels contain unit roots | | | | | |

**Source**: StataCorp. 2021
**Note**: The test shows that the panels do not contain unit roots, indicating stationary at a 1% level.

All countries and the different income levels are significant. We reject the null hypothesis that there are unit-roots present in the panels, and we have enough evidence to say that the variables are stationary.

# 6. Results

## 6.1 Overall model results

We test the overall model in steps, adding one regressor after the other. We use this method to capture the potential effect of omitted variable bias on the coefficients and the robustness of the model.

**Table 14: Regression output**

| Log Cereal yield | Model 1 | Model 2 | Model 3 | Model 4 |
|---|---|---|---|---|
| Log Disbursement | 0.039*** | 0.036*** | 0.026*** | 0.024*** |
| | (0.009) | (0.008) | (0.007) | (0.006) |
| Log Fertlizer | | 0.051*** | 0.024* | 0.021* |
| | | (0.015) | (0.013) | (0.012) |
| Log Secondary school | | | 0.315*** | 0.329*** |
| | | | (0.067) | (0.070) |
| Population growth | | | | 0.047** |
| | | | | (0.022) |
| Constant | 7.851*** | 7.585*** | 6.413*** | 6.309*** |
| | (0.023) | (0.060) | (0.279) | (0.301) |
| $R^2$-Within groups | 0.041 | 0.065 | 0.137 | 0.147 |
| Observations | 1,890 | 1,887 | 1,784 | 1,784 |

**Source**: StataCorp. 2021
**Note**: Clustered standard errors are reported in parentheses. *, **, *** indicates significance at the 90%, 95%, and 99% level. See Appendix 3 for the full list of countries and respective income categories. Within-groups-$R^2$, explain the variation in the dependent variable within the respective countries captured by the model over time.

Model 1 in Table 14 shows the fixed effects regression with clustered standard errors, only including the main variable of interest, log disbursement to agriculture. There is a significant positive relationship at the 99% level, and we reject the null hypothesis. Adding the other independent variables through model 2 - 4 shows some reduction in the main independent variable of interest, but the effect is significant through all steps. We reject the null hypothesis stated in 4.2 Model:

$$H_0: \beta_1 = 0 \ (Disbursement\ to\ agriculture\ \ has\ no\ effect\ on\ agricultural\ productivity)$$

Reservations about the robustness of the results are discussed in 6.3. The result is interpreted as an elasticity since both cereal yield and disbursement are on a logarithmic scale. Thus, a one percent increase in disbursement to agriculture leads to a 0.024 percent increase in cereal yield, ceteris paribus.

Model 2 in Table 14, fertilizer is significant at a 95% level, not including secondary school and population growth. In model 3 and 4, adding secondary school and population growth shows a positive omitted variable bias in model 2, driving the coefficient for fertilizer to be larger in absolute terms than it actually is.

Secondary school is significant at a 99% level, rejecting the null hypothesis. Interpreting the effect, a one percent increase in secondary schooling enrollment leads to a 0.329 percent increase in cereal yield, ceteris paribus.

Population growth is significant at a 95% level, rejecting the null hypothesis. Since the variable is not on a logarithmic scale, the interpretation is that a one percent increase in the annual population growth leads to a 4.7 percent increase in cereal yield, ceteris paribus.

## 6.2 Results by income category

We include a stratified fixed effects regression model containing the income categories from the UN to enhance the understanding of at what income level disbursement to agriculture has an effect.

**Table 15: Regression output - Income categories**

| Log Cereal yield | Low income | Lower middle income | Upper middle income | High income |
|---|---|---|---|---|
| Log Disbursement | 0.044* | 0.024** | 0.025*** | 0.011 |
| | (0.022) | (0.011) | (0.008) | (0.009) |
| Log Fertlizer | -0.002 | 0.008 | 0.075*** | 0.024* |
| | (0.013) | (0.018) | (0.020) | (0.054) |
| Log Secondary school | 0.216** | 0.429*** | 0.291** | 0.865** |
| | (0.089) | (0.088) | (0.130) | (0.337) |
| Population growth | -0.036 | -0.026 | 0.080** | 0.048* |
| | (0.050) | 0.065 | (0.047) | (0.024) |
| Constant | 6.456*** | 5.902*** | 6.146*** | 4.187** |
| | (0.388) | (0.413) | (0.551) | (1.540) |
| $R^2$-Within groups | 0.275 | 0.267 | 0.179 | 0.087 |
| Observations | 221 | 497 | 491 | 575 |

**Source**: StataCorp. 2021
**Note**: Clustered standard errors are reported in parentheses. *, **, *** indicates significance at the 90%, 95%, and 99% level. See Appendix 3 for the full list of countries and respective income categories. Within-groups-$R^2$, explain the variation in the dependent variable within the respective countries captured by the model over time.

We reject the null hypothesis for disbursement to agriculture for lower and upper-middle-income categories since it is significant at 95% and 99% levels, respectively. Therefore, rejecting the null hypothesis in 4.2 Model. Neither low nor high income are significant, and we fail to reject the null hypothesis in 4.2 Model.

Fertilizer is significant for the upper-middle-income category; the rest is non-significant. Secondary school enrollment is significant for all income levels. Population growth is only significant for upper-middle-income countries.

## 6.3 Reservations

The tests for normality show that the distributions are significantly non-normal. There is some evidence to support that linear regression models are robust to violations of normal

distributions given large sample sizes. However, there is some reservation about the robustness of the results.

We exclude some countries from the thesis due to missing values. This may affect the result because it does not yield a general finding but a relationship between cereal yield and disbursement to agriculture for the included countries.

The within-groups-$R^2$ is low for the model. From the specification Table 12, there are omitted variables that explain the low determination coefficient. Since we focus on causal analysis, we cannot add independent variables endlessly due to multicollinearity problems.

# 7. Analysis

The scope of the thesis is to investigate the relationship between microfinance and productivity in agriculture. Examining cereal yield from 2000-to 2018 in Appendix 4, all countries' average total factor productivity per cultivated land has increased.

## 7.1 Disbursement to Agriculture

Inspecting disbursement to agriculture in Appendix 4 over 2000-2018 indicates a positive trend, indicating more funds going to agriculture. The increasing disbursements ought to increase farmers' possibility of adopting new technologies and therefore, positively affect cereal yield. Previous studies show that farmers need to access sufficient amounts of credit to be able to acquire better and newer technologies that can lead to higher output (Osabohien et al., 2022).

The results for the overall model give evidence to support a causal relationship between higher disbursement to agriculture and higher cereal yield, kg per hectare. Since cereal yield is a productivity measurement, it is interpreted as disbursement to agriculture generating a more effective agricultural process. Consistent with Mariyonos's (2018a & 2018b) results, we find evidence that disbursement to agriculture significantly affects agricultural productivity. Disbursement acts as an enabling factor, promoting the adoption of new technology, which leads to higher productivity and profit in agriculture. It is also supported by the findings of de Janvry and Sadoulet (2002), Osabohien et al. (2022), and Peng et al. (2020).

We find evidence to support what Weber (2013) and de Janvry and Sadoulet (2021) observe: In the poorest countries disbursement and agricultural productivity show no significant relationship. We find a significant effect for lower and upper-middle-income countries. This result indicates that a certain level of development is necessary.

High cognitive load is primarily a consequence of being subject to extreme poverty. The results should display that disbursement to agriculture has a significant effect in lower-income countries to be consistent with the theory of cognitive load. Since there is no effect for low-income countries, the results do not support the theory. There may be some effect from disbursement in alleviating the negative cognitive effects of being poor. However, we

do not find evidence to support it. The lack of evidence may be because it is difficult to ascertain at a macro-level at which we investigate the problem. A smaller, micro-level study, such as Mani et al. (2013), may be more efficient at explaining the relationship between microfinance and its effect on cognitive load and technology adoption.

Weber (2013), de Janvry and Sadoulet (2021), and Nakano and Magezi (2020) suggest a mismatch exists between income from harvests and short-term microfinance loans. The design of microfinance is to have a fast repayment pace. In contrast, the time from using the acquired loan in the agricultural process and then receiving profit from selling the produce might take up to several years. The significant effect from disbursement to agriculture to cereal yield shows that agricultural productivity may gain from redesigning microfinance loans customized for farmers.

## 7.2 Fertilizer

Even though performed tests show no multicollinearity problems, there may be some multicollinearity affecting the results from fertilizer. The correlation tables in Table 7, and Appendix 5 show a high correlation between secondary school enrollment and fertilizer but not for the upper-middle-income countries, where it is close to zero. Implications for the result are that fertilizer and secondary school enrollment explain similar aspects of the effect on cereal yield. In line with the correlation tables, we see an exception for upper-middle-income countries where fertilizer has a significant effect on the cereal yield. Another possible explanation is that countries need certain factors in place in order to use fertilizer effectively. (Nakano & Magezi, 2020) For the results to support the theory, there should be a significant relationship for high-income countries, which it is not.

That fertilizers in our results show no effect except for upper-middle-income countries is in line with what Nakano and Magezi (2020) find in their results. They find no increased use of fertilizers due to the availability of microfinance for the farmers who already use a recommended level of fertilizers. However, they find an increase in the use of fertilizers for the farmers that did not use much fertilizer before increased access to credit. Nevertheless, the increase in fertilizer use does not lead to an increase in agricultural productivity. It shows the importance of understanding that there is no simple answer to how increased credit to agriculture can positively affect productivity growth. Increased access to credit needs to

complement other supportive factors, such as information and a certain level of human capital.

In induced innovation theory, technological innovations are the key drivers in lowering the cost of labor or land in agricultural productivity. There is no evidence to support that disbursement to agriculture has a more significant effect on one over the other from the results. A plausible alternative explanation is that the general increase in productivity combines labor and land saving effects.

## 7.3 Secondary school enrollment

Secondary school enrollment has a significant effect on cereal yield for the overall model and all income categories. Mariyono (2018b & 2018a), de Janvry and Sadoulet (2021), Foster and Rosenzweig (1996), and Schultz (1964) (referred to in Kaldor, 1964) all point to the importance of human capital for farmers to have access to microfinance. de Janvry & Sadoulet (2021) argues that a lack of information and education obstructs the adoption of new technologies. Foster & Rosenzweig (1996) and Theodore W. Schultz also emphasize the importance of education and its connection to technical change. Technological innovations are more likely to have a significant effect in countries with a higher level of education.

## 7.4 Population growth

There is a significant effect from population growth in the overall model, indicating an increase in cereal yield as population growth increases, all else equal, in line with the demographic transition model. As population growth increases, there is a temporary increase in the labor force, increasing the productivity of the whole population and, therefore, higher productivity in agriculture.

Upon closer inspection, population growth is only significant for upper-middle-income countries; this may be consistent with the model. Countries in the upper-middle-income category may have experienced a considerable labor force increase, increasing productivity. Observing population growth by category shows that Low and lower middle income has a higher growth rate than upper-middle and high-income. There is a long time lag for a new generation to be old enough to join the labor force and increase productivity. (de Janvry &

Sadoulet, 2021) Therefore, there may be a significant relationship in the future between population growth and lower-income countries.

## 7.5 Limitations

From the Ramsey Reset misspecification test, we conclude that our model exhibits omitted variables. Mariyono (2018a) writes that the main argument for structural equation modeling is to address models with endogeneity problems when analyzing the relationship between agricultural productivity and technology. Our goal was that by using fixed effects regression, we would eliminate some of the mentioned endogeneity. However, analyzing at a macro level for many different countries proved too much variation in the unobserved variable.

Other variables for a technological proxy might better describe the relationship between technology level and agricultural productivity. Such variables may be seeds, irrigation systems, prices on animal feeds, or agricultural machinery. Due to data limitations, we could not go forth to test those alternatives.

# 8. Conclusion

This thesis finds evidence that disbursement to agriculture leads to higher cereal yields and microfinance has an effect on agricultural productivity. However, there are some questions about the robustness of our result and potential omitted variable bias. We find little evidence that increased productivity depends on cash alleviating cognitive load when analyzing countries by income categories. We do not find evidence that fertilizer plays a significant role in explaining the cereal yield but rather that enrollment in secondary school is significant in explaining the relationship for higher productivity. The results contribute to current literature because, to the best of our knowledge, there has not been a study covering microfinance and agricultural productivity at an aggregated macro and country-level. The hope is that more extensive studies will establish a deeper and broader understanding of the subject.

An argument based on the results supports the effort to make microfinance more efficient at targeting farmers and minimizing the mismatch and time lag between harvest and repayment schedule. Weber (2013) provides evidence that microfinance has better farmer participation when the loans are more flexible and adapted to agriculture.

As discussed in methodology 4.2, Mariyono (2018a) used Structural Equation Modeling to better handle endogenous models and used Structural Equation Modeling to explain the causal relationship between microfinance and agricultural productivity. Using Structural Equation Modeling would be interesting to investigate how such a model would transfer from a micro-level, as Mariyono uses it, to a macro-level. In theory, it would allow us to use multiple regression models and, therefore, more variables than was possible for this fixed effects panel data regression. Then more omitted variables would be accounted for to establish a stronger causal relationship.

In conclusion, developing countries have a larger share of their economies from agriculture. Increasing productivity in agriculture should be an essential aspect of policies to reduce poverty. This thesis finds evidence supporting that credit disbursement to agriculture has a positive effect on cereal yield. Therefore, more effort to increase the efficiency of microfinance to agriculture is desirable.

# 9. References


Ankrah Twumasi, M., Jiang, Y., Ntiamoah, E. B., Akaba, S., Darfor, K. N., & Boateng, L. K. (2022). Access to credit and farmland abandonment nexus: The case of rural Ghana. *Natural Resources Forum*, *46*(1), 3–20. https://doi.org/10.1111/1477-8947.12233

Atteridge, A., Savvidou, G., Sadowski, S., Gortana, F., Meintrup, L., & Dzebo, A. (2019). *All Donors to All Recipients for Agriculture, Forestry, Fishing during 2002–2019*. Aid Atlas. https://aid-atlas.org (Accessed: 12 May 2022)

Alimukhamedova, N. (2019). The microfinance promise—can it be kept? A macro perspective. *Development Policy Review*, *37*(6), 812–842. https://doi.org/10.1111/dpr.12412

Athanasopoulos, G., & Hyndman, R.J. (2018) *Forecasting: principles and practice*, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2. https://otexts.com/fpp2/stationarity.html#fnref14 (Accessed: 1 May 2022)

Bruns, S., Dalheimer, B., & Musshoff, O. (2022). The effect of cognitive function on the poor's economic performance: Evidence from Cambodian smallholder farmers. *Agricultural Economics*, *53*(3), 468–480. https://doi.org/10.1111/agec.12686

de Janvry, A., & Sadoulet, E. (2002). World Poverty and the Role of Agricultural Technology: Direct and Indirect Effects. *Journal of Development Studies*, *38*(4), 1–26. https://doi.org/10.1080/00220380412331322401

de Janvry, A., & Sadoulet, E. (2021). "Development Economics". Second edition. New York: Routledge.

Fehr, E., & Haushofer, J. (2014). On the psychology of poverty. *Science*, *344*(6186), 862–867. https://doi.org/10.1126/science.1232491

Food and Agriculture Organization. (n.d.). *FAOSTAT*. https://www.fao.org/faostat/en/#data/IC (Accessed: 13 April 2022).

Foster, A. D., & Rosenzweig, M. R. (1996). Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution. *The American Economic Review*, *86*(4), 931–953.

Grote, U., Fasse, A., Nguyen, T. T., & Erenstein, O. (2021). Food Security and the Dynamics of Wheat and Maize Value Chains in Africa and Asia. *Frontiers in Sustainable Food Systems*, *4*. https://www.frontiersin.org/article/10.3389/fsufs.2020.617009 (Accessed: 24 April 2022).

Gujarati, D., & Porter, D. (2009). *Basic econometrics*. 5th edition. Boston: McGraw-Hill.

Hazell, P., & Herdt, R. (1987). Agricultural development: An international perspective. Journal of Development Economics, 26(1), 197-200. doi:10.1016/0304-3878(87)90069-1

Islam, K., & O'Gorman, M. (2019). Microcredit contract design: A macroeconomic evaluation. *World Development*, *124*, 104634. https://doi.org/10.1016/j.worlddev.2019.104634

Kaldor, D. R. (1964). Agricultural Economics [Review of *Transforming Traditional Agriculture*, by T. W. Schultz]. *Science*, *144*(3619), 688–689. http://www.jstor.org/stable/1712902

Khandker, S. R., & Koolwal, G. B. (2016). How has microcredit supported agriculture? Evidence using panel data from Bangladesh. *Agricultural Economics*, *47*(2), 157–168. https://doi.org/10.1111/agec.12185

Köhn, D. (Ed.). (2013). *Microfinance 3.0*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-41704-7

The Education Life. (2019, September 19). *Linear Interpolation Formula.* https://theeducationlife.com/interpolation-formula/ (Accessed: 25 May 2022)

Mani, A., Mullainathan, S., Shafir, E., & Zhao, J. (2013). Poverty Impedes Cognitive Function. *Science*, *341*(6149), 976–980. https://doi.org/10.1126/science.1238041

Mariyono, J. (2018a). Microcredit and technology adoption: Sustained pathways to improve farmers' prosperity in Indonesia. *Agricultural Finance Review*, *79*(1), 85–106. https://doi.org/10.1108/AFR-05-2017-0033

Mariyono, J. (2018b). Micro-credit as catalyst for improving rural livelihoods through agribusiness sector in Indonesia. *Journal of Entrepreneurship in Emerging Economies*, *11*(1), 98–121. https://doi.org/10.1108/JEEE-06-2017-0046

Statista. (n.d.). *Most produced food commodities worldwide 2019*. https://www.statista.com/statistics/1003455/most-produced-crops-and-livestock-products-worldwide/. (Accessed: 3 May 2022).

Nakano, Y., & Magezi, E. F. (2020). The impact of microcredit on agricultural technology adoption and productivity: Evidence from randomized control trial in Tanzania. *World Development*, *133*, 104997. https://doi.org/10.1016/j.worlddev.2020.104997

Osabohien, R., Mordi, A., & Ogundipe, A. (2022). Access to credit and agricultural sector performance in Nigeria. *African Journal of Science, Technology, Innovation and Development*, *14*(1), 247–255. https://doi.org/10.1080/20421338.2020.1799537

Peng, Y., Latief, R., & Zhou, Y. (2021). The Relationship between Agricultural Credit, Regional Agricultural Growth, and Economic Development: The Role of Rural Commercial Banks in Jiangsu, China. *Emerging Markets Finance and Trade*, *57*(7), 1878–1889. https://doi.org/10.1080/1540496X.2020.1829408

Schmidt, A. F., & Finan, C. (2018). Linear regression and the normality assumption. *Journal of Clinical Epidemiology*, *98*, 146–151. https://doi.org/10.1016/j.jclinepi.2017.12.006

StataCorp. 2021a. *Stata Statistical Software: Release 17*. College Station, TX: StataCorp LLC.https://www.stata.com/manuals13/rswilk.pdf. (Accessed: 1 May 2022)

StataCorp. 2021b. *Stata Statistical Software: Release 17*. College Station, TX: StataCorp https://www.stata.com/manuals/rksmirnov.pdf (Accessed: 1 May 2022)

Stock, J.H. & Watson, M.W. (2020). *Introduction to Econometrics*. 4th global edition. Harlow: Pearson.

Torres-Reyna, O. (2007). "Panel Data Analysis Fixed and Random Effects using Stata". Princeton university. December.

UCLA. (n.d.). *Regression with Stata Chapter 2 – Regression Diagnostics*. https://stats.oarc.ucla.edu/stata/webbooks/reg/chapter2/stata-webbooksregressionwith-statachapter-2-regression-diagnostics/ (Accessed: 1 May 2022).

Unesco. (2020, June 22). *Gross enrolment ratio*. http://uis.unesco.org/en/glossary-term/gross-enrolment-ratio (Accessed: 25 April 2022)

United Nations. (2022). World Economic Situation and Prospects 2022. *WORLD ECONOMIC SITUATION AND PROSPECTS, 37*. Doi:https://www.un.org/development/desa/dpad/wp-content/uploads/sites/45/WESP2022_ANNEX.pdf.

Weber, R. (2013). Microfinance Beyond the Standard? Evaluating Adequacy and Performance of Agricultural Microcredit. In D. Köhn (Ed.), *Microfinance 3.0: Reconciling Sustainability with Social Outreach and Responsible Delivery* (pp. 139–153). Springer. https://doi.org/10.1007/978-3-642-41704-7_7

Wooldridge, J.M. (2012). *Introductory Econometrics: A Modern Approach*. Fifth Edition. Mason: Cengage learning.

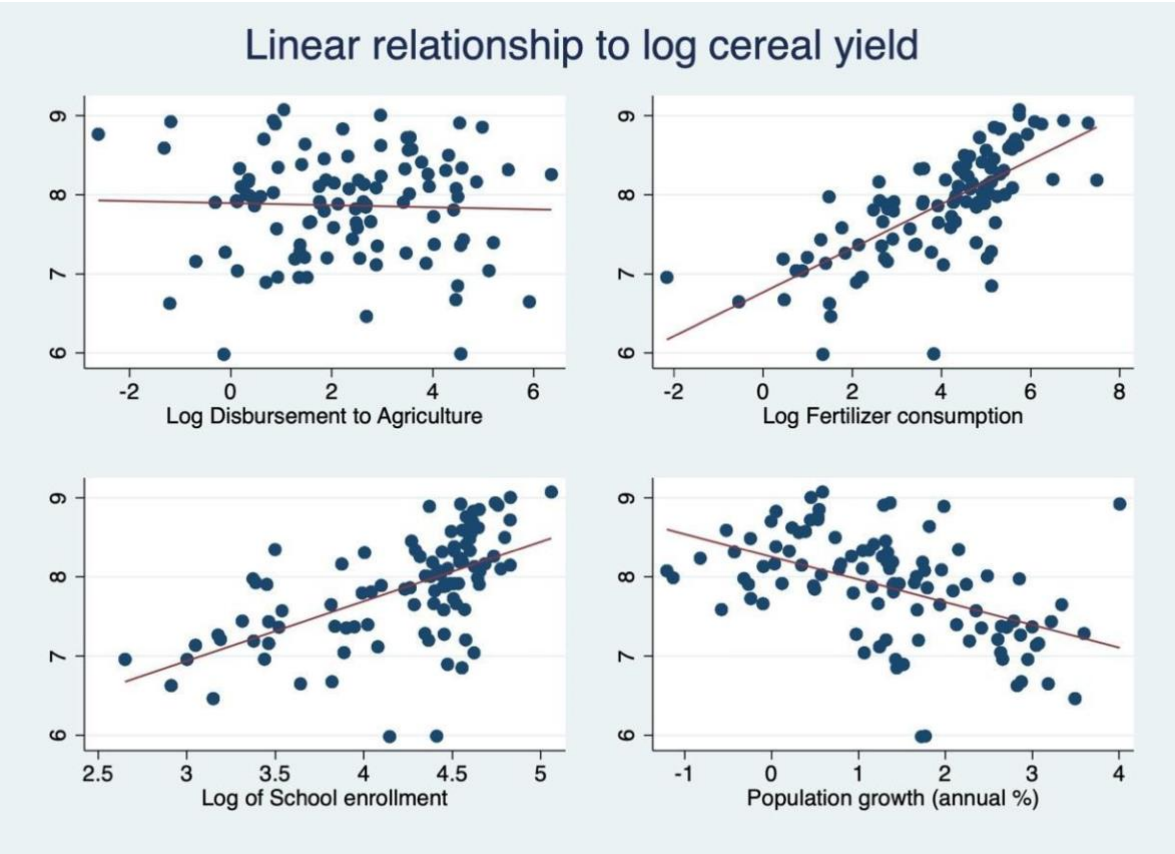
World Bank. (n.d.). *World Development Indicators | DataBank* . Available at: https://databank.worldbank.org/reports.aspx?source=2&country=VNM (Accessed: 13 April 2022).

World Bank. (2008). *World Development Report 2008: Agriculture for Development*. World Bank. https://doi.org/10.1596/978-0-8213-6807-7

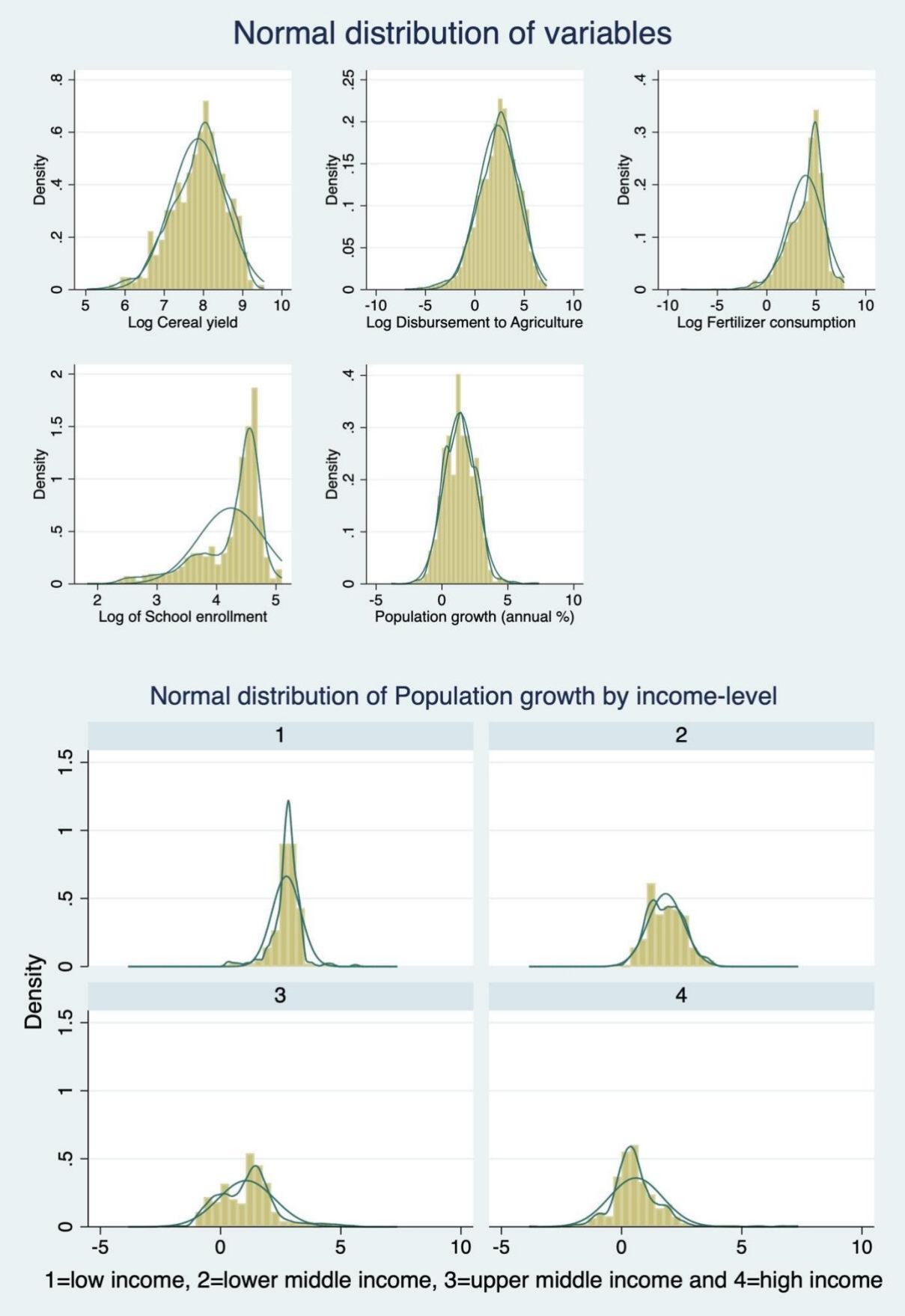
Young, R., & Johnson, D. R. (2015). Handling Missing Values in Longitudinal Panel Data With Multiple Imputation. *Journal of Marriage and the Family*, *77*(1), 277–294. https://doi.org/10.1111/jomf.12144

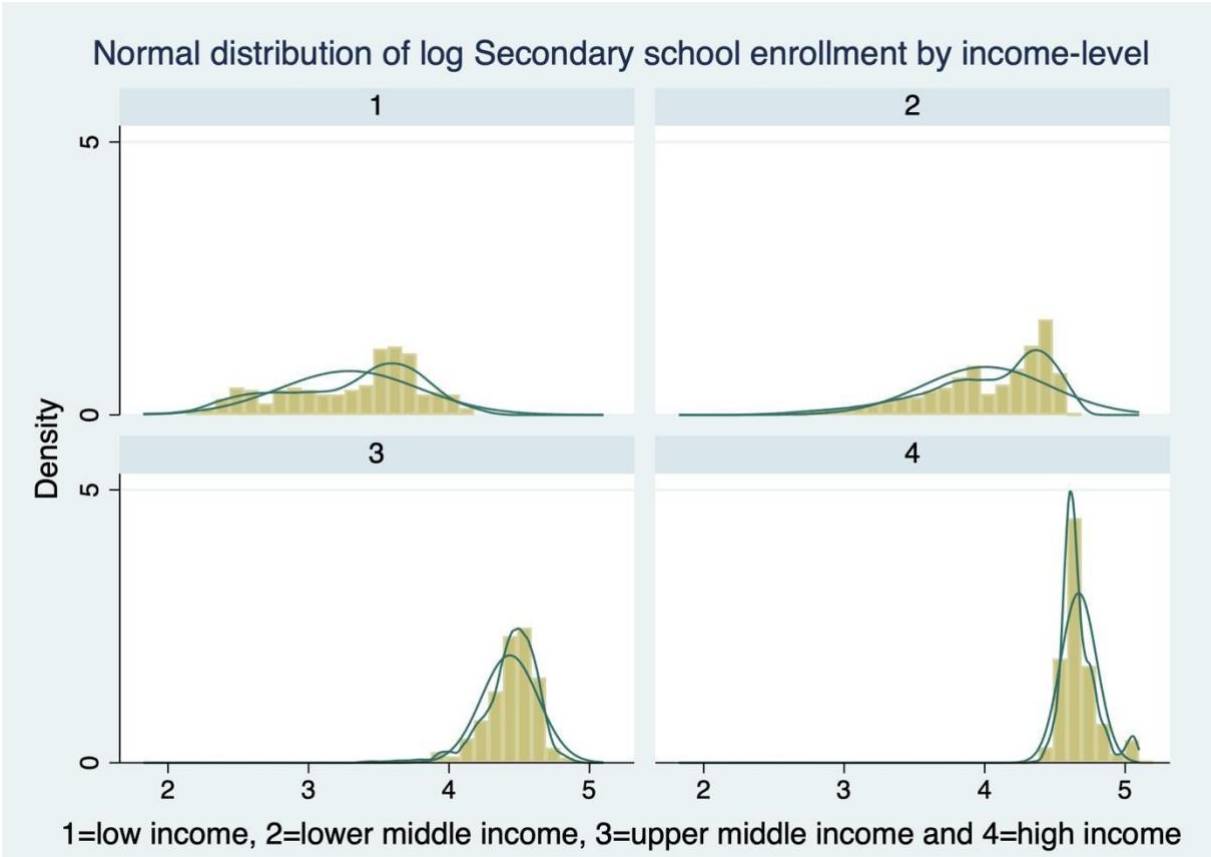
# 10. Appendix

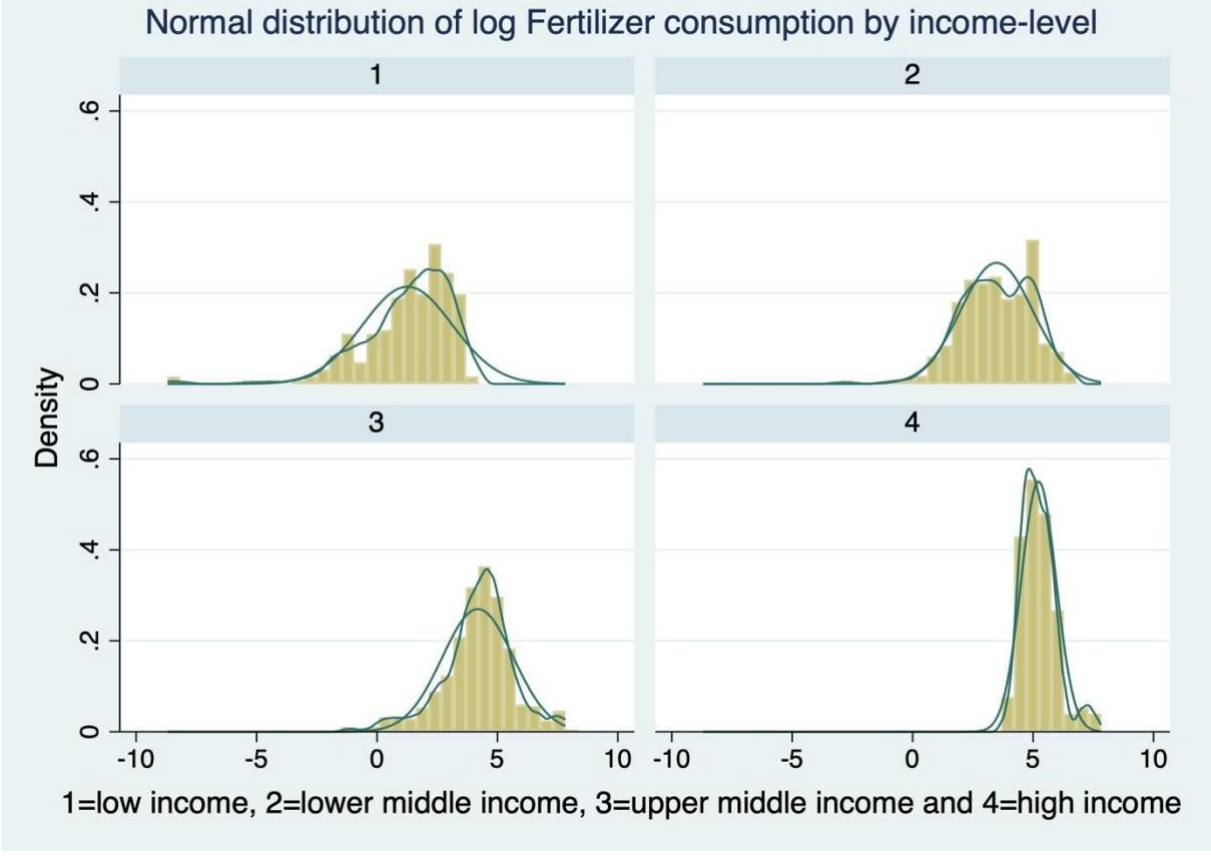
## Appendix 1: Linear relationship

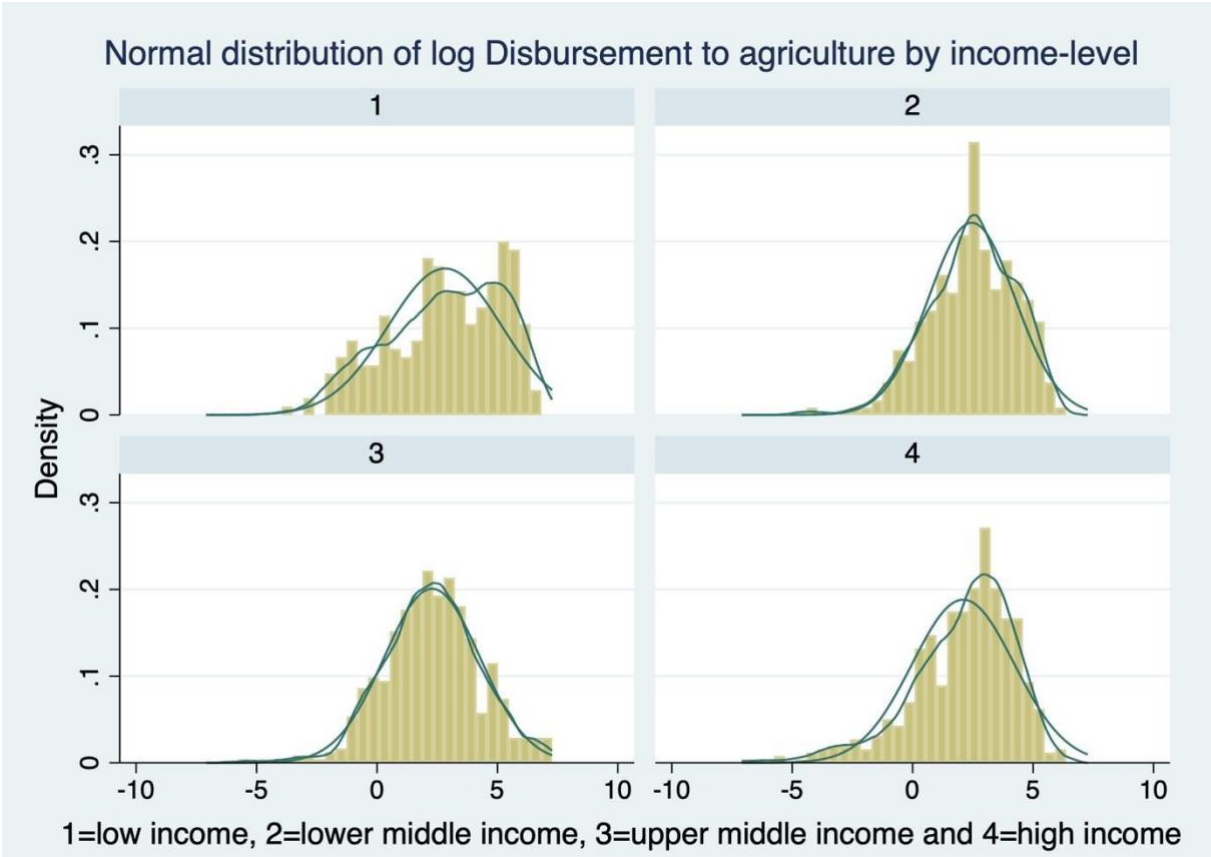


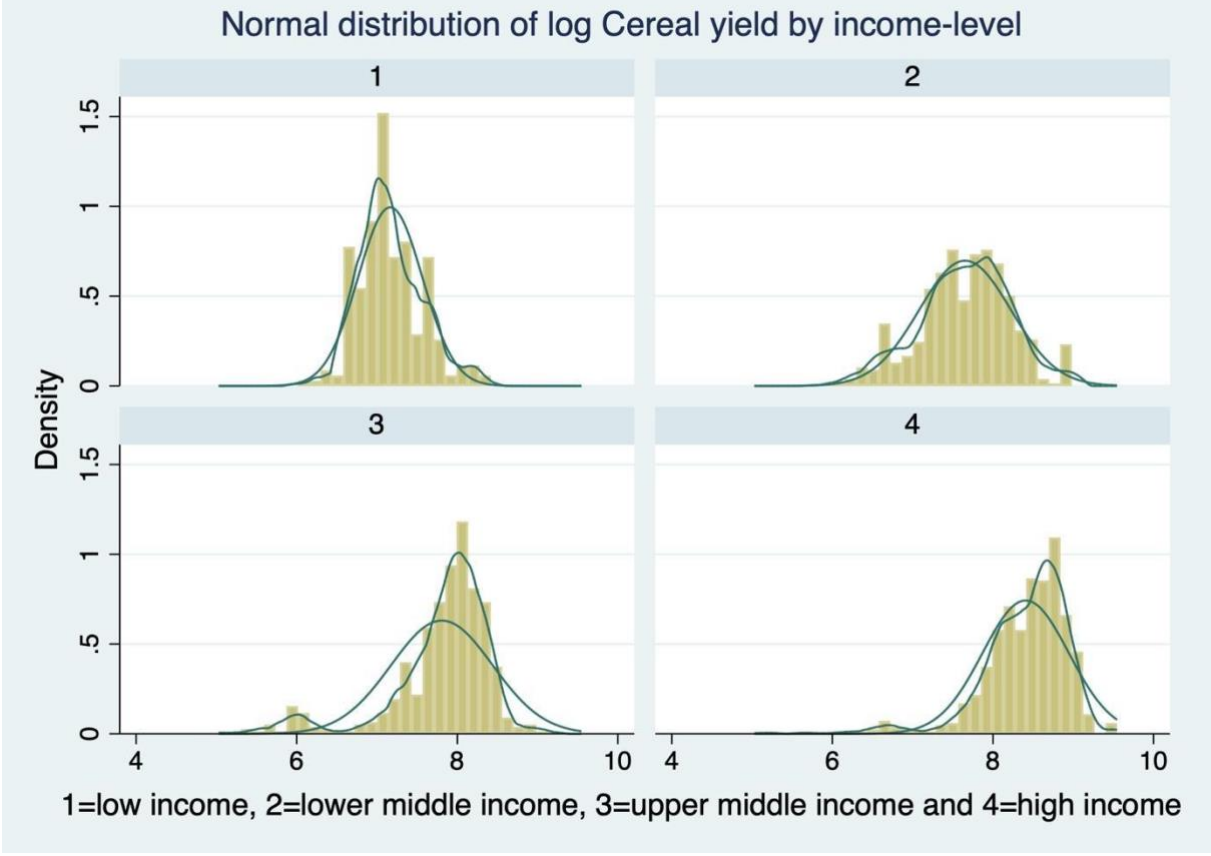
## Appendix 2: Normal distribution




Normal distribution of log Secondary school enrollment by income-level
1
2
3
4
Density
0
5
2
3
4
5
1=low income, 2=lower middle income, 3=upper middle income and 4=high income


Normal distribution of log Fertilizer consumption by income-level
1
2
3
4
Density
0
.2
.4
.6
-10
-5
0
5
10
1=low income, 2=lower middle income, 3=upper middle income and 4=high income

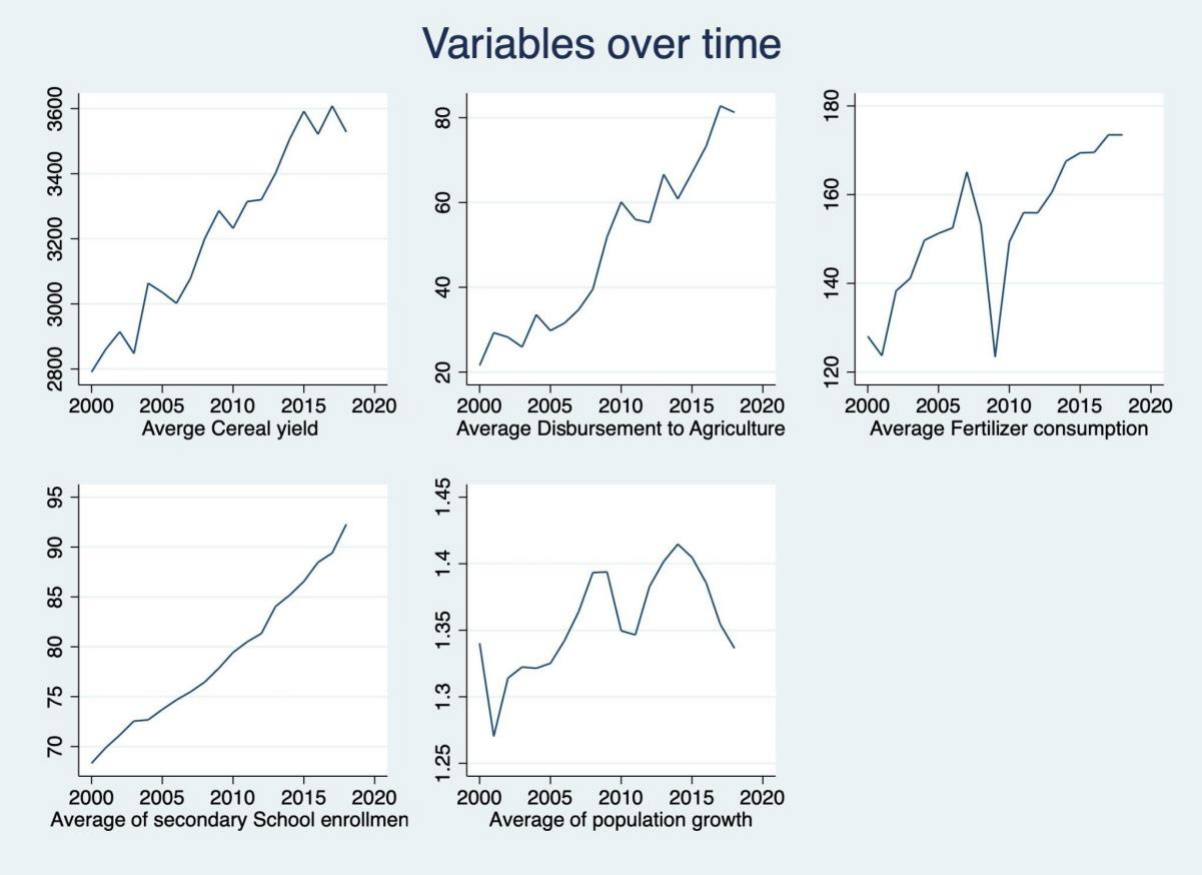

Normal distribution of log Disbursement to agriculture by income-level
1
2
3
4
Density
0
.1
.2
.3
-10
-5
0
5
10
1=low income, 2=lower middle income, 3=upper middle income and 4=high income


Normal distribution of log Cereal yield by income-level
1
2
3
4
Density
0
.5
1
1.5
4
6
8
10
1=low income, 2=lower middle income, 3=upper middle income and 4=high income

# Appendix 3: Country and income level

| Country | Country | Low income | Lower middle income | Upper middle income | High income |
|---|---|---|---|---|---|
| Afghanistan | Ireland | Afghanistan | Algeria | Albania | Austria |
| Albania | Israel | Burkina Faso | Angola | Argentina | Belgium |
| Algeria | Italy | Burundi | Bangladesh | Armenia | Brunei Darussalam |
| Angola | Japan | Central African Republic | Belize | Azerbaijan | Canada |
| Argentina | Jordan | Congo, Dem. Rep. | Bhutan | Belarus | Croatia |
| Armenia | Kazakhstan | Ethiopia | Bolivia | Botswana | Cyprus |
| Austria | Kenya | Guinea | Cambodia | Bulgaria | Denmark |
| Azerbaijan | Korea, Rep. | Madagascar | Congo, Rep. | Costa Rica | Estonia |
| Bangladesh | Kyrgyz Republic | Malawi | Cote d'Ivoire | Cuba | Finland |
| Belarus | Latvia | Mali | Egypt, Arab Rep. | Dominican Republic | France |
| Belgium | Lithuania | Mozambique | El Salvador | Ecuador | Germany |
| Belize | Luxembourg | Rwanda | Ghana | Fiji | Hungary |
| Bhutan | Madagascar | Togo | Honduras | Gabon | Ireland |
| Bolivia | Malawi | | India | Georgia | Israel |
| Botswana | Malaysia | | Indonesia | Guatemala | Italy |
| Brunei Darussalam | Maldives | | Iran, Islamic Rep. | Guyana | Japan |
| Bulgaria | Mali | | Kenya | Jordan | Korea, Rep. |
| Burkina Faso | Malta | | Kyrgyz Republic | Kazakhstan | Latvia |
| Burundi | Mauritius | | Mongolia | Malaysia | Lithuania |
| Cambodia | Mexico | | Morocco | Maldives | Luxembourg |
| Canada | Mongolia | | Myanmar | Mauritius | Malta |
| Central African Republic | Morocco | | Nepal | Mexico | Netherlands |
| Congo, Dem. Rep. | Mozambique | | Pakistan | Namibia | New Zealand |
| Congo, Rep. | Myanmar | | Papua New Guinea | Panama | Norway |
| Costa Rica | Namibia | | Philippines | Russian Federation | Oman |
| Cote d'Ivoire | Nepal | | Senegal | South Africa | Poland |
| Croatia | Netherlands | | Tajikistan | Suriname | Portugal |
| Cuba | New Zealand | | Tanzania | Thailand | Slovak Republic |
| Cyprus | Norway | | Tunisia | Turkey | Slovenia |
| Denmark | Oman | | | | Spain |
| Dominican Republic | Pakistan | | | | Sweden |
| Ecuador | Panama | | | | |
| Egypt, Arab Rep. | Papua New Guinea | | | | |
| El Salvador | Philippines | | | | |
| Estonia | Poland | | | | |
| Ethiopia | Portugal | | | | |
| Fiji | Russian Federation | | | | |
| Finland | Rwanda | | | | |
| France | Senegal | | | | |
| Gabon | Slovak Republic | | | | |
| Georgia | Slovenia | | | | |
| Germany | South Africa | | | | |
| Ghana | Spain | | | | |
| Guatemala | Suriname | | | | |
| Guinea | Sweden | | | | |
| Guyana | Tajikistan | | | | |
| Honduras | Tanzania | | | | |
| Hungary | Thailand | | | | |
| India | Togo | | | | |
| Indonesia | Tunisia | | | | |
| Iran, Islamic Rep. | Turkey | | | | |
| | Total = 102 | Total = 13 | Total = 29 | Total = 29 | Total = 31 |

## Appendix 4: Variables over time



# Appendix 5: Correlation table

## Low income

| Variable | Log Cereal yield | Log Disbursement | Log Fertlizer | Log Secondary school | Population growth |
|---|---|---|---|---|---|
| Log Cereal yield | **1.000** | | | | |
| Log Disbursement | 0.259 | **1.000** | | | |
| Log Fertlizer | 0.301 | -0.055 | **1.000** | | |
| Log Secondary school | 0.320 | 0.493 | 0.297 | **1.000** | |
| Population growth | -0.012 | 0.136 | 0.313 | 0.125 | **1.000** |

## Lower middle income

| Variable | Log Cereal yield | Log Disbursement | Log Fertlizer | Log Secondary school | Population growth |
|---|---|---|---|---|---|
| Log Cereal yield | **1.000** | | | | |
| Log Disbursement | -0.042 | **1.000** | | | |
| Log Fertlizer | 0.659 | -0.010 | **1.000** | | |
| Log Secondary school | 0.327 | 0.055 | 0.342 | **1.000** | |
| Population growth | -0.396 | 0.085 | -0.316 | -0.533 | **1.000** |

## Upper middle income

| Variable | Log Cereal yield | Log Disbursement | Log Fertlizer | Log Secondary school | Population growth |
|---|---|---|---|---|---|
| Log Cereal yield | **1.000** | | | | |
| Log Disbursement | 0.224 | **1.000** | | | |
| Log Fertlizer | 0.499 | 0.192 | **1.000** | | |
| Log Secondary school | 0.212 | 0.004 | -0.031 | **1.000** | |
| Population growth | -0.287 | -0.149 | 0.020 | -0.414 | **1.000** |

## High income

| Variable | Log Cereal yield | Log Disbursement | Log Fertlizer | Log Secondary school | Population growth |
|---|---|---|---|---|---|
| Log Cereal yield | **1.000** | | | | |
| Log Disbursement | -0.103 | **1.000** | | | |
| Log Fertlizer | 0.430 | -0.139 | **1.000** | | |
| Log Secondary school | 0.299 | 0.140 | 0.144 | **1.000** | |
| Population growth | 0.093 | 0.022 | 0.406 | -0.018 | **1.000** |

# Appendix 6: Stata syntax

```
/* Bachelor Thesis
Name: Albert Wallgren and Simon Andersson
Content: Bachelor Thesis: Data management
*/
// Setting working-directory
cd /Users/albertwallgren/Documents/BachelorThesis-Data
// Clearing memory, any priviously used data.
clear all
// set maxvar
set maxvar 32767
// Selecting the dataset i wan't to use
import excel "/Users/albertwallgren/Documents/BachelorThesis-Data/developing  ALL countries.xlsx",
sheet("Data") firstrow case(lower) clear
// By the same logic as above i close any previously used logs.
capture log close
log using BachelorThesisdofileWW, replace
************************** Working space
// compress for ease of use.
compress
// Destring all variables to use them for analysis.
destring time cerealyieldkgperhectarea fertilizerconsumptionkilogram fertilizerconsumptionoffer
schoolenrollmentsecondary schoolenrollmentprimarygr schoolenrollmenttertiaryg lifeexpectancyatbirthtotal
mortalityrateinfantper100 agedependencyratioofworki gdppercapitaconstant2015us gdppercapitaconstantlcun
cerealproductionmetrictons populationtotalsppoptotl depthofcreditinformationinde gdppercapitagrowthannual
populationgrowthannualsp s schoolenrollmentprimaryne , replace force
// Encode countryname from string to float and drop the stringversion.
encode countryname, gen(country)
drop countryname timecode countrycode
// Rename variablename that was wrongfully imported and compress names for ease of use.
rename cerealyieldkgperhectarea ceryiel
rename fertilizerconsumptionkilogram fertikg
rename fertilizerconsumptionoffer fertcon
rename schoolenrollmentsecondary enrlsec
rename schoolenrollmentprimarygr enrlpri
rename s enrlprin
rename schoolenrollmentprimaryne enrlsecn
rename schoolenrollmenttertiaryg enrlter
rename lifeexpectancyatbirthtotal lifeexp
rename mortalityrateinfantper100 infmort
rename agedependencyratioofworki agedepd
rename gdppercapitaconstant2015us gdppcdo
rename gdppercapitaconstantlcun gdppclc
rename cerealproductionmetrictons cerprod
rename populationtotalsppoptotl poptotl
rename depthofcreditinformationinde credinf
rename gdppercapitagrowthannual gdpgrow
rename populationgrowthannualsp popgrow
rename time year
```

```
// Sort to view it as paneldata.
sort country year
save "BachelorThesisdatasettomergeWW", replace
// Load and fix disbursement dataset
clear all
import excel "/Users/albertwallgren/Documents/BachelorThesis-Data/Disbursement Whole world .xls",
sheet("Sheet1") firstrow case(lower) clear
destring year value , replace force
rename value disagri
encode recipientcountry, gen(country)
drop domaincode domain donorcodefao donor recipientcountrycodefao recipientcountry elementcode element
itemcode item purposecode purpose yearcode unit flag flagdescription note
sort country year
save "disbursementWW.dta", replace
// Load master dataset and merge disbursement dataset.
clear all
use "BachelorThesisdatasettomergeWW"
sort country year
merge country using "disbursementWW.dta"
drop _merge
label variable disagri "Disbursement to Agriculture, forestry, fishing (Value US$, 2019 prices,millions)"
compress
xtset country year
// Interpolate missing values
bysort country: ipolate ceryiel year, gen(iceryiel)
bysort country: ipolate disagri year, gen(idisagri)
bysort country: ipolate fertikg year, gen(ifertikg)
bysort country: ipolate enrlsec year, gen(ienrlsec)
bysort country: ipolate popgrow year, gen(ipopgrow)
label variable iceryiel "Cereal yield (kg per hectare)"
label variable idisagri "Disbursement to Agriculture, forestry, fishing (Value US$, 2019 prices,millions)"
label variable ifertikg "Fertilizer consumption (kilograms per hectare of arable land)"
label variable ienrlsec "School enrollment, secondary (% gross)"
label variable ipopgrow "Population growth (annual %)"
drop enrlpri enrlter lifeexp infmort agedepd gdppcdo gdppclc cerprod poptotl credinf enrlprin enrlsecn gdpgrow
// Categorization variable of countries
// Generate 1 for low income
gen catgory = 1 if country == 1
replace catgory = 1 if country == 31
replace catgory = 1 if country == 32
replace catgory = 1 if country == 38
replace catgory = 1 if country == 45
replace catgory = 1 if country == 65
replace catgory = 1 if country == 82
replace catgory = 1 if country == 119
replace catgory = 1 if country == 120
replace catgory = 1 if country == 123
replace catgory = 1 if country == 135
```

```
replace catgory = 1 if country == 163
replace catgory = 1 if country == 195
// Generate 2 for lower middle income
replace catgory = 2 if country == 3
replace catgory = 2 if country == 6
replace catgory = 2 if country == 16
replace catgory = 2 if country == 20
replace catgory = 2 if country == 23
replace catgory = 2 if country == 24
replace catgory = 2 if country == 34
replace catgory = 2 if country == 46
replace catgory = 2 if country == 48
replace catgory = 2 if country == 59
replace catgory = 2 if country == 60
replace catgory = 2 if country == 75
replace catgory = 2 if country == 86
replace catgory = 2 if country == 90
replace catgory = 2 if country == 91
replace catgory = 2 if country == 92
replace catgory = 2 if country == 102
replace catgory = 2 if country == 108
replace catgory = 2 if country == 132
replace catgory = 2 if country == 134
replace catgory = 2 if country == 136
replace catgory = 2 if country == 139
replace catgory = 2 if country == 150
replace catgory = 2 if country == 153
replace catgory = 2 if country == 156
replace catgory = 2 if country == 168
replace catgory = 2 if country == 191
replace catgory = 2 if country == 198
replace catgory = 2 if country == 192
// Generate 3 for upper middle income
replace catgory = 3 if country == 2
replace catgory = 3 if country == 8
replace catgory = 3 if country == 9
replace catgory = 3 if country == 13
replace catgory = 3 if country == 18
replace catgory = 3 if country == 26
replace catgory = 3 if country == 30
replace catgory = 3 if country == 47
replace catgory = 3 if country == 50
replace catgory = 3 if country == 57
replace catgory = 3 if country == 58
replace catgory = 3 if country == 67
replace catgory = 3 if country == 71
replace catgory = 3 if country == 73
replace catgory = 3 if country == 81
```

```
replace catgory = 3 if country == 84
replace catgory = 3 if country == 100
replace catgory = 3 if country == 101
replace catgory = 3 if country == 121
replace catgory = 3 if country == 122
replace catgory = 3 if country == 127
replace catgory = 3 if country == 128
replace catgory = 3 if country == 137
replace catgory = 3 if country == 152
replace catgory = 3 if country == 162
replace catgory = 3 if country == 178
replace catgory = 3 if country == 187
replace catgory = 3 if country == 193
replace catgory = 3 if country == 199
// Generate 4 for high income
replace catgory = 4 if country == 12
replace catgory = 4 if country == 19
replace catgory = 4 if country == 29
replace catgory = 4 if country == 36
replace catgory = 4 if country == 49
replace catgory = 4 if country == 52
replace catgory = 4 if country == 54
replace catgory = 4 if country == 63
replace catgory = 4 if country == 68
replace catgory = 4 if country == 69
replace catgory = 4 if country == 74
replace catgory = 4 if country == 88
replace catgory = 4 if country == 94
replace catgory = 4 if country == 96
replace catgory = 4 if country == 97
replace catgory = 4 if country == 99
replace catgory = 4 if country == 110
replace catgory = 4 if country == 116
replace catgory = 4 if country == 117
replace catgory = 4 if country == 124
replace catgory = 4 if country == 140
replace catgory = 4 if country == 142
replace catgory = 4 if country == 148
replace catgory = 4 if country == 149
replace catgory = 4 if country == 157
replace catgory = 4 if country == 158
replace catgory = 4 if country == 105
replace catgory = 4 if country == 174
replace catgory = 4 if country == 175
replace catgory = 4 if country == 180
replace catgory = 4 if country == 188
label variable catgory "1=low income, 2=lower middle income, 3=upper middle income and 4=high income"
// Drop countries, variables and set time range
```

```
drop if year < 2000
drop if year == 2019
keep if (catgory ==1|catgory ==2 |catgory ==3 |catgory ==4)
drop ceryiel fertcon fertikg enrlsec popgrow disagri
gen liceryiel = log(iceryiel)
gen lidisagri = log(idisagri)
gen lifertikg = log(ifertikg)
gen lienrlsec = log(ienrlsec)
label variable liceryiel "Log Cereal yield (kg per hectare)"
label variable lidisagri "Log Disbursement to Agriculture, forestry, fishing (Value US$, 2019 prices,millions)"
label variable lifertikg "Log Fertilizer consumption (kilograms per hectare of arable land)"
label variable lienrlsec "Log of secondary School enrollment, (% gross)"
compress
xtset country year
// Tests on our data
// Variables over time
egen aveiceryiel = mean(iceryiel), by(year)
egen aveidisagri = mean(idisagri), by(year)
egen aveifertikg = mean(ifertikg), by(year)
egen aveienrlsec = mean(ienrlsec), by(year)
egen aveipopgrow = mean(ipopgrow), by(year)
label variable aveiceryiel "Average Cereal yield (kg per hectare)"
label variable aveidisagri "Disbursement to Agriculture (2019 million USD)"
label variable aveifertikg "Average Fertilizer consumption"
label variable aveienrlsec "Average of secondary School enrollment"
label variable aveipopgrow "Average of population growth"
sort year
twoway (line aveiceryiel year), name(overtime1, replace) legend(off) xtitle("Averge Cereal yield ") ytitle("")
twoway (line aveidisagri year) , name(overtime2, replace) legend(off) xtitle("Average Disbursement to
Agriculture") ytitle("")
twoway (line aveifertikg year) , name(overtime3, replace) legend(off) xtitle("Average Fertilizer consumption")
ytitle("")
twoway (line aveienrlsec year) , name(overtime4, replace) legend(off) xtitle("Average of secondary School
enrollment") ytitle("")
twoway (line aveipopgrow year) , name(overtime5, replace) legend(off) xtitle("Average of population growth")
ytitle("")
graph combine overtime1 overtime2 overtime3 overtime4 overtime5 , title("Variables over time")
//linear relationship
egen avliceryiel = mean(liceryiel), by(country)
egen avlidisagri = mean(lidisagri), by(country)
egen avlifertikg = mean(lifertikg), by(country)
egen avlienrlsec = mean(lienrlsec), by(country)
egen avipopgrow = mean(ipopgrow), by(country)
label variable avliceryiel "Averge Log Cereal yield (kg per hectare)"
label variable avlidisagri "Average Log Disbursement to Agriculture"
label variable avlifertikg "Average Log Fertilizer consumption"
label variable avlienrlsec "Average Log of secondary School enrollment"
label variable avipopgrow "Average Population growth (annual %)"
```

```
twoway (scatter avliceryiel avlidisagri) (lfit avliceryiel avlidisagri), legend(off) name(linear1, replace)
xtitle("Log Disbursement to Agriculture")
twoway (scatter avliceryiel avlifertikg) (lfit avliceryiel avlifertikg),legend(off) name(linear2, replace)
xtitle("Log Fertilizer consumption")
twoway (scatter avliceryiel avlienrlsec) (lfit avliceryiel avlienrlsec),legend(off) name(linear3, replace)
xtitle("Log of School enrollment")
twoway (scatter avliceryiel avipopgrow) (lfit avliceryiel avipopgrow),legend(off) name(linear4, replace)
xtitle("Population growth (annual %)")
graph combine linear1 linear2 linear3 linear4, title("Linear relationship")
// Correlation
pwcorr liceryiel lidisagri lifertikg lienrlsec ipopgrow
by catgory, sort : pwcorr liceryiel lidisagri lifertikg lienrlsec ipopgrow
// Descriptives
sum liceryiel lidisagri lifertikg lienrlsec ipopgrow
// Normality
hist liceryiel, kdensity normal xtitle("Log Cereal yield") name(graph1,replace)
hist lidisagri, kdensity normal xtitle("Log Disbursement to Agriculture") name(graph2,replace)
hist lifertikg, kdensity normal xtitle("Log Fertilizer consumption") name(graph3,replace)
hist lienrlsec, kdensity normal xtitle("Log of School enrollment") name(graph4,replace)
hist ipopgrow,  kdensity normal xtitle("Population growth (annual %)") name(graph5,replace)
graph combine graph1 graph2 graph3 graph4 graph5, title("Normal distribution of variables")
hist liceryiel, kdensity normal by(, title(Normal distribution of log Cereal yield by income-level, size(medium)))
by(catgory, note("")) by(, legend(off)) xtitle("1=low income, 2=lower middle income, 3=upper middle income
and 4=high income")
hist lidisagri, kdensity normal by(, title(Normal distribution of log Disbursement to agriculture by income-level,
size(medium))) by(catgory, note("")) by(, legend(off)) xtitle("1=low income, 2=lower middle income, 3=upper
middle income and 4=high income")
hist lifertikg, kdensity normal by(, title(Normal distribution of log Fertilizer consumption by income-level,
size(medium))) by(catgory, note("")) by(, legend(off)) xtitle("1=low income, 2=lower middle income, 3=upper
middle income and 4=high income")
hist lienrlsec, kdensity normal by(, title(Normal distribution of log Secondary school enrollment by income-
level, size(medium))) by(catgory, note("")) by(, legend(off)) xtitle("1=low income, 2=lower middle income,
3=upper middle income and 4=high income")
hist ipopgrow, kdensity normal by(, title(Normal distribution of Population growth by income-level,
size(medium))) by(catgory, note("")) by(, legend(off)) xtitle("1=low income, 2=lower middle income, 3=upper
middle income and 4=high income")
// Skew-kurtosis
sktest liceryiel lidisagri lifertikg lienrlsec ipopgrow
sktest liceryiel lidisagri lifertikg lienrlsec ipopgrow if catgory == 1
sktest liceryiel lidisagri lifertikg lienrlsec ipopgrow if catgory == 2
sktest liceryiel lidisagri lifertikg lienrlsec ipopgrow if catgory == 3
sktest liceryiel lidisagri lifertikg lienrlsec ipopgrow if catgory == 4
// Shapiro-wilk
swilk liceryiel lidisagri lifertikg lienrlsec ipopgrow
swilk liceryiel lidisagri lifertikg lienrlsec ipopgrow if catgory == 1
swilk liceryiel lidisagri lifertikg lienrlsec ipopgrow if catgory == 2
swilk liceryiel lidisagri lifertikg lienrlsec ipopgrow if catgory == 3
swilk liceryiel lidisagri lifertikg lienrlsec ipopgrow if catgory == 4
// Kolmogorov-Smirnov
// liceryiel
tab catgory, sum(liceryiel)
```

```
ksmirnov liceryiel = normal((liceryiel- 7.8640023)/.69298647)
ksmirnov liceryiel = normal((liceryiel- 7.1647206)/.4009096)  if catgory == 1
ksmirnov liceryiel = normal((liceryiel- 7.6527519)/.57184835) if catgory == 2
ksmirnov liceryiel = normal((liceryiel- 7.8075578)/.63238698) if catgory == 3
ksmirnov liceryiel = normal((liceryiel- 8.4076737)/.53711934) if catgory == 4
//lidisagri
tab catgory, sum(lidisagri)
ksmirnov lidisagri = normal((lidisagri- 2.3436238)/2.0392072)
ksmirnov lidisagri = normal((lidisagri- 2.8221516)/ 2.362631) if catgory == 1
ksmirnov lidisagri = normal((lidisagri- 2.4542433)/1.7991691) if catgory == 2
ksmirnov lidisagri = normal((lidisagri- 2.2879885)/1.9868086) if catgory == 3
ksmirnov lidisagri = normal((lidisagri- 2.0979214)/2.1201753) if catgory == 4
// lifertikg
tab catgory, sum(lifertikg)
ksmirnov lifertikg = normal((lifertikg- 3.9327349)/ 1.8313818)
ksmirnov lifertikg = normal((lifertikg- 1.2649367)/ 1.8662164) if catgory == 1
ksmirnov lifertikg = normal((lifertikg- 3.4865293)/1.4981078) if catgory == 2
ksmirnov lifertikg = normal((lifertikg- 4.1858952)/1.4775319) if catgory == 3
ksmirnov lifertikg = normal((lifertikg- 5.2328245)/.72495599) if catgory == 4
// ldienrlsec
tab catgory, sum(lienrlsec)
ksmirnov lienrlsec = normal((lienrlsec- 4.2453003)/ .49659854)
ksmirnov lienrlsec = normal((lienrlsec- 4.0080473)/ .45404809) if catgory == 1
ksmirnov lienrlsec = normal((lienrlsec- 3.4865293)/1.4981078) if catgory == 2
ksmirnov lienrlsec = normal((lienrlsec- 4.4328231)/.20246171) if catgory == 3
ksmirnov lienrlsec = normal((lienrlsec- 4.66969) / .12818721) if catgory == 4
// ipopgrow
tab catgory, sum(ipopgrow)
ksmirnov ipopgrow = normal((ipopgrow- 1.3559788)/ 1.2152157)
ksmirnov ipopgrow = normal((ipopgrow- 2.7362533)/ .6007097) if catgory == 1
ksmirnov ipopgrow = normal((ipopgrow- 1.8340865)/.74540535) if catgory == 2
ksmirnov ipopgrow = normal((ipopgrow- 1.0709437)/1.1745918) if catgory == 3
ksmirnov ipopgrow = normal((ipopgrow- .59653763)/1.1113849) if catgory == 4
// Hausman test
xtreg liceryiel lidisagri lifertikg lienrlsec ipopgrow, fe
estimates store fixed
xtreg liceryiel lidisagri lifertikg lienrlsec ipopgrow, re
estimates store random
hausman fixed random
// Heteroscedasticity
// Breusch-Pagan test
xi:reg liceryiel lidisagri lifertikg lienrlsec ipopgrow i.country
hettest lidisagri lifertikg lienrlsec ipopgrow i.country
xi:reg liceryiel lidisagri lifertikg lienrlsec ipopgrow i.country if(catgory == 1)
hettest lidisagri lifertikg lienrlsec ipopgrow i.country
xi:reg liceryiel lidisagri lifertikg lienrlsec ipopgrow i.country if(catgory == 2)
hettest lidisagri lifertikg lienrlsec ipopgrow i.country
xi:reg liceryiel lidisagri lifertikg lienrlsec ipopgrow i.country if(catgory == 3)
```

```
hettest lidisagri lifertikg lienrlsec ipopgrow i.country
xi:reg liceryiel lidisagri lifertikg lienrlsec ipopgrow i.country if(catgory == 4)
hettest lidisagri lifertikg lienrlsec ipopgrow i.country
// White's test
xi:reg liceryiel lidisagri lifertikg lienrlsec ipopgrow i.country
estat imtest, white
xi:reg liceryiel lidisagri lifertikg lienrlsec ipopgrow i.country if(catgory == 1)
estat imtest, white
xi:reg liceryiel lidisagri lifertikg lienrlsec ipopgrow i.country if(catgory == 2)
estat imtest, white
xi:reg liceryiel lidisagri lifertikg lienrlsec ipopgrow i.country if(catgory == 3)
estat imtest, white
xi:reg liceryiel lidisagri lifertikg lienrlsec ipopgrow i.country if(catgory == 4)
estat imtest, white
// VIF Multicolliniarity
xi:reg liceryiel lidisagri lifertikg lienrlsec ipopgrow, vce(cluster country)
estat vif
xi:reg liceryiel lidisagri lifertikg lienrlsec ipopgrow if(catgory == 1), vce(cluster country)
estat vif
xi:reg liceryiel lidisagri lifertikg lienrlsec ipopgrow if(catgory == 2), vce(cluster country)
estat vif
xi:reg liceryiel lidisagri lifertikg lienrlsec ipopgrow if(catgory == 3), vce(cluster country)
estat vif
xi:reg liceryiel lidisagri lifertikg lienrlsec ipopgrow if(catgory == 4), vce(cluster country)
estat vif
// Serial correlation (Wooldridge test)
gen lidisagri2 = (lidisagri)^2
gen lifertikg2 = (lifertikg)^2
gen ipopgrow2 = (ipopgrow)^2
net from http://www.stata-journal.com/software/sj3-2/
net describe st0039
net install st0039
xtserial liceryiel lidisagri* lifertikg* lienrlsec ipopgrow*, output
xtserial liceryiel lidisagri* lifertikg* lienrlsec ipopgrow* if catgory == 2, output
xtserial liceryiel lidisagri* lifertikg* lienrlsec ipopgrow* if catgory == 3, output
xtserial liceryiel lidisagri* lifertikg* lienrlsec ipopgrow* if catgory == 4, output
// Stationarity Fisher
xtunitroot fisher liceryiel , dfuller lags(0)
xtunitroot fisher liceryiel if catgory == 1 , dfuller lags(0)
xtunitroot fisher liceryiel if catgory == 2 , dfuller lags(0)
xtunitroot fisher liceryiel if catgory == 3 , dfuller lags(0)
xtunitroot fisher liceryiel if catgory == 4 , dfuller lags(0)
xtunitroot fisher lidisagri , dfuller lags(0)
xtunitroot fisher lidisagri if catgory == 1 , dfuller lags(0)
xtunitroot fisher lidisagri if catgory == 2 , dfuller lags(0)
xtunitroot fisher lidisagri if catgory == 3 , dfuller lags(0)
xtunitroot fisher lidisagri if catgory == 4 , dfuller lags(0)
xtunitroot fisher lifertikg , dfuller lags(0)
```

```
xtunitroot fisher lifertikg if catgory == 1 , dfuller lags(0)
xtunitroot fisher lifertikg if catgory == 2 , dfuller lags(0)
xtunitroot fisher lifertikg if catgory == 3 , dfuller lags(0)
xtunitroot fisher lifertikg if catgory == 4 , dfuller lags(0)
xtunitroot fisher lienrlsec, dfuller lags(0)
xtunitroot fisher lienrlsec if catgory == 1, dfuller lags(0)
xtunitroot fisher lienrlsec if catgory == 2, dfuller lags(0)
xtunitroot fisher lienrlsec if catgory == 3, dfuller lags(0)
xtunitroot fisher lienrlsec if catgory == 4, dfuller lags(0)
xtunitroot fisher ipopgrow, dfuller lags(0)
xtunitroot fisher ipopgrow if catgory == 1, dfuller lags(0)
xtunitroot fisher ipopgrow if catgory == 2, dfuller lags(0)
xtunitroot fisher ipopgrow if catgory == 3, dfuller lags(0)
xtunitroot fisher ipopgrow if catgory == 4, dfuller lags(0)
// Misspecification test
xi:reg liceryiel lidisagri lifertikg lienrlsec ipopgrow i.country, vce(cluster country)
estat ovtest
// Tests
xtreg liceryiel lidisagri, fe vce (cluster country)
xtreg liceryiel lidisagri lifertikg, fe vce (cluster country)
xtreg liceryiel lidisagri lifertikg lienrlsec, fe vce (cluster country)
xtreg liceryiel lidisagri lifertikg lienrlsec ipopgrow , fe vce (cluster country)
xtreg liceryiel lidisagri lifertikg lienrlsec ipopgrow if catgory ==1 , fe vce (cluster country)
xtreg liceryiel lidisagri lifertikg lienrlsec ipopgrow if catgory ==2 , fe vce (cluster country)
xtreg liceryiel lidisagri lifertikg lienrlsec ipopgrow if catgory ==3 , fe vce (cluster country)
xtreg liceryiel lidisagri lifertikg lienrlsec ipopgrow if catgory ==4 , fe vce (cluster country)
************************** Working space
// We'll send the dataset upon request, email: albertwallgren123@gmail.com
// Saving the changed dataset under to another file so as to not change the original file.
save BachelorThesisdatasetWW, replace
// Close the log-file.
log close
```