# ESSAYS IN ECONOMIC DEVELOPMENT

By

Maksym Ivanyna

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Economics – Doctor of Philosophy

2017

**ABSTRACT**

ESSAYS IN ECONOMIC DEVELOPMENT

By

Maksym Ivanyna

The dissertation consists of three essays: "Benefits of diversification in agriculture: Evidence from Malawi", "Corruption and economic growth revisited" and "How close is your government to its people? Worldwide indicators on localization and decentralization". All three essays focus on issues, which are particularly important for developing countries: food security, corruption, decentralization.

In the first essay I use data from a farmers' survey in Malawi to compare two agricultural technologies: monoculture maize and crop diversification (maize-legume intercrop). I match farmers locations with data on rainfall and air temperature to test whether more biodiverse agriculture is better at absorbing weather shocks. The data make it possible to compare variation not only over time, but also over different plots within the same time period, which helps reduce omitted variable bias. The instrumental variable method is used to eliminate rainfall measurement error. For a number of specifications, and controlling for fertilizer use, crop diversification is both more productive than monoculture maize and more resistant to weather shocks. Although I am not able to identify the average population effect, I show that the effect I identify is likely to prevail if the Malawian government decides to shift the focus of its agricultural subsidy at the margin from fertilizer to legume seeds and education.

The focus of the second essay is on corruption. While literature finds many channels through which corruption can hurt economic growth, the link proved hard to establish in empirical cross-country studies. In this paper I show that part of the explanation of this puzzle is that there is a reverse causality: everything else equal, exogenously-driven economic growth can increase corruption. The reason is that the boost to output increases tax revenue, and hence pool of resources that corrupt public officials can embezzle. I show the workings of this channel in a

simple stylized model, which is then accompanied by numerical simulations in a dynamic general equilibium overlapping-generations model, which allows for corruption and tax evasion. I also present empirical evidence, which supports my findings.

The third chapter assembles and analyzes a unique data base on local governance that provides a first approximation of the institutional architecture of local governance that has emerged as a result of the silent revolution (decentralization reforms) of the last three decades on moving governments closer to people. An important feature of this data set is that, for comparative purposes, it measures government decision making at the local level i.e. the order of government that is closest to the people and hence providing a better indicator of decentralized decision making as compared to the "sub-national governments" that also include intermediate tiers of provinces and states, used by the existing literature. This unique data set for 182 countries attempts to capture institutional dimensions of political, fiscal and administrative autonomy enjoyed by local governments under diverse multi-order governance regimes using a common framework. These dimensions are aggregated to develop a "decentralization index" and then adjusted for heterogeneity to develop a "government closeness index" that provides an overall ranking of countries on the closeness of their government to the people.

## ACKNOWLEDGEMENTS

First of all, I would like to express my sincere gratitude to the Chair of my Dissertation Committee Christian Ahlin. It is his persistence and encouragement, and at the same time firm adherence to high quality standards, that made this thesis possible. I also want to thank the Committee members, Sieglinde Snapp, Jay Wilson, and Jinhua Zhao, for their openness and willingness to help at any moment. I have been a "hard" student for the department, but I could always count on helping hand of our Graduate Secretary Lori Jean Nichols, and Director of Graduate Studies Steve Haider. Last but not least, my wife Kateryna, my parents, sister and niece believed in me much more that I did myself, and this belief has kept me going.

**TABLE OF CONTENTS**

## LIST OF TABLES

## LIST OF FIGURES

## KEY TO ABBREVIATIONS

**AISS** Agricultural Input Subsidy Survey

**CMAP** Climate Prediction Center Merged Analysis of Precipitation

**CoC** Control of Corruption Index

**DI** Decentralization Index

**EA** enumeration area

**F** fertilizer

**FAO** Food and Agriculture Organization

**FE** fixed effects

**GCI** Government Closeness Index

**GDP** Gross Domestic Product

**GG** General Government

**GMM** General Method of Moments

**ha** hectare

**HR** human resources

**IHS** Integrated Household Survey

**IMF** International Monetary Fund

**IPCC** International Panel on Climate Change

**kCal** kilo calories

**L** legume

**LER** land equivalency ratio

**LG** local governments

**LIC** low income countries

**M** monocropped maize

**math** Mathematica Weather Data

**MK** Malawian Kwacha

**ML** maize-legume intercrop

**MNVAC** Malawi National Vulnerability Assessment Committee

**OECD** Organization of Economic Cooperation and Development

**OLS** Ordinary Least Squares

**OPV** open pollinated variety

**PPP** Purchasing Power Parity

**RFE** Rainfall Estimate

**RRC** resource-rich countries

**SG** state governments

**SSA** Sub-Saharan Africa

**TA** traditional authority

**TI** Transparency International

**ToT** terms-of-trade

**UN** United Nations

**USD** United States Dollar

**W** weather

**WBES** World Bank Enterprise Survey

**WEO** World Economic Outlook

**WGI** Worldwide Governance Indicators

# CHAPTER 1

# INTRODUCTION

The dissertation consists of three essays: "Benefits of diversification in agriculture: Evidence from Malawi", "Corruption and economic growth revisited" and "How close is your government to its people? Worldwide indicators on localization and decentralization" (with Anwar Shah). All three essays focus on issues, which are particularly important for developing countries. The first essay focuses on food security, agricultural productivity, and ability of farmers to withstand weather shocks in Sub-Saharan Africa. The second essay is about corruption and its relationship with economic growth. The third essay is about measuring and analysis of the closeness of government to its people, or in other words decentralization.

In the first essay I explore the ways to improve food security in Sub-Saharan Africa. People there are extremely food-insecure. In 2011-13 25% of the SSA population were undernourished, down only eight percentage points from 1990. Improvements in food security in SSA have been pushed off track by frequent and severe weather shocks. Because over 90% of the region's cropland is rain-fed, and insurance markets are underdeveloped, yields and hence food security are highly dependent on rainfall. An additional risk is that agriculture in SSA relies on a supply of cheap fertilizer, which is not always available.

Given Africa's poor soils and dry climate, an often-proposed way to improve food security is crop diversification: growing different crops together on one plot. In Malawi the often suggested technology is an intercrop of maize and legumes. While the technology has been tested in field trials many times, the question that remains is how crop diversification performs against monoculture maize in real life conditions of smallholder farmers on a large scale, and whether crop diversification is better at absorbing shocks confronting farmers.

The first essay of this dissertation uses data on Malawian rural households to test whether crop diversification, in particular maize-legume intercropping, is indeed more productive and resilient than monoculture maize. Controlling for the use of fertilizer, I do find crop diversification to be more productive than monoculture maize. This result is robust to variety of estimation specifications and variable definitions, although for some of them the result is not statistically significant. I also find that Malawian farmers are highly vulnerable to fertilizer supply and weather shocks, and that more biodiverse agricultural technology likely works as a shock absorber, although again the results lose significance in some specifications.

In the second essay I switch from food security to analysis of corruption. Social scientists came up with many channels of how corruption can hurt economic growth and efficiency. Despite the large number of theoretical channels, many of which supported by micro-empirical evidence, at a macro level the negative link between corruption and growth has been hard to demonstrate. The literature offers several potential explanations for the weak link between corruption and growth: measurement error, omitted variables, or the explanation that the link is actually weak.

My essay offers an alternative explanation. Given the multitude of theoretical arguments, it is highly likely that corruption has large negative effect on long-term potential economic growth. However, the empirical relationship between the two is contaminated by the reverse causality from economic growth to corruption in short and medium term. During economic booms tax revenue is higher and it is easier for government to borrow. This blows up the budgets that bureaucrats are in charge of, and consequently increases embezzlement and corruption.

Corruption is often studied in the context of decentralization policy - the main topic of my third essay. A silent revolution (the so-called "decentralization" reforms) has been sweeping the globe since the 1980s. The main thrust of this revolution has been to move decision making closer to people. While there has been monumental literature dealing with various aspects of this revolution, there has not been any systemic study providing a time capsule of the changed world as a result of this revolution. Defining and measuring decentralization is the main objective of the third essay. We look at various dimensions of decentralization - fiscal, political and administrative. We also differentiate between local and state governments. By combining these measurements, an aggregate indicator of decentralization/localization is developed for each country. This index is then adjusted for population size, area and heterogeneity to arrive at the government closeness index.

## CHAPTER 2

## BENEFITS OF DIVERSIFICATION IN AGRICULTURE: EVIDENCE FROM MALAWI

### 2.1 Introduction

People in Sub-Saharan Africa (SSA) are extremely food-insecure. In 2011-13 25% of the SSA population were undernourished, down only eight percentage points from 1990. The region has been a laggard in meeting the Millennium Development Goals (UN, 2014).

Improvements in food security in SSA have been pushed off track by frequent and severe weather shocks. Because over 90% of the region's cropland is rain-fed, and insurance markets are underdeveloped, yields and hence food security are highly dependent on rainfall. Droughts and dry spells during a short growing season can have dreadful consequences. Sporadic rainfall in the 2001-02 growing season in Malawi brought on a severe food crisis, which caused as many as several thousand hunger-related deaths (ActionAid, 2006).

An additional risk is that agriculture in SSA relies on a supply of cheap fertilizer, which is not always available. Many countries in the region offer fertilizer subsidies, which are expensive for the government budget and the country's external position, and often require financial aid from international donors. In 2012 about 10% of government spending went to Malawi's Farm Input Subsidy Program; of the total 12% was directly covered by foreign aid (Chirwa & Dorwards, 2014). Fertilizer prices are extremely volatile: annual swings of 20-30% are not uncommon. This undermines either the sustainability of subsidy programs or adequate supply of fertilizer to the field.

In many SSA countries monoculture maize has been the dominant technology used by smallholder farmers. In Malawi its share of agricultural land increased from 27% in 2004 to 37% in 2009, despite slow long-run growth in productivity and vulnerability to weather and fertilizer supply shocks. Many scholars hail intensive monoculture as the only viable path to satisfy the ever-growing demand for food (Borlaug, 2000; Morris et al., 2007; Tilman et al., 2011). Economists mainly deal with questions like effectiveness of subsidy programs (Rickert-Gilbert et al., 2010; Dorwards & Chirwa, 2011), or why smallholder farmers do not take up fertilizer or new hybrid seeds on a large scale (Duflo et al., 2011; Suri, 2011).

Given Africa's poor soils and dry climate, an often-proposed alternative way to improve food security is crop diversification: growing different crops together on one plot. In Malawi the often suggested technology is an intercrop of maize and legumes. Farm trials and participatory farming demonstrate that even with less fertilizer this technology can be as productive as monoculture maize (Searle et al., 1981; Snapp et al., 2010; Sileshi et al., 2010; One Acre Fund, 2015). The technology now has the attention of the Gates Foundation and the Malawian Government, which decided in 2013 to subsidize legume seeds as well as maize seeds and fertilizer. The question that remains is how crop diversification performs against monoculture maize in real life conditions of smallholder farmers on a large scale, and whether crop diversification is better at absorbing shocks confronting farmers.[1] [2]

This paper uses data on Malawian rural households to test whether crop diversification, in particular maize-legume intercropping, is indeed more productive and resilient than monoculture maize. I match farmers' locations with multiple weather measures, which makes it possible to use the IV method to reduce measurement error. Also, the unique structure of the dataset allows me to observe not only farmers over time but also the performance of both technologies managed by the same farmer at a single point in time. This is important for reducing the omitted variable bias, though it is still not possible to identify an average population effect. However, I argue that, looking from a perspective of a farmer's technology choice, the effect which I do identify is likely to prevail if the government decides to shift the subsidy focus at the margin from maize seeds and fertilizer to legume seeds and education about crop diversification.

Controlling for the use of fertilizer, I find crop diversification to be more productive than

[1] As shown by Duflo et al. (2016), controlled experiments are not always informative about the policy implications in real life

[2] In few studies crop diversification is included as one of the regressors, and is shown to improve yields. See Sheahan et al. (2013). The variable is, however, never the focus of the study, and its interaction with weather and fertilizer shocks is not explored

monoculture maize. This result is robust to variety of estimation specifications and variable definitions, although for some of them the result is not statistically significant. According to the preferred specification, a ten percentage points shift in land use from maize to maize-legume intercrop is likely to raise average calorie yield by 2.4 percent. The better performance of the maize-legume intercrop seems to be driven by high productivity of legumes and the fact that they do not seem to interfere much with the growth of maize even when planted densely together. At the same time, on average the intercrop is not more productive than maize-legume rotation.

I also find that Malawian farmers are highly vulnerable to fertilizer supply and weather shocks, and that more biodiverse agricultural technology likely works as a shock absorber, although again the results lose significance in some specifications. A half historical standard deviation negative shock to rainfall is expected to reduce the average historical calorie yield by almost four percent. It would cost government about 0.9 percent of GDP to compensate the losses to farmers at 2013 maize price and the exchange rate. A ten percentage points increase in land used for maize-legume intercrop is likely to reduce the yield loss to 3.5 percent, which is equivalent to a reduction in the compensation cost of 0.1 percent of GDP. A similar pattern is observed for rainfall variance, temperature, and fertilizer supply shocks. Reforming the subsidy program by reducing fertilizer use by 10 percent and increasing use of maize-legume intercrop by nine percent would be yield-neutral but would make Malawi's agriculture more resilient and more sustainable. Finally, I find that maize-legume intercrop is also more weather-resistant than maize-legume rotation or growing maize and legume separately.

## 2.2 Background

### 2.2.1 About Malawi

Malawi is a land-locked country in SSA. Despite sustained economic growth of about 7% per year for the past five years, the country's GDP per capita in 2010 was only 925 PPP units, which was in the lowest 5% of the world's distribution. Malawi's population is 13 million people, 40% of which were living on less than USD 1.25 per day in 2010. Farmers represent 78% of the population.

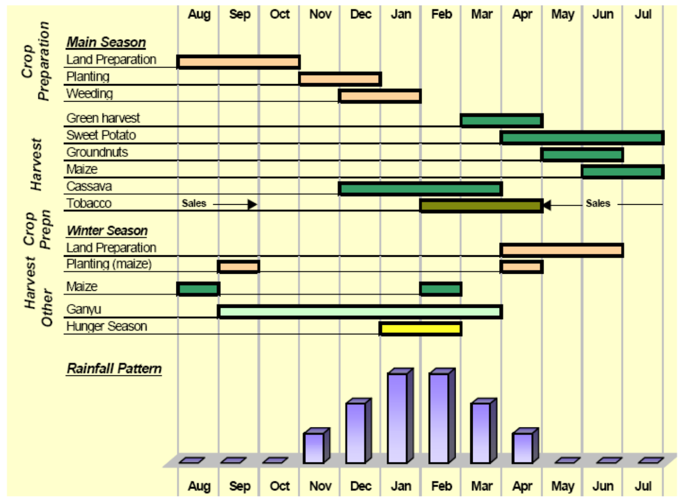


Figure 2.1: Usual cropping season of a Malawian household

*Note* Copied from MNVAC (2005)

According to Dorwards et al. (2010), only 10% of them are net sellers, i.e. produce surplus over their own consumption. Maize is by far the main staple crop in the country: it is grown by 97% of the farmers, and it accounts for 60% of total calorie consumption in the country.

### 2.2.2 A year in life of Malawian subsistence farmer

Maize is grown in Malawi without irrigation during the single rainy season between October and April. The usual cropping season is depicted on Figure 2.1. Weather can cause significant fluctuations of maize harvest. The plant has relatively shallow root system, which makes it dependent on soil moisture and consequently regular rainfall. Low precipitation between January and March is harmful to the harvest. High rainfall variation is damaging as well. Especially dangerous are dry spells of more than a week in January to March during maize's flowering and early grain filling. For example, the Malawian aggregate maize harvest decreased by about 40% in 2004/05 most likely as a result of two weeks without rain in February. In addition, Lobell et al. (2011) provides empirical evidence on negative impact of (too) high temperature on maize productivity (up to 1.7% decrease for each day above 86F (30C)). At the same time, moderate increases in temperature are expected to affect the harvest favorably.

### 2.2.3 Legumes and maize-legume intercrop

In addition to being vulnerable to rainfall shocks maize can deplete soils of nutrients - in particular carbon and nitrogen - if grown continuously without fallow periods. Nitrogen is an essential component of a plant's nutrition. Without sufficient quantity of nitrogen in the soil, yields can fall significantly. One solution to this problem is to use inorganic fertilizer. An alternative is to intercrop (grow at the same time) or rotate (grow one after another) maize with nitrogen-fixing plants, such as legumes.

Even though the supply of nitrogen in nature is practically boundless, most of it is contained in our atmosphere (70% of its volume). Most plants are not able to fix nitrogen directly from the atmosphere. They use ammonia ($NH_3$) in the soil, which is being produced by bacteria using air. Legumes are able to attract these bacteria on their roots by supplying them other nutrients. In exchange the bacteria fix nitrogen from the atmosphere and produce ammonia for the plant.

Growing legumes with maize in an intercrop can affect average yield in several ways (see Figure 2.2 for a photo of a plot with maize-legume intercrop). Since legumes fix nitrogen, it can increase nitrogen available for maize too (so-called nitrogen sharing). It can also improve the soil quality over the long term.[3] Intercropping also allows for denser planting (f.e. a row of legume between two regular rows of maize), and hence more efficient use of land, nutrients and water. A field with more diversified agriculture is likely to attract less insect pests (Wetzel et al., 2016). At the same time, planting maize and legumes in close proximity to each other may lead to inter-species competition (so called competition depression), but given that the plants have different height, root systems and growing cycles, the impact of this competition on yield is likely to be limited.

Additional potential benefit of growing maize and legumes together is that the system may become more resistant to weather shocks. Legumes provide more cover for the ground. They also shed leaves, which improves the organic matter of the soil. Together these factors increase the

[3]This argument also works for the maize-legume rotation

Figure 2.2: A plot with the maize-legume intercrop

*Note* Taken from www.cirad.fr. Accessed October 20, 2016.

capacity of soil to store water. Most legumes also have deeper root systems than maize, which makes them more weather-resistant. The usual risk diversification argument is also at work here - since the plants have different growing cycles, weather shocks are likely to have differential impact on the yields, and hence more diversified system is likely to be more stable. Even though this argument works in general for growing more than one crop in one season (f.e. maize and legume on different plots), intercropping gives an additional momentum to it. In case of an adverse weather shock, a failure of one of the crops would reduce inter-species competition on the plot and hence increase the yield of the other crop.

### 2.2.4 Agricultural Input Subsidy Programs

Malawian agricultural input subsidy was first implemented in 2005/06 after poor harvest season of 2004/05, and then every year afterward. It consisted of distribution of vouchers to roughly 50% of farmers to receive fertilizers for maize production at a 90% discount, and further distribution of vouchers for improved maize seeds and fertilizer for tobacco.

According to the Malawian government, the main objectives of the input subsidy are national and personal food self-sufficiency and food security, as well as income increase among the poorest. Poverty and food insecurity are seen as major market failures and causes for other negative externatilities, thus warranting government intervention. An additional reason for input subsidy is potential information externality. It is frequently argued that farmers do not realize the full payoff to the fertilizer, and therefore they need to be encouraged to use it and learn.

Information externality argument also applies to maize-legume intercropping. While the technology is not new, [4] it has not been traditionally used in all regions of the country, and it has been gradually crowded out by monoculture maize in recent decades, not least due to maize seed subsidy and government extension. Besides with climate rapidly changing and becoming more prone to droughts, farmers increasingly need to learn how to cope with the weather shocks. Maize-legume intercrop is one of the strategies. At the same time, some recent studies suggest that farmers likely do not fully realize the potential benefits of this technology. For example, after on-farm trials by the One Acre Fund in Kenya the adoption rate of the maize-legume intercrop increased by up to 40 percentage points (One Acre Fund, 2015). Another example, a so-called On-Farm Cropping Verification Trial - a 1998-2000 experiment of planting and educating Malawian farmers about six different technologies - demonstrated that of those who did not use the maize-pigeon pea intercrop before, 67% decided to use it after being shown how to properly use it.[5]

An additional potential obstacle to the wider spread of the maize-legume intercropping is the absence of necessary infrastructure. For example, in 2009 pigeon pea was being sold only in 25% of the villages, the rest of the farmers had to walk long distances to purchase the seeds.

At the same time, the technology could be more labor and seed intensive than monocropped maize, which would also stall its dissemination, and it would be the case against the subsidy. However, field trials suggest that the labor requirements are not too different among the two technologies, at least for legume species that are grown in Malawi (Snapp et al., 2003). Intercrop may require more labor at seeding and harvesting, but then it may require less weeding in between. Even if

[4] Growing maize and legumes together has been practiced as early as 3000 years ago in Americas - a system known as “the three sisters” along with squash.

[5] See Gilbert et al. (2002), and also Kerr et al. (2007)

the intercrop is more seed and labor intensive, but its yield is sufficiently high, the government might still be interested in subsidizing it, since it would increase agricultural productivity per unit of scarce arable land.

The input subsidy is costly, however. For example, 2008/09 subsidy program accounted for 15% of country's budget. Therefore, if it is to be continued, it is of utmost importance to maximize the program's efficiency, i.e. yielding maximum food security and minimizing the risks at the smallest cost possible. This includes the choice of inputs to subsidize, as well as the choice of extension activities.

## 2.3 Data

I use a panel of three surveys of Malawian households, which were implemented and used to track the progress of agricultural input subsidy programs in 2006-2009. The first wave - Integrated Household Survey Round 2 (IHS2) - was conducted from March 2004 to March 2005, and served as a baseline survey for the subsequent waves. A total of 9494 households were asked about a wide variety of topics, such as health, expenditures, time use, agricultural practices, etc. The next two waves - Agricultural Input Subsidy Survey 2007 (AISS07) and AISS09 - were conducted in May-July 2007 and February-July 2009 correspondingly. Due to the resource constraints these surveys were of much smaller scale than IHS2: only 3169 and 1918 households were surveyed correspondingly, and the set of questions was restricted mainly to such topics as agricultural practices, and in particular fertilizer use. In all three waves, the households were randomly sampled by district (tier 1 of Malawian local government), traditional authority of sub-chief unit (tier 2), and enumeration area (local sampling unit - EA), and most of the attrition between wave 1 and waves 2 and 3 comes from the random jettisoning of entire EA's. The number of observations and the attrition rates are given in Table 2.1. Note that the actual number of households participating in the three surveys is higher than the one reported and used in this paper, as I dropped those households, which did not participate in agricultural activity during the last cropping season, and households, which did not report their harvest.

Table 2.1: Malawian households survey: Available observations and attrition

| Indicator | Wave 1 (IHS2) | Wave 2 (AISS07) | Wave 3 (AISS09) |
|---|---|---|---|
| N households, total | 9753 | 3184 | 1918 |
| N households, all waves | 1242 | 1242 | 1242 |
| N households, at least two waves | 2752 | 2935 | 1489 |
| N households, common with prev. wave | n.a. | 2720 | 1457 |
| Attrition from previous wave | n.a. | 71% | 54% |
| Exogenous attrition (due to sampling design),% of attrition | n.a. | 92% | 81% |
| N households replaced/dropped | n.a. | 557 | 320 |
| N plots, total | 19703 | 5195 | 5022 |

*Note* Abbreviations: IHS2 - Integrated Household Survey Round 2 (1994-1995), AISS07 - Agricultural Input Subsidy Survey Round 1 (2007), AISS09 - Agricultural Input Subsidy Survey Round 2 (2009). Population of households is restricted to those engaged in agricultural activity (growing crops) and who reported their harvests. Definitions of indicators: *N households, total* - total number of interviewed households in a current wave of survey; *N households, all waves* - number of households, which were interviewed in all three waves; *N households, at least two waves* - number of households in a current wave, which were interviewed in at least one additional wave; *N households, common with prev. wave* - number of households, which were interviewed in current and previous waves; difference with the previous indicator is that some households were interviewed in Wave 1 and Wave 3 only; *Attrition from previous wave* - households which were not interviewed in a current wave but were interviewed in a previous wave, as a percentage of total number of interviewed households in a previous wave; *Exogenous attrition (due to sampling design)* - households from enumeration areas, which were jettisoned in a current wave, but were sampled in a previous wave, as a percentage of total number of households subject to attrition; *N households replaced/dropped* - number of households from EAs, which were sampled in both current and previous waves, but were subject to attrition in a current wave; *N plots, total* - number of plots owned by households interviewed in a current wave

A unique and useful feature of the data is that, where applicable, it is disaggregated by plot for each household. Plot is an area in which a uniform, consistent crop management system is used. The household-plot disaggregated data includes inputs and outputs of farming: use of fertilizer, seeds, number and timing of weeding, harvest by each crop grown on a plot.[6] General information about each household is also reported: demographic structure, alternative sources of income, etc.

The final dataset represents a "two-dimensional" panel of Malawian households. First, as in usual panel, I am able to observe farming households over time, which allows me to control for time-constant characteristics of the households. The second dimension is geographical - within one period I am able to observe farmers using different technologies at different plots. This allows me to control for all characteristics, including the time-changing ones, which are common to plots cultivated by the same household. I am not able to identify plots across years, as farmers numbered them differently during each wave, but I do track the technology use over time.

[6]Data on seeds and weeding are not available for all waves of surveys. Also data on seeds are not reported by crop

### 2.3.1 Measuring crop diversification: agricultural technologies in Malawi

As much as 50 different crops were grown in Malawi in 2003-2009. They are listed in Table A.1 in Appendix. Some crops are allocated a separate plot, some are grown together on the same plot. Farmers were asked to name up to five most important crops for each plot. On average, 1.76 crops are reported per each plot. Fifty two percent of plots in the sample grow only one crop, 25% grow two crops, 14% - three, 6% - four. Only 2.6% plots grow five crops, so the upper boundary does not seem too binding. Altogether 2109 different combinations of crops are grown.

Local maize and hybrid maize are the most popular crops. Either one of them is grown on 62% of all plots in the sample. The group of legumes includes groundnut, ground bean, bean, soybean, pigeon pea, peas, cow peas, mucuna, and hyacinth bean. At least one of the legumes is grown on 44% of the plots. Non-food crops, tobacco and cotton, which are primarily grown to get cash, comprise 8% of the plots. Other popular crops are sorghum, nkhwani, cassava, sweet potato.

The summary of the most popular agricultural technologies is presented in Table 2.2. For each technology I list shares of plots used by it - total and for each wave separately.

Two technologies that I compare in my main specifications are maize (M) and maize-legume intercrop (ML). A technology is M if at least one of the crops listed on the plot is local maize, hybrid maize or open pollinated variety (OPV) maize, and no legumes are grown on the plot. A technology is ML if both maize and at least one of the legumes are grown on the plot. Thirty percent of the plots use intercrop of maize with one of the legumes, maize without legumes is cultivated on 32% of plots.

I chose to take the broadest definition for M and ML in my main specifications. In particular, other crops except maize and legumes are allowed to grow on both M and ML plots. Also, even if very little legume or maize is grown on a plot, it would still make it to the estimation.[7] I opt for the broadest definition for two reasons. First, even though I am likely to pick up some noise in the data, I am also sure not to dismiss any relevant observations. As Table 2.3 shows, in absolute majority

[7] As a result, some of the ML plots could, in some sense, be "less intercropped" than some M plots. For example, a plot with 96% maize and 4% cassava (non-legume plant) would be considered M plot, while a plot with 97% maize and 3% pigeon pea would be considered ML

of ML plots harvest shares of maize and legumes are dominant (average is 96%). However, if one goes too strict and requires only maize and legumes grown on ML plots, one would loose 34% of observations. Second, even though one could base the definition on certain thresholds of maize and legume harvest shares, any such definition would be arbitrary and potentially could dismiss relevant observations. For example, Table 2.3 shows that there are many plots with a small (less than 5%) harvest share of legumes. This could be because legumes are simply not dominant crops on these plots, in which case the plots are not maize-legume intercrops. But this could also be due to a failure of a dominant crop, and one of the advantages of crop diversification is that it provides insurance against such failure (harvest of the other dominant crop). Hence, one would want to include these plots in the sample.[8] Finally, even though M plots include few intercrops of maize with non-legume plants, I do want to separate out intercrops of maize and legumes for the reasons listed in Section 2.2.3.

As a robustness check, I also check other definitions of M and ML: maize and intercrop when maize and legumes are listed among three most important crops, maize and intercrop without any other crops on the plot, maize and intercrop only on plots with significant harvest shares of both. I also check specifications with dummies for plots with only maize or legumes to see presence of other crops significantly affects my results. Finally, I also apply a "X vs. 100-X" rule, which requires the harvest share of legumes on ML plots to be at least X%, and the harvest share of maize on M plots to be at least (100-X)%.[9] Qualitatively the results are similar, although as expected due to reduced number of observations, some of them become statistically insignificant. [10]

Figure 2.3 breaks down the M and ML plots by crops. Both local and hybrid (supposedly more productive) maize are used to almost equal extent on monoculture and intercropped plots. Quite a

[8] I do not observe seed input or area seeded by crop, so I cannot use these to define ML

[9] This rule makes sure that in some sense M plots are not more "intercropped" than ML plots. For example, under the broad definition, a plot with 95% maize and 5% cassava (non-legume crop) would be defined as M plot. A plot with 97% maize and 3% legume would be defined as ML plot. In this case, the M plot is actually more diversified than the ML plot, which would be desirable to avoid, even though I do want to specifically test for maize-legume intercrops because of their peculiar properties (see Section 2.2.3). I test the rules with X=3 and X=5

[10] The results are reported in Tables A.7 - A.11

Table 2.2: Technology use by Malawian farmers, shares by plots

| Technology | Definition | wave 1 (IHS2) | wave 2 (AISS07) | wave 3 (AISS09) | total |
|---|---|---|---|---|---|
| M | maize, no legume | 28.68 | 40.51 | 37.06 | 32.26 |
| H | hybrid maize, no legume | 16.38 | 21.79 | 18.98 | 17.8 |
| ML | maize and legume | 25.58 | 39.26 | 34.65 | 29.61 |
| M13 | maize - one of 3 most important crops | 28.68 | 40.51 | 37.06 | 32.26 |
| ML13 | maize and legume - one of 3 most important crops | 23.52 | 38.23 | 33.87 | 27.96 |
| M only | maize - single crop on plot | 21.61 | 31.07 | 21.35 | 23.26 |
| ML only | maize and legume - single crops on plot | 12.92 | 22.28 | 17.48 | 15.39 |
| MPp | maize and pigeon pea | 12.51 | 21.11 | 15.43 | 14.56 |
| MB | maize and beans | 9.21 | 8.56 | 11.64 | 9.52 |
| MGn | maize and ground nuts | 6.63 | 13.29 | 9.63 | 8.35 |

*Note* Abbreviations: IHS2 - Integrated Household Survey Round 2 (1994-1995), AISS07 - Agricultural Input Subsidy Survey Round 1 (2007), AISS09 - Agricultural Input Subsidy Survey Round 2 (2009). Legumes include groundnut, ground bean, bean, soybean, pigeon pea, peas, cow peas, mucuna, hyacinth bean. Maize include local maize, OPV and hybrid. The numbers in the table are shares of corresponding plots in total.

Table 2.3: Maize (M) and maize-legume intercrop (ML): Harvest shares by crops

| crop | mean | sd | p(10) | p(25) | p(50) | p(75) | p(90) |
|---|---|---|---|---|---|---|---|
| **M plots, harvest shares of ...** | | | | | | | |
| Maize | 0.97 | 0.12 | 0.96 | 1 | 1 | 1 | 1 |
| **ML plots, harvest shares of ...** | | | | | | | |
| Maize | 0.73 | 0.24 | 0.36 | 0.6 | 0.79 | 0.92 | 0.98 |
| Legumes | 0.23 | 0.23 | 0.01 | 0.06 | 0.15 | 0.33 | 0.58 |
| Maize + Legumes | 0.96 | 0.1 | 0.85 | 0.97 | 1 | 1 | 1 |

*Note* Legumes include groundnut, ground bean, bean, soybean, pigeon pea, peas, cow peas, mucuna, hyacinth bean. Maize include local maize, OPV and hybrid. Harvest shares are calculated based on the caloric value of crops.

few plots mix the two, both M and ML. Most popular legumes are pigeon pea, bean and ground nut.

An important feature of the agriculture technology use in Malawi is that many farmers apply two and more technologies (on different plots) within one growing season. Table 2.4 reports the distribution by wave. Most farmers still use only one technology, but as much as 800 of them over three waves use two. Low number for Wave 1 is because the harvests in the IHS2 survey are reported only as aggregates across all plots. So, even though types of crops grown on each plot are reported, I cannot differentiate between the two technologies in my estimation if they involve same crops.[11] Plots, where it is impossible to identify harvests, are dropped from the estimation.

[11] For example, if a farmer grows local maize on one plot, and local maize plus some legume

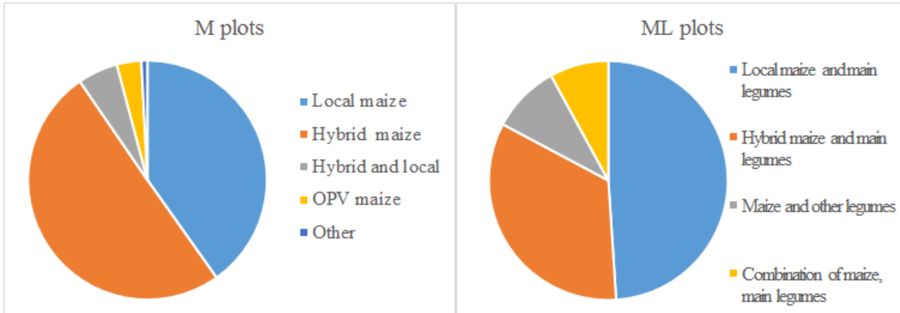


Figure 2.3: Composition of M and ML plots by crops

*Note* Main legumes: pigeon pea, bean, ground nut. Included are only the plots used in the estimation.

Table 2.4: Technology use and diversification by Malawian farmers, number of households

| | wave 1 (IHS2) | | wave 2 (AISS07) | | wave 3 (AISS09) | |
|---|---|---|---|---|---|---|
| | M | ML | M | ML | M | ML |
| M | 4183 | 208 | 1261 | 217 | 694 | 414 |
| ML | | 3656 | | 1266 | | 707 |

*Note* Abbreviations of technologies: M - maize, ML - maize-legume intercrop. Diagonals are numbers of households using only one corresponding technology at a given year. Above diagonals are numbers of households using two corresponding technologies at a given year. Reported numbers are after all adjustments (i.e numbers used in estimation).

### 2.3.2 Measuring output

My objective in this paper is to measure real agricultural output, as relates to the food security. This is not straightforward when several crops are harvested. Therefore I use several measures.

My main measure is the total energy output of the plot. The unit of measurement is kilo calorie. The measure makes sense since both maize and legumes are staple crops, and it is their nutritional value that the farmers look up to. Besides food security is primarily defined as the access to a minimum required number of calories per day (FAO, 2008). I use data from Food Composition Tables of Food and Agriculture Organization of United Nations (FAO) to transform multi-crop harvest into its energetic value. The conversion rates are reported in Table A.1.

I use four additional measures in my estimations. The first one is similar to the real GDP of a

on another then harvests by plots are impossible to identify. At the same time, some M and ML combinations are possible to identify, as I observe harvests by each variety of maize, and I do observe types of crops grown by plot. For example, if a farmer grows hybrid maize on one plot, and local maize plus some legume on another then I can identify harvests by plots

Table 2.5: Measures of harvest and fertilizer: Summary statistics

| | count | mean | sd | p10 | p50 | p90 |
|---|---|---|---|---|---|---|
| energy yield, mln kCal/ha | 12606 | 2.89 | 2.65 | 0.55 | 2.05 | 6.28 |
| cash yield, thousand MK/ha | 12606 | 10.98 | 11.84 | 1.99 | 7.85 | 23.69 |
| maize eqiuv. yield (Liu-Myers), tonnes/ha | 12606 | 1.21 | 1.45 | 0.16 | 0.76 | 2.79 |
| grain yield, tonnes/ha | 12606 | 0.83 | 0.76 | 0.16 | 0.60 | 1.83 |
| protein yield, kg/ha | 12606 | 85.34 | 79.26 | 15.14 | 62.52 | 183.11 |
| maize yield, tonnes/ha | 12606 | 0.71 | 0.71 | 0.11 | 0.46 | 1.61 |
| fertilizer applied, tonnes/ha | 12606 | 0.08 | 0.11 | 0 | 0.04 | 0.25 |

*Note:* All summary statistics are calculated at a household level. Only farmers who use M or ML (or both) are included. Exchange rate as of 2007 is 141MK per USD.

country: it is cash value of harvests of all crops at 2007 (constant) national prices. National price of a crop is defined as the median of the crop's price distribution over the whole country, as reported in community surveys, which were conducted parallel to the household surveys during each wave. The prices that I used are reported in Table A.1. The second measure is similar. It was proposed by Liu & Myers (2009) and is defined as the cash value of the harvest in current prices divided by the current price of maize. The measure takes into account the relative change of maize price, but it is not clear why should it be used if maize productivity is not sole focus of the research. The third measure is the amount of grain produced - both maize and legume in case of intercrop. In other words, this is just the weight of harvest. It is directly observable by farmers, and probably best understood by them. Finally, the fourth measure is the total protein output measured in grams. Again, I use FAO's conversion rates for transformation.

All measures of output are converted into yields per hectare. On household level (over all plots) yields are computed as geometrical averages of yields by plots, where weights are the corresponding land shares. [12] Yields summary statistics are reported in Table 2.5.

### 2.3.3 Measuring weather conditions

I match farmers' locations on TA (traditional authority - third tier of government) level with the measurements of several weather conditions, which are likely to affect maize and legumes harvest - rainfall, rainfall variation, and temperature. All weather conditions are measured during the

[12] The rationale for geometric average are explained in section on Identification

Table 2.6: Weather indicators dataset: Descriptions

| Name | Full name | Provider | Measured indicators | Method | Cover. | Resolution | Start. from | Freq. |
|---|---|---|---|---|---|---|---|---|
| CMAP | Climate Prediction Center Merged Analysis of Precipitation | NOAA/OAR/ESRL PSD, Boulder, Colorado, USA | Precipitation | Satellite | world | 2.5x2.5 degrees | 1979 | 5 days |
| RFE | Rainfall Estimate | FEWS NET (Famine Early Warning System Network) | Precipitation | Satellite | SSA | 8x8 km | 1995 | 10 days |
| math | Mathematica WeatherData | Wolfram's Mathematica | Precipitation, temperature | Weather stations | world | 17000 weather stations | 1949 | 0.1-10 days |

*Note:* Web-addresses: CMAP - *www.esrl.noaa.gov/psd/*, RFE - *earlywarning.usgs.gov/fews/index.php*, math - *reference.wolfram.com/language/ref/WeatherData.html*

growing season (January to March or April).

Measurement of historical rainfall (and hence rainfall variation) in SSA is not a trivial undertaking, as the weather stations are sparse and using outdated equipment, and so most measurements are proxies and interpolations. To reduce the influence of a measurement error in my regressions, I collect three rainfall measures: CMAP, RFE, and MATH. They are described in Table 2.6. CMAP and RFE use satellite-based data and algorithms to derive the precipitation data (see Xie & Arkin (1997) for more details), while MATH uses the data from weather stations. To measure rainfall and rainfall variation at a traditional authority X I take the corresponding indicators from the data point (point in the grid for CMAP and RFE, weather station for MATH), which is the closest to the geographical center of X.[13]

The summary statistics for the weather indicators are reported in Table 2.7. The average rainfall in Malawi in 2003-09 is around 180 mm per month from January to March, and the average temperature is around 23°C. The correlations between all three rainfall measures are positive, which is reassuring, but never greater than 0.4, which suggests non-trivial measurement error. Indeed, CMAP and RFE rely on indirect rainfall measurement, and their grid is relatively coarse. MATH relies on often untrustworthy data from the weather stations, which are also very scarce in the region. For instance, the mean distance from a TA to the closest weather station is 28 km

[13] The coordinates of geographical center of X are derived as: latitude=(max(latitude)+min(latitude))/2; longitude=(max(longitude)+min(longitude))/2

Table 2.7: Weather indicators: Summary statistics

| | count | mean | sd | p10 | p50 | p90 |
|---|---|---|---|---|---|---|
| CMAP rainfall, January-March, mm per day | 18293 | 5.75 | 0.91 | 4.44 | 5.68 | 6.76 |
| CMAP rainfall variance, January-March | 18293 | 4.56 | 1.17 | 3.41 | 4.55 | 6.36 |
| RFE rainfall, January-March, mm per day | 18293 | 6.44 | 1.13 | 4.78 | 6.70 | 7.76 |
| RFE rainfall variance, January-March | 18293 | 3.91 | 1.04 | 2.75 | 3.84 | 5.35 |
| math rainfall, January-March, mm per day | 18293 | 6.44 | 1.39 | 4.93 | 5.96 | 8.98 |
| math rainfall variance, January-March | 18293 | 3.07 | 1.26 | 1.65 | 2.84 | 4.92 |
| math temperature, January-April, C | 18293 | 22.38 | 2.48 | 20.18 | 21.38 | 27.46 |

*Note:* All summary statistics are calculated at a household level. For the rainfall and temperature - the average over the indicated period; for the rainfall variance - the variance over the indicated period.

(the distance to the closest CMAP grid data point is 95 km). Combining several weather indicators measures helps to reduce the effect of the measurement error in the estimation, as one can see in the next sections.

## 2.4 Identification strategy

### 2.4.1 General framework

Assume, the general production function of an average Malawian subsistence farmer is Cobb-Douglas with two essential factors needed for non-zero output - labor and (seeded) land:

$$Y_{ixt} = A_{ixt} L_{ixt}^{\gamma_1} D_{ixt}^{\gamma_2} e^{\epsilon_{ixt}}, \tag{2.1}$$

where $Y_{ixt}$ is the output of household $i$ using technology $x$ at time $t$, $L$ is labor, and $D$ is land. $A$ is the total factor productivity, which includes all the other factors that may affect output, including use of fertilizer and weather conditions. $e^{\epsilon_{ixt}}$ is the multiplicative productivity shock, which is independent of other factors.

Dividing both sides of (2.1) by $D$ we get the production function in per unit of area terms. Taking logarithm and denoting per unit of area terms by the corresponding small letters, and assuming that the production function is constant returns to scale, we get:

$$\log y_{ixt} = \log A_{ixt} + \gamma_1 \log l_{ixt} + \epsilon_{ixt} \tag{2.2}$$

Assume $\log A_{ixt}$ has the following form:

$$\log A_{ixt} = \beta_0 + \beta_1 T_{ix} + \beta_2 V_{ixt}^{(2)} + \beta_3 V_{it}^{(3)} + \beta_4 V_{ix}^{(4)} + \beta_5 V_i^{(5)}, \tag{2.3}$$

where $T$ is the technology dummy - monoculture maize vs. crop diversification. Technologies use varies by farmer. Also, farmers can use both technologies at the same time (on different plots).[14] $V^{(2)}_{ixt}$ represents factors that vary by time and by technology for each household: fertilizer use, farmer's experience and knowledge about the technology, etc. $V^{(2)}_{ixt}$ can also include interactions of $T_{ixt}$ with variables like weather or fertilizer. $V^{(3)}_{it}$ represents a wide group of factors that vary by household and by time, but not by technology: current farmer's abilities and education, health, demography of a household, preference for consumption and leisure, current wealth and credit constraints, weather conditions, soil quality, geology, and in general any contemporary farmer's or environment characteristics. $V^{(4)}_{ix}$ and $V^{(5)}_{i}$ represent factors that are time-constant, and only vary by household and technology, or only by household: for instance, technology's labor or capital intensity, its other time-constant characteristics (potential for drought resistance, needs for certain composition of soil, etc.)

Many of the factors included in $V$'s are not measured in the dataset, or simply not observable, and can be correlated with the farmer's choice of technology. "Two-dimension" panel structure of the dataset - the fact that I observe farmers both over time and over multiple technologies - helps to put many potential biases in check.

First, since we observe same farmers using different technologies at the same time, we can include household-time fixed effects to effectively control for $V^{(3)}_{it}$ and $V^{(5)}_{i}$ - a wide set of technology-constant but potentially time-changing factors. In the household-time fixed effects regression the yields are compared across plots using different technologies, within the same household and year.

Second, we can also include household-technology fixed effects, which will allow us to control for $V^{(4)}_{ix}$ - the set of factors, which vary by household and technology but constant over time. Note that the technology dummy $T_{ix}$ in (2.3) is one of these factors, so $\beta_1$ cannot be identified if we include household-technology fixed effects. Only farmers who use both technologies at the same time and for more than one period would be included in a regression with both household-time and

[14]In the data there are cases when same technology is used on more than one plot. Since I cannot identify plots over time, I simply join all plots that use the same technology into a single plot-technology observation.

household-technology fixed effects.

An alternative way to control for $V_{ix}^{(4)}$, but not for $V_{it}^{(3)}$, is to use household fixed effects in a traditional "one-dimension" household-year panel, where all plots of the household are assembled and analyzed as a single field.

### 2.4.2 Main specification and summary of identification issues

The main specification that I use for (2.3) is the following:

$$\log y_{ixt} = \beta_0 + \beta_1 T_{ix} + \beta_{31} * W_{it} + \beta_{21}(T_{ix} * W_{it}) + \\ + \beta_{22} * F_{ixt} + \beta_{23}(T_{ix} * F_{ixt}) + \epsilon_{ixt} \quad (2.4)$$

Here $x$ is either maize (M) or maize-legume intercrop (ML), see definition in Section 2.3.1. $y$ is yield as defined in Section 2.3.2. $T$ is one if $x$ is ML, and zero otherwise. $W$ is a vector of weather indicators: rainfall, rainfall variation within growing season, temperature, temperature squared, and temperature cubed. All weather indicators are demeaned using the sample averages (reported in Table 2.7).[15] The latter two indicators are added to test possible non-linearities in the response of maize yield to temperature, as recently argued in Lobell et al. (2011).[16,17] $T * W$ is the interaction of the technology dummy and weather indicator. $F$ is a vector of fertilizer use indicators: fertilizer applied and fertilizer applied squared to reflect possibility of diminishing returns. This variable is also demeaned using the sample average (reported in Table 2.5).[18] The demeaning of weather and fertilizer indicators is done to provide for a reasonable interpretation of $\beta_1$ - an effect of using

[15] Sample averages are very close to historical average in 2001-2010

[16] Temperature indicator is first demeaned and then squared and cubed, so that temperature squared and temperature cubed are both zero at sample average

[17] As a robustness check, I also try non-linear (quadratic and cubic) specifications of rainfall and rainfall variance. The results are qualitatively similar to the main specification, even though I do find statistically significant non-linear effects of rainfall and its variance on the yields. In this paper I decided to stick to conventional specifications from the literature, while exploring the non-linearities in more detail is left for further research

[18] Like with the temperature, fertilizer indicator is first demeaned and then squared, so that $F^2 = 0$ at the sample average

ML versus M when weather and fertilizer use are at the sample averages (because at the sample averages W*T=F*T=0).

The coefficients of interest are $\beta_1$, $\beta_{21}$, and $\beta_{23}$. If $\beta_1 > 0$ then it means that, everything else equal, maize-legume intercrop is more productive than maize. If ML is also more weather shocks-absorbing than M then I expect that $\beta_{21}$ and $\beta_{31}$ would have the opposite signs, and $\beta_{21} < \beta_{31}$.[19] For example, rainfall is likely to affect yield positively. For M the effect is $\beta_{31} > 0$, while for ML the effect is $\beta_{31} + \beta_{21}$. It is also positive but smaller than $\beta_{31}$, which means that ML is less responsive to the weather shocks. The interpretation is similar in case of fertilizer use.

I run (2.4) using three different estimation techniques. The first one is pooled OLS: the unit of observation is household-technology, and no fixed effects are included. It means that household X growing maize and the same household growing maize-legume intercrop are considered two independent observations. The next two techniques are household-time FE and household-time and household-technology FE as described in the section above. There again the unit of observation is household-technology (i.e. up to two observations per household per year), but with the corresponding fixed effects included. [20]

I also use an auxiliary specification in a traditional "one-dimensional" household-year panel, where the observations are identified over household and year, and technologies (plots) are lumped

[19] Generally speaking, $\beta_{21}$ and $\beta_{31}$ are (1x5) vectors. The conjecture is that this relationship between the coefficients holds element by element

[20] Household-time FE are run the following way. The dataset is declared a panel with cross-section dimension identified by household-year i.d., and "time" dimension identified by technology. Then a simple FE regression is run. The household-time and household-technology FE are more complicated to run. First, I demean each variable by technology, i.e. I manually remove household-time fixed effects. For any variable $a$, denote $\ddot{a}$ its demeaned value, where mean is taken by technology for each household and year. Then $\ddot{a}_{ixt} = a_{ixt} - \sum_{x} a_{ixt}$. Households at any given year are dropped if they used only one technology at that year. Second, I declare the dataset of the demeaned variables to be a household-technology/year panel, and then run the FE regression. The standard errors are correct since I cluster them by household in every regression

together:

$$\log y_{it} = \alpha_0 + \alpha_1\theta_{it} + \alpha_{31} * W_{it} + \alpha_{21}(\theta_{it} * W_{it}) + \\ + \alpha_{22} * F_{it} + \alpha_{23}(\theta_{it} * F_{i1t}) + \epsilon_{it} \quad (2.5)$$

The specification is similar to (2.4), but instead of technology dummy $T_{ix}$ there is a land share of ML - $\theta_{it}$. $y_{it}$ is a total yield of a household $i$ at year $t$. To make $\alpha$'s in (2.5) be directly comparable with the $\beta$'s in (2.4) $y$ should be a geometric average of yields by the two technologies, and the weights are the corresponding land shares. [21]

Specification (2.5) is run using two estimation techniques. The first is traditional household FE. Now the unit of observation is household-year, all plots within the same household-year are lumped together. Household FE allow comparing same households over time. The second estimation method is Arellano-Bond GMM,[22] is added to reduce the bias in case the strict exogeneity assumption is violated in the household FE. The chances are that this assumption is indeed violated as most farmers in Malawi are credit constraint and the harvest they collect likely affects the next year's harvest through health, productivity, and ability to buy inputs.[23] Note that these factors are taken care of in household-time FE as they are common to all technologies used by the farmer within one year.

Table 2.8 summarizes the possible factors that affect yield and how various estimation methods cope with controlling for these factors. There are two threats to the identification of betas in (2.4) and (2.5). First, even though some estimation methods, especially household-time FE and household-technology FE, allow to control for plenitude of factors, none of the methods does it all. Factors like (potentially time changing) technology-specific education, experience, labor and seed input are likely to be correlated both with the yield and with the choice of technology. The second threat is that the household-time FE and household-technology FE, the two most robust methods, are by definition applied only to a sample of farmers, which use both technologies at the same time.

[21] For details, see Section A.2 in Appendix

[22] See Arellano & Bond (1991) for more details

[23] See Foster & Rosenzweig (2010)

Table 2.8: Summary of estimation techniques: Identification and potential biases

| Est. method: | pooled OLS | hhold-time FE | hhold-time & techn. FE | hhold FE | GMM |
|---|---|---|---|---|---|
| | **Identification comes from comparing...** | | | | |
| | plots | plots over hholds | plots over hholds and time | hholds over time | hholds over time |
| | **Factors (potential biases):** | | | | |
| Technology | + | + | + | + | + |
| Weather | + | + | + | + | + |
| Fertilizer | + | + | + | + | + |
| Geology | -/+[a] | +[b] | + | + | + |
| General abilities, education | - | + | + | -/+[c] | -/+[c] |
| Technology-specific education, experience | - | - | -/+[c] | -/+[c] | -/+[c] |
| Labor | - | -/+[d] | -/+[c,d] | -/+[c] | -/+[c] |
| Seed[e] | - | -/+[d] | -/+[c,d] | -/+[c] | -/+[c] |
| Credit constraints | - | + | + | - | + |
| Selection bias | no | yes | yes | no | no |

"+" - controlled for; "-" - not controlled for; "-/+" - partially controlled for

[a] To the extent geology is similar within region (when regional dummies are included)

[b] To the extent that farmers do not systematically subject one of the technologies to better local conditions. The (sparse) evidence from the surveys is that they are not. See Table A.4 in Appendix

[c] Only time-constant factors are controlled for

[d] Only factors common to both technologies: preference for leisure, labor supply

[e] I do have data on seed input by plot, but this data is likely unreliable. See Section 2.7 for details

The number of such farmers in the sample is quite large, but they are still minority and they could be a special crowd, which raises concerns about the sample selection bias.

In what follows I provide few remedies to the issues above. First, the five estimators that I use allow to control for different sets of factors, and hence potentially yield different magnitudes and directions of biases. So if all estimators produce similar results then the biases are likely to be small-scale.

Second, in the next section I provide some considerations about how farmers are likely to choose which technologies to use. The self-selection into different technologies, or their combination, is linked to the technology-specific factors, many of which I do not control for in my regressions - education, experience, labor, seeds. None of the methods I use consistently estimates the average treatment effect. But my argument is that, even if I do not control for some technology-specific factors, household-time fixed effects regression consistently estimates the likely effect of government policy changes (e.g. technology-specific education or input subsidy) if the changes are small. This is because the likely compliers to such policy are those who already use both technologies (e.g. the sample for household-time fixed effects regression) or those who are "close" to using the two.

For large policy changes my results are less reliable. The actual effect will depend on specific policies, and on systemic difference in labor and seeds requirements between the two technologies. In Section 2.2.4 I argue that the requirements are not too different. In addition, I control for the seed input as one of the robustness checks, and I find no qualitative change in results.[24] The time-constant labor and seed technology requirements are also controlled for in household-time and technology fixed effects, but this specification is not the main, because it does not allow to estimate $\beta_1$, and the estimation sample is too small to yield any statistically significant results.

### 2.4.3 Technology choice by farmers: Selection bias and omitted technology-specific factors

Consider a farmer who owns a unit of land. The choice is which agricultural technologies to use on this land. It can be one dominant technology, or it can be a combination of several technologies. All technologies are subject to weather (or other kinds of) risks, which are not perfectly correlated with each other. Like in finance, it is natural to hedge productivity risks by diversifying if two or more technologies are available. The likely reason why it often does not happen is that for some farmers the difference in net returns between the technologies is too large. The definition of a net return here is very broad, and takes into account all possible factors that affect farmer's choice. Farmer may not use a technology either because they know very little about how to use it, or they perceive it to be less productive and/or more risky than it actually is, or the technology's cost is prohibitively high. For example, the seed is not available at a local market. Such situation is not

[24] Seed input is not used in the main specification as the data is incomplete and likely unreliable. See Section 2.7

uncommon in Malawi. See Section 2.2.4.[25]

Technology use by farmers is not random, and it is affected by many variables, which I cannot include in my regressions. Therefore, neither household-time FE nor any other estimation method consistently identify the average population effect of technology on yield. Sample selection and technology-specific biases are not removed.

A question one may ask though is whether we actually need to identify the average population effect if our goal is to analyze government policies that intend to affect the current state of technology use. To assess the results of such government policy we need to identify the effect only on compliers with this policy rather then on the whole population. Unless the government can directly force particular farmers to use this or other technology, the set of compliers will also be subject to selection bias.

Suppose the government wants to induce farmers to use more of a certain technology. First, it can reduce/subsidize the cost of the technology. Second, the government can provide technical assistance/education. Both policies will automatically increase the technology's net returns for each farmer. The question is which farmers will comply with the reform and actually change their land use, i.e. increase the share of land used for the subsidized technology. The reform will first of all affect those farmers, who already use both technologies (subsidized one and alternative). This is because for these farmers the difference between the technologies is not too large, the perceived net returns are comparable. The group of diversifiers may also by joined by farmers, who used only alternative technology before, but where on the margin of using the subsidized technology. Some farmers may switch to specializing only in subsidized technology.

The effect of the subsidy reform on yield should then be very similar to what we estimate in household-time FE regression, in particular if policy changes are small. The compliers of the subsidy are those, for whom the change in technology's net return matters. These are the

[25] In Malawi, the incentives for the diversification by individual farmers or for simply growing a dominant crop (maize) are stronger, because the country-wide harvest insurance mechanisms are not available. Inter-regional or even inter-village trade is undeveloped too due to the lack of infrastructure

farmers with moderate perceived difference between technologies - those that already use both technologies or those that were on the margin before the subsidy was introduced. That is why it is their productivity rather than productivity of an average farmer in the population that is important at identifying the effect of the reform. For small policy changes, these are farmers which are targeted by the household-time FE.

## 2.5 Estimation results

The main estimation results are summarized in Table 2.9. For all nine columns the variables specification is always as in (2.4). My main dependent variable is log energy yield (see Table 2.5). The technologies I compare are M (maize) and ML (maize-legume intecrop) - see Table 2.2 for the definition. I use five main estimation methods, which are described in Section 2.4.2. In columns (1), (3), (5), and (7) the rainfall and rainfall variance measures are taken from CMAP (see Table 2.6 for more information). The possibility of a large measurement error in weather indicators was discussed above. I deal with this problem using the instrumental variable approach. In columns (2), (4), (6), (8), and (9) CMAP measures of rainfall and rainfall variance are instrumented by the equivalent measures from RFE and MATH (see Table 2.6 for more information). The measurement error did prove to have a significant role in a number of specifications, especially for the household-time FE, so in what follows I stick to the results from IV estimations. As a robustness check, Table A.2 in Appendix reports the results when variables are added one by one.

The magnitudes and significance of the coefficients differ by the specification, but the overall results seem to be in favor of the maize-legume intercrop: everything else equal, ML seems to be more productive, and it seems to absorb shocks better. The consistency of the results throughout the estimation methods and specifications is reassuring, as each method uses different sample of farmers and potentially insulates against (or vulnerable to) different kinds of biases. Importantly, household-time FE, which is the preferred method according to the reasoning and the model above, is consistent with the rest of the results. Household-time and technology FE, which is even more

Table 2.9: Estimation results: Energy yield

| | pooled OLS | | hhold-time FE | | hhold-time & techn. FE | | hhold FE | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| | no IV | IV | no IV | IV | no IV | IV | no IV | IV | GMM |
| ML technology | 0.18*** | 0.18*** | 0.31*** | 0.27*** | | | 0.15*** | 0.05 | 0.26*** |
| (dummy or landshare) | (0.02) | (0.02) | (0.05) | (0.05) | | | (0.05) | (0.06) | (0.07) |
| rainfall, cm | 0.36*** | 0.40** | | | | | 0.82*** | 2.30*** | 2.07*** |
| per day | (0.13) | (0.20) | | | | | (0.27) | (0.40) | (0.32) |
| ML X rainfall | -0.69*** | -0.80*** | -0.05 | -1.25* | -1.95 | -2.33 | -1.11*** | -2.42*** | -2.04*** |
| | (0.19) | (0.30) | (0.39) | (0.70) | (1.73) | (2.51) | (0.41) | (0.69) | (0.53) |
| rainfall | 6.03*** | 5.87** | | | | | 1.50 | -30.60*** | -4.58 |
| variance, daily | (1.19) | (2.39) | | | | | (2.79) | (8.02) | (3.07) |
| ML X rainvar | 9.45*** | 10.33*** | 0.47 | 8.69* | 28.89** | 20.57 | 20.58*** | 42.51*** | 14.69*** |
| | (1.46) | (2.39) | (3.10) | (4.63) | (14.07) | (24.07) | (3.19) | (5.11) | (4.45) |
| temperature, C | 2.35** | 2.36** | | | | | 11.10*** | 16.68*** | 23.55*** |
| /100 | (1.02) | (1.02) | | | | | (2.87) | (3.35) | (3.15) |
| temperature | 1.46*** | 1.47*** | | | | | 0.75 | -2.08* | 3.86*** |
| ^2 /100 | (0.29) | (0.32) | | | | | (0.75) | (1.14) | (0.60) |
| temperature | -0.33*** | -0.33*** | | | | | -0.60*** | -0.57*** | -1.15*** |
| ^3 /100 | (0.07) | (0.07) | | | | | (0.18) | (0.19) | (0.17) |
| ML X temp | -1.97 | -2.00 | -5.59** | -7.41*** | -24.99*** | -23.25** | -9.77*** | -14.55*** | -16.63*** |
| | (1.25) | (1.25) | (2.66) | (2.84) | (8.81) | (9.15) | (3.18) | (3.37) | (4.18) |
| ML X temp_sq | -0.68* | -0.66 | -1.32** | -1.06 | 3.08 | 2.15 | -1.66** | 0.02 | -3.80*** |
| | (0.39) | (0.42) | (0.66) | (0.72) | (3.15) | (4.21) | (0.81) | (0.90) | (0.99) |
| ML X temp_cub | 0.12 | 0.12 | 0.33** | 0.31* | 1.34** | 1.19** | 0.69*** | 0.53** | 1.06*** |
| | (0.09) | (0.09) | (0.15) | (0.17) | (0.58) | (0.59) | (0.20) | (0.21) | (0.23) |
| fertilizer, | 3.37*** | 3.38*** | 2.93*** | 2.86*** | 0.28 | 0.32 | 3.00*** | 2.98*** | 2.80*** |
| tonnes per ha | (0.14) | (0.14) | (0.44) | (0.45) | (0.63) | (0.63) | (0.28) | (0.29) | (0.37) |
| fertilizer | -2.50*** | -2.50*** | -1.34* | -1.22 | 0.47 | 0.41 | -3.57*** | -3.52*** | -2.93*** |
| ^2 | (0.35) | (0.35) | (0.75) | (0.85) | (1.36) | (1.37) | (0.69) | (0.70) | (0.76) |
| ML X Txfert | 1.06*** | 1.05*** | -0.99** | -1.05** | -1.61 | -0.99 | 0.91** | 0.97** | 0.29 |
| | (0.19) | (0.19) | (0.47) | (0.49) | (1.80) | (1.97) | (0.39) | (0.41) | (0.57) |
| ML X Txfert_sq | -1.57*** | -1.56*** | 2.90** | 3.12* | 3.62 | 1.58 | -0.98 | -1.01 | 0.35 |
| | (0.48) | (0.48) | (1.41) | (1.67) | (7.14) | (7.72) | (1.12) | (1.15) | (1.56) |
| L.log_energy_yield | | | | | | | | | 0.08** |
| | | | | | | | | | (0.03) |
| Constant | 14.60*** | 14.60*** | 14.30*** | 14.30*** | -0.02 | -0.02 | 14.69*** | 14.81*** | 13.14*** |
| | (0.02) | (0.02) | (0.02) | (0.02) | (0.04) | (0.04) | (0.05) | (0.06) | (0.46) |
| Observations | 13305 | 13305 | 1676 | 1676 | 196 | 196 | 5696 | 5696 | 3095 |
| R-squared | 0.17 | 0.17 | 0.17 | | 0.16 | | 0.18 | | |
| R-squared overall | | | 0.13 | 0.13 | 0.03 | 0.03 | 0.16 | 0.10 | |

*Note:* Standard errors (clustered by household) in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable in each regression - log_energy_yield. Technologies - M (maize) vs. ML (maize-legume intercrop). OLS, household FE, and GMM all include time fixed effects, which are also effectively controlled for in household-time FE. The results without time fixed effects are similar. GMM regression assumes all variables except time fixed effects potentially violate strict exogeneity assumption (so all variables are instrumented).

robust,[26] produces similar signs and magnitudes of the coefficients, but very few of them are significant. The effective sample for this method is very small, while the weight of a measurement error (in yield in particular) soars after "‘double demeaning"’.[27]

The coefficient on ML - dummy in columns (1)-(6) and land share in columns (7)-(9) - is positive and statistically significant, except for the household FE.[28] This is, in particular, true for the househodel-time FE, which is the preferred estimation method. The estimates in IV regressions - columns (2), (4), and (9) - tell that calorie yields on plots with ML are around 15-30% higher than the yields of M.[29] Note that the coefficient on ML cannot be identified in household-time and technology FE, because the technology dummy does not change with time.

The results on the fertilizer use and its interaction with ML are also suggestive, but with a couple of caveats. Even though the coefficients on *fertilizer* and *fertilizer^2* are of the opposite signs, the overall response of the yield to the fertilizer is positive for both technologies in all specifications and for reasonable quantities of the fertilizer. [30] Even though the coefficient on *fertilizer* is positive, the response to the fertilizer seems to flatten out with more fertilizer applied, as can be seen from negative coefficients on *fertilizer^2*. The results of different estimation methods diverge on the interaction of ML and the fertilizer. Household-time FE says that ML is less responsive to the fertilizer - the coefficient on *ML X fert* is negative, while in OLS and GMM the coefficient is positive. In case of household-time FE the difference in the responsiveness is not large though - ML remains significantly more productive than M at any reasonable quantity of fertilizer applied, as shown in Figure 2.4.[31] The gap between the two technologies narrows down to around 15%

[26] In particular, household-time and technology FE controls for time-constant technology-specific labor and seed inputs

[27] See Griliches & Hausman (1986) for details on measurement error in FE regressions

[28] It is also not statistically significant in some specifications of the GMM regression. See Table A.14

[29] All other variables in the regressions are demeaned, so the coefficient on ML means the effect of using ML at sample average fertilizer use and weather

[30] 500-600 kg per ha in most specifications, 200-300 kg per ha in case of ML in GMM specification

[31] In GMM specification the difference is also positive at reasonable quantity of fertilizer, but not always statistically significant. See Figure A.1

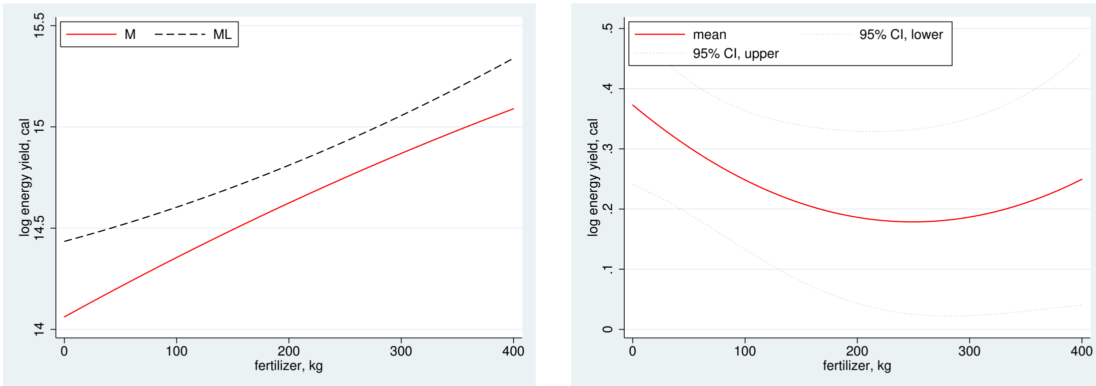


Figure 2.4: Estimated response to fertilizer in M vs. ML (left); Difference between ML and M with confidence interval (right)

*Note:* Based on the estimation in Table 2.1, column 4 - household-time FE. Figures A.1 and A.2 in Appendix use other estimates and yield measures. Standard errors for the confidence interval are bootstrapped.

at 200 kg of fertilizer per hectare, but then starts to widen again. Starting from a sample average of around 80 kg per hectare, the gap between ML and M gets wider if fertilizer is used less, so according to the household-time FE ML would be more effective at absorbing adverse fertilizer use shocks for most of the farmers. The efficiency of the fertilizer for both technologies is around 13.7-14 thousand kCal per kg of nitrogen.[32] The flip side of being more resistant to the fertilizer use shocks, the efficiency of fertilizer for ML is generally lower than for M - 4% lower at the sample average of 80 kg per tonne.

Rainfall and rainfall variance, as expected, are important determinants of harvest in Malawi, and ML seems to be an effective absorbent of the weather shocks. Coefficient on rainfall is positive and significant in OLS and GMM (in household-time FE it is not identified). Controlling for the average rainfall, rainfall variance affects the yield negatively. The coefficients on ML interactions with rainfall and rainfall variance have the opposite signs, which means that ML likely smooths the weather's effect on yields.[33] Note that, as mentioned above, coefficients on rainfall

[32] A calorie equivalent of around 3.8-3.9 kg of maize. Assuming 30% of nitrogen in a kg of fertilizer. The efficiency is calculated at the average fertilizer use among all farmers - 80 kg per hectare.

[33] For example, using the results from column 9, a decrease in rainfall of 0.1cm per day would decrease average M yield by about 2.07*0.1=20%, but will reduce average ML yield by only (2.07-

and rainfall variance are not identified in household-time FE (my preferred specification), so the implicit assumption here is that their signs and magnitudes are similar to those in OLS or GMM; otherwise the interpretation of the coefficients on ML interactions can be different. [34] This result is consistent across all specifications, although the magnitudes of the coefficients differ, and in some specifications the result is not significant.[35] For the household-time FE it proved to be crucial to use IV for my rainfall measures. This result means that ML is less responsive to the weather shocks, droughts and sporadic rainfall in particular.

The relationship between the yield and the temperature is more complex. At the current range of observed temperatures in Malawi, the relationship is positive both for M and ML, but becomes negative after a threshold of around 26°C for M and 29°C for ML. This result should be taken with caution as the maximal average temperature observed in Malawi over the estimation period is 28.5°C, but it is consistent with Lobell et al. (2011). As with rainfall and rainfall variance, ML seems to be better than M at absorbing the temperature shocks, although it is harder to see just looking at the coefficients. Figure 2.5, which is based on the estimation in column (9), demonstrates this result. ML's response to the temperature follows a similar path to M: it first decreases, then starts to increase, and then decreases again. But ML's response curve is much smoother than M's curve.[36]

---

2.04)*0.1=0.3%. In some specifications the absolute value of the coefficient on ML interaction with the rainfall is actually larger than the coefficient on rainfall. Taken at a face value, this means that ML responds negatively to rainfall. This result is never statistically different from zero, and it is hardly plausible, but even if it is true, ML still smooths the weather effects, because the absolute value of the coefficient on ML interaction is never twice as high as the coefficient on rainfall

[34]OLS and (especially) household FE and GMM are frequently used in the literature to estimate the effect of random weather shocks on yield, which adds to plausibility of the assumption. See Section 2.6 for details. Important caveat here is that the coefficients are effectively estimated together with the technology use, which is endogenous

[35]Household-time and technology FE. Also some specifications of the GMM regression (see Table A.14)

[36]Note that yield-depressing effect of extremely high temperature does not show up in OLS

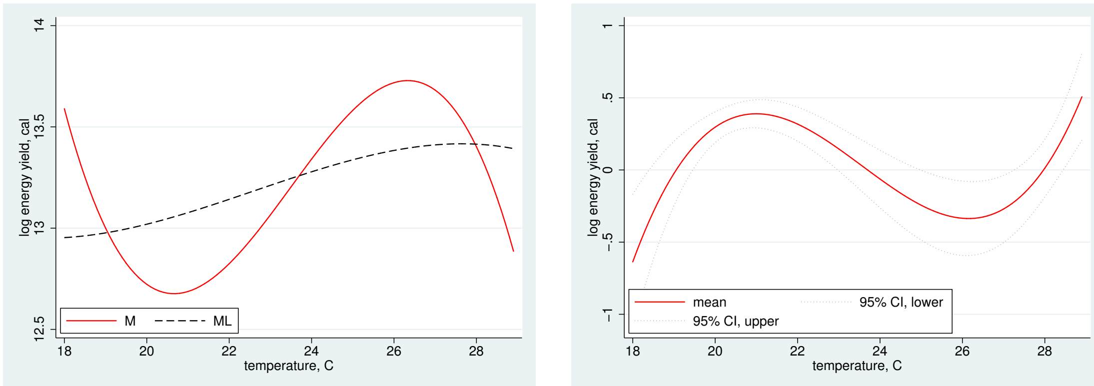


Figure 2.5: Estimated effect of temperature on yield: M vs. ML (left); Difference between ML and M with confidence interval (right)

*Note:* Based on the estimation in Table 2.1, column 9 - GMM. Using household-time FE estimates where possible yields similar results (see Figure A.3 in Appendix). The maximal average temperature observed in Malawi over the estimation period is 28.5°C. For higher temperatures the predictions are out of sample. Standard errors for the confidence interval are bootstrapped.

## 2.6 Economic significance of results: Macroeconomic scenarios

The previous section discusses the economic rationale and statistical significance of the results, but how economically significant are they? Are plausible weather or fertilizer use shocks macroeconomically important for Malawi, and how effective is ML in smoothing them? I build a range of shock scenarios for Malawian agriculture to answer these questions. For each scenario I consider three stances of reform. The first stance is a baseline - the land used by M and ML remains as it is was in 2009. The second stance is a program - the government implements a subsidy or education program that increases land used for ML by 10 percentage points. The third stance is a hypothetical situation, when all land is used for ML. Note that the coefficients I identify in Table 2.9 are likely to be close to the true coefficients only in the case of the second stance (program), when changes in the land use are not that large. The results for the third stance are less robust and for demonstration only.

In each scenario I look at two main indicators. The first is the average change in yields as a result of the shock. The energy (or grain) loss of the yield can further be expressed in terms of GDP in case of a hypothetical situation that the government decides to compensate the loss by importing an equivalent amount of maize from abroad. As of 2013, 84% of Malawian households are rural, and of them around 86% grow either maize or maize-legume intercrop. So for example, at the 2013 Malawian kwacha to US dollar exchange rate, and the price of maize at 35 US cents per kg, a 5% average drop in yield translates into 1.15% of GDP of compensation cost.[37]

The second indicator that I look at is the change in the share of self-sufficient households, i.e. households that produce enough calories to cover their own energy needs throughout the year. At the FAO's recommended minimal personal dietary requirement of 1700 kCal per day,[38] an average self-sufficient Malawian household is supposed to harvest about 6.2 mln kCal per year.[39] In 2009 only 17% of households reached self-sufficiency. Many households have alternative sources of revenue: growing other crops (e.g. tobacco), animal farming, or supplying labor ("'ganyu"') to richer households. However, changes in food self-sufficiency are likely to be correlated with the changes in food security.[40] The results on the self-sufficiency must be taken with caution, as the implicit assumption that I make here is that the impact of the shock is similar for all households, i.e. the shock does not affect the form of the distribution curve, it only shifts it.

I consider four medium-term and three long-term scenarios. Medium-term scenarios model the effect of four large but plausible adverse shocks, one shock per scenario: negative shocks to fertilizer use, rainfall, and temperature, positive shock to rainfall variance. The size of the fertilizer use shock is 10% decline from the current use. This shock can result from a corresponding increase in fertilizer price, which is not uncommon at all in the last two decades,[41] given that the overwhelming majority of Malawian farmers are credit-constraint.[42] The shock can also be

[37] A cost which carves into already shaky external position of the country

[38] FAO (2008)

[39] 1700kCal per day is the average requirement per person, so in households with many children the requirement may be lower

[40] Besides most of the alternative sources of revenue for Malawian households are likely to be pro-cyclical, i.e. move together with the staple crop yield. For example, adverse weather shock impacts not only the particular farmer, but the whole village, which likely reduces demand for ganyu labor.

[41] According to National Agricultural Statistics Service, USDA, the average absolute annual change in the price of Urea fertilizer in 1994-2013 is 24%

[42] In 2009 only 6% of them used loans to buy agricultural inputs

the result of a change in the government fertilizer subsidy, which is subject to various fiscal risks (e.g. political instability and hence volatile foreign aid, relatively high level of external debt). The weather shocks are equal to half average standard deviation from the historical (2004-2009) mean, where average is taken over traditional authorities. This is roughly 30th percentile of the corresponding historical distribution.[43]

The long term scenarios involve only weather shocks, and are based on the long-term climate projections for Malawi summarized in IPCC (2007). I consider two standard scenarios: a "‘mild"’ optimistic B1, median of the nine climate models reported, and less optimistic A2, 10th percentile of the nine models (ensemble low). I also consider the same A2, supplemented by an assumption of severe positive temperature and negative rainfall shocks (which are, however, not that unrealistic in this scenario),[44] to test the effect of extreme weather conditions.

I first construct the baseline, i.e. when the technology use is the same as in 2009 (columns 1 and 4 of Table 2.11). To do this I estimate the effect of weather on yield in the baseline, by employing the framework used in Deschenes & Greenstone (2007), Deschenes & Greenstone (2011), Fisher et al. (forthcoming), and Burke et al. (2011). I simply fit the log of household yield to the weather indicators - rainfall, rainfall variance, and temperature.[45] As in my main estimation results, I instrument rainfall and rainfall variance by alternative measures to get rid of the measurement error. In Table 2.10 I report the results using OLS, FE, and GMM, but in my projections I rely on GMM. As argued by the papers above, OLS does not control for various long-term factors such as ability of farmers to adapt to slow changes in the weather, whereas FE supposedly provides the reaction of yield to a short-term weather shock. At the same time, and as discussed above, it is

[43]If the distribution is close to normal

[44]The exact shock (temperature or rainfall) is two standard deviations of the 90th percentile of TAs distributions. In other words, the shock has a 5% chance to materialize in at least 10% of TAs

[45]I am reluctant to use the GMM estimates from Table 2.9 because they are likely to be inconsistent. At the same time, the household-time FE, my preferred specification, does not identify the coefficients on weather for M users. Therefore, the most reliable option in this case is likely to estimate the regression of yield on weather variables without controlling for the technology choice. At the same time, household-time FE does allow estimation of coefficients on both fertilizer and its interaction with the technology dummy, so I use this specification to calculate the baseline for the fertilizer shock

Table 2.10: Energy yield and weather shocks: Mixed technologies

| | pooled OLS | | hhold FE | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| | no IV | IV | no IV | IV | GMM |
| | b/se | b/se | b/se | b/se | b/se |
| rainfall, cm | 0.55*** | 0.18 | 0.63*** | 1.77*** | 1.72*** |
| per day | (0.10) | (0.16) | (0.22) | (0.32) | (0.31) |
| rainfall | 0.32 | -14.51*** | -6.62*** | -43.75*** | -4.22* |
| variance, daily | (0.77) | (1.34) | (1.75) | (3.31) | (2.48) |
| temperature, C | 2.67*** | 3.16*** | 6.67*** | 10.33*** | 16.06*** |
| /100 | (0.66) | (0.67) | (1.85) | (2.05) | (1.93) |
| temperature | -0.24 | -1.37*** | -2.08*** | -7.13*** | 2.02*** |
| ^2 /100 | (0.22) | (0.23) | (0.58) | (0.73) | (0.54) |
| temperature | -0.06 | 0.13*** | 0.02 | -0.29** | -0.57*** |
| ^3 /100 | (0.05) | (0.05) | (0.12) | (0.14) | (0.11) |
| fertilizer, | 3.42*** | 3.51*** | 2.87*** | 3.77*** | 4.89*** |
| tonnes per ha | (0.10) | (0.10) | (0.22) | (0.25) | (1.09) |
| fertilizer | -2.65*** | -2.87*** | -3.17*** | -4.41*** | -13.84** |
| ^2 | (0.26) | (0.25) | (0.58) | (0.63) | (6.50) |
| L.log_energy_yield | | | | | -0.04 |
| | | | | | (0.06) |
| Constant | 14.55*** | 14.59*** | 14.59*** | 15.08*** | 14.93*** |
| | (0.01) | (0.01) | (0.04) | (0.06) | (0.89) |
| Observations | 13305 | 13305 | 5696 | 5696 | 3095 |
| R-squared | 0.12 | 0.07 | 0.06 | | |
| R-squared overall | | | 0.08 | 0.00 | |

*Note:* Standard errors (clustered by household) in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable in each regression - log_energy_yield .

important in case of Malawi to take previous period's harvest into account, since many farmers are credit-constraint. GMM should take care of this. The results are as expected. They repeat the pattern of Table 2.9, and the magnitudes of the coefficients are mostly between those of M and ML users in Table 2.9.

The coefficients from Table 2.10, column 5 are then used to construct the baseline (column 1 of Table 2.11), where shocks are departures from the historical averages or current 2009 use in case of fertilizer. The baseline for the fertilizer shock scenario is constructed using coefficients from Table 2.9, household-time FE, since this (preferred) specification does allow estimation of coefficients on both fertilizer and its interaction with the technology dummy (unlike in case of weather variables).

The program (used for columns 2,3 of Table 2.11) is then constructed in the following way:

$$\log y_{prog} = \log y_{bas} + \Delta T_{ixt} * (\beta_1 + \beta_{21} * W_{it} + \beta_{23} * F_{ixt}), \quad (2.6)$$

where $\beta$'s are taken from Table 2.9, household-time FE. $y_{bas}$ is the corresponding baseline value,

$\Delta T_{ixt}$ is the change in the use of ML intercrop: $\Delta T_{ixt} = 0.1$ in the second stance of reform (columns 2 and 5 of Table 2.11), and $\Delta T_{ixt} = 0.52$ in the third stance of reform (columns 3 and 6 of Table 2.11). The corresponding after-shock values of W and F are used for each shock scenario.

To calculate the self-sufficiency values (for columns 4,5,6 of Table 2.11) the corresponding yields are transformed into harvests at a household level and then compared to the minimal dietary requirements.

Table 2.11 does not report $y_{prog}$ or $y_{bas}$ themselves, but the changes. In raw 1 of the table (line *No shocks*), changes are with respect to the baseline; in the rest of scenarios changes are with respect to the corresponding cell in the *No shocks* line. To summarize, for example the number in column 1, raw 3 (rainfall shock) is the percent change of $y_{bas}$ in case of reduced rainfall compared to $y_{bas}$ in case of no shock, both are means across all households from 2009 panel of the survey. In column 2 of raw 3, the number is percent change of $y_{prog}$ in case of reduced rainfall compared to $y_{prog}$ in case of no shock (the one computed for column 2, row 1).[46]

The results are reported in Table 2.11. The first line compares the baseline and the program without shocks. A 10 percentage points increase in ML land share is likely to increase average energy yield by 2.4%, and increase the share of self-sufficient households by 0.66 percentage points (around 3.8% increase). ML smooths fertilizer shocks, although not so effectively as it does with the weather shocks. A 10% fertilizer use shock reduces yield by 2.17% in the program, and 2.18% in the baseline. Note that the program's yield with the fertilizer shock is higher than the baseline's yield without the shock. This means, for example, that a cut to the government fertilizer subsidy program, which causes a 10% drop in fertilizer use, would be yield-neutral if simultaneously a land share of ML increases by around nine percentage points. At the same time, the agricultural system would become more resistant to weather and fertilizer shocks.

[46]This means for example that a rainfall shock will reduce average yield by 3.8% in case of baseline, and by 3.4% in case of second stance of the reform (increase of ML use by 10 p.p. The overall effect of the reform on yield in case of rainfall shock is approximately the sum of -3.4% and +2.4% (column 2, raw 1), which is around -1%. This is an improvement of around 2.8% compared to the baseline - 2.4% comes from the average increase in productivity, and 0.4% comes from smoother reaction to the shock

Looking at the difference between the baseline and the program, ML is more effective at absorbing weather shocks rather than fertilizer shocks. The rainfall shock, which is particularly harmful,[47] reduces the baseline's average yield by 3.84%, which would cost 0.9% of GDP to compensate. Using more ML, as in program, does not eliminate the shock completely, but reduces its impact by 0.4 percentage points, which saves government around 0.1% of GDP.

The long-term climate projections are actually quite favorable for Malawi, though not in all aspects. The mild scenario B1 predicts a moderate increase in rainfall (8mm per month), substantial increase in temperature (+2.1°C), but also high rainfall variance. The less favorable scenario A2 foresees even bigger increase in temperature, which is still likely to be favorable for the yield, but lower rainfall and much higher rainfall variation. With these projections, yields are likely to go up both in the baseline and the program. However, the increase is larger in case of the program - a better ability to withstand adverse shocks in rainfall and increased rainfall variation allows better harnessing the potential of increased temperatures. Although the average predictions are favorable, and the median projected temperature increase is still within the range where the yield responds positively, there is now a non-zero probability of an extreme event (e.g. increase in temperature of 8°C from the current average level), which is likely to devastate the agricultural sector, as demonstrated in the last two lines of Table 2.11. As before, more use of ML helps to mitigate the catastrophe.[48]

## 2.7 Robustness checks

I provide several robustness checks of my main results. First, as mentioned in Section 2.5, I check the results using alternative measures of yield: cash yield, maize equivalent yield (Liu-Myers), grain or protein yield. The results are similar when it comes to the signs of coefficients, although the magnitudes differ for some specifications (Table A.3).

[47] Note that the shocks are of the same magnitude if measured in standard deviations from the mean

[48] Note that this result relies on extrapolation of the temperature outside of historically observed data, so it shouls be taken with caution

Table 2.11: Effect of shocks on energy yield under different scenarios

| Scenario | %Δ average yield | | | Δ self-sufficiency | | |
|---|---|---|---|---|---|---|
| | Baseline | Δ ML +10p.p. | all use ML | Baseline | Δ ML +10p.p. | all use ML |
| No shocks | . | 2.428 | 13.711 | . | 0.661 | 4.187 |
| **Medium term scenario** | | | | | | |
| Fertilizer use: -10% | -2.181 | -2.172 | -2.316 | -0.331 | -0.275 | -0.165 |
| Rainfall: -0.228 mm per day | -3.840 | -3.480 | -2.042 | -1.212 | -0.826 | -0.331 |
| Rainfall variance: + 0.314 | -1.316 | -0.958 | 0.464 | -0.441 | -0.496 | 0.661 |
| Temperature: -0.307 degrees | -0.958 | -0.873 | -0.585 | -0.220 | -0.110 | 0.110 |
| **Long term scenario** | | | | | | |
| Scenario B1, median | 14.205 | 14.232 | 14.366 | 4.959 | 5.289 | 4.077 |
| Scenario A2, ensemble low | 3.160 | 4.432 | 10.149 | 2.314 | 2.755 | 3.030 |
| Scenario A2, ensemble low, ext. temperature | -23.924 | -22.289 | -14.603 | -5.014 | -4.738 | -4.463 |
| Scenario A2, ensemble low, ext. drought | -31.302 | -28.362 | -14.418 | -7.934 | -6.777 | -3.140 |

*Note:* Three states of reform considered: *Baseline* - land shares of M and ML remain as they were in 2009; Δ *ML +10p.p.* - land share of ML increases by 10 p.p.; *all use ML* - land share of ML is 1. *No shocks* - computed are changes with respect to the baseline; in the rest of scenarios changes are with respect to the corresponding cell in the *No shocks* line. Seven shock scenarios considered. *Medium term scenarios*: negative shocks on fertilizer use, rainfall and temperature, positive shock on rainfall variance. The size of the latter three is equal to half standard deviation - mean by TAs over 2004-2009. *Long term scenarios* are described in IPCC (2000), data from the World Bank Climate Change Data Portal. *Scenario B1, median* - median of nine climate models used in IPCC (2007) - rainfall +8mm per month, temperature +2.1°C, rainfall variation +20% from current. *Scenario A2, ensemble low* - 10th percentile of the nine climate models - rainfall -16mm per month, temperature +2.9°C, rainfall variance +37% from current. *Scenario A2, ensemble low, ext. temperature* - rainfall and rainfall variation as in scenario above, extreme scenario for temperature +8.1°C. Temperature shock is formed as the sum of projected expected increase (+4.5°C) and two projected standard deviations (90th percentile over TAs) (3.6°C). *Scenario A2, ensemble low, ext. drought* - temperature and rainfall variation are as in scenario *A2, ensemble low*, extreme scenario for rainfall (-87mm per month) - formed as the sum of projected expected fall (-16mm per month) and two projected standard deviations (90th percentile over TAs) (-53mm per month).

Second, I check different definitions of technologies, and various modifications of the sample or specification. In particular, I check how hybrid maize performs against the maize-legume intercrop. I also repeat estimation in case maize and legumes are the only crops grown on a plot, in case they are recorded as one of three most important crops on a plot, and in case their harvest share is significant (50, 75, 95, or 100%). I also test other definitions of M and ML. See Tables A.7 - A.11. To reduce measurement error I then drop 1st and 99th percentiles of farmers' distribution by yield. I also check what happens if I use only plots with conventional units of measurement, i.e. whether errors in transformation rates affect my results. I also drop farms, which are larger than 2 ha, as they are less likely to belong to subsistence farmers. I check if the results are robust when I control for total seed input or land size, and when I add quadratic and cubic terms of rainfall and rainfall variance.[49] Qualitatively the results do not change: ML intercrop does seem to be more productive

[49]For the rainfall and rainfall variance I do find a non-linear pattern similar to the one of temperature - significant and large coefficients on the cubic terms, which makes extreme weather

and more weather-resistant than monoculture maize, although in some specifications the results lose statistical significance. [50]

Another robustness check concerns the properties of the plots. It could be the case that farmers use the ML intercrop, which requires less fertilizer, at plots with a better soil quality and other properties, and then use fertilized maize at worse plots to make the most of them. This would bias the results on *ML technology* upwards. Note that the argument could go the other way around: farmers could use ML intercrop on the worse plots given the technology is productive anyway. In this case the bias would be downwards, which only reinforces my results.[51] My data on plot properties is very sparse, and therefore I cannot use it in my main specification. However, I can provide several arguments in the defense of my results. First, household-time FE effectively control for the soil quality at a region or district level, so what is uncontrolled for is the local variation, which is not likely to be high. Second, as Table A.2 reports, even without controlling for fertilizer, ML performs better than M. In case of household-time FE, the coefficients on *ML technology* in columns (7) and (8) are identical, which leaves no room for the "plot properties bias" as implied by the argument above. Third, I use an auxilliary regression, where the dependent variable is the choice of technology (0 - M, 1 - ML) and independent variables are various plot properties, for which the data are available: distance from dwelling, slope, texture, number of weedings performed.[52] The results are reported in Table A.4. They are inconclusive. If anything, I find that ML intercrop is used on plots with more adverse conditions - further away from a dwelling, and on plots with larger slope. The results are often economically insignificant. For example, increasing the distance of a plot from dwelling by 1 km increases the change of using ML intercrop by only 1 to 6%.

---

events especially harmfull for the yield. I also do find that, like with temperature, using the ML technology significantly reduces the impact of these extreme events. To conform with the existing literature, and to make the analysis more tractable, I decided to go with the linear specification for the rainfall and its variance in main specification. The non-linearity of the weather-yield relationship is left for further research

[50] Some of the results, including other robustness checks, are reported in Section A.3

[51] None of the strategies is consistent with the linear specification of the yield as in (2.4) - farmers should be indifferent on where to use ML intercrop

[52] The number of weedings can be considered both as a labor input and as a property of a plot - its propensity to produce weeds

### 2.7.1 What drives the performance of ML intercrop?

What drives the superior performance of the ML intercrop over M? Is it due to higher productivity of legumes, in which case the policy recommendation would be to encourage growing this crop regardless whether it is intercropped with maize, rotated or grown separately. Or are there also efficiency gains from growing maize and legume together, due to nitrogen sharing or denser planting (and despite competition depression)? I cannot answer these questions definitively because my data lacks critical inputs - seed input and area by crop.[53] The available evidence though suggests that the better performance of ML is both due to the productivity of legumes and the efficiency gains. The evidence is presented in Tables 2.12 and 2.13.

First, on average legumes do seem to have higher yield than maize, and hence productivity of ML intercrop could be driven by their presence in the mix of crops. Table 2.13, column 1 shows that in the cross-section of farmers (OLS), and controlling for weather and fertilizer use, legume yields are around 25% higher than those of maize. Table 2.12, columns 3 and 5 show that increased harvest share of legumes is associated with higher average yield.[54]

Second, as suggested by the indirect evidence, ML performance is also likely driven by the efficiency gains. Table 2.12, column 4 shows that ML is on average as productive as L, so the performance of ML cannot be only driven by the presence of legumes. Table 2.12, columns (1)-(3) demonstrate how ML performs compared to other technologies, which combine growing of maize and legume - M-L rotation and growing M and L separately (M+L). M-L rotation means that growing M is followed by growing L on the same plot next season, and so on. As I do not observe plots over time, and I do not observe farmers in consecutive years, I am not able to identify M-L

[53]The data on seed input per plot is available for waves 2 and 3 of the survey, however it is not clear whether farmers reported seeds only for maize or for all crops on plot. Total seed input is included as an explanatory variable in one of the robustness checks. The coefficient on it is positive, but the rest of the results are generally unchanged.

[54]Note that controlling for the harvest share of legumes in ML intercrops makes the coefficient on ML intercrop dummy insignificant. This is expected because the interpretation of the dummy is the average productivity differential when the harvest share of legumes approaches zero, at which point ML intercrop ceases to be intercrop (with an exception of intercrop plots, where legumes harvest completely failed, but the number of such plots is unlikely to be high)

Table 2.12: Estimation results: ML intercrop vs. other technologies

| | (1) ML rotation b/se | (2) M+L b/se | (3) M+L, legume share b/se | (4) L b/se | (5) M, legume share b/se |
|---|---|---|---|---|---|
| ML X rainfall | -3.06** | -4.91** | -4.83** | -2.02** | -1.28* |
| | (1.39) | (2.11) | (2.06) | (0.92) | (0.68) |
| ML X rainvar | 50.38*** | 40.21* | 40.76* | 7.34 | 5.57 |
| | (16.43) | (22.25) | (21.70) | (11.41) | (4.52) |
| =1 if ML | -0.00 | -0.28 | -0.06 | 0.06 | 0.06 |
| technology | (0.12) | (0.19) | (0.20) | (0.20) | (0.06) |
| ML X temp | 5.99 | 9.60 | 8.17 | 2.28 | -4.98* |
| | (8.41) | (11.95) | (11.71) | (6.04) | (2.80) |
| ML X temp_sq | 3.11** | 3.86** | 3.78** | 2.80** | -0.79 |
| | (1.30) | (1.80) | (1.76) | (1.18) | (0.71) |
| ML X temp_cub | -0.96* | -0.95 | -0.89 | -0.43 | 0.24 |
| | (0.55) | (0.76) | (0.74) | (0.38) | (0.16) |
| fertilizer, | -0.69 | 0.80 | 1.86 | -8.45*** | 2.86*** |
| tonnes per ha | (0.95) | (1.92) | (1.92) | (3.06) | (0.44) |
| fertilizer | 4.21** | 4.01 | 1.50 | 14.06** | -1.39* |
| ^2 | (2.14) | (6.51) | (6.42) | (5.97) | (0.83) |
| ML X fert | -1.12 | -1.62 | -2.03 | 9.29*** | -0.70 |
| | (1.13) | (1.78) | (1.74) | (3.10) | (0.48) |
| ML X fert_sq | 8.01** | 4.99 | 5.15 | -8.98 | 2.58 |
| | (3.55) | (7.25) | (7.08) | (6.35) | (1.64) |
| harvest share | | | 0.79*** | | 0.80*** |
| of L | | | (0.30) | | (0.13) |
| Constant | 14.58*** | 14.72*** | 14.32*** | 14.53*** | 14.31*** |
| | (0.06) | (0.07) | (0.17) | (0.19) | (0.02) |
| Observations | 809 | 284 | 284 | 1599 | 1676 |
| R-squared overall | 0.01 | 0.05 | 0.03 | 0.04 | 0.10 |

*Note* Standard errors (clustered by household) in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Estimation method for all columns: household-time FE, IV used for *rainfall*. Dependent variable in all columns - log of energy yield. ML intercrop is compared with: column (1) - M-L rotation; (2) - M+L (growing M and L on separate plots); (3) - M+L, controlling for the harvest share of L; (4) - L (growing only L); (5) - M, controlling for the harvest share of L in ML.

rotators in the data. What I can do is to come up with a proxy. In particular, I define a farmer to be M-L rotator if she/he has a record of growing both M (and no L) and L (and no M) on separate plots in at least one of the three waves of the survey. M+L is then a special case of M-L rotation, when M and L are grown on separate plots within the same season (and then likely rotated). As the proxy for M-L rotation is likely to be very noisy, the results must be taken with caution. The results suggest, however, that ML intercrop is on average as productive as M-L rotation or M+L (after controlling for the legume share of harvest). At the same time, I am able to estimate the maize yield for farmers who rotate maize with legumes versus those who do not rotate, even though it is only possible in the cross-section of farmers, e.g. via OLS.[55] The results in Table 2.13, column 3 suggest that maize

[55]For M-L rotation, unlike ML intercrop, I directly observe areas under M and L, as they are grown on separate plots (possibly during different periods of time). However, by definition of my M-L rotation proxy, I cannot have farmers growing maize with and without rotation, i.e. no

Table 2.13: Estimation results: Maize and legume productivity for different technologies

| | **M** vs. L | | **M only** vs. M in M-L rotation | M yield, **ML** vs. M | L yield, **ML** vs. L |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| | OLS | hhold-time FE | OLS | hhold-time FE | hhold-time FE |
| | b/se | b/se | b/se | b/se | b/se |
| rainfall, cm | 1.07*** | | 2.31*** | | |
| per day | (0.35) | | (0.36) | | |
| ML X rainfall | -0.04 | 0.86 | -1.80*** | -1.10 | -0.80 |
| | (0.39) | (0.57) | (0.42) | (0.70) | (1.12) |
| rainfall | 0.51 | | -63.81*** | | |
| variance, daily | (4.59) | | (5.15) | | |
| ML X rainvar | -20.57*** | -42.44*** | 51.49*** | -0.18 | 80.23*** |
| | (4.87) | (7.55) | (5.54) | (4.66) | (14.31) |
| =1 if | -0.25** | -0.08 | -0.11** | -0.10** | -1.78*** |
| technology in **bold** | (0.10) | (0.23) | (0.04) | (0.05) | (0.24) |
| temperature, C | -0.08 | | 1.74 | | |
| /100 | (2.27) | | (2.16) | | |
| temperature | -3.19*** | | -3.39*** | | |
| ^2 /100 | (0.49) | | (0.45) | | |
| temperature | 0.23 | | 0.59*** | | |
| ^3 /100 | (0.16) | | (0.16) | | |
| ML X temp | 3.61 | 4.03 | -1.81 | -0.91 | 23.00*** |
| | (2.44) | (3.61) | (2.46) | (2.85) | (7.62) |
| ML X temp_sq | 2.10*** | 1.39* | 2.47*** | -0.57 | 9.41*** |
| | (0.55) | (0.80) | (0.60) | (0.72) | (1.44) |
| ML X temp_cub | -0.13 | -0.09 | -0.44** | 0.15 | -1.90*** |
| | (0.17) | (0.26) | (0.18) | (0.17) | (0.48) |
| fertilizer, | -7.09*** | -10.25*** | 1.43*** | 3.06*** | -13.12*** |
| tonnes per ha | (1.40) | (3.03) | (0.28) | (0.45) | (3.71) |
| fertilizer | 12.38*** | 27.90** | -0.78 | -1.80** | 18.20** |
| ^2 | (2.74) | (11.76) | (0.60) | (0.85) | (7.29) |
| ML X fert | 10.17*** | 12.38*** | 2.10*** | -0.16 | 12.34*** |
| | (1.41) | (3.04) | (0.33) | (0.49) | (3.77) |
| ML X fert_sq | -14.36*** | -29.24** | -1.98*** | 1.44 | -12.77* |
| | (2.77) | (11.76) | (0.75) | (1.67) | (7.74) |
| Constant | 14.65*** | 14.29*** | 14.60*** | 14.25*** | 14.24*** |
| | (0.10) | (0.23) | (0.04) | (0.02) | (0.23) |
| Observations | 9533 | 3718 | 7058 | 1667 | 1558 |
| R-squared | 0.13 | | 0.04 | | |
| R-squared overall | | 0.07 | | 0.12 | 0.58 |

*Note* Standard errors (clustered by household) in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. In all columns, IV used for *rainfall*. Maize and legume productivity for different technologies (technology dummy is one for technologies in bold): columns (1) and (2) - M vs. L, log of energy yield, OLS and hhold-time FE; (3) - log of energy yield, maize only (no rotation or intercrop) vs. maize in M-L rotation; (4) - M only yield in M vs. ML; (5) - L only yield in L vs. ML.

is significantly more productive when it is rotated with legumes. Likewise, Table 2.13, column 2 shows that maize is only slightly less productive than legumes when comparing among farmers who grow both (e.g. M+L - a subset of M-L rotators).[56] These results suggest that performance of M-L rotation is not only due to performance of legumes, maize becomes more productive too, and since ML intercrop is as productive as M-L rotation, there have to be efficiency gains there too.

It is not clear whether the efficiency gains stem from the nitrogen sharing, possibility of denser planting without too much of competition depression, or the combination of the two. To answer we need to know areas under maize and legumes, as well as the seed inputs by each crop. The answer probably also depends on the structure of the intercrop.[57] These data are not available. Table 2.13, columns 4 and 5 suggests though that it is probably the combination of the two. The columns report the yields of only maize and only legumes on plots with ML intercrop versus plots where these crops are grown separately. In both cases the yields at ML intercrop plots are lower - by about 10% for maize, and much more (about 73%) for legumes. This could be the result of less denser planting when compared to monocropped plots (e.g. smaller area under each crop at ML intercrop plots), but it can also reflect the competition depression, e.g crops crowding out each other. Combining the two yields though gives the land equivalency ratio of more than one, which means that growing the same harvest, in the same proportions of M and L, would take more land if we use monocropped plots instead of ML intercrop.[58] This means that either the competition depression is not strong enough, and hence denser planting of maize and legumes is possible, or there is nitrogen sharing (or a combination of the two).

While on average ML intercrop is not the most productive technology (although still more productive than monocroped maize), it seems to perform best when it comes to weather-resistance.

household-time fixed effect is possible

[56]This is what household-time FE imply

[57]F.e. one can grow M and L in separate rows or in the same row, etc.

[58]Land equivalency ratio (LER) is defined as follows: $LER = \sum_{I=M,L} \frac{I\ yield\ on\ ML\ intercrop}{I\ yield\ on\ monocrop}$. From Table 2.13, columns 4 and 5, in our case LER=exp(-0.1)+exp(-1.78)=1.07. This is not necessarily different from 1 in a statistical sense, but LER is generally larger for other specifications (f.e. when M and L are the only crops grown on plot)

Table 2.12 shows that ML intercrop likely withstands rainfall shocks better than either M-L rotation or monocropped legumes, in addition to monocropped maize (as in my main specification). All coefficients on ML intercrop interaction with rainfall are negative and significant.[59] What drives this performance? Partly it might be due to the improved resilience of maize and/or legumes on the ML intercrop plots - Table 2.13, columns 3 and 4 show the negative coefficients on *ML x rainfall*, although they are both statistically insignificant.

Another possible part of the explanation is that growing the two crops together serves as an insurance - similar to the portfolio diversification in finance. A negative shock to one crop's yield is (partly) mitigated by the less affected yield of the other crop. This point is consistent with the evidence in Table A.5. The left panel shows summary statistics of total, maize and legume grain yields on the ML intercrop plots. The mean total yield is higher than either maize or legume, but its standard deviation is the lowest. The right panel shows a simple regression of log maize yield on a ML intercrop plot on log legume yield. The two are positively correlated, which is expected, but as demonstrated by low $R^2$ in both specifications, the correlation is very low.

The portfolio diversification does not seem to explain the whole story though. If it did then ML intercrop were not more resilient than M+L, i.e. growing M and L separately within the same season. Table 2.12, columns 2 and 3 suggest that ML intercrop does perform better. The explanation could be that the insurance works more intensively at ML intercrop plots: one crop's failure due to an adverse weather shock means less inter-species competition for the other crop.

## 2.8 Conclusions

Using more biodiverse agricultural technologies is a viable way to improve food security in Sub-Saharan Africa. Using the evidence from Malawi I show that maize-legume intrecrop is both more productive and better at absorbing weather and fertilizer shocks. I do not identify the average population effect. The estimated coefficients are subject to technology-specific factors

[59] Some coefficients seem implausibly large in absolute value, which is probably driven by smaller samples. All estimates are very imprecise, upper bounds of 95% confidence intervals almost reach zero in most specifications

bias. However, I show that if the government implements a cost or education subsidy to induce the maize-legume intercrop usage the effect is likely to be the one I identify. This is because the complying farmers will be the group, which is very similar to my estimation sample.

According to my results, a reform that induces a nine percentage point increase in land used for the intercrop can be yield-neutral even if the fertilizer used drops by 10%. At the same time, a system like this is likely to be more resilient to weather and fertilizer use shocks.

This paper has its limitations, either due to data unavailability or due to a risk of turning into a book. First, I do not specify what kind of reforms would induce farmers to use more intercrop, and what their potential cost and effectiveness are. The space for reforms is clearly large though. Infrastructure for the legume markets and education are probably the most important reform directions, as demonstrated by few examples in Section 2.2.4.

The second limitation of the paper is that I do not take into account potential systematic differences between the labor and seed requirements of the two technologies, which may stall ML's take off despite higher yield. If the ML generally requires more labor or seed inputs then its net benefit may turn out to be lower than that of M. However, as mentioned in Section 2.2.4, field trials suggest that the labor requirements are not too different, so the potential "‘excess"’ labor could stem purely from the lack of education. The scarce evidence on seed inputs in Section 2.7 suggests that the seed requirements are not likely to be the binding constraint either. In Section 2.4.3 I show that, even if we do not control for labor and seed inputs, for small policy changes household-time fixed effects regression still consistently estimates the effect of the policy on agricultural productivity. For larger policy changes this limitation has to be further investigated.

There are also factors, which I did not analyze in this paper, and which can add to the benefits of the intercrop. First, growing legumes can improve long-term quality of the soil, because of nitrogen fixation and legume biomass incorporation after the growing season. Second, consuming more biodiverse diet, e.g. maize vs. maize and legumes, improves nutrition, food security, and hence long-term health and productivity of the farmers (Kerr et al., 2011).

**CHAPTER 3**

**CORRUPTION AND ECONOMIC GROWTH REVISITED**

## 3.1 Introduction

Social scientists came up with many channels of how corruption can hurt economic growth and efficiency.[1] Bribery is a secretive and highly uncertain payment to a bureaucrat, which increases cost of investment, erodes confidence, and distorts competition, while allowing dishonest businesses to avoid efficiency-enhancing government regulation. Embezzlement reduces efficiency of government spending, and hence undermines the fiscal sustainability, as well as government's ability to provide public goods.

Despite the large number of theoretical channels, many of which supported by micro-empirical evidence, at a macro level the negative link between corruption and growth has been hard to demonstrate. A seminal study on the topic, Mauro (1995), finds the relationship between corruption and growth to be unrobust.[2] Svensson (2005) replicates the Mauro's regressions at an updated sample, uses alternative estimation methods, but finds no significant results. He calls this finding "a puzzle". In a meta-study Ugur (2014) analyzes 29 peer-reviewed studies on the topic and concludes that corruption and per-capita GDP growth are negatively linked, but the relationship is weak and unrobust.

The literature offers several potential explanations for the weak link between corruption and growth. Svensson (2005) points to a measurement error, omitted variables in the growth regressions, and the fact that "corruption takes many forms, and there is no reason to believe that all types of corruption are equally harmful for growth" (p.39). Huntington (1968) and a number of subsequent studies argue that corruption can "grease the wheels of business", and hence actually improve growth in highly bureaucratic economies. However, to the extent that the inefficient bureaucracy is

[1] See Olken & Pande (2012) for overview

[2] The relationship of corruption and investment is more robust. Also bureaucratic efficiency seems to perform better than corruption in the growth regressions

also likely a result of corruption, this argument is not likely to hold in general equilibrium setting. Ivanyna et al. (2016b) argue that in general equilibrium setting the effect of corruption on growth is negative but can be small in closed economies and when the government's borrowing is constrained. It is still unclear though why the empirical link between corruption and growth is weak across all countries.

This paper offers an alternative explanation of the weak empirical link between corruption and growth. Given the multitude of theoretical arguments,[3] it is highly likely that corruption has large negative effect on long-term potential economic growth. However, the empirical relationship between the two is contaminated by the reverse causality from economic growth to corruption in short and medium term. During economic booms tax revenue is higher and it is easier for government to borrow. This blows up the budgets that bureaucrats are in charge of, and consequently increases embezzlement and corruption. Hence what we empirically observe is the two offsetting effects: high-corruption countries growing slower on average over long-term, and at the same time higher corruption in countries that grow above potential in short and medium term. The two effects result in a weak empirical link between corruption and growth, especially when the relationship is tested over shorter time spans. Even for longer time spans the "growth-to-corruption" effect may still be present as long as corruption is driven by commodity price and financial cycles - both of much lower frequency than the business cycles.

I demonstrate my point using three approaches. First, I build a simple stylized model with corruption in an endowment economy. The goal is to provide a tractable framework to comprehend the workings of the "growth-to-corruption" effect. I show that higher endowments cause higher corruption. Second, I continue with much less tractable but much more comprehensive general equilibrium model, calibrated to an average developing economy. The model is a modified version of Ivanyna et al. (2016b). The price of being comprehensive is that I can only solve the model numerically. The numerical computations go in line with the simple model. Third, I put my findings to the data, and my results conform with the models.

[3]Especially taking into account the general equilibrium nature of corruption, which includes not only petty bribery but also policymaking at a grand level

## 3.2 Stylized model of corruption

The goal of the model in this section is to demonstrate the workings of the "growth-to-corruption" channel in a simplified tractable way. In the next section I show that the results of this simple model also carry through in a more comprehensive framework.

The model is significantly simplified version of Ivanyna et al. (2016b). It is static (one time period), and there is no production. There are two types of agents - private households and public officials. Each agent receives an exogenous endowment. Private households consume the endowment plus transfers from government subject to income tax they are supposed to pay. They can also choose to evade part of the tax payment. Public officials redistribute the tax revenue, and may choose to embezzle part of it. Collectively, they also set the income tax rate.

### 3.2.1 Private households

There are $N$ private households, each is exogenously endowed with income $w$. They pay tax $\tau$ on their endowment and consider evading part of the tax. They also receive transfer $\hat{g}$ from the government. Tax evasion is costly, part of the hidden income is lost when trying to conceal it. In addition, the households are averse to illegal activity, which is expressed in a loss of utility when there is tax evasion. The utility of the representative household is the following:

$$U = c - \frac{\phi}{2}v^2, \tag{3.1}$$

which they maximize subject to the budget constraint:

$$c = (1-\tau)(1-v)w + \theta^{\tau} vw + \hat{g}. \tag{3.2}$$

Here $c$ is consumption, and $v$ is tax evasion - a fraction of $w$ that is concealed. The second part of $U$ reflects the aversion to tax evasion, the "guilt". $\phi$ is the aversion parameter, the larger is $\phi$ the more there is disutility.

The total household income, the right hand side of (3.2), consists of three parts. First part is the after-tax income from endowment, which was not concealed. The second part is the income that was concealed and is available for private use - fraction $v$ of $w$, adjusted by $\theta^{\tau}$. $\theta^{\tau}$ represents government checks on tax evasion. It ranges between 0 (no concealed income is available for use at all) and 1 (no resource cost of hiding income). The more difficult it is to hide income from the government, the smaller is $\theta^{\tau}$, the less of it can be used, thus lowering the benefit of evasion. [4] The third part of the income is government transfer $\hat{g}$.

The first-order conditions imply that the optimal level of tax evasion is:[5]

$$v = \frac{(\theta^{\tau} - (1-\tau))\, w}{\phi} \tag{3.3}$$

Everything else equal, tax evasion increases with weaker checks on evasion and higher $\tau$. It also increases if private households are less averse to illegal activity (lower $\phi$).

### 3.2.2 Public officials

There is fixed number $\epsilon N$ of public officials in the economy. They are exogenously selected from the population of private sector households, and have preferences that are identical to the private households. Each public official is paid salary $w^{g}$, which is exogenously fixed, and takes charge of distributing government transfers to the private households. So the role of the government in this model is simply to collect income tax and redistribute it back to private households. This could be justified by concerns about income inequality and poverty. For simplicity, there is no production of public goods in this model.

Public officials set income tax rate $\tau$ collectively, while the redistribution is decentralized. Each public officials is allocated an equal share of tax revenue $G/\epsilon N$ after public salaries are paid. Part of the budget can be embezzled. As in case with the private households, embezzlement is costly, not all of it can be recovered for private use, and public officials are averse to illegal activity. The utility of the representative household is the following:

[4] In other words, $\theta^{\tau}$ here is a pure waste or deadweight loss from having tax evasion

[5] For simplicity, here and below, assume that values of parameters are such that there are no corner solutions to the model

$$U^g = c^g - \frac{\phi}{2}u^2, \tag{3.4}$$

which they maximize subject to the budget constraint:

$$c^g = (1-\tau)w^g + \theta^g u \frac{G}{\epsilon N}. \tag{3.5}$$

$c^g$ is consumption, and $u$ is the level of corruption - a fraction of the budget that is embezzled. The total income of a public official consists of after-tax salary and the embezzled funds, adjusted for $\theta^g$. $\theta_g$ represent checks on corruption. The lower it is the less of stolen funds are available for private use. It ranges from 0 (no stolen funds are available for use) to 1 (no checks, so no cost of embezzlement).

Individually, public officials take $\tau$ and $G$ as given, and select optimal $u$. From the first-order conditions:

$$u = \frac{\theta^g}{\phi}\frac{G}{\epsilon N}. \tag{3.6}$$

Corruption increases with lower checks on corruption, lower aversion to illegal activity. Importantly, it also increases with the size of the budget that each public official gets. When $\epsilon$ is fixed, the size of individual budget depends on the total budget $G$. In this simple setting corruption is not related to the public salary $w^g$, but this depends on the functional specification of utility.[6]

### 3.2.3 Government budget constraint

Government collects income tax revenue and spends it on public salaries and transfers to the private households. The government budget constraint is the following:

$$G = \left(\tau(1-v)w + \tau w^g \epsilon\right) N - w^g \epsilon N \tag{3.7}$$

The income tax revenue consists of tax paid by private households, that is after part of $w$ is concealed, and tax paid by public officials.

[6] In this model I hold $w^g$ fixed, so relationship between $u$ and $w^g$ is less important. In the more comprehensive model this relationship is analyzed in more detail

The budget per public official is then:

$$g \equiv \frac{G}{\epsilon N} = \frac{\tau(1-v)w}{\epsilon} - (1-\tau)w^g. \tag{3.8}$$

Then the government transfer that is effectively received by a representative private household:

$$\hat{g} = \frac{(1-u)G}{N} \tag{3.9}$$

The key question I attempt to answer in this model is if a positive shock to $w$ increases $u$. Since $u$ depends on $G$, the answer depends on whether increase in $w$ increases $G$, which in turn depends on what happens with the tax paid by private households - $\tau(1-v)w$. Taking into account (3.3):

$$R \equiv \tau(1-v)w = \tau w - \tau \frac{\theta^{\tau} - (1-\tau)}{\phi} w^2. \tag{3.10}$$

The tax paid by the private households $R$ is a quadratic function of $w$, with a negative quadratic term, so whether it increases or decreases with $w$ depends on the government's choice of $\tau$.

### 3.2.4 Choice of $\tau$

Collectively public officials choose $\tau$. They do it by maximizing the utility of a representative public official and taking into account the optimal responses of private households and public officials when it comes to the choice of $v$ and $u$. The utility of a public official can be rewritten:

$$U^g = c^g - \frac{\phi}{2}u^2 = (1-\tau)w^g + \theta^g ug - \frac{\phi}{2}u^2 =$$

$$= (1-\tau)w^g + \frac{\theta^{g2}}{2\phi}g^2. \tag{3.11}$$

In choosing the tax rate public officials want to maximize the budget per public official $g$, which depends on $\tau$ directly and indirectly through $v$, and at the same time they weigh in the fact that they pay the tax as well. In addition, corruption increases the disposable income but brings in the disutility.

Differentiating (3.11) with respect to $\tau$ yields the first-order conditions:

$$-w^g + \frac{\theta^{g2}}{\phi} g \frac{\partial g}{\partial \tau} \;\Rightarrow\; \frac{\partial g}{\partial \tau} = \frac{w^g}{\theta^g u}, \tag{3.12}$$

where the last equality uses the first-order condition (3.6). In the optimum, public officials set the tax rate to the left of the peak of the Laffer curve - $\frac{\partial g}{\partial \tau} > 0$. The trade-off is between paying higher tax themselves and increasing the allocated budget per official $g$, and stealing part of it. This marginal benefit is $\theta^g u$ - smaller than one by construction.

### 3.2.5 Does corruption go up if private endowment increases?

What happens to corruption if private endowment $w$ receives a positive shock? Such a shock would be a simulation of the "growth-to-corruption" effect - an exogenously driven economic growth, which affects the behavior of public officials.

From (3.6) it follows that corruption $u$ changes in the same direction as $g$ - the budget per official allocated for government transfers. So the key question is how an increase in $w$ affects $g$.

In short run, when the tax rate $\tau$ is not yet changed and the tax evasion behavior, and hence $v$, remain the same, increased endowments means larger tax revenue means large budget for the transfers. Larger $g$ means more opportunities for corruption, and hence more corruption.

In longer run, private households optimally respond to the shock by adjusting their tax evasion, and in principle $g$ could end up lower than before. This will not happen however, because increased tax evasion will not offset the shock completely, and hence tax revenue will still increase.

To see this, let us first analyze medium run - a likely situation when tax evasion responds optimally, but the tax rate remains unchanged. The first step is to demonstrate that $g$ is maximized when tax evasion $v = \frac{1}{2}$. With unchanged $\tau$:

$$\frac{\partial g}{\partial w} = \tau \frac{1-v}{\epsilon} - \frac{\tau w}{\epsilon} \frac{\partial v}{\partial w} =$$

$$= \tau \frac{1-v}{\epsilon} - \tau \frac{v}{\epsilon} = \frac{\tau}{\epsilon}(1-2v) = 0 \Leftrightarrow v = \frac{1}{2}. \quad (3.13)$$

The consequence of (3.13) is that $\frac{\partial g}{\partial w} > 0$ if $v < \frac{1}{2}$.

The second step is to show that at the pre-shock optimum $v$ is actually smaller than 0.5. Using

(3.10) we can derive the tax rate that maximizes tax paid by private households:

$$\frac{\partial R}{\partial \tau} = w - \frac{\theta^{\tau} - 1}{\phi} w^2 - 2\tau \frac{w^2}{\phi} \Rightarrow$$

$$\Rightarrow \tau^{max} = \frac{1 - \frac{\theta^{\tau} - 1}{\phi} w}{2\frac{w}{\phi}}. \quad (3.14)$$

Then the tax evasion when $\tau^{max}$ is charged:

$$v^{max} = \frac{\theta^{\tau} - (1 - \tau^{max})}{\phi} w = \frac{1}{2} + \frac{1}{2} \frac{\theta^{\tau} - 1}{\phi} w < \frac{1}{2}, \quad (3.15)$$

because $\theta^{\tau} < 1$ by definition.

The tax rate chosen by public officials in the optimum is even smaller than $\tau^{max}$. Using (3.12):

$$\frac{\partial g}{\partial \tau} = \frac{1}{\epsilon} \frac{\partial R}{\partial \tau} + w^g = \frac{w^g}{\theta^g u} \Rightarrow$$

$$\Rightarrow \frac{1}{\epsilon} \frac{\partial R}{\partial \tau} = w^g \left(\frac{1}{\theta^g u} - 1\right) > 0 \quad (3.16)$$

At the optimal pre-shock $\tau$ tax paid by private households increases, which means that $\tau < \tau^{max}$. Therefore $v$ is smaller that $v^{max}$, and so smaller than 0.5. Hence $g$ will increase as a result of positive shock in $w$ and so will corruption.

Corruption will still remain elevated in the longer run, when $\tau$ is also allowed to react to the shock. This is because in this case $\frac{\partial g}{\partial w} = \frac{\partial g}{\partial \tau} \frac{\partial \tau}{\partial w}$. From (3.12), $\frac{\partial g}{\partial \tau} > 0$, so the main question is what happens to the tax rate. $\tau$ should increase too, because, as shown above, with unchanged tax $g$ goes up. Then from (3.12) $\frac{\partial g}{\partial \tau}$ goes down, which can only be the case if $\tau$ increases.

Intuitively, public officials always set tax rate so that the tax paid by the private households is to the left of the Laffer curve peak. Increase in taxable income increases tax evasion by the private households, but the tax paid nevertheless increases too. Increased evasion and taxable income induce public officials to raise the tax rate, but again the tax revenue remains increased. Hence, corruption goes up.

This stylized model is a basic demonstration of how exogenously driven economic growth can cause corruption to increase. The next step is to relax some of the assumptions, which I did to get the analytical solution, and see if the result still holds. The cost of making the model more comprehensive is that I am only able to get the numerical solutions with reasonably calibrated parameters.

## 3.3 Main model

Here I demonstrate that the "growth-to-corruption" effect takes place also in a dynamic general equilibrium setting with investment, production, and productivity-enhancing role of the government. The model I use is a gentle modification of Ivanyna et al. (2016b). I introduce two main differences. First, the aggregate productivity growth is now stochastic and receives shocks every period. Second, I introduce sluggishness in the public wages - they are not as responsive to productivity shocks as private wages are.

The model is an overlapping-generations model of private capital accumulation. Similarly to the stylized model of the previous section, private households are allowed to evade taxes, public official are allowed to embezzle public funds. Both illegal activities are costly because resources are lost in attempting to conceal the actions. The stronger are the government's mechanisms for detection, the more resources are lost in avoiding detection. Households also experience a loss in utility, "‘guilt"’ from violating a social norm, when evading taxes or embezzling public funds. Furthermore the strength of the guilt associated with tax evasion varies inversely with the average level of corruption by government officials. In addition, the cultural effect extends to the government officials themselves - individual government officials are more likely to engage in corrupt behavior the higher is the level of corruption around them. See Ivanyna et al. (2016b) for justification of these assumptions.

### 3.3.1 Private choices

There are $N$ young private households in each period, and fixed number $\epsilon N$ of public officials. The households and public officials are standard two-period life-cycle savers. They work to earn wages ($w_t$), consume ($c_{1t}$), and save ($s_t$) in the first period to finance second period retirement-consumption ($c_{2t+1}$). Same with subscript $g$ for the public officials. In addition to their own consumption, households also care about the general state of the economy - the average level of worker productivity during both periods of their lives $(y_t, y_{t+1})$. The last assumption is a form of altruism, which is introduced to allow for the possibility that households who become public officials have concerns about the current and future state of the economy and not only their private consumption.

The preferences of private households and public officials are written as

$$U_{y,t} = \ln c_{1t} + \beta \ln c_{2t+1} + \gamma(\ln y_t + \beta \ln y_{t+1}) - \frac{\phi}{2\bar{u}_t} v_t^2 \tag{3.17}$$

and

$$U^g_{y,t} = \ln c^g_{1t} + \beta \ln c^g_{2t+1} + \gamma(\ln y_t + \beta \ln y_{t+1}) - \frac{\phi}{2\bar{u}_t} u_t^2, \tag{3.18}$$

The illegal activity of private households is measured by $v$, the fraction of their income that is not reported for tax purposes. The illegal activity of public officials is measured by $u$, the fraction of the public investment budget that is diverted for private use. The last term in each expression captures the "‘guilt"’ or direct disutility of engaging in illegal activity.

Higher values of $\phi$ imply a stronger distaste for illegal activity. The disutility of illegal activity is also affected by the average level of corruption among government officials. The greater is the average level of corruption the less disutility an individual experiences from their own illegal activity. Ivanyna et al. (2016b) refers to this as the "‘culture of corruption"’ (COC) effect, and shows that it is essential element of the model if it is to replicate key features of the data on corruption.

The private household maximizes utility subject to the lifetime budget constraint

$$c_{1t} + \frac{c_{2t+1}}{(1 + r_{t+1})} = (1 - \tau_t) w_t (1 - v_t) + \theta^\tau w_t v_t, \tag{3.19}$$

where $\theta^\tau$ is a parameter, that lies between zero and one, reflecting the fraction of unreported income that the household can recover for private use.

The maximization problem generates the following equation for tax evasion and private household saving

$$v_t = \frac{1}{2}\left(\sqrt{T^2 + \frac{4(1+\beta)\bar{u}_t}{\phi}} - T\right), \tag{3.20}$$

where $T \equiv \frac{1-\tau_t}{\theta^\tau-(1-\tau_t)}$.

$$s_t = \frac{\beta}{1+\beta}\left(1-\tau_t+(\theta^\tau-1+\tau_t)v_t\right)w_t. \tag{3.21}$$

Evasion is increasing in $\tau_t$ and in $\theta^\tau$. Evasion is also increasing in $\bar{u}$. The term $(1+\beta)/\phi$ is a measure of "'greed"' because it is a measure of the value of consumption relative to the disutility of being dishonest. Tax evasion is increasing in greed, other things constant.

Next, we move to the behavior of the public official. In the case of uncoordinated or decentralized corruption, each public official takes the average level of corruption, the tax rate, and the total public investment budget as given when making their private choices. The public official's private choices include what fraction of their project budget to divert for their own private use. The budget allocated to each public official is $\hat{G}_{t+1}/\epsilon N$, where $\hat{G}_{t+1}$ is the amount of recorded or planned investment and not the actual investment.

The officials maximize utility subject to the public budget and their private lifetime budget constraint,

$$c^g_{1t} + \frac{c^g_{2t+1}}{(1+r_{t+1})} = (1-\tau_t)\eta w_t + \theta^g u_t\left(\frac{\hat{G}_{t+1}}{\epsilon N}\right), \tag{3.22}$$

where $\theta^g$ is a parameter, that lies between zero and one, reflecting the fraction of diverted public funds that the official can recover for private use.

The wage paid to public officials is proportional to the private sector wage, i.e. the public official's wage is $\eta w_t$. $\eta$ is one of the key variables in this model. It reflects two features. First is how are wages in private and public sectors related in the long-run. Second is how sluggish are public wages in the short run. In other words, how responsive they are to the temporary productivity shocks or changes in the private sector wages. $\eta$ is defined in the following way:

$$\eta = \frac{\eta_1}{1+\eta_2\xi}, \tag{3.23}$$

where $\eta_1$ characterizes the long-term relation to the private sector wages, $\eta_2$ characterizes the sluggishness, and $\xi$ is the size of a temporary shock or a temporary deviation of private sector wages from equilibrium.[7] If $\eta_1 = 1$ then public wages track private sector wages one-to-one in the

[7] $\xi$ is discussed in the next subsection

long run (when $\xi = 0$). $\eta_2$ varies from 0 to 1. If $\eta_2 = 0$ then public wages perfectly track private sector wages also in the short run. If $\eta_2 = 1$ it means public wages do not react to temporary shocks at all.For example, suppose private sector wage receives a multiplicative temporary shock $1 + \xi$. So in the current period the wage is $w(1 + \xi)$, where $w$ is the long-run value. Then public wage is $\eta w(1 + \xi) = \eta_1 w \frac{1+\xi}{1+\eta_2 \xi}$.

The maximization problem generates the following equations for corruption and the public official's private saving

$$u_t = \frac{1}{2}\left(\sqrt{\Gamma^2 + \frac{4(1+\beta)\bar{u}_t}{\phi}} - \Gamma\right), \tag{3.24}$$

where $\Gamma \equiv \frac{1-\tau_t}{\theta^g \frac{\hat{G}_{t+1}/\epsilon N}{\eta w_t}}$.

$$s_t = \frac{\beta}{1+\beta}\left((1-\tau_t) + \theta^g u_t \frac{\hat{G}_{t+1}/\epsilon N}{\eta w_t}\right)\eta w_t. \tag{3.25}$$

As with evasion, corruption is increasing in $\tau_t$ and in $\theta^g$ . The larger is the budget that the official manages, relative to his official after-tax wage, the more tempting it is to be corrupt. This is also why corruption is decreasing in $\eta\epsilon$ - the larger is the official wage (increasing in $\eta$) relative to the official's budget (decreasing in the number of officials or $\epsilon$), the lower is corruption. An increase in the official's wage raises consumption and lowers the value of additional consumption gained by diverting public funds. However, the larger is the size of the public budget, the greater is the benefit of diverting a higher fraction of it. In particular, if the budget increases faster than the wage corruption increases.

### 3.3.2 Firms

Production takes place within standard neoclassical firms that combine physical capital and human capital to produce output from a Cobb-Douglas technology

$$Y_t = K_t^{\alpha}(D_t N)^{1-\alpha}. \tag{3.26}$$

However, the productivity index ($D$) is a function of disembodied technology ($A$) and public capital per adult worker ($G/((1+\epsilon)N)$) and is given by

$$D_t = A_t^{1-\mu}(G_t/(1+\epsilon)N)^{\mu}, \tag{3.27}$$

where $0 < \mu < 1$ is a constant parameter. This specification captures the idea that public infrastructure raises the productivity of the private sector.

Disembodied technology $A$ progresses at an exogenous rate $d$ every period on average. In addition the growth rate of productivity is subject to a random temporary shock $\xi$ with mean zero. So the overall growth of A every period is $d+\xi$. $\xi$ here reflects all temporary factors that can affect the output. This includes both supply and demand shocks - business cycle developments, as well as developments over financial and asset price cycles, which have lower frequency.

Firms operate in perfectly competitive factor and output markets. This implies the profit-maximizing factor mix must satisfy

$$\delta + r_t = \alpha g_t^{\mu(1-\alpha)} k_t^{\alpha-1}, \tag{3.28}$$

$$w_t = (1-\alpha) A_t g_t^{\mu(1-\alpha)} k_t^{\alpha}, \tag{3.29}$$

where $\delta$ is the rate of depreciation on physical capital, which we take to be one for simplicity, $g \equiv G_t/A(1+\epsilon)N$, and $k \equiv K/AN$.

### 3.3.3 Capital market equilibrium and government budget constraint

The capital stock rented to firms in period $t$ must be accumulated as retirement savings by the private households and government officials,

$$K_{t+1} = Ns_t + \epsilon N s_t^g. \tag{3.30}$$

The government budget constraint is:

$$\tau_t(w_t(1-v_t)N + \epsilon\eta w_t N) = \eta w_t \epsilon N + \hat{G}_{t+1}, \tag{3.31}$$

which implies that $\hat{G}_{t+1}/w_t\epsilon N = \tau_t\left(\frac{1-v_t}{\epsilon} + \eta\right) - \eta$.

The actual investment in public capital is the accounting measure $\hat{G}_{t+1}$ minus the budget funds consumed by the government officials. Subtracting the portion of the capital budget that is consumed by government officials from (3.31), and de-trending by dividing by $A_{t+1}$, gives us the transition equation for public capital intensity in the presence of corruption and evasion,

$$g_{t+1} = (1 - u_t)(\tau_t(1 - v_t + \epsilon\eta) - \eta\epsilon)\frac{(1 - \alpha)g_t^{\mu(1-\alpha)}k_t^{\alpha}}{(1 + d + \xi)(1 + \epsilon)}. \tag{3.32}$$

For a given tax rate, corruption and evasion both serve to shift the transition equation for public capital downward.

The private saving functions for private households and public officials, and (3.30) can be used to derive the transition equation for private capital,

$$\begin{aligned} k_{t+1} = \frac{\beta}{1 + \beta} * & \\ & * \left[ \left(1 - \tau_t + (\theta^{\tau} - 1 + \tau_t)v_t\right) + \eta\epsilon \left((1 - \tau_t) + \theta^{g} u_t \left(\tau_t \left(\frac{1 - v_t}{\eta\epsilon} + 1\right) - 1\right)\right)\right] * \\ & * \frac{(1 - \alpha)g_t^{\mu(1-\alpha)}k_t^{\alpha}}{1 + d + \xi}. \end{aligned} \tag{3.33}$$

While corruption and evasion reduce funds available for public investment, for a given tax rate, they increase funds available for private investment. Thus, the overall effect of corruption and evasion on growth is not clear. In addition, the presence of corruption and evasion affects the tax rate chosen by the public officials.

### 3.3.4 Corruption, evasion, and the tax rate

Collectively public officials choose the tax rate, which maximizes the representative public official's welfare. The optimal tax rate takes in account tax rate effects on private choices, whether made by private households or public officials. This includes the effects of the tax rate on both corruption and evasion.

The representative government official's preferences for generation-$t$, including only those terms that are influenced by the choice of the current period tax rate:

$$
\begin{aligned}
(1+\beta)\ln\left((1-\tau_t)+\theta^g u_t\left(\tau_t\left(\frac{1-v_t}{\eta\epsilon}+1\right)-1\right)\right)-\frac{\phi}{2}u_t+ \\
+\beta\mu(1-\alpha)(1+\gamma)\ln\left((1-u_t)\left(\tau_t\left(\frac{1-v_t}{\eta\epsilon}+1\right)-1\right)\right)+\beta(\alpha(1+\gamma)-1)* \\
*\ln\left((1-\tau_t+(\theta^\tau-1+\tau_t)v_t)+\eta\epsilon\left((1-\tau_t)+\theta^g u_t\left(\tau_t\left(\frac{1-v_t}{\eta\epsilon}+1\right)-1\right)\right)\right). \quad (3.34)
\end{aligned}
$$

The first term determines the effect of tax rates and tax revenue on the private income and consumption of the government official. The second term is the disutility of being corrupt. The third term is the effect of taxation on public investment. Next period's public capital raises the welfare of a generation-$t$ official because it (i) raises the marginal product of private capital and the rate of return to private capital and (ii) increases next period's worker productivity, which is valued by individuals in the economy under our assumptions. The last term is the effect of taxation on private investment. Private capital has two opposing effects on the public official's welfare. Next period's private capital stock lowers welfare because it lowers the rate of return to private capital, but also raises welfare because it increases next period's worker productivity.

It is not possible to derive an analytical expression for the optimal tax rate, and hence corruption and tax evasion. I calibrate the model and find a numerical solution. The focus is on an average developing economy without much institutional checks on corruption and evasion. The key question that I ask is how aggregate productivity shock $\xi$ affects corruption, everything else equal.

### 3.3.5 Calibrating the model

The model is calibrated to an average developing economy. I take values for most of the parameters from Ivanyna et al. (2016b). The output elasticities of private and public capital are conventional estimates: $\alpha = 0.33$, $\mu = 0.3$. Assuming that each period in the model lasts 20 years and the average annualized growth in labor productivity due to exogenous technological change is 2 percent we have $d = (1.02)^{20} - 1 = 0.4859$. I set share of public officials in the economy $\epsilon = 0.1429$, and I assume public wages track private sector wages one-to-one in the long-run - $\eta_1 = 1$. In the short-run I start with the assumption that public wages are fully sluggish, i.e. they do not respond to temporary shocks in private sectors wages - $\eta_2 = 1$. Values for parameters in the utility functions are also taken from Ivanyna et al. (2016b): $\beta = 0.198$ and $\gamma = 4.756$. Finally, I start with the assumption of no institutional checks on corruption or tax evasion: $\theta^{\tau} = \theta^{g} = 1$, and I calibrate $\phi$ to target the value of tax evasion $v$ to 1/3 ($\phi = 1.07$) The target is based on the summary of estimates for the relative size of the shadow economy reported by Porta & Shleifer (2008, Table I).

Under these values of parameters the steady state is the same as in Ivanyna et al. (2016b). What is left is to choose the size of aggregate productivity shocks $\xi$ which hit the economy every period. The size of $\xi$ depends on the length of period, which is assumed to be 20 years in the model. In general, the shorter is the period the larger is $\xi$, as temporary shocks would tend to offset each other over longer periods of time. For example, annual output gaps in developing economies can be as large as 10%, but over 20 years they would average almost to zero. At the same time, some asset price cycles can result in large $\xi$'s even over longer periods of time. Ivanyna et al. (2016a) identify public debt cycles in upper middle income countries with an average amplitude of 17 percentage points of GDP, and an average length of 13 years. Commodity prices are also very volatile and persistent. For example, the price of oil was mostly around USD30-40 in the 90s, down from USD50-90 in 70s and 80s, and then went up to USD60-100 in the 2000s. This results in high volatility of output growth in resource-rich countries. I take parsimonious approach in calibrating $\xi$, and solve the model under a wide range of possible outcomes. $\xi$ varies from -0.95 to 2.4, which corresponds to the departure of the annual output growth from the potential (2%) of -5 percentage points to 5 percentage points (and 20 years in one period).

#### 3.3.6 Simulations: Do aggregate productivity shocks increase corruption?

Based on the calibration above I do a number of simulations. I start with the baseline without shocks, which is identical to Ivanyna et al. (2016b). Then I hit the economy with aggregate productivity shock of different magnitude and see what happens to the equilibrium values of output, corruption, and tax evasion. The shocks vary from -0.95 to 2.4, and I initially assume public sector salary does

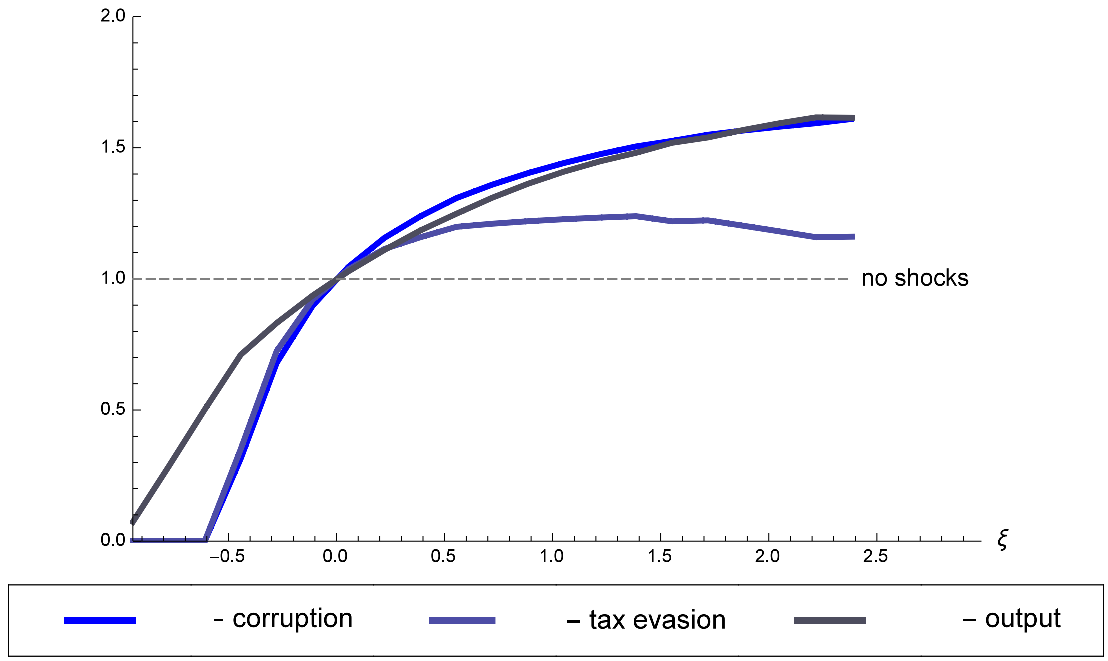


Figure 3.1: Aggregate productivity shocks: Response of corruption, tax evasion and output

*Note* Figure shows response of corruption, tax evasion and output for a range of aggregate productivity shocks $\xi$. All variables are expressed as ratio to the corresponding value in the baseline without shocks ($\xi = 0$). The computation is based on the following values for the model's parameters: $\theta^{gov} = 1$, $\theta^{\tau} = 1$, $\phi = 1.07$, $\eta_1 = 1$, $\eta_2 = 1$, $\epsilon = 0.14$, $\alpha = 0.33$, $\mu = 0.3$, $d = 0.49$, $\beta = 0.2$, $\gamma = 4.76$.

not react to them ($\eta_2 = 1$).

Figure 3.1 shows the results. The first and main observation is that positive shocks generate both larger output and more corruption. The larger are the shocks the larger are the output and corruption as compared to the baseline with no shocks. A positive productivity shock increases the private sector wages and hence income tax revenue. Public sector salary does not adjust, which means that each public official gets a larger budget for public investment relative to her/his salary. Hence corruption increases. Because of the culture-of-corruption effect, tax evasion may also respond positively to the increased corruption. Public officials compensate for this by choosing a lower tax rate. As a result, when economy is hit by a positive aggregate productivity shock, we observe large increases in corruption and output, followed by a small or no increase in tax evasion.

Resulting from the productivity shocks, the positive correlation between output and corruption can explain why empirically the link between economic growth and corruption is very weak. From the one side, weak institutional checks on corruption lead to higher corruption and lower steady state output per worker. This is the relationship, which researchers are trying to estimate by using the growth regressions. From the other side, aggregate productivity shocks temporarily increase both corruption and output, which biases the regression coefficients. If shocks are persistent, the bias can be present even if the data is averaged over longer time spans. Figure 3.2 simulates the growth regressions from the model to demonstrate the points above. I take 100 observations ("countries"), which are identical to each other, i.e. all structural parameters are the same, except institutional checks on corruption $\theta^g$ and aggregate productivity shocks $\xi$ that they face. $\theta^g$ is randomly drawn from uniform distribution $U(0.45, 1)$,[8] $\xi$ is drawn from $U(-0.95, 2.4)$. $\xi$ and $\theta^g$ are drawn independently. I run the model for each of the 100 observations, and find the equilibrium corruption and economic growth, which is defined as the growth of output relative to the baseline with no shocks and $\theta^g = 1$.[9] The right panel of Figure 3.2 shows the scatter plot of corruption and growth in case there are no shocks (or alternatively, if we manage to control for shocks in regressions). Here the pattern is as expected: lower corruption is associated with higher growth. The left panel shows the scatter plot when shocks are included. With shocks included, the relationship between corruption and growth turns inside out - now lower corruption is actually associated with lower growth. The "growth-to-corruption" effect dominates in this simulation. In reality the relationship between growth and corruption is usually not positive. Its sign and magnitude depend on the size of the shocks, the relationship between checks on corruption and steady-state output, and the sluggishness of public wages, which in this simulation is assumed to be full ($\eta_2 = 1$). But one thing is clear, the "growth-to-corruption" effect biases upwards the coefficient on corruption in growth regressions, and hence we likely underestimate the harm that corruption makes to the economy.

Another implication of the general model is that the "growth-to-corruption" effect is smaller if public officials' salary is more responsive to the productivity shocks. This is demonstrated by Figure 3.3. The figure shows the response of corruption to the shocks for different values of $\eta_2$.

[8]The lower value of the support is chosen so that there is only small amount of countries with zero corruption

[9]So the growth rate in country with $\theta^g = 1$ and $\xi = 0$ is zero.

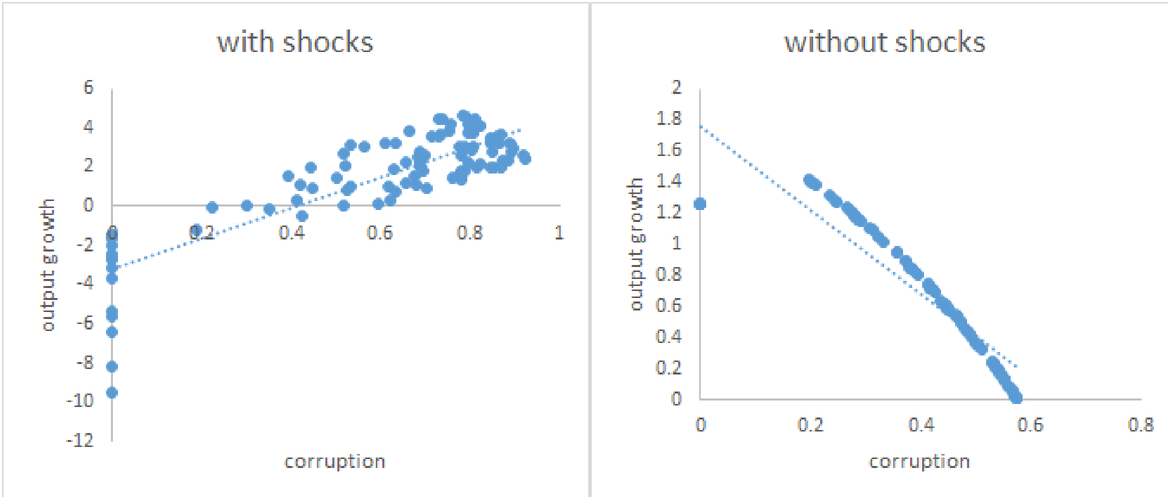


Figure 3.2: Growth regression simulations with and without shocks

*Note* Figure simulates from the model the growth regressions with corruption included as the independent variable. Sample of countries is formed using 100 independent draws from uniform [0.45,1] distribution for $\theta_g$ and uniform [-0.95,2.4] distribution for $\xi$. All other variables for all countries are left the same as in the baseline: $\theta^{\tau} = 1$, $\phi = 1.07$, $\eta_1 = 1$, $\eta_2 = 1$, $\epsilon = 0.14$, $\alpha = 0.33$, $\mu = 0.3$, $d = 0.49$, $\beta = 0.2$, $\gamma = 4.76$. Output growth is computed as the ratio to the corresponding value in the baseline without shocks ($\xi = 0$ and $\theta_g = 1$). Corruption is generated from the model. Left panel shows the scatter-plot and the fitted line for the "world" with both government checks on corruption and aggregate productivity shocks. Right panel shows the same for the "world" with no shocks.

Lower values of this parameter mean the salary is more responsive. For each shock, corruption is closer to the baseline the lower is $\eta_2$. If public wages adjust quicker then budget-to-salary ratio does not increase as much, and so is corruption. At the same time, even if $\eta_2$ is low, the departure of corruption from the baseline can be significant. For example, $\eta_2 = 0.33$ means that 2/3 of the temporary shock is absorbed by the public wages. Still the corruption increase is generally more than half of the increase in case the wages do not adjust at all ($\eta_2 = 1$).

Though it is not directly in the model, one of its policy implications is that it might be a bad idea to use public wage bill as a tool for counter-cyclical fiscal policy, i.e. to manage aggregate demand. If public wage is counter-cyclical it means that it is more sluggish in the terminology of the model, and hence corruption may increase with its long-term negative consequences.[10] On a contrary, counter-cyclical purchase of goods and services or public investment reduce the budget

[10]This concerns only the average public wage, not the public employment. If employment falls this may actually decrease corruption.

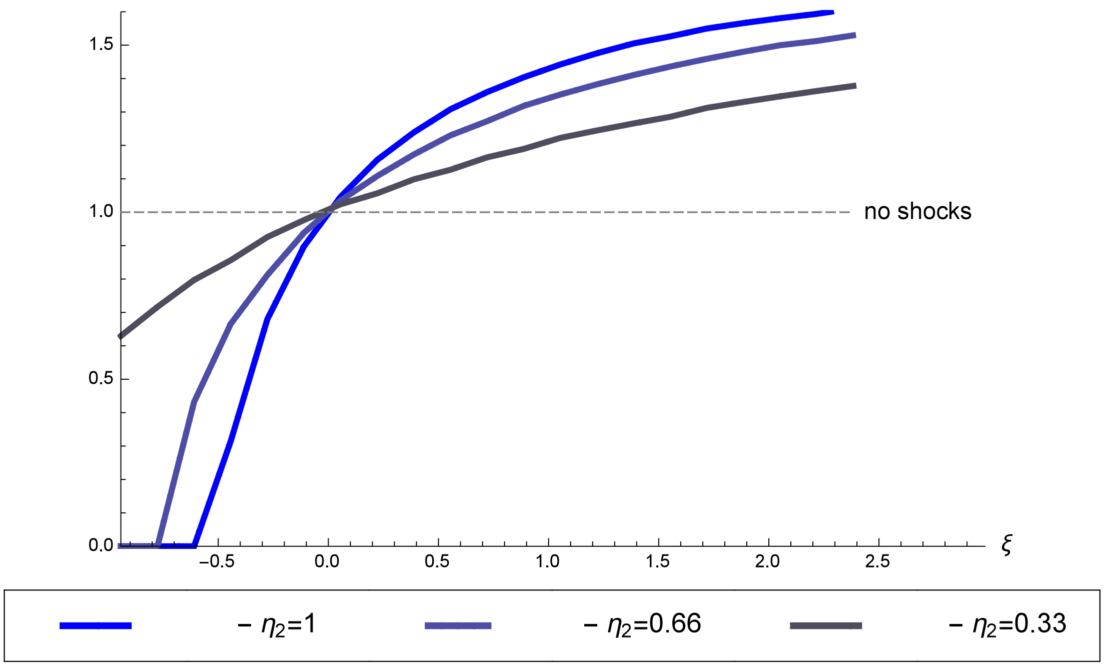


Figure 3.3: Corruption, shocks, and sluggishness of public officials' salaries

*Note* Figure shows response of corruption for a range of aggregate productivity shocks $\xi$ and when responsiveness of public officials' salary to these shocks varies. $\eta_2 = 0$ means $w^g$ fully responds to temporary shocks (no sluggishness), $\eta_2 = 1$ means $w^g$ does not respond at all. All variables are expressed as ratio to the corresponding value in the baseline without shocks ($\xi = 0$). The computation is based on the following values for the model's parameters: $\theta^{gov} = 1$, $\theta^{\tau} = 1$, $\phi = 1.07$, $\eta_1 = 1$, $\epsilon = 0.14$, $\alpha = 0.33$, $\mu = 0.3$, $d = 0.49$, $\beta = 0.2$, $\gamma = 4.76$.

allocated to each public official, and as a result, provided public officials' salary remains the same, corruption may decrease (or increase more moderately). Deregulation and structural reform may also decrease the budget per official, and consequently have similar effect.

Figure 3.1 also demonstrates another simple but important point. The aggregate productivity shocks increase output in equilibrium, but not by as much as they would in case corruption remained unchanged. For example, if corruption remained as in the baseline, a shock of 2.5 would increase output by 2.5 times. Instead, the output increases only by 1.5 times. The difference is lost due to corruption. What we observe in reality is the joint effect of corruption and shocks on output. An important implication of this is that it is hard to pin down the actual productivity shocks from the GDP fluctuations alone.

Since corruption decreases output during boom times (and vice-versa during recessions) one may say that it is a counter-cyclical tool itself. Technically this is so, but it is by far not the most efficient stabilizer. First, increased corruption is a pure efficiency loss. Instead of the output loss due to corruption, favorable macroeconomic environment could be used to accumulate buffers (fiscal space or international reserves), which will be needed when the environment becomes more adverse. Second, corruption is likely to have long-lasting effect also on steady-state productivity. It reduces stock of public capital (as opposed to counter factual with the baseline corruption) and hence productivity of labor and private capital. It may also worsen the perception of the government policies by households and firms, including international investors. These perceptions tend to be persistent.[11] It also worsens income inequality, which may undermine sustainability of economic growth.

## 3.4 Empirical evidence

I now take the models to the data and check if there is any empirical evidence that the "growth-to-corruption" effect is present. The empirical work on the topic is challenging, in particular because it is really hard to identify exogenous productivity shocks in the data. I take parsimonious approach and use only annual cross-country data. There is a significant space for further empirical exploration of this issue (for example, using firm or sector-level data, and also on identifying the shocks). It is left for future research.

I present few simple pieces of empirical evidence in Figure 3.4 and Tables 3.2-3.6, which combined together support my theoretical conjectures.

The first piece of evidence is presented in Figure 3.4. It depicts average levels of bribery across the world against three time periods: 2006-2008, when most of the economies in the world were booming (a period of positive productivity shocks); 2009-2010, when most of the economies where in recession (a period of negative productivity shocks); and 2011-2016, which is a period of partial and slow recovery from the Great Recession. One can call 2011-2016 a period of small/moderate productivity shocks. Bribery is measured from the World Bank Enterprise Surveys as a percent of

[11]The perceptions are not in the model, but the consideration is staightforward

Table 3.1: Variables in regressions: Summary Statistics

| Name | Explanation | Source | N | mean | s.d. | p(10) | p(50) | p(90) |
|---|---|---|---|---|---|---|---|---|
| bribery, total | % firms paid a bribe during period | WBES | 240 | 19.9 | 16.9 | 3.5 | 15 | 44 |
| bribery, procurement | % firms asked for a bribe in procurement | WBES | 301 | 29.9 | 21.7 | 6.5 | 26.2 | 60.3 |
| bribery, tax inspection | % firms asked for a bribe by tax inspector | WBES | 308 | 20.9 | 20.3 | 1.4 | 14.4 | 55.1 |
| WGI CoC | Control of Corruption Index (higher = better) | WGI | 3015 | 0 | 1 | -1.1 | -0.3 | 1.4 |
| TI CPI | Corruption Perception Index (higher = better) | TI | 2633 | 4.3 | 2.2 | 2.1 | 3.6 | 8 |
| shadow economy | % firms competing against informal sector | WBES | 55.1 | 19 | 29.7 | 55.6 | 77.9 | |
| GDP per capita | in thousands 2011 PPP | WB WDI | 4803 | 15.1 | 18.5 | 1.4 | 8.2 | 38.6 |
| GDP per capita growth | geometric average over a period, % | own | 4608 | 2.1 | 6.4 | -3 | 2.2 | 7.2 |
| output gap | % of potential GDP (HP-filter on annual data) | own | 5352 | 0 | 3.1 | -2.3 | 0 | 2.4 |
| ToT gap | net barter terms-of-trade index demeaned (2000=100, higher = better) | WB WDI | 4838 | 0 | 32.6 | -32.4 | -1.3 | 31.4 |

*Notes:* All summary statistics over annual data. Abbreviations: WBES - World Bank Enterprize Surveys, WGI - Worldwide Governance Indicators, TI - Transparency International, WB WDI - World Bank's World Development Indicators. WBES data are available only for 2006-2016.

surveyed firms in the economy, which paid a bribe during the year of the survey. The sample of participating countries consists of 140 mostly low and middle income economies, years covered are 2006-2016. In most countries the survey was done more than once. See Table 3.1 for definition and summary statistics.

According to Figure 3.4, bribery was at its highest on average during the boom of 2006-2008. During the crisis of 2009-2010 it went sharply down from 22% to 17%. Then, when economies started to recover, bribery went up too, but it has not yet reached the pre-crisis levels on average. We can observe similar pattern by also looking at only resource-rich countries. In 2011-2014, when commodity prices recovered to the pre-crisis levels and sometimes even higher, the bribery in RRCs was more than 30%. When the prices sharply fell in 2015-16, corruption followed too. These patterns are consistent with the model, and without "growth-to-corruption" effect they would suggest that corruption actually helps the economy.[12]

[12] The mean differences between the periods are not always statistically significant, but still, one may ask why don't we observe the opposite relationship

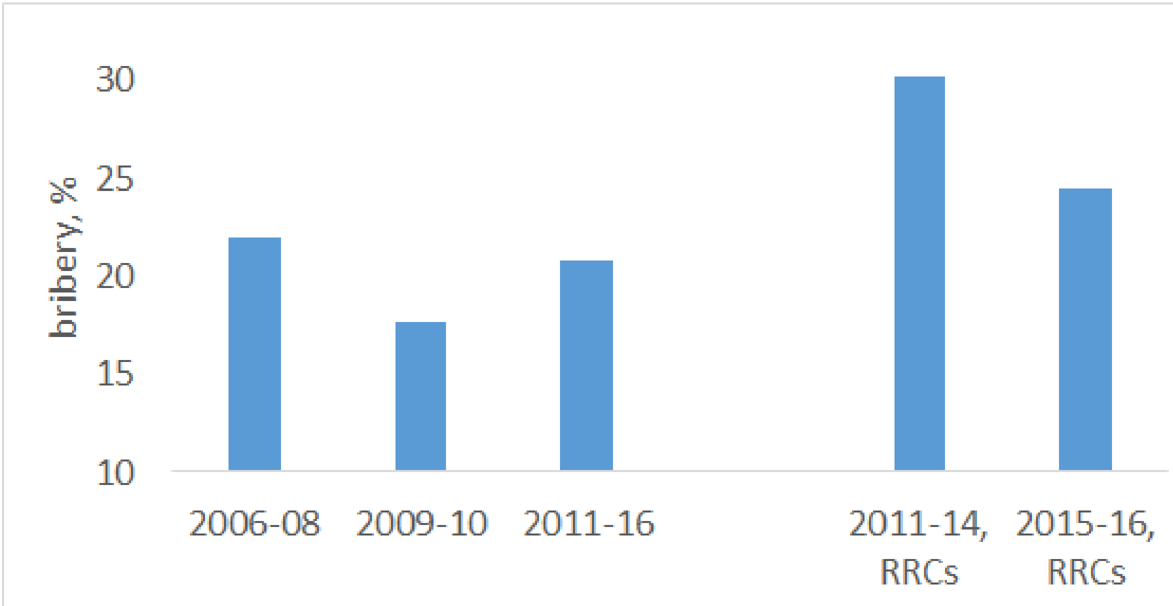


Figure 3.4: Corruption during booms and busts: Cross-country averages

*Note* Average total bribery levels during selected time periods. Datasource: WBES. See definition in Table 3.1. RRCs - resource-rich countries according to the IMF's definition (International Monetary Fund, 2012).

The second piece of evidence are several regressions that are reported in Table 3.2. The general specification that I use is:

$$< corruption >_{it} = \beta_0 + \beta_1 * < GDP\ per\ capita >_{it} + \\ + < output\ gap >_{it} + < terms - of - trade\ gap >_{it} + \psi_{it} \quad (3.35)$$

In Table 3.2 I regress bribery on the GDP per capita (proxy for the level of development) and two proxies for the productivity shocks. The first proxy is the output gap, which I obtain from country GDP data by simply running a Hodrick-Prescott filter with $\lambda = 6.25$ for each country and by using IMF WEO's forecasts to reduce the end-point bias in 2016. The output gap reflects the business cycles developments, the demand shocks which hit the productivity in short run. The second proxy is the terms-of-trade gap, which is simply demeaned net barter terms-of-trade index as measured by the World Bank (2000=100). Higher values of the index mean better ratio of export to import prices for the country. The ToT-gap reflects mostly the developments of commodity prices, which is also an important source of productivity shocks. The idea here is that commodity price cycles are usually longer than business cycles, so I simply remove the long-term (sixteen year) average. I run several specifications - adding variables one by one, then adding year fixed effects, then checking the dynamic relationship, and then splitting the sample on RRCs vs. non-RRCs.

The results in Table 3.2 are also suggestive of the "growth-to-corruption" effect. First, as expected higher level of development (higher GDP per capita) is associated with lower corruption. This is consistent through all specifications. Controlling for the level of development, larger output gap (economy being above potential) is associated with higher bribery. This is direct evidence of the "growth-to-corruption" effect, and this result is also consistent through all specifications. Taking it as causal, one percentage point increase in output gap is expected to increase bribery by 1.3-2 percentage points, which is about a 10% increase from the world's average (see Table 3.1). The evidence on terms-of-trade gap is less conclusive. Although the sign of the coefficient is as expected (in columns (2) and (3)), the statistical significance is marginal and only in case year fixed effects are added. Adding the dynamics (column 4) does not change much for the output gap (its first lag is insignificant), but suggests that ToT gap affects corruption with a lag (the 1st lag is significant). The latter effect seems to be driven exclusively by the resource-rich countries (column 5 vs. column 6), whereas the result on output gap survives for both RRCs and non-RRCs. I tried adding other potential determinants of corruption to the regressions (for example, proxies for education), but it did not change the overall pattern.[13]

In Table 3.3 I explore bribery in different sectors of the economy, as well as the response of shadow economy to the shocks, while the general specification is as in (3.35). My model suggests that corruption is mainly driven by the embezzlement, or misallocation of public funds by public officials. This is primarily corruption in procurement. At the same time, tax evasion is not as

[13] I also tried running fixed effects, but the results were inconclusive. Similarly to OLS, the signs of coefficients on output gap and ToT gap were positive, but neither of the coefficients were significant. Bribery is a slowly changing variable. In addition, both bribery and shocks are measured with significant measurement error, which is exacerbated under fixed effects. At the same time, the sign on GDP per capita changes the sign, which suggests that this variable partly captures the "growth-to-corruption" effect, and it actually dominates the relationship for a relatively short time span of 2006-2016.

Table 3.2: Corruption and aggregate productivity shocks

| | all countries | | | | RRCs | non-RRCs |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| log GDP per capita | -6.94*** | -7.29*** | -7.81*** | -7.94*** | -5.39** | -7.15*** |
| | (1.03) | (1.09) | (1.14) | (1.13) | (2.36) | (1.26) |
| output gap, % GDP | 1.34*** | 1.33*** | 1.53*** | 1.59*** | 2.21** | 1.34** |
| | (0.50) | (0.48) | (0.57) | (0.56) | (1.08) | (0.52) |
| terms of trade gap | | 0.06 | 0.08* | -0.08 | -0.01 | -0.65*** |
| | | (0.05) | (0.05) | (0.09) | (0.10) | (0.19) |
| L.output gap, % GDP | | | | 0.06 | -0.44 | 0.69 |
| | | | | (0.51) | (1.16) | (0.63) |
| L.terms of trade gap | | | | 0.19** | 0.18* | 0.52*** |
| | | | | (0.08) | (0.11) | (0.20) |
| Constant | 80.07*** | 83.05*** | 73.82*** | 74.92*** | 72.10*** | 75.73*** |
| | (9.37) | (9.87) | (11.00) | (11.30) | (24.80) | (12.00) |
| year effects | No | No | Yes | Yes | Yes | Yes |
| Observations | 238 | 218 | 218 | 218 | 71 | 138 |
| R-squared | 0.24 | 0.26 | 0.31 | 0.33 | 0.29 | 0.47 |

*Note:* Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable - bribery, total. See definitions in Table 3.1. Method of estimation - OLS. Regressions with additional controls (f.e. education) have been tried. Results (signs and magnitudes of coefficients of interest) are qualitatively similar.

responsive to shocks as corruption. This may also mean that bribery to tax inspectors should be less responsive. Columns 5 and 6 of Table 3.3 clearly demonstrate that the shadow economy, also measured from WBES, is not responsive to output gap or ToT gap. The evidence on bribery in procurement vs. tax inspection is less clear, but also supportive of the model. The coefficient on ToT gap is positive and significant in case of procurement, but insignificant and even negative in case of tax inspection. The coefficient on output gap is only marginally significant and positive for both types of bribery, although the magnitude of the point estimates is larger for procurement.[14].

Table 3.4 looks at other measures of corruption as dependent variables - WGI's Control of Corruption Index and TI's Corruption Perception Index. As opposed to the objective (actually experienced) measurement of bribery by WBES, these two measures are based mostly on subjective data - experts' perception about corruption in a country. While the coefficient on ToT gap is significant and the sign is as expected,[15] neither of the measures links corruption to the output gap. The reason might be that faster growing economies are (sometimes wrongly) perceived by experts

[14]Note that bribery in procurement and tax inspection are measured differently from the total bribery. For total bribery the firms are asked if they actually paid bribes. For bribery in procurement and tax inspection the firms were asked if they were expected to pay bribes. This may increase the measurement errors in the latter, and hence worsen the preciseness of the estimation

[15]Note that for both CoC and CPI higher value means less corruption

Table 3.3: Corruption in procurement vs. shadow economy and corruption in tax inspection

| | bribery procurement | | bribery tax inspection | | shadow economy | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| log GDP per capita | -6.73*** | -6.41*** | -5.22*** | -8.43*** | -7.18*** | -7.03*** |
| | (1.25) | (1.34) | (1.21) | (1.03) | (1.17) | (1.12) |
| output gap, % GDP | 1.21* | 1.16 | 0.71 | 1.06* | -0.34 | 0.10 |
| | (0.66) | (0.82) | (0.58) | (0.59) | (0.48) | (0.52) |
| terms of trade gap | 0.13** | 0.14** | -0.08* | 0.04 | 0.03 | 0.01 |
| | (0.06) | (0.06) | (0.05) | (0.04) | (0.05) | (0.05) |
| Constant | 88.65*** | 83.08*** | 67.41*** | 119.69*** | 116.36*** | 109.71*** |
| | (11.26) | (12.60) | (10.97) | (10.00) | (10.25) | (10.48) |
| year effects | No | Yes | No | Yes | No | Yes |
| Observations | 277 | 277 | 284 | 284 | 208 | 208 |
| R-squared | 0.14 | 0.19 | 0.09 | 0.48 | 0.14 | 0.30 |

*Note:* Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variables indicated in the first raw. See definitions in Table 3.1. Method of estimation - OLS.

Table 3.4: Bribery vs. corruption perception

| | WGI CoC | | TI CPI | |
|---|---|---|---|---|
| | (1) all observations | (2) WBES sample | (3) all observations | (4) WBES sample |
| log GDP per capita | 0.56*** | 0.56*** | 1.30*** | 1.30*** |
| | (0.01) | (0.01) | (0.03) | (0.03) |
| output gap, % GDP | -0.00 | -0.00 | -0.01 | -0.01 |
| | (0.01) | (0.01) | (0.01) | (0.01) |
| terms of trade gap | -0.00*** | -0.00*** | -0.01*** | -0.01*** |
| | (0.00) | (0.00) | (0.00) | (0.00) |
| Constant | -5.09*** | -5.07*** | -7.62*** | -7.61*** |
| | (0.10) | (0.10) | (0.24) | (0.24) |
| year effects | No | Yes | No | Yes |
| Observations | 2552 | 2552 | 2384 | 2384 |
| R-squared | 0.50 | 0.50 | 0.53 | 0.54 |

*Note:* Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variables indicated in the first raw. See definitions in Table 3.1. Method of estimation - OLS. Columns (1) and (3) use all observations available; columns (2) and (4) use only observations, which are also available for estimation in Table 3.2.

as less corrupt, unless the growth is driven by easily observed improvement in terms-of-trade.

The last piece of empirical evidence that I present in Tables 3.5 and 3.6 is a set of elementary growth regressions, where corruption is an explanatory variable:

$$< Av.\ GDP\ per\ capita\ growth >_{it} = \beta_0 + \beta_1 * < Av.\ corruption > it + \\ + \beta_2 * < GDP\ per\ capita >_{i,t-1} + \psi_{it} \quad (3.36)$$

The main goal here is to explore the relationship between growth and corruption in settings, where unaccounted exogenous productivity shocks have various degrees of importance. The expectation is that in settings with smaller importance of shocks the growth-corruption relationship should be less biased by the "growth-to-corruption" effect, and hence I expect negative significant coefficient with increasing absolute magnitude. I use two ways to vary the importance of shocks. First, I average the data over different time spans. The longer is the time span the smaller should be the influence of temporary shocks as they offset each other in the long run. Second, I restrict the estimation sample to countries, which are likely to be less exposed to the exogenous shocks.

Table 3.5 shows the results on restricted and unrestricted samples, and for three time spans - 20 years, 10 years, and 5 years. Column 1 in all three sub-tables shows the growth regressions on unrestricted sample. The coefficient on control of corruption is negative (opposite to expected) and insignificant. For 5-year span even the convergence result does not hold - the coefficient on initial log GDP per capita is insignificant. In column 2 I exclude low income countries as classified by the World Bank in 2010. LICs are more likely to be more sensitive to exogenous shocks due to their reliance on undiversified exports, and often absence of buffers to counteract the shocks. The coefficient on CoC is now positive for 10-year and 20-year span, but it is still not significant. Excluding the RRCs, as I do in column 3, increased the coefficient on CoC further, but it remains insignificant. Finally, if also small countries (those with population less than 0.5 mln) are excluded, the relationship between corruption and growth becomes as expected by most economists.

Longer time span also improves the results of the growth regressions. In Table 3.5 the coefficients under the 20-year span are always larger in magnitude than those under smaller time spans. In Table 3.6 the time spans are compared when only one time period is used for the estimation, to equalize the number of observations, but the specification is the same as in (3.36).[16] When only one period is used for the estimation the advantage of the 20-year span becomes much clearer. It is the only time span under which the growth-corruption is as we expect. Neither 10-year nor 5-year spans produce satisfactory results. This result holds both when the period ends in 2015 and when the period ends in 2010. Ugur (2014) also describes the better performance of longer time spans

[16] In Table 3.5 I used all available observations. For example, if the estimation period is 1995-2015, for a 20-year span this would still mean only one period for estimation. For a 10-year span this means two periods - one ending in 2015 and one ending in 2005. For a 5-year span this means four periods. In Table 3.6 only one period is used for all time spans.

Table 3.5: Corruption in growth regressions: Importance of time span and sample restrictions

| | (1) all countries | (2) no LICs | (3) no LICs, no RRCs | (4) no LICs, no RRCs, large |
|---|---|---|---|---|
| WGI CoC, 20y average | -0.09 | 0.06 | 0.39 | 0.70** |
| | (0.26) | (0.26) | (0.27) | (0.27) |
| L20.log GDP per capita | -0.45** | -1.15*** | -1.23*** | -1.78*** |
| | (0.20) | (0.26) | (0.27) | (0.29) |
| Constant | 6.41*** | 13.13*** | 13.68*** | 18.93*** |
| | (1.78) | (2.37) | (2.46) | (2.59) |
| Observations | 174 | 129 | 95 | 76 |
| R-squared | 0.08 | 0.23 | 0.25 | 0.42 |

| | (1) all countries | (2) no LICs | (3) no LICs, no RRCs | (4) no LICs, no RRCs, large |
|---|---|---|---|---|
| WGI CoC, 10y average | -0.21 | 0.03 | 0.37 | 0.72*** |
| | (0.23) | (0.25) | (0.24) | (0.26) |
| L10.log GDP per capita | -0.32* | -1.12*** | -1.24*** | -1.84*** |
| | (0.18) | (0.25) | (0.25) | (0.28) |
| Constant | 5.32*** | 13.10*** | 13.95*** | 19.79*** |
| | (1.58) | (2.29) | (2.32) | (2.58) |
| Observations | 362 | 270 | 193 | 154 |
| R-squared | 0.04 | 0.13 | 0.16 | 0.27 |

| | (1) all countries | (2) no LICs | (3) no LICs, no RRCs | (4) no LICs, no RRCs, large |
|---|---|---|---|---|
| WGI CoC, 5y average | -0.28 | -0.15 | 0.23 | 0.61** |
| | (0.19) | (0.20) | (0.22) | (0.24) |
| L5.log GDP per capita | -0.21 | -0.87*** | -1.11*** | -1.76*** |
| | (0.15) | (0.21) | (0.23) | (0.26) |
| Constant | 4.39*** | 10.79*** | 12.89*** | 19.19*** |
| | (1.37) | (1.92) | (2.10) | (2.44) |
| Observations | 741 | 554 | 388 | 310 |
| R-squared | 0.02 | 0.07 | 0.09 | 0.17 |

*Note:* Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is GDP growth over a corresponding period (time span): 20 years in first sub-table, 10 years in second, 5 years in third. See definitions of variables in Table 3.1. Method of estimation - OLS. Gradual restrictions on sample are applied: column (1) - all countries; (2) - all except low income countries (LICs); (3) - no LICs and no resource-rich countries; (4) - no LICs, no RRCs, and no small countries (population less than 0.5mln). Period of estimation - 1995-2015, e.g. one period is used for 20 years time span, two periods for 10 years, four periods for 5 years.

when estimating growth-corruption relationship, but he does not explain this phenomenon.

The conclusion of the empirical section is that the "growth-to-corruption" effect is likely present in the data, and it biases the growth-corruption relationship making corruption look less harmful. Therefore one has to be extremely cautious when using growth regressions to estimate the effect of corruption on growth unless aggregate productivity shocks are properly identified and controlled for. Averaging data over longer time spans may help, although even 10-year spans do not perform particularly well. Finally, even a 20-year span does not perform well if the sample is not restricted to countries, which are likely less exposed to the shocks.

Table 3.6: Corruption in growth regressions: Longer vs. shorter time span

| | last year - 2015 | | | last year - 2010 | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | 20y | 10y | 5y | 20y | 10y | 5y |
| WGI CoC, 20y average | 0.70** | | | 0.83*** | | |
| | (0.27) | | | (0.28) | | |
| L20.log GDP per capita | -1.78*** | | | -1.49*** | | |
| | (0.29) | | | (0.32) | | |
| L10.log GDP per capita | | -1.75*** | | | -1.36*** | |
| | | (0.32) | | | (0.35) | |
| L5.log GDP per capita | | | -1.66*** | | | -1.73*** |
| | | | (0.42) | | | (0.42) |
| WGI CoC, 10y average | | 0.38 | | | -0.07 | |
| | | (0.29) | | | (0.33) | |
| WGI CoC, 5y average | | | 0.59* | | | -0.00 |
| | | | (0.35) | | | (0.38) |
| Constant | 18.93*** | 18.87*** | 17.83*** | 15.66*** | 15.77*** | 19.02*** |
| | (2.59) | (3.03) | (3.96) | (2.87) | (3.22) | (3.94) |
| Observations | 76 | 78 | 78 | 65 | 78 | 78 |
| R-squared | 0.42 | 0.40 | 0.21 | 0.27 | 0.40 | 0.37 |

*Note:* Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable is GDP growth over a corresponding period (time span): 20, 10 or 5 years. See definitions of variables in Table 3.1. Method of estimation - OLS. In all regressions the sample excludes LICs, RRCs, small countries. Only one year is used in estimation: 2015 in columns (1)-(3), 2010 in (4)-(6).

## 3.5 Conclusions

This paper provides an explanation of weak empirical relationship between economic growth and corruption. The potential reason is the “growth-to-corruption” effect. When economy is hit by exogenous productivity shocks this likely changes corruption and output in the same direction. For example, a favorable shock increases tax revenue and hence opportunities for corruption. It also makes the economy grow faster. This positive correlation biases the results in growth regressions and weakens the effect of corruption (sometimes making it even positive).

To demonstrate my point I build two models - an illustrative but parsimonious stylized model and comprehensive but complex general model. In both cases an exogenous productivity shock increases tax revenue, because tax evasion does not respond to the shock much. Increased tax revenue is allocated to public officials, which increases the share that they embezzle. This result depends on how responsive is public salary to productivity shocks. The more sluggish is the salary the larger is the increase in corruption. This makes public wage bill not very efficient countercyclical fiscal policy tool, as opposed to purchase of goods and services and public investment.

I also support my findings by the empirical investigation. I find evidence that is consistent with the presence of “growth-to-corruption” effect in the data, although more work is needed to properly identify the productivity shocks and use less aggregate data. The “growth-to-corruption” effect biases the results in growth regressions, so has to be cautious when using them. Their performance can be improved by averaging data over longer time spans and restricting the sample to countries, which are less exposed to shocks.

## CHAPTER 4

## HOW CLOSE IS YOUR GOVERNMENT TO ITS PEOPLE? WORLDWIDE INDICATORS ON LOCALIZATION AND DECENTRALIZATION

This essay is coauthored with Anwar Shah, International Consultant in Public Economics and Governance, World Bank, Washington, DC, Brookings Institution, Washington, DC, and Center for Public Economics,SWUFE, Chengdu/Wenjiang, P.R. China (email - shah.anwar@gmail.com).
It has been published in an open-access peer-reviewed journal. The full citation is the following:

Ivanyna, Maksym and Anwar Shah, 2014 "How close is your government to its people? Worldwide indicators on localization and decentralization", *Economics: The Open-Access, Open-Assessment E-Journal*, vol.8, 2014-3

The coauthor does not object me including this essay as one of the chapter in my dissertation.

## 4.1 Introduction

A silent revolution (the so-called "decentralization" reforms) has been sweeping the globe since the 1980s. Hugely complex factors such as political transition in Eastern Europe, the end of colonialism, the globalization and information revolution, assertion of basic rights of citizens by courts, divisive politics and citizens' dissatisfaction with governance and their quest for responsive and accountable governance have been some of the contributing factors in gathering this storm. The main thrust of this revolution has been to move decision making closer to people to establish fair, accountable, incorruptible and responsive (F.A.I.R.) governance. The revolution has achieved varying degrees of success in government transformation across the globe due to inhibiting factors such as path dependency accentuated by powerful political, military and bureaucratic elites. While there has been monumental literature dealing with various aspects of this revolution, there has not been any systemic study providing a time capsule of the changed world as a result of this revolution. Such an assessment is critical to providing a comparative world perspective on government responsiveness and accountability and to have an informed debate on the impact of these reforms. This paper takes an important first step in this direction by providing a framework for measuring closeness of the government to its people, developing a unique data base on local governance and providing a worldwide ranking of countries using this framework. It must be noted at the outset that the paper is only concerned with formal institutions of government and governance and informal institutions of governance that exist in many countries around the globe are not within the scope of this paper and also not relevant for the specific objectives of this paper. The importance of informal institutions to deal with government failures is, nevertheless, acknowledged by the authors and also well recognized by a large body of the literature.

The paper is organized in four sections as follows. Section 4.2 is concerned with highlighting the conceptual underpinnings and developing a framework to measure closeness of the government to people. It presents a brief overview of conceptual underpinnings of moving governments closer to people. This is followed by a discussion of basic concepts in measuring government closeness to its people. It calls into question the methodologies followed by the existing literature and argues for a focus on the role and responsibilities of local governments as opposed to sub-national governments where intermediate order governments typically dominate. It is the first paper that advocates and treats various tiers of local governments (below the intermediate order of government) as the unit of comparative analysis for multi-order governance reforms.

Section 4.3 presents highlights of the unique dataset compiled for this study. It presents summary statistics on structure, size, tiers of local governments and security of their existence. It also presents summary statistics on the various sub-components of political, fiscal and administrative decentralization.

Section 4.4 is concerned with empirical implementation of the framework presented in Section 4.2. It begins by highlighting the relative importance and significance of local governments. This is followed by providing country rankings on various aspects of political, fiscal and administrative decentralization. By combining these measurements, an aggregate indicator of decentralization/localization is developed for each country. This index is then adjusted for population size, area and heterogeneity to arrive at the government closeness index. For comparative purposes, we also provide correlations of these indexes with the corruption perceptions index, human development index, GDP growth, size of the government and the ease or difficulty of doing business in the country.

Section 4.5 provides concluding observations highlighting the strength and limitations of the constructed indexes.

## 4.2 Moving Governments Closer to People: Conceptual Underpinning of the Rationale and an Empirical Framework for Comparative Analysis

### 4.2.1 Why Closeness of Government to Its People Matters: Conceptual Underpinnings

Several accepted theories provide a strong rationale for moving decision making closer to people on the grounds of efficiency, accountability, manageability and autonomy. Stigler (1957) argued that that the closer a representative government is to its people, the better it works. According to the decentralization theorem advanced by Oates (1972), "each public service should be provided by the jurisdiction having control over the minimum geographic area that would internalize benefits and costs of such provision", because:

- local governments understand the concerns of local residents;
- local decision making is responsive to the people for whom the services are intended, thus encouraging fiscal responsibility and efficiency, especially if financing of services is also decentralized;
- unnecessary layers of jurisdictions are eliminated;
- inter-jurisdictional competition and innovation are enhanced.

An ideal decentralized system ensures a level and combination of public services consistent with voters' preferences while providing incentives for the efficient provision of such services. The subsidiarity principle originating from the social teaching of the Roman Catholic Church and later adopted by the European Union has argued for assignment of taxing, spending and regulatory functions to the government closest to the people unless a convincing case can be made for higher level assignment. Recent literature have further argued that such local jurisdictions exercising such responsibilities should be organized along functional lines while overlapping geographically so that individuals are free to choose among competing service providers (see the concept of functional, overlapping and competing jurisdictions (FOCJ) by Frey & Eichenberger, 1999).

Moving government closer to people has also been advanced on the grounds of creating public value. This is because local governments have the stronger potential to tap some of the resources that come as free goods - namely, resources of consent, goodwill, good Samaritan values, community spirit (see Moore, 1996).

Moving government closer to people also matters in reducing transactions costs of individuals to hold the government to account for incompetence or malfeasance - a neo-institutional economics perspective advanced by Shah & Shah (2006). Finally, a network form of governance is needed to forge partnership of various stakeholders such as interest based network, hope based network, private for profit or for non-profit providers and government providers to improve economic and social outcomes. Such network form of governance is facilitated by having an empowered government closer to people that plays a catalytic role in facilitating such partnerships (see Dollery & Wallis, 2001).

The case for moving public decision making closer to people, however, has invited much controversy and debate. A number of influential scholars (Prud'homme, 1994; Tanzi, 1996) have questioned the merits of decentralization reforms. They have highlighted a multitude of potential "dangers of decentralization". They have argued that a decentralized fiscal and political system will result in macro instability and in a race towards bottom in public service provision as a consequence of wasteful competition for local investment promotion. According to them such a system will undermine regional equity and will result in a fragmentation of internal common market. They are also concerned that such a system will breed corruption due to personalism, weak monitoring and vertical controls, overgrazing by politicians and bureaucrats (Treisman, 2007), lack of fiscal discipline and interest group capture.

The debate over the net impact of moving decision making closer to the government on government performance, however, remains unsettled due to the non-availability of basic data that provides pertinent information on the closeness of public decision making to people in various countries. To inform this debate and to have empirical testing of various alternative hypotheses, the development of a methodology for a comparative global assessment of a government's closeness to its people is critically needed. This is precisely the focus of research in the next section.

### 4.2.2 Measuring a Government's Closeness to Its People: An Empirical Framework

A government is closer to its people if it encompasses a small geographical area and population, and it enjoys home rule and cannot be arbitrarily dismissed by higher level governments. This requires an understanding of the structure, size and significance of local governments including legal and constitutional foundation of their existence. An empirical framework for a comparative assessment of local governments must incorporate these factors. The following paragraphs elaborate on the methodology adopted in this paper to capture these elements.

*Unit of analysis.* The literature to-date without exception takes sub-national governments as a unit of analysis for measuring closeness to people. This viewpoint is hard to defend. Intermediate orders of government in large federal countries may be farther removed from people than the central governments in smaller unitary states. Therefore it would be inappropriate to compare provinces in Canada or states in Brazil, India, or the USA with municipalities, say, in Greece. This approach also vitiates against small countries such as Liechtenstein and Singapore as these countries would be mistakenly rated as having decision making far removed from people. In view of these considerations, local governments are the appropriate unit for measuring closeness to people as implemented here.

*Local government tiers.* Local government administrative structure varies across countries and the number of administrative tiers varies from 1 to 5. This has also a bearing on the closeness of the government and must be taken into consideration.

*Local government size.* Average size of local government in terms of population and area also varies across countries and it has a bearing on potential participation of citizens in decision making.

*Significance of local government.* Whether or not local governments command a significant share of national expenditures indicates their respective role in multi-order public governance. This is important in terms of their roles and responsibilities. For example, a local government may have autonomy but only a limited and highly constrained role as in India. This needs to be taken into consideration while making judgment on closeness of government decision making to people.

*Security of existence of local governments.* If local governments do not have any security of existence then their autonomy can be a hollow promise. Thus safeguards against arbitrary dismissal of local governments must be examined. This is to be assessed both by de-jure, the legal and or constitutional foundations of local government creation, and also de-facto working of such provisions.

*Empowerment of local government.* This is to be assessed on three dimensions - political, fiscal and administrative (see Boadway & Shah, 2009; Shah et al., 2004).

*Political or democratic decentralization* implies directly elected local governments thereby making elected officials accountable to local residents.. Political decentralization is to be assessed using the following criteria: direct popular elections of council members and the executive head; recall provisions for elected officials; popular participation in local elections and the contestability and competition in local elections.

*Fiscal decentralization* ensures that all elected officials weigh carefully the joys of spending someone else's money as well as the pain associated with raising revenues from the electorate and facing the possibility of being voted out. Fiscal decentralization is to be evaluated using the following criteria: range of local functions; local government autonomy in rate and base setting for local revenues; transparency and predictability and unconditionality of higher level transfers; whether finance follows function or whether revenue means match local responsibility; degree of self-financing of local expenditures; responsibility and control over municipal and social services; autonomy in local planning, autonomy in local procurement; ability to borrow domestically and from foreign sources; ability to issue domestic and foreign bonds; and higher level government assistance for capital finance.

*Administrative decentralization* empowers local governments to hire, fire and set terms of reference for local employment without making any reference to higher level governments, thereby making local officials accountable to elected officials. This is to be assessed using indicators for: freedom to hire, fire and set terms of reference for local government employment; freedom to contract out own responsibilities and forge public-private partnerships; and regulation of local activities by passing bye-laws.

## 4.3 Description of the Data

To implement the above framework, we have developed a unique and comprehensive dataset for 182 countries using data for mid 2000s (mostly 2005) on the relative importance of local governments, their security of existence and various dimensions of their empowerment. The following sections introduce and analyze various dimensions of these data.

### 4.3.1 Local Government - Basic Definitions

General government (GG) consists of 3 parts: Central Government (CG), State or Provincial Government (SG), and Local Government (LG). Each part consists of governmental units (in case of CG - only 1 unit), which are united into one or more tiers (in case of CG - 1 tier). As far as data permits, Social Security Funds are consolidated with an appropriate part of GG. We use commonly accepted definitions of LG and SG as provided by the IMF Government Finance Statistics (GFS). These definitions are quite vague which results into countries deciding for themselves and reporting corresponding data. This sometimes leads to inconsistencies. For example, France with three sub-national tiers of government reports all of them as LG, whereas Spain - which in many ways has the same administrative structure as France - reports one tier of SG, and two tiers of LG. Giving more precise definitions for LG and SG, which could be applied to all countries, and especially collecting data according to these definitions are difficult tasks. In constructing a comparative data set, we nevertheless attempted to correct for these self-reporting biases by using country specific research studies where available to make a distinction between SG and LG tiers.

### 4.3.2 Administrative structure and Size of Local Government

Our dataset contains detailed information about administrative structure of every country. In particular, we report which tiers of GG are ascribed to a local government, and number of governmental units at each tier. Tiers are needed to calculate the average population of LG administrative unit as follows:

$$LG\_pop = \frac{T * P}{\sum_{i=1}^{T} X_i}, \tag{4.1}$$

where $LG\_pop$ is the average population of an LG unit, $T$ is the number of tiers in the country, $P$ is its population, and $X_i$ is the number of LG units at the $i$'th tier. Equation (4.1) means that countries with additional tiers of LG, everything else equal, have higher average population of LG.

Of the sample of 182 countries only 20 have state governments (SG), while the rest of the countries have only local and central governments. 26 percent of the countries have one tier of local government, 46 percent have two tiers, while 23 percent and 6 percent have three and four tiers respectively.

Table 4.1 reports the summary statistics of number of LG tiers in countries by geographic region and by per capita income. On average a country has two LG tiers. Countries in South Asia and East Asia are above this average. High income countries tend to have lower number of LG tiers as compared to lower income countries.

The average tiers-adjusted population of a local government unit ranges from about several thousand people (Equatorial Guinea, Switzerland, Czech Republic, Austria) to several hundred thousand people (Somalia, DR Congo, Indonesia, Korea), with the country-average population of 101,000 people. As shown in Table 4.1 local governments in European and North American countries are significantly smaller in population size than the ones in the rest of the world, while the LG in Sub-Saharan Africa and East Asia are on average more than five times larger. Lower income countries have significantly larger population size governments.

Table 4.1: Local Government Administrative Structure and Size by Region and Income Class of Countries

| | # tiers | | av. pop., th | | av. area, tsk | |
|---|---|---|---|---|---|---|
| | mean | sd | mean | sd | mean | sd |
| Total | 2.03 | 0.8 | 101.06 | 175.47 | 2.13 | 6.95 |
| **By region:** | | | | | | |
| Southern Asia | 2.43 | 0.98 | 79.76 | 75.5 | 0.32 | 0.58 |
| Europe and Central Asia | 2 | 0.74 | 29.49 | 56.28 | 0.29 | 0.4 |
| Middle East and North Africa | 2 | 0.86 | 111.79 | 116.41 | 5.14 | 15.68 |
| Sub-Saharan Africa | 2.02 | 0.76 | 171.64 | 178.56 | 4.09 | 8 |
| Latin America and Caribbean | 1.74 | 0.63 | 63.16 | 51.88 | 1.12 | 1.73 |
| East Asia and Pacific | 2.5 | 1 | 171.4 | 379.83 | 1.22 | 2.53 |
| North America | 2 | 0 | 11.6 | 6.79 | 1.32 | 1.72 |
| **By income:** | | | | | | |
| high income | 1.69 | 0.67 | 72.51 | 119.35 | 1.13 | 2.71 |
| middle upper income | 1.76 | 0.72 | 67.3 | 78.76 | 4.09 | 13.25 |
| middle lower income | 2.35 | 0.76 | 93.92 | 246.42 | 1.12 | 2.32 |
| low income | 2.26 | 0.82 | 162.25 | 178.02 | 2.58 | 5.45 |

*Source:* Authors' calculations based upon data sources reported in Annex Table B.1.
*Note* The classification of the countries is according to the World Bank. *th* - thousand people, *tsk* - thousand square kilometers.

Table 4.2: Local Government Administrative Structure and Size : Summary Statistics

| Name | N | Min | Max | Mean | S.d. | Countries, min.value | Countries, max.value |
|---|---|---|---|---|---|---|---|
| LG # of tiers | 177 | 1 | 4 | 2 | .8 | 1 tier - 47 countries | Timor-Leste(4) Iran(4) Bangladesh(4) China(4) Madagascar(4) |
| LG average population | 177 | 1.1 th | 0.9 mln | 93 th | 128 | East-Timor (1.1 th) Eq. Guinea (1.4 th) Laos (1.5 th) Cyprus (1.6 th) Switzerland (2.7 th) | Somalia (0.9 mln) Congo DR (0.6 mln) UAE (0.5 mln) Burundi (0.5 mln) Indonesia (0.5 mln) |
| LG average area | 177 | 0.01 tsk | 70.4 tsk | 2.1 tsk | 6.9 | Czech Rep (0.01 tsk) France (0.01 tsk) Lebanon (0.01 tsk) India (0.01 tsk) Phillipines (0.01 tsk) | Lybia (70.4 tsk) Botswana (42.9 tsk) Somalia (35.4 tsk) Namibia (13.7 tsk) Congo DR (11.6 tsk) |

*Source:* Authors' calculations based upon data sources reported in Annex Table B.1.
*Note* Units of measurement: *th* - thousand people, *mln* - million people, *tsk* - thousand square kilometers.

The average area of a local government unit ranges from 0.01 thousand square kilometers (TSK) in Czech Republic to 70 TSK in Libya, with the cross-country average of 2.1 TSK (see Table 4.2). European and South Asian countries have relatively much smaller area size local government units, while Africa and Middle East have average LG areas of up to 14 times larger. LG in higher income countries are generally smaller in average area than the ones in lower income countries (see Table 4.1).

The overall pattern observed here is that higher income countries on average tend to have smaller size (both in terms of population and area) local governments with fewer tiers than lower income countries. This result reflects the fact that local governments in both Europe and North America have lower than average number of local government tiers, smaller jurisdictional area and smaller population size.

Figures B.4, B.5, B.6 depict the corresponding world maps.

### 4.3.3 The Significance of Local Government: Relative Importance and Security of Their Existence

Measurement of relative importance of local government and constitutional safeguards regarding arbitrary disbandment are critical to reaching a judgment about closeness of the government to its people. The following paragraphs highlight the variables used in this measurement.

#### 4.3.3.1 Relative Importance of Local Governments

The relative importance of local governments is measured by share of LG expenditures ($lg\_expdec$) in consolidated general government expenditures for all orders of government (GG) (see Table 4.3). This is obviously an imperfect measure of relative importance of local governments as a significant part of local government expenditures may simply be in response to higher level government mandates with little local discretion. However, data on autonomous local government expenditures are simply not available.

LG share of GG expenditures varies greatly over our sample - from virtually zero percent in a number of countries (Guyana, Mozambique, Haiti, etc.) to 59 percent in Denmark, and have near chi-square distribution with one degree of freedom. A large majority of countries (63 percent) have local government expenditure shares less than the sample average of 13 percent, and only 11 percent of the countries have LG expenditures shares higher than 30 percent (see Table 4.4, also Figure B.7 in Appendix). Only in Europe, East Asia and North America, local governments are important players in the public sector.

An alternate variable that could serve as a proxy for the relative importance of LG is LG employment ($lg\_empl$): share of LG employment in GG employment. The available data on this variable are however much less reliable and shows a great deal of year to year volatility for most developing nations. In view of this, we are left with no alternative but the use of expenditure shares

Table 4.3: Definitions of Variables for Measuring Relative Importance and Security of Existence of Local Government

| Name | Type | Definition |
|---|---|---|
| LG expenditures (Importance of LG) | Continuous: 0-100 | LG expenditures as % of GG expenditures |
| LG independence (Security of LG existence) | Discrete: 0, 0.25, 0.5, 0.75, 1 | 1 - legislative safeguards against dismissal of LG council by CG; 0.5 - LG can be dismissed under certain circumstances (prescribed by law or constitution); 0 - LG can be dismissed in an arbitrary situation. 0.25 or 0.75 - if LG are treated asymmetrically |

*Note:* By asymmetric treatment of LG we mean a situation when LGs in one country are subject to differing regulations with respect to a given decentralization variable.

Table 4.4: LG Independence and Their Relative Significance: Summary Statistics

| Name | N | Min | Max | Mean | S.d. | Countries, min.value | Countries, max.value |
|---|---|---|---|---|---|---|---|
| LG independence (Security of existence) | 182 | 0 | 1 | .25 | .28 | 0 - 89 countries | 1 - Denmark, Brazil, Austria, Norway, Sweden, Switzerland 0.75 - Poland, Iceland, Canada, Ethiopia, Germany, Belgium, Estonia, USA, Finland, Japan, Korea |
| LG expenditures | 158 | 0 | 59.4 | 15 | 14 | <0.02 - 39 countries | Denmark(59.4) Uzbekistan(55) China(51.4) Sweden(44.2) Japan(41.4) |

*Source:* Authors' calculations based upon data sources reported in Annex Table B.1.
*Note* City states are excluded from the rankings

as the only variable to measure the relative importance of local governments. LG employment is used in calculation of administrative decentralization index.

#### 4.3.3.2 Security of Existence of Local Governments

Local government security of existence is measured by LG independence($lg\_indep$). This measure attempts to capture the constitutional and legal restraints on arbitrary dismissal of local governments (see Table 4.3).

Only in 6 out of 182 countries, local governments have significant safeguards against arbitrary dismissal. LG in 48 percent of the countries have limited independence and for the remaining 49 percent of countries in our sample, local governments can be arbitrarily dismissed by higher order governments. Europe, North America and Brazil receive relatively higher scores on this indicator whereas local governments in Africa and the Middle East have almost no security of existence.

### 4.3.4 Local Government Empowerment

Local government empowerment is measured on fiscal, political, and administrative dimensions as discussed below.

#### 4.3.4.1 Fiscal Decentralization

The following variables are used to assess local government fiscal autonomy.

- *LG vertical fiscal gap($lg\_vergap$).* Vertical fiscal gap refers to the fiscal deficiency arising from differences in expenditure needs and revenue means of local government. These deficiencies are partially or fully overcome by higher level financing. Therefore, vertical fiscal gap is a measure of fiscal dependence of local government on higher level financing. The design and nature of higher level financing has implications for fiscal autonomy of local governments. It must therefore be recognized that vertical fiscal gap while being a useful concept cannot be looked in isolation of a number of related indicators to have a better judgment on local fiscal autonomy as done here. The average vertical gap in the world is 52 percent. It is somewhat higher in African and Latin American countries. However, in all regions there are local governments with high share of expenditures and high reliance on financing from above (e.g. Brazil), as well as almost non-existent LG governments that rely solely on their own financing (Togo, Niger).

- *LG taxation autonomy ($lg\_taxaut$).* This measure reflects upon a local government's empowerment and access to tools to finance own expenditures without recourse to higher level governments. It measures its ability to determine policy on local taxation (determining bases and setting rates) and as well as autonomy in tax collection and administration. Only 16 percent of the countries in our sample grant significant taxation autonomy to their LGs, while the rest grant limited or no tax autonomy to their local governments.

- *LG unconditional transfers ($lg\_transf$).* Unconditional, formula based grants preserve local autonomy. Such grants are now commonplace yet conditional grants still dominate. Europe and North America, Latin America and Southern Asia regions have high percentage of countries with high scores on this indicator.

- *LG Expenditure Autonomy($lg\_expaut$)*. As pointed out above, our main variable - Relative Importance of LG, measured by share of LG expenditures in total GG expenditures, - does not fully reflect the actual expenditure discretion that local governments have. First, LG may be simple distributors of the funding transferred to them from an upper-tier government, and have little choice over how the money in their budget should be spent. If the LG vertical gap (difference between LG expenditures and LG non-transfer revenues) is wide, and if the transfers from upper-tier governments are earmarked and discretionary, the actual spending power of LG may be much lower than it would be indicated by $lg\_expdec$. Second, even the own revenues of LG (tax revenues or borrowed funds) may strongly depend on CG policy. If LG are not allowed to regulate taxes without CG interference (usually in such cases they receive a revenue-share of a tax, which is regulated by CG), then they cannot fully rely on the revenues from these taxes, and their policy would still be partly dependent on CG.

  We adjust for the first argument - that the real LG expenditure autonomy depends on the vertical gap and the structure of intergovernmental grants - by defining LG expenditure autonomy variable ($lg\_expaut$):

$$lg\_expaut = 1 - lg\_vergap * ((1-\gamma) - (1-2\gamma) * lg\_transf), \tag{4.2}$$

  where $\gamma$ is a smoothing parameter. From (4.2), even if a country has the widest possible vertical gap ($lg\_vergap = 1$), and the smallest possible share of unconditional formula-based transfers ($lg\_transf = 0$) it still keeps $\gamma$ share of its original expenditure autonomy. This is to reflect the fact that discretionary conditional grant from CG still gives more autonomy to the LG than the direct spending of CG. At the same time, a country with a positive vertical gap and the best possible set of transfers still gets $lg\_expaut$ smaller than 1 ($lg\_expaut = 1 - \gamma * lg\_vergap$). This is to reflect the fact that even the best set of transfers does not give LG as much fiscal independence as its own revenues.

Table 4.5: Fiscal Decentralization Variables

| Name | Type | Definition |
|---|---|---|
| LG vertical gap | Continuous: 0-100 | Grants from other govt's (same- or upper-tier, also from other countries) as % of LG revenues |
| LG taxation autonomy | Discrete: 0, 0.25, 0.5, 0.75, 1 | 1 - LG regulates fully (sets base and rate) at least one major tax (property, income, or sales tax); 0.5 - LG partly regulates (sets rate or base in CG defined boundaries, or only after CG approval) at least one major tax, or fully regulates some fees and minor taxes; 0 - no administration of major taxes, partial administration of minor taxes; 0.25 or 0.75 - LG are treated asymmetrically |
| LG unconditional transfers | Discrete: 0, 0.25, 0.5, 0.75, 1 | 1 - at least half of transfers (to LG budgets from same- or upper-tier governments) are unconditional and formula-based; 0.5 - quarter to half of transfers are unconditional and formula-based; 0 - all transfers are either conditional or discretionary; 0.25 or 0.75 - LG are treated asymmetrically |
| LG expenditure autonomy | Continuous: 0 - 1 | Derivative of *LG unconditional transfers* and *LG vertical gap*. See formula (4.2) |
| LG borrowing freedom | Discrete: 0, 0.25, 0.5, 0.75, 1 | 1 - borrowing is not regulated by CG; 0.5 - borrowing only from CG or under CG approval or regulation; 0 - borrowing is not allowed; 0.25 or 0.75 - LG are treated asymmetrically |

*Note:* By asymmetric treatment of LG we mean a situation when LGs in one country are subject to differing regulations with respect to a given decentralization variable.

Table 4.6: Fiscal Decentralization: Summary Statistics

| Name | N | Min | Max | Mean | S.d. | Countries, min.value | Countries, max.value |
|---|---|---|---|---|---|---|---|
| LG vert. gap | 123 | 0 | 100 | 66 | 32 | Niger(0) Togo(0) Iran (5.9) Iceland(9.2) Romania(9.5) | Syria(100) Uganda(90) Burundi(90) India(90) Burkina Faso(90) |
| LG tax. autonomy | 158 | 0 | 1 | .34 | .35 | 0 - 71 country | 1 - 25 countries |
| LG uncond. transfers | 159 | 0 | 1 | .43 | .4 | 0 - 56 countries | 1 - 45 countries |
| LG exp. autonomy | 182 | 0.25 | 1 | .63 | .26 | 0.25 - 39 countries | Togo(1), Niger(1), Hong Kong SAR, China(1), Singapore(1), Iceland(0.98) |
| LG borr. freedom | 160 | 0 | 1 | .27 | .36 | 0 - 89 countries | 1 - 22 countries |

*Source:* Authors' calculations based upon data sources reported in Annex Table B.1.

- *LG borrowing freedom ($lg\_borrow$).* Can LG borrow money to satisfy their capital finance needs? Can the borrowing be done without consent or regulation of CG? 89 of 160 countries in our sample forbid any kind of borrowing by LGs, while only in 22 countries LGs are allowed to borrow without any restrictions. Local borrowing rules are more accommodating in Europe and Latin America.

The descriptions, definitions and sample statistics of fiscal decentralization variables are reported in Tables 4.5 and 4.6. Frequency distributions are reported in Figures B.1, B.2, and B.3. Figure B.8 depicts the corresponding world maps.

#### 4.3.4.2 Political Decentralization

Political decentralization refers to home rule for local self-governance. This is examined using the following criteria.

- *LG legislative election($lg\_legel$).* Are legislative bodies at the local level elected or appointed? Is the truth somewhere in between? (For example, a part of council is appointed, another part is elected, or members of councils are elected from a list pre-approved by CG.)

  Elected local councils are now commonplace around the world with only 34 percent of the countries in the sample having any restraints on popular elections of legislative councils at the local level, and only 14 countries appoint local councils . Middle East and Sub-Saharan Africa are lagging behind the rest of the world in permitting directly elected local councils.

- *LG executive election($lg\_exel$).* Are executive heads (mayors) at the local level elected - directly or indirectly - or appointed? Direct elections of mayors are not yet commonplace with some restrictions on direct elections in 79 percent of the countries. Thirty-six countries have no restrictions, while in 36 countries mayors are appointed at all LG tiers. While Africa and Middle East are traditionally lagging behind, European countries also receive relatively low scores on this indicator as most of the countries have some tiers of local government with appointed or indirectly elected mayors.

- *Direct democracy provisions($lg\_dirdem$).* Are there legislative provisions for obligatory local referenda for major spending, taxing and regulatory decisions, recall of public officials, and requirement for direct citizen participation in local decision making processes?

  Only three countries in our sample (Switzerland, Japan and USA) have direct democracy provisions (as defined in Table 4.7) prescribed in their national or state constitutions. About 40 percent of countries in the sample do not allow any kind of direct citizen participation in decision making at the local level. North American, European and Latin American countries have in recent years introduced isolated provisions for direct democracy, while in Africa and Middle East such people empowerment is virtually non-existent.

Table 4.7: Political Decentralization Variables

| Name | Type | Definition |
|---|---|---|
| LG legislative election | App. continuous: 0-1 | Final value: average over all tiers considered; for each tier: 1 - whole council is directly elected; 0.5 - council is partly elected, partly appointed, council is elected indirectly, LG are treated asymmetrically; 0 - council is appointed, or does not exist |
| LG executive election | App. continuous: 0-1 | Final value: average over all tiers considered; for each tier: 1 - mayor is directly elected; 0.5 - mayor is indirectly elected, does not exist, coexist with an appointed executive, LG are treated asymmetrically; 0 - major is appointed |
| Direct democracy | Discrete: 0, 0.25, 0.5, 1 | 1 - obligatory referendum in case of certain gov't decisions (prescribed by law or constitution); 0.5 - obligatory public approval in case of certain gov't decisions (public hearings, citizen assemblies); 0.25 - leg. provisions for other forms of citizen participation (civil councils, open LG sessions, possibility to submit petition or initiate referendum); 0 - no leg. provisions for direct democracy |

*Note:* By asymmetric treatment of LG we mean a situation when LGs in one country are subject to differing regulations with respect to a given decentralization variable.

Table 4.8: Political Decentralization: Summary Statistics

| Name | N | Min | Max | Mean | S.d. | Countries, min.value | Countries, max.value |
|---|---|---|---|---|---|---|---|
| LG leg. election | 173 | 0 | 1 | .79 | .33 | 0 - 14 countries (incl. Haiti, Malaysia, Lesotho, Liberia, Oman) | 1 - 113 countries |
| LG exec. election | 169 | 0 | 1 | .48 | .36 | 0 - 36 countries | 1 - 36 countries |
| Direct democracy | 147 | 0 | 1 | .16 | .2 | 0 - 60 countries | Switzerland(1) USA(1) Japan(1) |

*Source:* Authors' calculations based upon data sources reported in Annex Table B.1.

The descriptions, definitions and sample statistics of political decentralization variables are reported in Tables 4.7 and 4.8. Frequency distributions are reported in Figures B.1, B.2, and B.3. Figure B.9 depicts the corresponding world maps.

#### 4.3.4.3 Administrative Decentralization

Our concern here is to measure the ability of local governments to hire and fire and set terms of employment of local employees as well as having regulatory control over own functions. As the latter data are not available, we are constrained to measure administrative decentralization simply by the first set of variables as follows.

- *LG HR policies (lg_hrpol).* Are LG able to conduct their own policies regarding hiring, firing and setting terms of local employment? Only 43 of 158 countries allow their LGs full discretion regarding whom and at what terms to hire or fire. Europe, North America,

Table 4.9: Administrative Decentralization Variables

| Name | Type | Definition |
|---|---|---|
| LG employment | Continuous: 0-100 | LG employment as % of GG employment (excluding health, education and police sectors) |
| LG HR policies | Discrete: 0, 0.25, 0.5, 0.75, 1 | 1 - full LG discretion over local employment (subject to general CG laws); 0.5 - partly LG discretion (hiring but terms for public employment are set by CG, hiring only to the minor posts, hiring from selected by CG candidates, hiring after CG examination); 0 - no LG discretion in hiring; 0.25 or 0.75 - LG are treated asymmetrically |

*Note:* By asymmetric treatment of LG we mean a situation when LGs in one country are subject to differing regulations with respect to a given decentralization variable.

Table 4.10: Administrative Decentralization: Summary Statistics

| Name | N | Min | Max | Mean | S.d. | Countries, min.value | Countries, max.value |
|---|---|---|---|---|---|---|---|
| LG employment | 144 | 0 | 92 | 23 | 23 | close to 0 - 7 countries | China(90) Albania(80) Norway(80) Finland(80) Sweden(80) |
| LG HR policies | 158 | 0 | 1 | .37 | .41 | 0 - 77 countries | 1 - 43 countries |

*Source:* Authors' calculations based upon data sources reported in Annex Table B.1.

Australia, and Latin America are leaders on this indicator. Many more countries (77) make this kind of decisions only at the central level even for local employees.

- *LG employment ($lg\_empl$)*: share of LG employment in GG employment. Country average for LG employment is estimated to be 26 percent. However, about 34 percent of the countries in our sample report more than 30 percent of public workforce to be employed at the local level.

The descriptions, definitions and sample statistics of administrative decentralization variables are reported in Tables 4.9 and 4.10. Frequency distributions are reported in Figures B.1, B.2, and B.3. Figure B.9 depicts the corresponding world maps.

## 4.4 Worldwide Ranking of Countries on Various Dimensions of Closeness of Their Governments to the People

Our main assumption is that decentralization of local governments matters only when local governments are important players in the public sector as measured by their share of general government expenditures, and have security of existence. Indeed, it is hard to believe that local governments - however politically or administratively independent they are from the center - have any ability to serve their residents if they do not command significant budgetary resources and if they can be dissolved at will by a higher order government. These two variables adjusted by the degree of political, fiscal and administrative decentralization form the basis of our aggregate country rankings on "closeness" or "decentralization" nexus.

In the following, political, fiscal and administrative decentralization sub-indexes are first constructed for sample countries. These indexes are then aggregated to develop a composite decentralization index (DI). Finally this index is adjusted for heterogeneity, area and population of LGs to develop an index of government closeness to its people - the so-called Government Closeness Index (CGI).

### 4.4.1 Fiscal Decentralization Index

The formula for our fiscal decentralization index (*fdi*) is the following:

$$fdi = lg\_expaut * (\gamma + (1-\gamma)/2 * (lg\_taxaut + lg\_borrow), \tag{4.3}$$

where $lg\_expaut$ is local expenditure autonomy, $lg\_taxaut$ is tax autonomy and $lg\_borrow$ represents legal empowerment for local borrowing. Again, $\gamma$ is a smoothing parameter. This index penalizes those countries, where LG do not have taxation autonomy nor borrowing freedom, however, it may still be positive for these countries (equal to $\gamma$ share of $lg\_expaut$) reflecting the fact that own revenues do grant some degree of discretion to LG. At the same time, countries with full taxation autonomy and borrowing freedom get an index, which is equal to $lg\_expaut$.

If there is no data on $lg\_taxaut$ or $lg\_borrow$ then the worst possible values are assumed: $lg\_taxaut = lg\_borrow = 0$.

### 4.4.2 Political Decentralization Index

This index is constructed by simply taking the average variables described in the earlier section:

$$pdi = \frac{1}{3}(lg\_legel + lg\_exel + lg\_dirdem) \tag{4.4}$$

Table 4.11: Indexes of Decentralization: Summary Statistics

| Name | N | Min | Max | Mean | S.d. | Countries, min.value |
|---|---|---|---|---|---|---|
| Fisc. dec | 158 | .06 | 1 | .33 | .25 | 0-0.01 - 69 countries |
| Pol. dec. | 182 | 0 | 1 | .47 | .24 | 0 - 18 countries |
| Adm. dec. | 182 | 0 | .9 | .3 | .28 | 0 - 32 countries |
| Dec. index | 158 | 0 | 34 | 2 | 4.5 | 0-0.01 - 56 countries |

*Source:* Authors' calculations based upon data sources reported in Annex Table B.1.

Table 4.12: Decentralization Indexes: Top Ten Leading Countries

| N | fiscal | political | administrative | overall |
|---|---|---|---|---|
| 1 | Hong Kong SAR, China (1) | Switzerland (1) | Finland(0.9) | Denmark(34) |
| 2 | Singapore (1) | Japan(1) | Norway(0.9) | Sweden(21) |
| 3 | Switzerland (0.96) | USA (1) | Denmark(0.9) | Switzerland (20) |
| 4 | USA (0.9) | Greece (0.83) | Sweden(0.9) | Hong Kong SAR, China (17) |
| 5 | Denmark(0.9) | Uruguay (0.83) | Albania (0.9) | Singapore (17) |
| 6 | Canada (0.9) | Brazil(0.83) | Switzerland (0.9) | Finland(16) |
| 7 | Luxembourg (0.89) | Canada(0.83) | Armenia (0.88) | Japan (15) |
| 8 | Iceland (0.79) | Mexico (0.83) | Moldova(0.84) | Norway (15) |
| 9 | New Zealand (0.78) | Italy (0.83) | Hungary (0.82) | USA (14) |
| 10 | Australia (0.78) | 0.75 - 23 countries | Canada (0.75) | Korea (12) |

*Source:* Authors' calculations based upon data sources reported in Annex Table B.1.

Every variable discussed above is an essential and independent part of political decentralization. Therefore, taking the average of all variables seems to be a reasonable measure.

The index is calculated for 182 countries.

### 4.4.3 Administrative Decentralization Index

Administrative decentralization index (*adi*) is constructed as follows:

$$adi = \frac{1}{2}(lg\_hrpol + lg\_empl). \tag{4.5}$$

The index is built for 182 countries.

### 4.4.4 The Aggregate Decentralization Index

The aggregate index (di) incorporates the relative importance of LG (measured by $lg\_expdec$), the security of existence of LG (measured by $lg\_indep$), and fiscal, political and administrative indexes. It is constructed as follows:

$$di = lg\_expdec * (\gamma + (1-\gamma) * lg\_indep) * \\ * fdi * (\gamma + (1-\gamma) * pdi) * (\gamma + (1-\gamma) * adi). \quad (4.6)$$

The index penalizes countries with low political and administrative decentralization, but even if pdi=adi=0 the index is still positive if LG have some fiscal autonomy and security of existence. It reflects the fact that even fully subordinated LG without any considerable administrative responsibilities still makes fiscal decisions in more decentralized way than the CG. Note that fdi from (4.6) is also effectively smoothed using $\gamma$, see (4.3).

This index is constructed for 158 countries worldwide, using smoothing parameter $\gamma$ equal to 0.25. Together these countries comprise 98% of the world's GDP, and 99% of the world's population. The Figure 4.1 depicts distribution of the decentralization index on the World map.[1] The darker the color of a country, the more decentralized it is. European countries, North America, Brazil, and China receive high scores on this index. Countries from Latin America, former Soviet Union, and East Asia receive average decentralization index, while Middle East and African countries are the least decentralized.

The summary statistics for each index are presented in Table 4.10. Table 4.11 reports top ten countries in each index.

### 4.4.5 Choice of Smoothing Parameter $\gamma$

There are two reasons for why we decided to smooth our sub-indexes in the calculation of *di*. First, as was noted earlier, $\gamma = 0$ (no smoothing) is not an acceptable assumption. For example, under such assumption a country, where share of LG expenditures is 50 percent, but the only source of LG revenue are earmarked transfers, would be equally decentralized with the country where there are no local governments at all. The second reason is that for many decentralization variables, which enter the calculation of *di*, we assign discrete values instead of what is ideally supposed to

[1]Distributions by each subindex are provided in Figure B.11 in Appendix

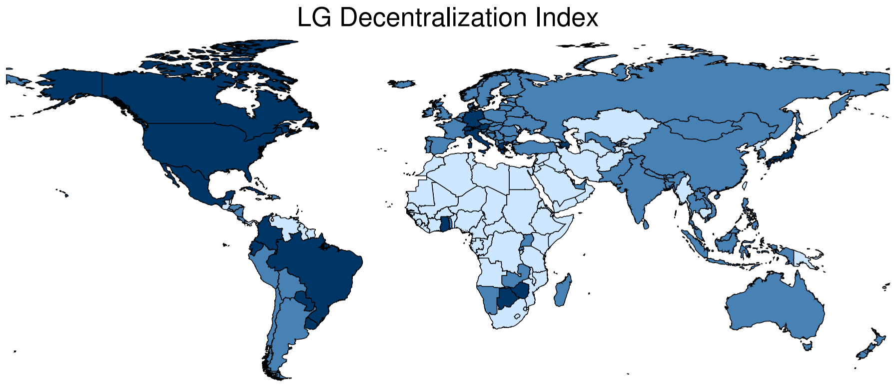


Figure 4.1: Index of Decentralization - World Map

*Source:* Authors' calculations based upon data sources reported in Annex Table B.1.
*Note* Each country is colored according to its index of decentralization, given by the formula (4.6). The lighter is the shade of color the smaller is the decentralization index. Shades of the color correspond to 0-25th, 25-50th, 50-75th, 75-100th percentiles of the corresponding world's distribution.

be a continuous measurement. For instance, LG taxation autonomy would ideally be measured by the share of LG tax revenue that comes from locally regulated taxes. Since we do not have such data for most of the countries, the best we can do is to characterize the LG taxation autonomy of each country by a discrete score. It does not mean, however, that any country with the score 0 has no taxation autonomy at all. Smoothing the score effectively means assigning a higher score - an average in a corresponding group (e.g. a group of countries with the score 0).

The choice of $\gamma = 0.25$ for our calculation of $di$, although somewhat arbitrary, is driven by several considerations. First, we assume each decentralization variable X, which we measure discreetly (usually, on a 3-score scale - 0, 0.5, 1), is in fact continuous and can be characterized by a latent unobservable variable $u \in [0, 1]$. The score X is then assigned the following way: X=0, if $u < 0.25$; X=0.5 if $0.25 < u < 0.75$; and X=1 if $u > 0.75$. Ideally, we would like our index to be adjusted by u, and not by X. While we do not observe u we can calculate its expected value given X. In the absence of any country-specific information and assuming that u is distributed uniformly the expected value of u is equal to 0.125 in case X=0; u=0.5 when X=0.5, and u=0.875 when X=1. Shifting u forward by 0.125 in order to set maximal adjustment equal to 1, we get the result that di adjustment by X, smoothed by $\gamma = 0.25$, is equivalent to the adjustment of di by expected value of

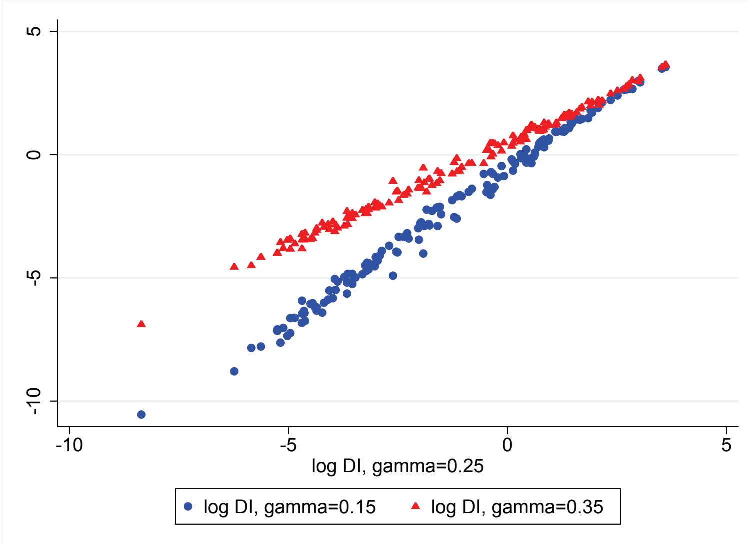


Figure 4.2: Decentralization Index under Different Gammas

u (with the shift).

A second consideration behind the somewhat arbitrary choice of $\gamma$ is that the choice itself does not significantly affect country rankings as long as $\gamma$ stays close to 0.25. Figure 4.2 shows decentralization index with $\gamma = 0.25$ against the ones with $\gamma = 0.15$ and $\gamma = 0.35$. As one can see from the figure all three indexes are very closely aligned, with the index with higher smoothing parameter assigning higher decentralization scores to countries on the lower end of distribution.

#### 4.4.6 Developing the Government "Closeness" Index by Adjusting the Decentralization Index for Heterogeneity of Size and Preferences

Our main premise is that the decentralization brings government decision making closer to the people. The decentralization indexes reported earlier indicate how significant local governments are in policymaking and public service delivery responsibilities in any country. These indexes, however, do not fully capture the actual closeness of local governments to people. This is because local governments vary widely in population, area and diversity of preferences of residents. For example, Indonesia has average LG unit population size of 0.5 mln people, while in Switzerland, for instance, the average local government population size is only 3 thousands. Population of such countries as Malta, Iceland, Belize, Maldives, etc. is lower than 0.5 mln people. It is obvious that in most aspects, e.g. accounting for heterogeneous preferences, being accountable and responsive to people, etc., even central governments in these countries are closer to people than the LGs in Indonesia. Therefore, the decentralization indexes need to be adjusted for LG population and area and other measures of a country's heterogeneity.

Our procedure of the adjustment is the following. Suppose we have a country with decentralization index $\beta$, average population of LG unit $N$, and heterogeneity index $\alpha$. Heterogeneity index is based on average area of LG unit, ethno-linguistic, age, income, urbanization composition of the country's population, as well as its geographical features (relief, versatility of climatic zones, etc.) Each resident of the country has different preferences regarding the level of governmental services provided. If an average LG provides $x$ units of the service then the disutility of a resident $i$ is $f(|x-i|, \alpha)$, where $f$ is some function of two arguments. Disutility increases with the distance between the decision of the government and the preference of the resident, and all things equal, disutility increases with heterogeneity of the country, i.e. residents are more distant in their preferences in more heterogeneous countries. Governments are assumed to be benevolent, and minimize the aggregate disutility of all residents in a region they are in charge of. Since we assume symmetric distribution of preferences in the region, benevolent government would provide $N/2$ units of the service - a level preferred by the median resident.

Given the assumptions above, the question we ask is what decentralization index $\beta'$ should $(\beta, N, \alpha)$-country have in order to produce a disutility of an average resident equal to the one in $(\beta, \bar{N}, \bar{\alpha})$-country - a country with the same decentralization index $\beta$, but some benchmark levels of average LG unit population $\bar{N}$ and heterogeneity index $\bar{\alpha}$? The answer to this question follows from the identity below:

$$\beta' AD(N, \alpha) = \beta AD(\bar{N}, \bar{\alpha}) \;\Rightarrow\; \beta' = \beta \frac{AD(\bar{N}, \bar{\alpha})}{AD(N, \alpha)}, \tag{4.7}$$

where $AD(N, \alpha)$ is the disutility of an average resident in LG with population $N$ and heterogeneity index $\alpha$, given that the government sets its service to satisfy the median resident. $AD$ can be found from the following expression:

$$AD(N,\alpha) = \frac{1}{N}\sum_{i=1}^{N} f(\left|\frac{N}{2} - i\right|, \alpha) \approx \frac{2}{N}\int_{0}^{N/2} f(\frac{N}{2} - i, \alpha)di, \tag{4.8}$$

where in last equality we use approximation of a sum with the integral (to simplify calculations), and our assumption about symmetric around median preferences.

For our calculation of decentralization index adjustment we take the following $f$ :

$$f(\frac{N}{2} - i, \alpha) = \ln(1 + \frac{A}{1-\alpha}\left(\frac{N}{2} - i\right)), \tag{4.9}$$

where parameter $A$ allows us to control the sensitivity of our results to large differences in average LG unit population. According to (4.9), disutility of an individual $i$ is increasing with population of LG unit, as well as with heterogeneity index $\alpha$ and sensitivity parameter A. At the same time, $f$ expresses diminishing returns to scale - the further the individuals are from the median resident the smaller is the difference between their disutilities. The median resident's disutility is zero.

Given $f$, the AD from (4.8) becomes:

$$\frac{2}{N}\int_{0}^{N/2} \ln(1 + \frac{A}{1-\alpha}\left(\frac{N}{2} - i\right))di = \left(\frac{1-\alpha}{A}\frac{2}{N} + 1\right)\ln(1 + \frac{A}{1-\alpha}\frac{N}{2}) - 1. \tag{4.10}$$

First, we assume there is no heterogeneity, i.e. $\alpha = 0$. By choosing different $A$'s we consider three scenarios: sensitive ($A = 0.01$), moderate ($A = 0.1$), and conservative ($A = 1$).[2] Then we introduce heterogeneity in the moderate scenario. At first, we base our $\alpha$ only on the average LG unit area. Then the heterogeneity index is extended to account for additional variables. These are age, residency, income, ethnic, religious, linguistic structure of population, country's area, relief heterogeneity (difference between highest and lowest points), and climate heterogeneity (difference between highest and lowest latitude). The reference parameters $\bar{N}, \bar{\alpha}$ are taken to be corresponding medians from the sample.

[2]One must note that sensitivity here concerns only average disutility - the higher A the greater is average disutility, everything else equal. The adjustment of decentralization index, however, may go either way, because it depends on the ratio of disutility in a given country and disutility in a reference country.

Table 4.13: Government Closeness Index: Adjusting decentralization index for population and heterogeneity

| N | no adj. | $A = 0.01$ | $A = 0.1$ | $A = 1$ | adj. area ($A = 0.1$) | heterogeneity ($A = 0.1$) |
|---|---|---|---|---|---|---|
| (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| 1 | Denmark (34) | Switzerland (46) | Switzerland (34) | Denmark (31) | Switzerland (34) | Switzerland (32) |
| 2 | Sweden (21) | Denmark (28) | Denmark (30) | Switzerland (29) | Denmark (30) | Denmark (31) |
| 3 | Switzerland (20) | USA (24) | Sweden (20) | Sweden (20) | USA (20) | Sweden (20) |
| 4 | Hong Kong SAR, China (17) | Finland (22) | Finland (20) | Finland (19) | Sweden (20) | Finland (20) |
| 5 | Singapore (17) | Iceland (22) | USA (19) | USA (20) | Finland (19) | USA (18) |
| 6 | Finland (16) | Sweden (19) | Norway (17) | Norway (17) | Norway (17) | Norway (17) |
| 7 | Japan (15) | Norway (18) | Iceland (16) | Iceland (17) | Iceland (16) | Iceland (17) |
| 8 | Norway (15) | Austria (12) | Japan (13) | Japan (13) | Japan (13) | Japan (13) |
| 9 | USA (14) | Japan (13) | Hong Kong SAR, China (13) | Hong Kong SAR, China (12) | Hong Kong SAR, China (13) | Hong Kong SAR, China (13) |
| 10 | Korea (12) | Hungary (12) | Austria (11) | Singapore (12) | Austria(11) | Singapore (11) |

*Source:* Authors' calculations based upon data sources reported in Annex Table B.1.
*Note* In each column top ten most decentralized countries are presented. Indexes are adjusted for: columns 3-5 - only for average LG unit population, column 6 - for average LG unit population and area, 7 - for average LG unit population and heterogeneity index. The corresponding sensitivities (defined by parameter $A$) are in the brackets of column titles. The benchmark country is a hypothetical country with median parameters from the sample: $\bar{N} = 43253$, $\bar{area} = 0.076, \bar{\alpha} = 0.359$. The original decentralization index is presented in column 1.

Figure 4.3 maps all countries using the indicators for the Government Closeness Index in the world. Countries in dark blue color represent the quartile where the government is closer to its people whereas at the other extreme, countries colored light blue indicates that the public decision making is highly centralized. A closer look at the world map using these indexes points to potential hot spots for internal strife. This is borne out by recent political events. For example almost all Middle East and North African countries are shaded light blue indicating that the governments are distant from people and people are not empowered to hold these government to account. This leaves people with no voice and exit options other than extra-constitutional steps for overthrow of their governments. That is why popular movements in these countries, commonly termed as the Arab Spring, have sought revolutionary changes to empower voiceless people.

Table 4.13 presents top ten leaders in each of the five new indexes (columns 3-7), each corresponding to adjustments presented above. The decentralization index without adjustments is presented in column 2. The scenario with population adjustment (A=1) results in the smallest changes. Yet, even under this scenario, Finland, Switzerland, USA, and Iceland move up the ladder

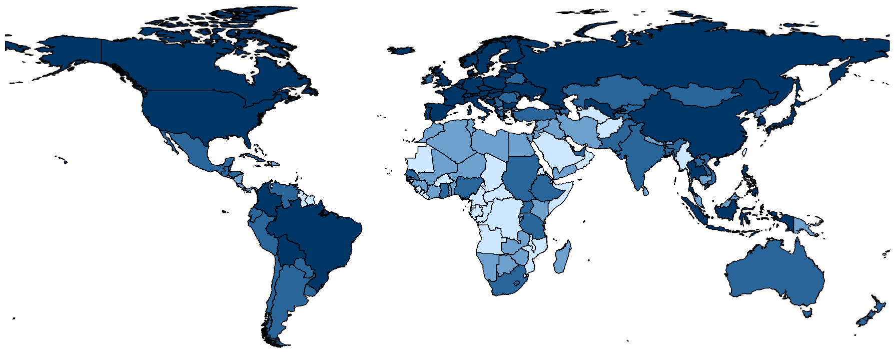


Figure 4.3: Government Closeness Index - World Map

*Source:* Authors' calculations based upon data sources reported in Annex Table B.1.
*Note* Each country is colored according to its Government Closeness Index, given by the equation (4.10) with A=0.1, and accounting for heterogeneity. The lighter is the shade of color the smaller is the index. Shades of the color correspond to 0-25th, 25-50th, 50-75th, 75-100th percentiles of the corresponding world's distribution.

as the countries with traditionally small local governments. On the other hand, countries with large average LG population e.g. China, Japan, and Republic of Korea have their rankings lowered. Moving from conservative to sensitive scenario, countries with small LG continue to get relatively higher indexes. Switzerland is the most decentralized country with this kind of adjustment, Iceland is the second. More European countries (Hungary, Georgia, Czech Republic) enter the list of leaders instead of Asian countries. Adjustment for area and heterogeneity do not change the ranking much, which may suggest that the adjustment procedure is too conservative. The only notable difference is that Switzerland gets lower index (moves down from 1st to 2nd place) because of its linguistic and ethnic heterogeneity.

Figures 4.4 and 4.5 compare the Government Closeness Index (last column of Table 4.13) with our Decentralization Index and LG expenditures, which is commonly used in the literature to measure decentralization. The difference between the latter and GCI is rather large for many countries. For example, in Uzbekistan more than 55 percent of government expenditures are local, but the country fares relatively poor on other aspects of fiscal decentralization - in particular, taxation autonomy, borrowing freedom. The LG vertical gap in Uzbekistan is 60 percent. China also has high share of local government expenditure, but fares relatively poor on political decentralization.

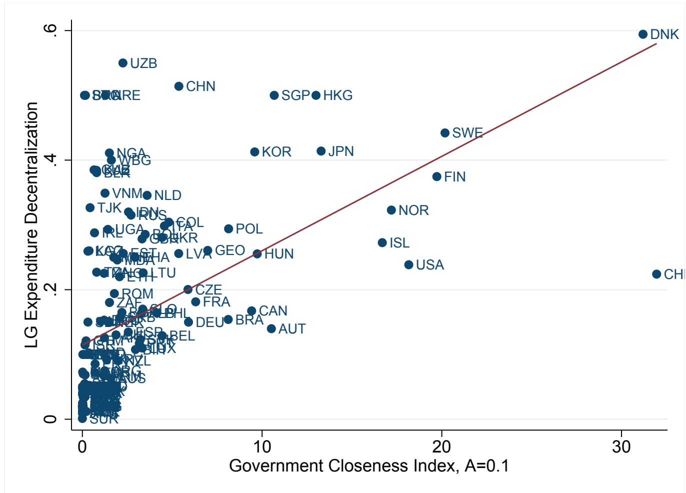


Figure 4.4: Government Closeness Index vs. LG Expenditures

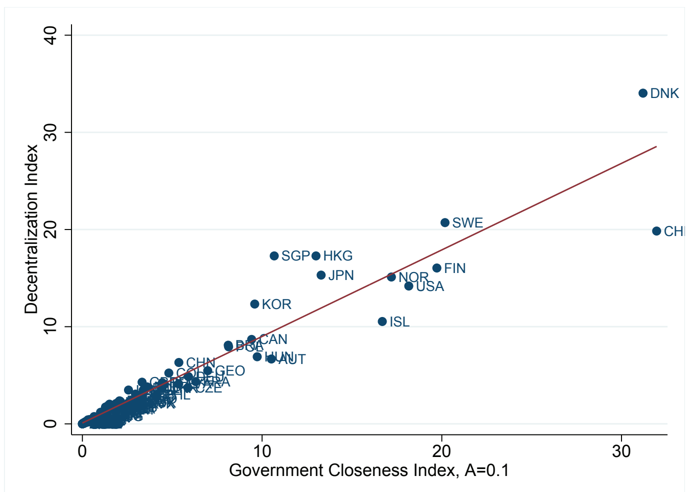


Figure 4.5: Government Closeness Index vs. Decentralization Index

Government Closeness Index and Decentralization Index are much more aligned. Nevertheless, countries with small local governments - Austria, Israel, USA, Canada, and especially Switzerland - significantly improve their ranking.

Table 4.14: Decentralization Indexes and Human Development, Corruption and Growth

| All vars as of 2005 | Human Development Index | | | Corruption Perception Index | | | Real GDP per Capita Growth, 2000-10 | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| log GCI | 0.00475** | | | 0.230*** | | | 0.122 | | |
| | (0.00217) | | | (0.0532) | | | (0.129) | | |
| log DI | | 0.00453** | | | 0.241*** | | | 0.104 | |
| | | (0.00217) | | | (0.0528) | | | (0.129) | |
| log LG expenditures | | | 0.00511 | | | 0.198** | | | 0.117 |
| | | | (0.00371) | | | (0.0960) | | | (0.218) |
| log population | -0.00379 | -0.00393 | -0.00205 | -0.269*** | -0.290*** | -0.181*** | 0.0546 | 0.0592 | 0.102 |
| | (0.00260) | (0.00267) | (0.00242) | (0.0648) | (0.0658) | (0.0637) | (0.155) | (0.158) | (0.143) |
| log GDP per capita | 0.0550*** | 0.0549*** | 0.0558*** | 0.804*** | 0.794*** | 0.856*** | -0.0784 | -0.0759 | -0.0557 |
| | (0.00380) | (0.00382) | (0.00381) | (0.0951) | (0.0946) | (0.0990) | (0.226) | (0.227) | (0.225) |
| openness | -8.87e-05 | -9.05e-05 | -8.06e-05 | 0.00123 | 0.00107 | 0.00153 | 0.000320 | 0.000307 | 0.000535 |
| | (7.23e-05) | (7.25e-05) | (7.29e-05) | (0.00180) | (0.00178) | (0.00189) | (0.00430) | (0.00431) | (0.00430) |
| literacy rate | 0.00263*** | 0.00264*** | 0.00270*** | -0.00699 | -0.00702 | -0.00345 | -0.0172 | -0.0166 | -0.0154 |
| | (0.000271) | (0.000271) | (0.000272) | (0.00666) | (0.00659) | (0.00696) | (0.0161) | (0.0161) | (0.0160) |
| 1 if in OECD | 0.00272 | 0.00312 | 0.00674 | 1.821*** | 1.826*** | 1.983*** | -3.170*** | -3.152*** | -3.069*** |
| | (0.0135) | (0.0135) | (0.0136) | (0.337) | (0.334) | (0.353) | (0.805) | (0.805) | (0.804) |
| constant | 0.118** | 0.120** | 0.111** | -1.075 | -0.861 | -1.644 | 7.140** | 7.077** | 6.890** |
| | (0.0465) | (0.0469) | (0.0480) | (1.136) | (1.137) | (1.223) | (2.764) | (2.790) | (2.830) |
| R-squared | 0.960 | 0.960 | 0.959 | 0.834 | 0.836 | 0.816 | 0.256 | 0.255 | 0.253 |
| observations | 143 | 143 | 143 | 141 | 141 | 141 | 143 | 143 | 143 |

*Note* * - significant at 10% level, ** - significant at 5% level, *** - significant at 1% level. Standard errors in parentheses. All regressions include regional dummies. Abbreviation: *GCI* - Government Closeness Index, *DI* - Decentralization Index.

### 4.4.7 Relationship of the Government Closeness Index and Decentralization Index with Government Size, Incidence of Corruption, Ease of Doing Business, Human Development, and Growth

We check the association of our government closeness (GCI) and decentralization indexes (DI) (and $lg\_expdec$ - a standard measure of decentralization in the literature) and a number of economic indicators: general level of human development (as measured by UN's Human Development Index), incidence of corruption, GDP per capita growth. We also check whether the decentralization is associated with higher government size and higher regulation burden (as measured by the number of procedures to start a business or enforce contract in a country). Corresponding OLS regressions are presented in the Tables 4.14 and 4.15. In each regression we control for income (GDP per capita), size of a country (by population and GDP), its openness to trade, state of human capital (literacy rate), and if the country is a member of OECD. Table B.2 presents the summary statistics for all variables used.

The regressions indicate that decentralized local governance as measured by GCI or DI is associated with higher human development, lower corruption, and higher growth, although in case

Table 4.15: Decentralization Indexes and Size of Government, Ease of Doing Business

| All vars as of 2005 | General Government Employment | | | Procedurec to Start Business | | | Procedures to Enforce Contracts | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| log GCI | -0.0400 | | | -0.0980 | | | -0.296 | | |
| | (0.234) | | | (0.184) | | | (0.325) | | |
| log DI | | -0.0305 | | | -0.112 | | | -0.332 | |
| | | (0.233) | | | (0.185) | | | (0.326) | |
| log LG expenditures | | | 0.317 | | | -0.314 | | | -0.145 |
| | | | (0.402) | | | (0.332) | | | (0.557) |
| log population | -0.283 | -0.286 | -0.346 | 0.829*** | 0.839*** | 0.842*** | -0.551 | -0.514 | -0.686* |
| | (0.249) | (0.253) | (0.231) | (0.230) | (0.233) | (0.220) | (0.389) | (0.397) | (0.360) |
| log GDP per capita | 0.254 | 0.254 | 0.250 | -0.401 | -0.392 | -0.402 | -0.0407 | -0.0172 | -0.122 |
| | (0.353) | (0.354) | (0.350) | (0.324) | (0.325) | (0.320) | (0.584) | (0.586) | (0.581) |
| openness | -0.00726 | -0.00726 | -0.00763 | -0.00155 | -0.00144 | -0.00146 | -0.0304*** | -0.0302*** | -0.0308*** |
| | (0.00617) | (0.00619) | (0.00611) | (0.00607) | (0.00608) | (0.00605) | (0.0108) | (0.0108) | (0.0108) |
| literacy rate | 0.0275 | 0.0272 | 0.0225 | -0.00681 | -0.00645 | -0.00570 | -0.0623 | -0.0618 | -0.0682* |
| | (0.0294) | (0.0294) | (0.0290) | (0.0215) | (0.0215) | (0.0212) | (0.0405) | (0.0404) | (0.0403) |
| 1 if in OECD | 1.692 | 1.681 | 1.636 | -2.559** | -2.563** | -2.649** | -4.685** | -4.697** | -4.856** |
| | (1.181) | (1.177) | (1.162) | (1.112) | (1.111) | (1.107) | (2.058) | (2.055) | (2.068) |
| constant | 2.134 | 2.159 | 3.448 | 7.999** | 7.837** | 7.174* | 50.13*** | 49.70*** | 51.41*** |
| | (4.199) | (4.240) | (4.284) | (3.678) | (3.719) | (3.804) | (6.930) | (6.988) | (7.114) |
| R-squared | 0.328 | 0.328 | 0.334 | 0.399 | 0.399 | 0.402 | 0.371 | 0.372 | 0.367 |
| observations | 87 | 87 | 87 | 127 | 127 | 127 | 140 | 140 | 140 |

*Note* * - significant at 10% level, ** - significant at 5% level, *** - significant at 1% level. Standard errors in parentheses. All regressions include regional dummies. Abbreviation: *GCI* - Government Closeness Index, *DI* - Decentralization Index.

of growth the coefficients on GCI and DI are not statistically significant. Table 4.15 also indicates that relatively more decentralized countries have smaller size governments and business friendly regulatory regimes but these results are not statistically significant either.

When decentralization is measured only by $lg\_expdec$ the statistical associations between decentralization and our selected economic indicators have generally lower statistical and economic significance (i.e. have lower t-statistics and magnitudes of coefficients). In case of general government employment there is also a change in the sign of coefficient. At the same time, the Government Closeness Index performs better (in terms of statistical and economical significance) than DI for Human Development Index and GDP per capital growth. For the rest of dependent variables both indexes seem to perform equally well.

## 4.5 Concluding Remarks

The silent revolution (decentralization reforms) of the past two decades has attracted strong policy and research attention worldwide. The assessment of the impact of this revolution in moving decision making closer to the people, however, remains an unanswered question. This paper takes an important first step in this direction by providing a framework of comparative measurement and developing worldwide ranking of countries on people empowerment on various aspects of government decision making. While there is a crying need for systematic collection of quality data needed for the application of the comparative framework presented here, the integration of available diverse dataset as done here has yielded promising results. For example, the government closeness indexes presented here show that one could have predicted well in advance with a fair degree of accuracy, the countries that were ripe for popular people revolt such as the one experienced through the Arab Spring or similar movements across the globe that seek revolutionary changes in the political and economic system to empower voiceless people. The ranking developed here is also helpful in understanding occasional popular revolts even in decentralized countries such as recent protests against government priorities that favored sports complexes over social services in public spending in Brazil. In this latter group of countries, citizen empowerment stimulates greater citizen activism to hold the government to account for delivering services consistent with their preferences. The indexes also provide useful barometers of the enabling environment for doing business or promoting growth and economic development and good governance. Overall they provide useful aggregate measures of government closeness to their people. We hope this paper will stimulate further research to improve upon the data and the methodology presented here as well as facilitate building common consensus in countries poorly ranked here for fundamental governance reforms. Further various indicators developed here offer scholars the opportunity for undertaking more rigorous research to further our understanding of the structure of multi-order governance and its relationship with social and economic outcomes that have a bearing on the quality of life of citizens.

# APPENDICES

**APPENDIX A**

**APPENDIX TO CHAPTER 1**

## A.1 Summary of Malawian crops

Table A.1: Malawian Crops: Cash, Energy, and Protein Value

| crop | national price in 2007, MK/kg | energy, kCal/kg | proteins, gr/Kg | legume |
|---|---|---|---|---|
| Local maize | 10.72 | 3560 | 95 | 0 |
| Composite/OPV maize | 12.5 | 3560 | 95 | 0 |
| Hybrid maize | 11.49 | 3560 | 95 | 0 |
| Cassava | 17.4 | 1090 | 9 | 0 |
| Sweet potatoes | 15.87 | 920 | 7 | 0 |
| Irish potatoes | 34.14 | 670 | 16 | 0 |
| G/Nuts | 15 | 5670 | 257 | 1 |
| Ground bean/Nzama | 16.16 | 3650 | 177 | 1 |
| Rice | 26.69 | 3570 | 75 | 0 |
| Finger millet | 34.65 | 3400 | 97 | 0 |
| Sorghum | 18.15 | 3430 | 101 | 0 |
| Pearl millet | 19.9 | 3400 | 97 | 0 |
| Beans | 54.01 | 3410 | 221 | 1 |
| Soyabean | 20.49 | 3350 | 380 | 1 |
| Pigeonpea (Nandolo) | 24.7 | 3430 | 209 | 1 |
| Burley tobacco | 156 | 0 | 0 | 0 |
| Tobacco-other | n/a | 0 | 0 | 0 |
| Cotton | n/a | 0 | 0 | 0 |
| Sugar cane | n/a | 300 | 20 | 0 |
| Cabbage | n/a | 190 | 10 | 0 |
| Tanaposi - Chinese cabbage | n/a | 190 | 10 | 0 |
| Nkhwani - pumpkin leaves | n/a | 0 | 0 | 0 |
| Okra | n/a | 310 | 16 | 0 |
| Tomato | n/a | 170 | 8 | 0 |
| Onion | n/a | 310 | 11 | 0 |
| Peas - Nsawawa | n/a | 3460 | 225 | 1 |
| Other | n/a | 0 | 0 | 0 |
| Sunflower | n/a | 3080 | 123 | 0 |
| Pepper - piri piri | n/a | 2760 | 107 | 0 |
| Coffee | n/a | 560 | 80 | 0 |
| Rape - mpiri wotuwa | n/a | 4940 | 196 | 0 |
| Paprika | n/a | 220 | 14 | 0 |
| Cowpeas - Khobwe - Nseula | n/a | 3420 | 234 | 1 |
| Mucuna (Kalongonda) - legume | n/a | 3430 | 234 | 1 |
| Bonongwe - leafy greens | n/a | 0 | 0 | 0 |
| Watermelon | n/a | 170 | 3 | 0 |
| Mpoza - fruit | n/a | 450 | 50 | 0 |
| Pumpkin | n/a | 190 | 9 | 0 |
| Buffalo bean (Mucuna variety) - legume | n/a | 3430 | 234 | 1 |
| Cucumber | n/a | 130 | 5 | 0 |
| Kabaifa - nut | n/a | 2620 | 70 | 0 |
| Hyacinth Bean - Nkhungudzu | n/a | 3430 | 234 | 1 |

*Note* Datasource: *national price* - community surveys in all three waves, median in the country's distribution of crop's prices; *energy, proteins* - Food Composition Tables, FAO. *legume* indicates whether the crop is nitrogen-fixing legume. MK - Malawian kwacha. Exchange rate as of September, 2011 - 165 MK for 1 USD.

## A.2 Household-time-technology panel vs. household-time panel

By definition of geometric average, the total household yield at time $t$ ($y_{it}$) is:

$$y_{it} = y_{i0t}^{1-\theta_{it}} y_{i1t}^{\theta_{it}} \Rightarrow \log y_{it} = \log y_{i0t} + \theta_{it}(\log y_{i1t} - \log y_{i0t}), \tag{A.1}$$

where $\theta_{it}$ is the land share of ML (technology 1). Expand (A.1) using (2.4) and the fact that $T_{i0} = 0$ and $T_{i1} = 1$:

$$\begin{aligned}
\log y_{it} = \beta_0 + \beta_{31} * W_{it} + \beta_{22} * F_{i0t} + \\
+ \theta_{it} \left(\beta_1 + \beta_{21} * W_{it} + \beta_{22} * (F_{i1t} - F_{i0t}) + \beta_{23} * F_{i1t}\right) = \\
= \beta_0 + \beta_1 \theta_{it} + \beta_{31} * W_{it} + \beta_{21}(\theta_{it} * W_{it}) + \\
+ \beta_{22} * F_{it} + \beta_{23}(\theta_{it} * F_{i1t}),
\end{aligned} \tag{A.2}$$

where $F_{it} = F_{i0t} + \theta_{it}(F_{i1t} - F_{i0t})$ - an arithmetic average of fertilizer use indicators over both technologies.

Therefore, the coefficient on $\theta_{it}$ in (2.5) is equal to the coefficient on $T_{ix}$ in (2.4) in the probability limit, in case both are consistent. Similarly equal are the coefficients on the other variables and their interactions.

## A.3 Additional estimation results

Table A.2: Estimation results: Specific to general

| | pooled OLS | | | | hhold-time FE | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | no IV | IV | IV | IV | no IV | IV | IV | IV |
| | b/se | b/se | b/se | b/se | b/se | b/se | b/se | b/se |
| ML technology | 0.14*** | 0.19*** | 0.19*** | 0.25*** | 0.28*** | 0.25*** | 0.28*** | 0.27*** |
| (dummy or landshare) | (0.02) | (0.02) | (0.02) | (0.02) | (0.03) | (0.03) | (0.05) | (0.05) |
| rainfall, cm | | 1.02*** | 0.99*** | 1.03*** | | | | |
| per day | | (0.21) | (0.22) | (0.20) | | | | |
| ML X rainfall | | -1.26*** | -1.61*** | -1.36*** | | -0.68 | -1.31* | -1.25* |
| | | (0.32) | (0.32) | (0.30) | | (0.69) | (0.73) | (0.70) |
| rainfall | | -19.64*** | -21.25*** | -20.07*** | | | | |
| variance, daily | | (1.96) | (2.14) | (2.00) | | | | |
| ML X rainvar | | 13.18*** | 14.13*** | 7.70*** | | 9.88** | 12.55*** | 8.69* |
| | | (2.42) | (2.65) | (2.48) | | (4.06) | (4.83) | (4.63) |
| temperature, C | | | 3.80*** | 3.53*** | | | | |
| /100 | | | (1.14) | (1.03) | | | | |
| temperature | | | -0.40 | -1.09*** | | | | |
| ^2 /100 | | | (0.31) | (0.29) | | | | |
| temperature | | | -0.07 | 0.09 | | | | |
| ^3 /100 | | | (0.08) | (0.07) | | | | |
| ML X temp | | | -5.23*** | -2.43* | | | -7.75*** | -7.41*** |
| | | | (1.40) | (1.28) | | | (2.97) | (2.84) |
| ML X temp_sq | | | -0.44 | -0.74* | | | -1.36* | -1.06 |
| | | | (0.45) | (0.42) | | | (0.75) | (0.72) |
| ML X temp_cub | | | 0.20** | 0.14 | | | 0.38** | 0.31* |
| | | | (0.10) | (0.09) | | | (0.17) | (0.17) |
| fertilizer, | | | | 3.00*** | | | | 2.86*** |
| tonnes per ha | | | | (0.15) | | | | (0.45) |
| fertilizer | | | | -1.98*** | | | | -1.22 |
| ^2 | | | | (0.36) | | | | (0.85) |
| ML X Txfert | | | | 1.07*** | | | | -1.05** |
| | | | | (0.20) | | | | (0.49) |
| ML X Txfert_sq | | | | -1.58*** | | | | 3.12* |
| | | | | (0.49) | | | | (1.67) |
| Constant | 14.42*** | 14.39*** | 14.43*** | 14.47*** | 14.35*** | 14.35*** | 14.35*** | 14.30*** |
| | (0.01) | (0.01) | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) |
| Observations | 13445 | 13305 | 13305 | 13305 | 1678 | 1676 | 1676 | 1676 |
| R-squared | 0.01 | . | . | 0.09 | 0.08 | | | |
| R-squared overall | | | | | 0.02 | 0.03 | 0.02 | 0.13 |

*Note* Standard errors (clustered by household) in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable in each regression - log_energy_yield.

Table A.3: Estimation results: Alternative yield measures

| | (1) energy b/se | (2) grain b/se | (3) cash b/se | (4) protein b/se | (5) maize equiv. b/se |
|---|---|---|---|---|---|
| ML X rainfall | -1.25$^{*}$ | -1.42$^{**}$ | -1.69$^{**}$ | -1.14 | -3.13$^{***}$ |
| | (0.70) | (0.69) | (0.79) | (0.74) | (0.83) |
| ML X rainvar | 8.69$^{*}$ | 11.96$^{***}$ | 14.08$^{***}$ | 9.29$^{*}$ | 32.98$^{***}$ |
| | (4.63) | (4.58) | (5.22) | (4.89) | (5.25) |
| ML technology | 0.27$^{***}$ | 0.19$^{***}$ | 0.36$^{***}$ | 0.53$^{***}$ | 0.10$^{*}$ |
| (dummy or landshare) | (0.05) | (0.05) | (0.06) | (0.06) | (0.06) |
| ML X temp | -7.41$^{***}$ | -6.86$^{**}$ | -8.07$^{**}$ | -9.01$^{***}$ | -15.21$^{***}$ |
| | (2.84) | (2.81) | (3.20) | (3.00) | (3.69) |
| ML X temp_sq | -1.06 | -1.03 | -0.73 | -1.46$^{*}$ | -1.18 |
| | (0.72) | (0.71) | (0.81) | (0.76) | (0.85) |
| ML X temp_cub | 0.31$^{*}$ | 0.24 | 0.16 | 0.38$^{**}$ | 0.39$^{*}$ |
| | (0.17) | (0.17) | (0.19) | (0.18) | (0.22) |
| fertilizer, | 2.86$^{***}$ | 2.46$^{***}$ | 2.44$^{***}$ | 2.91$^{***}$ | 2.64$^{***}$ |
| tonnes per ha | (0.45) | (0.44) | (0.50) | (0.47) | (0.54) |
| fertilizer | -1.22 | -0.63 | -0.30 | -1.16 | -1.08 |
| ^2 | (0.85) | (0.84) | (0.95) | (0.89) | (0.99) |
| ML X fert | -1.05$^{**}$ | -0.65 | -0.58 | -1.25$^{**}$ | -0.45 |
| | (0.49) | (0.48) | (0.55) | (0.51) | (0.58) |
| ML X fert_sq | 3.12$^{*}$ | 2.75$^{*}$ | 2.88 | 3.30$^{*}$ | 1.97 |
| | (1.67) | (1.65) | (1.88) | (1.76) | (2.01) |
| Constant | 14.30$^{***}$ | 6.20$^{***}$ | 8.68$^{***}$ | 10.66$^{***}$ | 2.20$^{***}$ |
| | (0.02) | (0.02) | (0.03) | (0.03) | (0.03) |
| Observations | 1676 | 1676 | 1676 | 1676 | 1480 |
| R-squared overall | 0.13 | 0.11 | 0.12 | 0.17 | 0.10 |

*Note* Standard errors (clustered by household) in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Estimation method for all columns: household-time FE, IV used for *rainfall*. Dependent variables: column (1) - log of energy yield (main measure, for comparison purposes); (2) - log of grain yield; (3) - log of cash yield; (4) - log of protein yield; (5) - log of maize equivalent yield (Liu-Myers).

Table A.4: Choice of technology: Role of plot properties

| | pooled OLS | | | | hhold-time FE | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) b/se | (2) b/se | (3) b/se | (4) b/se | (5) b/se | (6) b/se | (7) b/se | (8) b/se |
| Distance from | 0.01$^{**}$ | | | 0.02$^{**}$ | 0.06$^{**}$ | | | 0.06$^{**}$ |
| dwelling, km | (0.01) | | | (0.01) | (0.03) | | | (0.03) |
| Texture of soil | | 0.02$^{*}$ | | | | -0.24$^{**}$ | | |
| (1 - sand to 3 - clay) | | (0.01) | | | | (0.11) | | |
| Slope (1 - flat | | 0.09$^{***}$ | | | | 0.19$^{**}$ | | |
| to 4 - steep hilly) | | (0.01) | | | | (0.09) | | |
| Number of | | | 0.00 | -0.01 | | | 0.11 | 0.01 |
| weedings | | | (0.01) | (0.02) | | | (0.08) | (0.13) |
| Constant | 0.49$^{***}$ | 0.28$^{***}$ | 0.50$^{***}$ | 0.51$^{***}$ | 0.44$^{***}$ | 0.66$^{**}$ | 0.31$^{**}$ | 0.42$^{*}$ |
| | (0.01) | (0.02) | (0.02) | (0.03) | (0.02) | (0.29) | (0.15) | (0.24) |
| Observations | 2200 | 8237 | 5188 | 2198 | 822 | 416 | 1261 | 821 |
| R-squared | 0.00 | 0.02 | 0.00 | 0.00 | 0.01 | 0.03 | 0.00 | 0.01 |
| R-squared overall | | | | | 0.00 | 0.00 | 0.00 | 0.00 |

*Note* Standard errors (clustered by household) in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Dependent variable in each regression - tM_ML (0 if M, 1 if ML).

Table A.5: Maize and legume yields in intercrop: Summary statistics

| | mean | sd |
|---|---|---|
| log total yield | 6.41 | 0.87 |
| log maize yield | 6.00 | 1.06 |
| log legume yield | 4.79 | 1.22 |

| | (1) pooled OLS b/se | (2) hhold FE b/se |
|---|---|---|
| log legume yield | $0.13^{***}$ | $0.22^{***}$ |
| | (0.01) | (0.02) |
| Constant | $5.40^{***}$ | $4.95^{***}$ |
| | (0.05) | (0.12) |
| Observations | 6117 | 6117 |
| R-squared | 0.02 | 0.06 |
| R-squared overall | | 0.02 |

*Note Left panel:* Summary statistics: total, maize, and legume yields in tonnes of grain per hectare. Only ML intercrop plots included. *Right panel:* Regressions of log maize yield on log legume yield. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Only ML intercrop plots are included.

Table A.6: Estimation results: Disaggregate Household Fixed Effects

| | (1) pooled OLS | (2) hhold-time FE | (3) hhold-time & techn. FE | (4) hhold FE, disaggregated | (5) hhold FE, aggregated | (6) hhold FE GMM |
|---|---|---|---|---|---|---|
| ML technology | $0.25^{***}$ | $0.27^{***}$ | | $0.09^{**}$ | 0.01 | $0.30^{**}$ |
| (dummy or landshare) | (0.02) | (0.05) | | (0.04) | (0.06) | (0.12) |
| rainfall, cm | $1.03^{***}$ | | | $2.92^{***}$ | $3.48^{***}$ | $2.83^{***}$ |
| per day | (0.20) | | | (0.36) | (0.39) | (0.38) |
| ML X rainfall | $-1.36^{***}$ | $-1.25^{*}$ | -1.60 | $-2.28^{***}$ | $-3.26^{***}$ | $-2.22^{***}$ |
| | (0.30) | (0.70) | (2.47) | (0.55) | (0.77) | (0.68) |
| rainfall | $-20.07^{***}$ | | | $-62.17^{***}$ | $-69.72^{***}$ | $-14.27^{***}$ |
| variance, daily | (2.00) | | | (4.29) | (4.64) | (3.63) |
| ML X rainvar | $7.70^{***}$ | $8.69^{*}$ | 12.52 | $36.58^{***}$ | $50.38^{***}$ | $15.74^{***}$ |
| | (2.48) | (4.63) | (24.16) | (4.01) | (5.53) | (5.49) |
| temperature, C | $3.53^{***}$ | | | $19.32^{***}$ | $23.09^{***}$ | $21.80^{***}$ |
| /100 | (1.03) | | | (2.56) | (3.31) | (3.39) |
| temperature | $-1.09^{***}$ | | | $-6.14^{***}$ | $-6.16^{***}$ | $2.58^{***}$ |
| ^2 /100 | (0.29) | | | (0.77) | (0.83) | (0.69) |
| temperature | 0.09 | | | $-0.51^{***}$ | $-0.64^{***}$ | $-0.90^{***}$ |
| ^3 /100 | (0.07) | | | (0.16) | (0.21) | (0.20) |
| ML X temp | $-2.43^{*}$ | $-7.41^{***}$ | $-21.24^{**}$ | $-13.37^{***}$ | $-17.93^{***}$ | $-12.44^{***}$ |
| | (1.28) | (2.84) | (8.88) | (2.39) | (3.66) | (4.72) |
| ML X temp_sq | $-0.74^{*}$ | -1.06 | 3.29 | 0.71 | 1.30 | -1.53 |
| | (0.42) | (0.72) | (4.04) | (0.63) | (0.98) | (1.15) |
| ML X temp_cub | 0.14 | $0.31^{*}$ | 0.83 | $0.25^{*}$ | $0.39^{*}$ | $0.69^{***}$ |
| | (0.09) | (0.17) | (0.60) | (0.14) | (0.23) | (0.26) |
| fertilizer, | $3.00^{***}$ | $2.86^{***}$ | $2.82^{**}$ | $3.00^{***}$ | $2.88^{***}$ | $2.72^{*}$ |
| tonnes per ha | (0.15) | (0.45) | (1.39) | (0.26) | (0.31) | (1.59) |
| fertilizer | $-1.98^{***}$ | -1.22 | -1.00 | $-2.70^{***}$ | $-3.35^{***}$ | -3.91 |
| ^2 | (0.36) | (0.85) | (2.38) | (0.51) | (0.77) | (4.36) |
| ML X Txfert | $1.07^{***}$ | $-1.05^{**}$ | -2.40 | $0.77^{**}$ | $1.29^{***}$ | 2.41 |
| | (0.20) | (0.49) | (2.01) | (0.34) | (0.44) | (2.47) |
| ML X Txfert_sq | $-1.58^{***}$ | $3.12^{*}$ | 4.40 | -1.22 | -1.36 | -9.63 |
| | (0.49) | (1.67) | (7.48) | (0.93) | (1.27) | (9.15) |
| L.log_energy_yield | | | | | | -0.09 |
| | | | | | | (0.06) |
| Constant | $14.47^{***}$ | $14.30^{***}$ | -0.00 | $14.45^{***}$ | $14.95^{***}$ | $15.47^{***}$ |
| | (0.02) | (0.02) | (0.03) | (0.01) | (0.06) | (0.89) |
| Observations | 13305 | 1676 | 196 | 6818 | 5696 | 3095 |
| R-squared | 0.09 | | | . | | |
| R-squared overall | | 0.13 | 0.09 | | 0.01 | |

*Note* Standard errors (clustered by household) in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Estimation method: column 1 - pooled OLS, column 2 - household-time FE, column 3 - household-time and technology FE, column 4 - household FE, disaggregated (e.g. no aggregation of harvest over household), column 5 - household FE, aggregated, column 6 - household FE GMM. For all columns, IV used for *rainfall*. Dependent variable for all columns - log of energy yield.

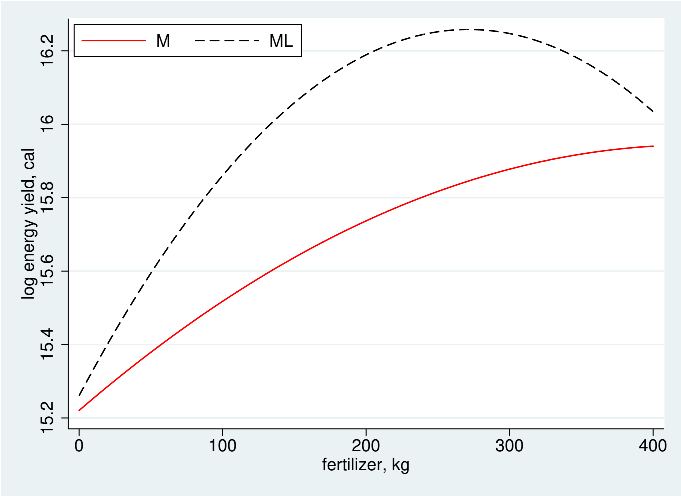


Figure A.1: Estimated response to fertilizer: M vs. ML - GMM

*Note:* Based on the estimation in Table 2.1, column 9 - GMM.

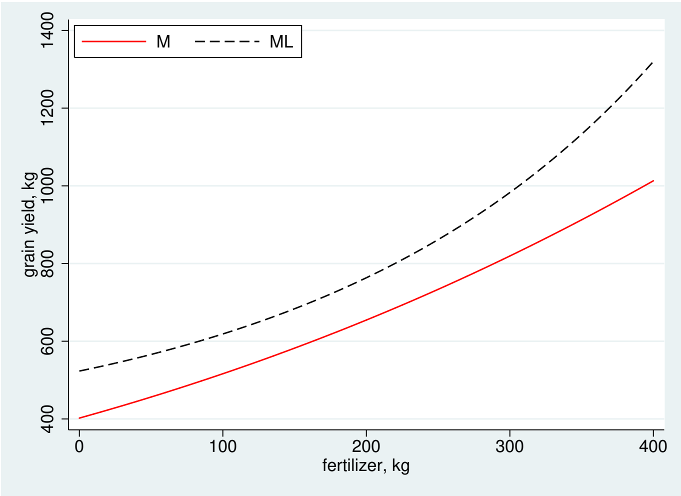


Figure A.2: Estimated response to fertilizer: M vs. ML - Grain yield

*Note:* Based on the estimation in Table 2.1, column 4 - household-time FE. Yield measure - log grain yield.

Table A.7: Estimation results: Alternative Definitions of M and ML - Part 1

| | ML | | ML13 | | ML only | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | OLS | hhold-time FE | OLS | hhold-time FE | OLS | hhold-time FE |
| =1 if ML | $0.25^{***}$ | $0.27^{***}$ | $0.24^{***}$ | $0.26^{***}$ | $0.21^{***}$ | $0.26^{***}$ |
| technology | (0.02) | (0.05) | (0.02) | (0.05) | (0.03) | (0.09) |
| rainfall, cm | $1.03^{***}$ | | $1.03^{***}$ | | $0.93^{***}$ | |
| per day | (0.20) | | (0.20) | | (0.24) | |
| ML X rainfall | $-1.36^{***}$ | $-1.25^{*}$ | $-1.42^{***}$ | $-1.19^{*}$ | $-1.35^{***}$ | -1.48 |
| | (0.30) | (0.70) | (0.30) | (0.70) | (0.37) | (1.10) |
| rainfall | $-20.07^{***}$ | | $-20.07^{***}$ | | $-21.33^{***}$ | |
| variance, daily | (2.00) | | (2.00) | | (2.40) | |
| ML X rainvar | $7.70^{***}$ | $8.69^{*}$ | $7.68^{***}$ | $9.87^{**}$ | -0.11 | 7.72 |
| | (2.48) | (4.63) | (2.52) | (4.73) | (3.55) | (8.98) |
| temperature, C | $3.53^{***}$ | | $3.53^{***}$ | | $3.91^{***}$ | |
| /100 | (1.03) | | (1.03) | | (1.19) | |
| temperature | $-1.09^{***}$ | | $-1.09^{***}$ | | $-0.91^{***}$ | |
| ^2 /100 | (0.29) | | (0.29) | | (0.33) | |
| temperature | 0.09 | | 0.09 | | 0.10 | |
| ^3 /100 | (0.07) | | (0.07) | | (0.08) | |
| ML X temp | $-2.43^{*}$ | $-7.41^{***}$ | -2.11 | $-7.60^{***}$ | -2.27 | $-14.47^{***}$ |
| | (1.28) | (2.84) | (1.31) | (2.89) | (1.77) | (4.64) |
| ML X temp_sq | $-0.74^{*}$ | -1.06 | -0.64 | -1.02 | -0.38 | $-2.48^{**}$ |
| | (0.42) | (0.72) | (0.43) | (0.72) | (0.51) | (1.03) |
| ML X temp_cub | 0.14 | $0.31^{*}$ | 0.12 | $0.31^{*}$ | 0.11 | $0.74^{***}$ |
| | (0.09) | (0.17) | (0.09) | (0.17) | (0.12) | (0.26) |
| fertilizer, | $3.00^{***}$ | $2.86^{***}$ | $2.99^{***}$ | $2.75^{***}$ | $2.88^{***}$ | $2.70^{***}$ |
| tonnes per ha | (0.15) | (0.45) | (0.14) | (0.45) | (0.16) | (0.70) |
| fertilizer | $-1.98^{***}$ | -1.22 | $-1.98^{***}$ | -1.07 | $-1.58^{***}$ | -0.32 |
| ^2 | (0.36) | (0.85) | (0.36) | (0.85) | (0.38) | (1.28) |
| ML X fert | $1.07^{***}$ | $-1.05^{**}$ | $0.95^{***}$ | $-0.97^{**}$ | $0.47^{**}$ | -0.70 |
| | (0.20) | (0.49) | (0.20) | (0.49) | (0.24) | (0.72) |
| ML X fert_sq | $-1.58^{***}$ | $3.12^{*}$ | $-1.41^{***}$ | $3.11^{*}$ | $-0.96^{*}$ | 3.65 |
| | (0.49) | (1.67) | (0.49) | (1.69) | (0.53) | (2.80) |
| Constant | $14.47^{***}$ | $14.30^{***}$ | $14.47^{***}$ | $14.31^{***}$ | $14.48^{***}$ | $14.32^{***}$ |
| | (0.02) | (0.02) | (0.02) | (0.03) | (0.02) | (0.04) |
| Observations | 13305 | 1676 | 12927 | 1636 | 8534 | 740 |
| R-squared overall | | 0.13 | | 0.12 | | 0.12 |

*Note* Standard errors (clustered by household) in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. For all columns, IV used for *rainfall*. Dependent variable for all columns - log of energy yield. Definitions of M and ML intercrop: ML - M and L grown on the same plot vs. M grown without L (other crops permitted); ML13 - both M and L are listed among three most important crops grown on a plot; ML only - M and L are the only crops grown on a plot.

Table A.8: Estimation results: Alternative Definitions of M and ML - Part 2

| | ML share>50% | | ML share>75% | | ML share>95% | | ML share=100% | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | OLS | hhold-time FE | OLS | hhold-time FE | OLS | hhold-time FE | OLS | hhold-time FE |
| =1 if ML | 0.18*** | 0.27*** | 0.16*** | 0.27*** | 0.11*** | 0.23*** | 0.16*** | 0.05 |
| technology | (0.03) | (0.07) | (0.03) | (0.07) | (0.03) | (0.08) | (0.05) | (0.16) |
| rainfall, cm | 1.01*** | | 1.00*** | | 1.06*** | | 1.08*** | |
| per day | (0.21) | | (0.21) | | (0.22) | | (0.22) | |
| ML X rainfall | -1.65*** | -1.52* | -1.71*** | -1.88** | -2.18*** | -2.22** | -2.60*** | -3.89* |
| | (0.35) | (0.87) | (0.35) | (0.94) | (0.38) | (1.11) | (0.59) | (2.09) |
| rainfall | -20.50*** | | -20.34*** | | -22.18*** | | -22.75*** | |
| variance, daily | (2.04) | | (2.09) | | (2.23) | | (2.26) | |
| ML X rainvar | 13.12*** | 9.71 | 13.30*** | 10.93* | 15.69*** | 12.72 | 18.32*** | 20.49 |
| | (2.80) | (6.23) | (2.87) | (6.61) | (3.30) | (8.18) | (4.89) | (12.54) |
| temperature, C | 3.61*** | | 3.69*** | | 3.43*** | | 3.59*** | |
| /100 | (1.04) | | (1.05) | | (1.08) | | (1.10) | |
| temperature | -1.17*** | | -1.17*** | | -1.22*** | | -1.17*** | |
| ^2 /100 | (0.30) | | (0.30) | | (0.31) | | (0.31) | |
| temperature | 0.11 | | 0.11 | | 0.14* | | 0.13* | |
| ^3 /100 | (0.07) | | (0.07) | | (0.07) | | (0.07) | |
| ML X temp | -0.45 | -12.29*** | -0.93 | -13.19*** | -0.74 | -17.43*** | -1.90 | -18.50** |
| | (1.43) | (3.71) | (1.46) | (3.97) | (1.60) | (4.56) | (2.34) | (8.16) |
| ML X temp_sq | -0.42 | -1.61* | -0.26 | -1.58 | 0.58 | -2.11* | 0.38 | -0.13 |
| | (0.53) | (0.98) | (0.53) | (1.03) | (0.56) | (1.13) | (0.89) | (2.04) |
| ML X temp_cub | 0.01 | 0.54** | 0.03 | 0.57** | -0.09 | 0.80*** | 0.03 | 0.55 |
| | (0.11) | (0.23) | (0.11) | (0.25) | (0.12) | (0.28) | (0.18) | (0.47) |
| fertilizer, | 3.01*** | 2.85*** | 3.03*** | 3.08*** | 2.98*** | 2.70*** | 2.94*** | 3.15** |
| tonnes per ha | (0.15) | (0.62) | (0.15) | (0.65) | (0.15) | (0.72) | (0.15) | (1.37) |
| fertilizer | -2.01*** | -0.42 | -2.07*** | -0.79 | -1.98*** | -0.64 | -1.93*** | -1.94 |
| ^2 | (0.36) | (1.56) | (0.36) | (1.62) | (0.36) | (1.75) | (0.37) | (2.85) |
| ML X fert | 0.73*** | -0.40 | 0.55** | -0.42 | 0.09 | -0.27 | 0.26 | -1.88 |
| | (0.26) | (0.70) | (0.26) | (0.73) | (0.27) | (0.81) | (0.41) | (1.35) |
| ML X fert_sq | -2.37** | 0.04 | -1.76* | -0.28 | -0.47 | -0.44 | 0.28 | 4.61 |
| | (1.02) | (2.81) | (1.02) | (2.99) | (1.06) | (3.26) | (1.55) | (6.11) |
| Constant | 14.46*** | 14.22*** | 14.47*** | 14.23*** | 14.46*** | 14.20*** | 14.46*** | 14.31*** |
| | (0.02) | (0.03) | (0.02) | (0.04) | (0.02) | (0.04) | (0.02) | (0.07) |
| Observations | 10720 | 1032 | 10374 | 968 | 9319 | 798 | 7001 | 274 |
| R-squared overall | | 0.08 | | 0.08 | | 0.08 | | 0.12 |

*Note* Standard errors (clustered by household) in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. For all columns, IV used for *rainfall*. Dependent variable for all columns - log of energy yield. Definitions of M and ML intercrop: ML share>X% - for ML, harvest shares of both M and L are larger than 10%, sum of the shares larger than X%; for M, its harvest share is larger than X%. X is either 50, 75, 95, or 100.

Table A.9: Estimation results: Alternative Definitions of M and ML - Part 3

| | (1) with M only & dummy | (2) with ML only & dummy | (3) 3 vs. 100-3 & rule | (4) 5 vs. 100-5 & rule |
|---|---|---|---|---|
| ML technology | 0.30*** | 0.26*** | 0.28*** | 0.29*** |
| (dummy or landshare) | (0.07) | (0.06) | (0.06) | (0.07) |
| ML X rainfall | -1.29* | -1.26* | -2.09** | -2.29*** |
| | (0.70) | (0.70) | (0.84) | (0.86) |
| ML X rainvar | 8.65* | 8.63* | 17.20*** | 17.56*** |
| | (4.64) | (4.62) | (6.30) | (6.42) |
| ML X temp | -7.55*** | -7.35** | -12.67*** | -13.55*** |
| | (2.85) | (2.86) | (3.53) | (3.58) |
| ML X temp_sq | -1.08 | -1.07 | -1.37 | -1.61* |
| | (0.72) | (0.72) | (0.87) | (0.90) |
| ML X temp_cub | 0.31* | 0.31* | 0.53** | 0.58*** |
| | (0.17) | (0.17) | (0.21) | (0.22) |
| fertilizer, | 2.84*** | 2.86*** | 3.14*** | 2.95*** |
| tonnes per ha | (0.45) | (0.45) | (0.55) | (0.56) |
| fertilizer | -1.22 | -1.22 | -1.82* | -1.54 |
| ^2 | (0.85) | (0.85) | (1.07) | (1.08) |
| ML X fert | -1.06** | -1.05** | -1.06* | -0.88 |
| | (0.49) | (0.49) | (0.62) | (0.64) |
| ML X fert_sq | 3.21* | 3.14* | 2.94 | 2.45 |
| | (1.68) | (1.67) | (2.33) | (2.44) |
| =1 if M only | 0.05 | | | |
| | (0.07) | | | |
| =1 if ML only | | 0.02 | | |
| | | (0.07) | | |
| Constant | 14.27*** | 14.30*** | 14.26*** | 14.23*** |
| | (0.05) | (0.02) | (0.03) | (0.03) |
| Observations | 1676 | 1676 | 1200 | 1150 |
| R-squared | | | | |
| R-squared overall | 0.13 | 0.13 | 0.13 | 0.12 |

*Note* Standard errors (clustered by household) in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. For all columns, IV used for *rainfall*. Dependent variable for all columns - log of energy yield. Definitions of M and ML intercrop: column (1) - as in the main specification, but regression includes "M only" dummy; (2) - as in the main specification, but regression includes "ML only" dummy; (3) - "3 vs. 100-3" rule: harvest share of M on M plots is at least 97%, harvest share of legumes on ML plots is at least 3%; (4) - "5 vs. 100-5" rule: harvest share of M on M plots is at least 95%, harvest share of legumes on ML plots is at least 5%.

Table A.10: Alternative definition of ML intercrop (ML shares > 75%) vs. other technologies (Analog of Table 2.12)

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| | ML rotation | M+L | M+L, legume share | L | M, legume share |
| =1 if ML | -0.04 | -0.39 | -0.15 | -0.39** | -0.09 |
| technology | (0.18) | (0.28) | (0.28) | (0.20) | (0.10) |
| rainfall, cm | | | | | |
| per day | | | | | |
| ML X rainfall | -4.75** | -4.37 | -5.04* | -1.52 | -1.65* |
| | (2.16) | (3.00) | (2.96) | (1.37) | (0.84) |
| rainfall | | | | | |
| variance, daily | | | | | |
| ML X rainvar | 52.98** | 38.89 | 52.99* | 29.44* | 7.17 |
| | (21.26) | (30.13) | (28.86) | (16.68) | (6.05) |
| temperature, C | | | | | |
| /100 | | | | | |
| temperature | | | | | |
| ^2 /100 | | | | | |
| temperature | | | | | |
| ^3 /100 | | | | | |
| ML X temp | 7.85 | 14.84 | 13.19 | 18.15** | -10.50*** |
| | (11.48) | (16.41) | (16.12) | (8.24) | (3.69) |
| ML X temp_sq | 3.23* | 4.23 | 4.67* | 4.89*** | -1.30 |
| | (1.82) | (2.65) | (2.59) | (1.70) | (0.97) |
| ML X temp_cub | -0.95 | -1.14 | -1.18 | -1.23** | 0.49** |
| | (0.75) | (0.99) | (0.97) | (0.53) | (0.23) |
| fertilizer, | -4.38*** | 1.87 | 3.49 | -5.89 | 3.24*** |
| tonnes per ha | (1.57) | (2.96) | (2.96) | (3.93) | (0.61) |
| fertilizer | 18.04*** | 4.29 | 1.23 | 9.96 | -1.23 |
| ^2 | (6.98) | (11.01) | (10.85) | (6.82) | (1.54) |
| ML X fert | 2.24 | -1.89 | -2.04 | 6.48 | -0.16 |
| | (1.89) | (2.93) | (2.86) | (3.98) | (0.69) |
| ML X fert_sq | -10.08 | -1.44 | -3.41 | -5.23 | -0.90 |
| | (8.79) | (12.84) | (12.53) | (7.82) | (2.77) |
| harvest share | | | 1.01** | | 1.00*** |
| of L | | | (0.43) | | (0.20) |
| Constant | 14.36*** | 14.67*** | 14.11*** | 14.71*** | 14.23*** |
| | (0.08) | (0.10) | (0.26) | (0.17) | (0.03) |
| Observations | 426 | 164 | 164 | 726 | 1032 |
| R-squared overall | 0.04 | 0.08 | 0.07 | 0.12 | 0.07 |

*Note* Standard errors (clustered by household) in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Estimation method for all columns: household-time FE, IV used for *rainfall*. Dependent variable in all columns - log of energy yield. ML intercrop is defined as all plots with harvest shares of both M and L larger than 10%, sum of the shares larger than 75%. ML intercrop is compared with: column (1) - M-L rotation; (2) - M+L (growing M and L on separate plots); (3) - M+L, controlling for the harvest share of L; (4) - L (growing only L); (5) - M, controlling for the harvest share of L in ML.

Table A.11: Maize and legume productivity for different technologies (with alternative definition of ML intercrop, ML shares >75%, analog of Table 2.13)

| | **M** vs. L | | **M only** vs. M in M-L rotation | M yield, **ML** vs. M | L yield, **ML** vs. L |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| | OLS | hhold-time FE | OLS | hhold-time FE | hhold-time FE |
| =1 if | -0.23** | -0.32 | -0.09** | -0.24*** | -1.46*** |
| technology in **bold** | (0.10) | (0.25) | (0.04) | (0.07) | (0.21) |
| rainfall, cm | 1.07*** | | 2.31*** | | |
| per day | (0.35) | | (0.36) | | |
| ML X rainfall | -0.07 | 0.90 | -1.77*** | -1.80** | -0.84 |
| | (0.39) | (0.58) | (0.43) | (0.86) | (1.48) |
| rainfall | 0.51 | | -63.81*** | | |
| variance, daily | (4.59) | | (5.15) | | |
| ML X rainvar | -20.85*** | -44.75*** | 51.30*** | 5.31 | 39.90** |
| | (4.91) | (7.76) | (5.57) | (6.26) | (17.98) |
| temperature, C | -0.08 | | 1.74 | | |
| /100 | (2.27) | | (2.16) | | |
| temperature | -3.19*** | | -3.39*** | | |
| ^2 /100 | (0.49) | | (0.45) | | |
| temperature | 0.23 | | 0.59*** | | |
| ^3 /100 | (0.16) | | (0.16) | | |
| ML X temp | 3.77 | 3.41 | -1.45 | -7.97** | 19.57** |
| | (2.44) | (3.64) | (2.47) | (3.77) | (8.98) |
| ML X temp_sq | 2.03*** | 1.37* | 2.36*** | -0.86 | 5.67*** |
| | (0.55) | (0.80) | (0.61) | (0.99) | (1.82) |
| ML X temp_cub | -0.12 | -0.05 | -0.42** | 0.39 | -1.29** |
| | (0.17) | (0.26) | (0.18) | (0.24) | (0.59) |
| fertilizer, | -7.08*** | -7.03** | 1.43*** | 3.32*** | -11.61*** |
| tonnes per ha | (1.39) | (3.27) | (0.28) | (0.63) | (4.36) |
| fertilizer | 12.38*** | 17.63 | -0.78 | -1.26 | 18.25** |
| ^2 | (2.74) | (12.30) | (0.60) | (1.57) | (7.36) |
| ML X fert | 10.18*** | 9.12*** | 2.12*** | -0.17 | 11.71*** |
| | (1.40) | (3.27) | (0.33) | (0.71) | (4.41) |
| ML X fert_sq | -14.45*** | -19.08 | -2.04*** | -1.10 | -13.68 |
| | (2.77) | (12.30) | (0.75) | (2.84) | (8.41) |
| Constant | 14.65*** | 14.52*** | 14.60*** | 14.17*** | 14.42*** |
| | (0.10) | (0.25) | (0.04) | (0.03) | (0.19) |
| Observations | 9229 | 3668 | 6804 | 1029 | 723 |
| R-squared | 0.13 | | 0.04 | | |
| R-squared overall | | 0.07 | | 0.08 | 0.48 |

*Note* Standard errors (clustered by household) in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. In all columns, IV used for *rainfall*, estimation method - household-time FE. Maize and legume productivity for different technologies (technology dummy is one for technologies in bold): columns (1) and (2) - M vs. L, log of energy yield, OLS and hhold-time FE; (3) - log of energy yield, maize only (no rotation or intercrop) vs. maize in M-L rotation; (4) - M only yield in M vs. ML; (5) - L only yield in L vs. ML. ML intercrop is defined as all plots with harvest shares of both M and L larger than 10%, sum of the shares larger than 75%.

Table A.12: Alternative functional specifications for weather variables

| | linear | | square | | cubic | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | pooled OLS | hhold-time FE | pooled OLS | hhold-time FE | pooled OLS | hhold-time FE |
| ML technology | 0.22*** | 0.24*** | -0.00 | 0.35*** | 0.09** | 0.28*** |
| (dummy or landshare) | (0.02) | (0.04) | (0.04) | (0.08) | (0.03) | (0.09) |
| fertilizer, | 2.97*** | 2.84*** | 2.96*** | 2.98*** | 3.15*** | 2.72*** |
| tonnes per ha | (0.14) | (0.45) | (0.15) | (0.45) | (0.15) | (0.47) |
| fertilizer | -1.98*** | -1.16 | -1.92*** | -1.24 | -2.07*** | -0.96 |
| ^2 | (0.37) | (0.85) | (0.38) | (0.85) | (0.36) | (0.88) |
| ML X fert | 1.00*** | -1.14** | 0.90*** | -1.03** | 0.57*** | -0.98* |
| | (0.19) | (0.49) | (0.20) | (0.49) | (0.21) | (0.51) |
| ML X fert_sq | -1.46*** | 3.50** | -1.29*** | 3.04* | -0.80 | 3.37** |
| | (0.50) | (1.67) | (0.50) | (1.67) | (0.50) | (1.71) |
| rainfall, cm | 1.13*** | | 0.66*** | | 4.71*** | |
| per day | (0.20) | | (0.25) | | (0.44) | |
| ML X rainfall | -1.38*** | -1.51** | -0.72** | -1.51** | -2.29*** | -5.08*** |
| | (0.30) | (0.74) | (0.34) | (0.70) | (0.67) | (1.60) |
| rainfall | -19.15*** | | -15.51*** | | 8.22* | |
| variance, daily | (1.95) | | (2.56) | | (4.34) | |
| ML X rainvar | 8.77*** | 11.44** | 7.12** | 19.39*** | 2.99 | 18.46* |
| | (2.38) | (4.57) | (3.28) | (6.12) | (5.73) | (9.96) |
| temperature, C | 2.34*** | | 4.16*** | | 8.31*** | |
| /100 | (0.51) | | (0.83) | | (1.36) | |
| ML X temp | -2.47*** | -3.94*** | -2.98*** | -2.21 | -4.73*** | -3.21 |
| | (0.63) | (1.39) | (1.05) | (1.97) | (1.76) | (3.43) |
| rainfall | | | -16.72*** | | 5.72** | |
| ^2 | | | (3.11) | | (2.36) | |
| ML X | | | 19.46*** | 2.33 | 5.35* | 6.36 |
| rainfall_sq | | | (4.02) | (5.22) | (2.90) | (5.11) |
| rainfall | | | 1.09 | | 6.50*** | |
| variance ^2 x 100 | | | (1.26) | | (1.02) | |
| ML X rainvar_sq | | | 1.44 | -7.32*** | 11.08*** | -10.13* |
| | | | (1.52) | (2.56) | (2.28) | (6.14) |
| temperature | | | -0.81*** | | -0.42 | |
| ^2 /100 | | | (0.21) | | (0.34) | |
| ML X temp_sq | | | 0.31 | -0.24 | -0.83 | -0.05 |
| | | | (0.26) | (0.45) | (0.52) | (0.81) |
| rainfall | | | | | -0.90*** | |
| ^3 x 100 | | | | | (0.11) | |
| ML X | | | | | 0.51*** | 1.03*** |
| rainfall_cube | | | | | (0.17) | (0.36) |
| rainfall | | | | | -44.68*** | |
| variance ^3 x 1000 | | | | | (6.79) | |
| ML X | | | | | -24.75** | 19.05 |
| rainvar_cube | | | | | (10.65) | (23.59) |
| temperature | | | | | -0.16* | |
| ^3 /100 | | | | | (0.08) | |
| ML X temp_cub | | | | | 0.26** | 0.04 |
| | | | | | (0.12) | (0.21) |
| Constant | 14.42*** | 14.30*** | 14.61*** | 14.30*** | 14.38*** | 14.30*** |
| | (0.01) | (0.02) | (0.03) | (0.02) | (0.02) | (0.03) |
| Observations | 13305 | 1676 | 13305 | 1676 | 13305 | 1676 |
| R-squared | 0.09 | | 0.08 | | 0.05 | |
| R-squared overall | | 0.13 | | 0.12 | | 0.10 |

*Note* Standard errors (clustered by household) in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. In all columns, IV used for *rainfall*, *rainfall^2*, and *rainfall^3*. Dependent variable in all specifications - log energy yield. Columns 1,2 - all weather variables (rainfall, rainfall variance, temperature) linear; column 3,4 - all weather variables linear and quadratic; column 5,6 - all weather variables linear, quadratic, and cubic.

Table A.13: M vs. ML controlling for seed input and land size

| | pooled OLS | | hhold-time FE | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| | log_energy_yield | log_energy_yield | log_energy_yield | log_energy_yield |
| ML technology | 0.21*** | 0.23*** | 0.33*** | 0.36*** |
| (dummy or landshare) | (0.05) | (0.05) | (0.07) | (0.07) |
| rainfall, cm | 2.67*** | 2.63*** | | |
| per day | (0.42) | (0.41) | | |
| ML X rainfall | -3.94*** | -3.89*** | -1.55* | -1.66* |
| | (0.56) | (0.56) | (0.91) | (0.89) |
| rainfall | -8.91*** | -9.18*** | | |
| variance, daily | (2.63) | (2.60) | | |
| ML X rainvar | 17.61*** | 17.96*** | 4.77 | 6.59 |
| | (3.51) | (3.47) | (5.48) | (5.40) |
| temperature, C | 19.65*** | 19.21*** | | |
| /100 | (1.98) | (1.96) | | |
| temperature | 2.85*** | 2.65*** | | |
| ^2 /100 | (0.39) | (0.39) | | |
| temperature | -0.85*** | -0.82*** | | |
| ^3 /100 | (0.10) | (0.10) | | |
| ML X temp | -14.50*** | -14.49*** | -8.52** | -8.75** |
| | (2.50) | (2.47) | (3.67) | (3.60) |
| ML X temp_sq | -1.63*** | -1.61*** | -1.63* | -1.48* |
| | (0.56) | (0.56) | (0.83) | (0.82) |
| ML X temp_cub | 0.64*** | 0.64*** | 0.39** | 0.36* |
| | (0.13) | (0.13) | (0.19) | (0.18) |
| fertilizer, | 3.18*** | 3.06*** | 2.66*** | 2.51*** |
| tonnes per ha | (0.22) | (0.22) | (0.49) | (0.48) |
| fertilizer | -2.47*** | -2.44*** | -1.23 | -1.29 |
| ^2 | (0.47) | (0.46) | (0.89) | (0.87) |
| ML X fert | 0.32 | 0.22 | -1.38** | -1.34** |
| | (0.30) | (0.30) | (0.58) | (0.57) |
| ML X fert_sq | -0.66 | -0.53 | 4.64** | 4.21** |
| | (0.71) | (0.70) | (1.83) | (1.80) |
| log seed input, | 0.10*** | 0.08*** | 0.06 | 0.04 |
| kg per ha | (0.01) | (0.01) | (0.05) | (0.05) |
| plot area, ha | | -0.17*** | | -0.29*** |
| | | (0.04) | | (0.06) |
| Constant | 13.89*** | 14.06*** | 14.07*** | 14.25*** |
| | (0.05) | (0.06) | (0.13) | (0.14) |
| Observations | 4954 | 4954 | 1215 | 1215 |
| R-squared | 0.20 | 0.21 | | |
| R-squared overall | | | 0.11 | 0.15 |

*Note* Standard errors (clustered by household) in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. In all columns, IV used for *rainfall*. Dependent variable in all specifications - log energy yield. Seed data is available only for waves 2 and 3 of the survey (hence, GMM estimation is not possible).

Table A.14: Estimation results: Various specification of GMM regression

| | (1) main specification | (2) all var's except yield strictly exogenous | (3) weather var's only strictly exogenous | (4) T and fert. vars lagged | (5) no time f.e. | (6) orthogonal projections |
|---|---|---|---|---|---|---|
| ML technology | 0.26*** | 0.27*** | 0.15* | 0.17 | 0.28*** | 0.26*** |
| (dummy or landshare) | (0.07) | (0.07) | (0.08) | (0.13) | (0.07) | (0.07) |
| rainfall, cm | 2.07*** | 1.32*** | 2.54*** | 1.88*** | 2.18*** | 2.07*** |
| per day | (0.32) | (0.31) | (0.34) | (0.50) | (0.32) | (0.32) |
| ML X rainfall | -2.04*** | -1.63*** | -2.09*** | -0.67 | -2.10*** | -2.04*** |
| | (0.53) | (0.53) | (0.57) | (0.83) | (0.54) | (0.53) |
| rainfall | -4.58 | -1.47 | -8.34** | -7.67* | -8.84*** | -4.57 |
| variance, daily | (3.08) | (3.00) | (3.44) | (4.36) | (3.03) | (3.07) |
| ML X rainvar | 14.69*** | 12.29*** | 17.47*** | 14.17** | 13.92*** | 14.75*** |
| | (4.45) | (4.71) | (5.06) | (6.50) | (4.47) | (4.45) |
| temperature, C | 23.55*** | 23.09*** | 25.24*** | 28.63*** | 21.01*** | 23.56*** |
| /100 | (3.16) | (3.10) | (3.17) | (4.06) | (3.13) | (3.15) |
| temperature | 3.86*** | 3.40*** | 2.82*** | 4.01*** | 3.72*** | 3.86*** |
| ^2 /100 | (0.60) | (0.56) | (0.56) | (0.83) | (0.60) | (0.60) |
| temperature | -1.15*** | -1.20*** | -1.13*** | -1.53*** | -0.99*** | -1.16*** |
| ^3 /100 | (0.17) | (0.17) | (0.17) | (0.25) | (0.17) | (0.17) |
| ML X temp | -16.63*** | -15.80*** | -17.44*** | -21.40*** | -13.82*** | -16.67*** |
| | (4.18) | (4.21) | (4.35) | (5.73) | (4.17) | (4.18) |
| ML X temp_sq | -3.80*** | -3.20*** | -2.42** | -4.69*** | -2.70*** | -3.80*** |
| | (0.99) | (0.93) | (0.95) | (1.41) | (0.98) | (0.99) |
| ML X temp_cub | 1.06*** | 1.09*** | 0.99*** | 1.54*** | 0.79*** | 1.07*** |
| | (0.23) | (0.23) | (0.23) | (0.33) | (0.22) | (0.23) |
| fertilizer, | 2.80*** | 2.88*** | 2.94*** | -2.73 | 2.97*** | 2.80*** |
| tonnes per ha | (0.37) | (0.31) | (0.39) | (2.04) | (0.37) | (0.37) |
| fertilizer | -2.93*** | -2.44*** | -3.19*** | 6.82 | -3.25*** | -2.93*** |
| ^2 | (0.76) | (0.64) | (0.74) | (5.97) | (0.78) | (0.76) |
| ML X Txfert | 0.29 | 0.49 | 0.47 | 0.40 | 0.32 | 0.29 |
| | (0.57) | (0.48) | (0.58) | (2.68) | (0.57) | (0.57) |
| ML X Txfert_sq | 0.35 | -0.47 | -0.00 | 9.36 | -0.09 | 0.34 |
| | (1.56) | (1.32) | (1.59) | (9.98) | (1.61) | (1.56) |
| L.log_energy_yield | 0.08** | 0.12*** | 0.11*** | 0.17*** | 0.03 | 0.08** |
| | (0.03) | (0.04) | (0.04) | (0.04) | (0.03) | (0.03) |
| (mean) wave=1 | 0.00 | 0.00 | 0.00 | 0.00 | | 0.00 |
| | (.) | (.) | (.) | (.) | | (.) |
| (mean) wave=2 | 12.91*** | 12.31*** | 0.00 | 11.64*** | | 12.89*** |
| | (0.46) | (0.58) | (.) | (0.63) | | (0.47) |
| (mean) wave=3 | 13.14*** | 12.54*** | 0.24*** | 12.04*** | | 13.13*** |
| | (0.46) | (0.57) | (0.04) | (0.61) | | (0.46) |
| Constant | 0.00 | 0.00 | 12.57*** | 0.00 | 13.72*** | 0.00 |
| | (.) | (.) | (0.52) | (.) | (0.46) | (.) |
| Observations | 3095 | 3095 | 3095 | 3095 | 3095 | 3095 |

*Note* Standard errors (clustered by household) in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. In all columns, IV used for *rainfall*, *rainfall^2*, and *rainfall^3*. Dependent variable in all specifications - log energy yield. Wave represents the wave of the survey (total of 3). Detailed description of the specifications: column 1 - the main specification from Table 2.9; column 2 - all variables except lagged log_energy_yield are assumed strictly exogenous; column 3 - only weather variables (rainfall rainfall variation and their interactions with T) are assumed strictly exogenous; column 4 - similar to column 3, but T and fertilizer variables are instrumented with second lags; column 5 - similar to column 1, but time fixed effects are not included; column 6 - similar to column 1, but orthogonal projections are used instead of differencing.

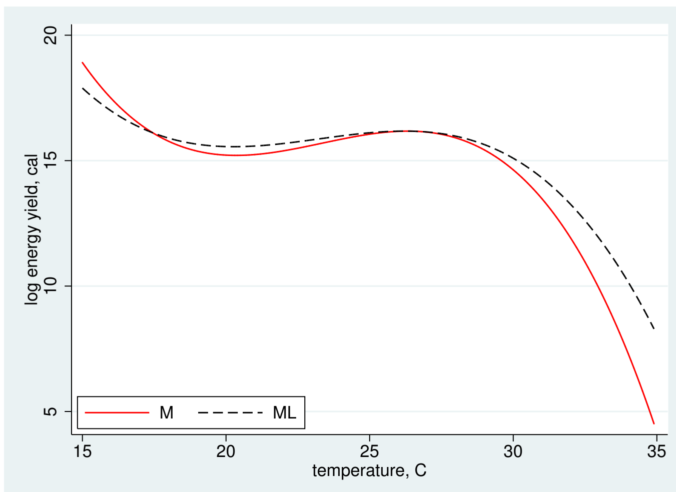


Figure A.3: Estimated effect of temperature on yield: M vs. ML - Household-time FE

*Note:* Based on the estimation in Table 2.1, column 4 (household-time FE) for ML-specific coefficients, column 9 (GMM) for non ML-specific coefficients. The maximal average temperature observed in Malawi over the estimation period is 28.5°C. For higher temperatures the predictions are out of sample.

# APPENDIX B

# APPENDIX TO CHAPTER 3

## B.1 Additional tables and figures

Table B.1: Data sources

| Variable | Source |
|---|---|
| | **Decentralization variables** |
| all variables | (in the order of frequency of use) United Cities and Local Governments (UCLG) (2008), IMF's Government Finance Statistics (GFS) (http://www2.imfstatistics.org/GFS/logon.aspx), Shah (2006), Commonwealth Local Government Initiative (CLGI) country profiles (http://www.clgf.org.uk), Eckardt & Shah (2008), Shah et al. (2004), UN Public Administration Program (UNPAN) public administration country profiles (http://www.unpan.org), Program on Governance in Arab Countries (POGAR) (http://www.pogar.org ), White & Smoke (2005), official web-sites of ministries of local government, ministries of finance |
| | **Variables for heterogeneity index** |
| age 0-14, 15-65, >65, % population | WB World Development Indicators (http://data.worldbank.org) |
| % urban population | WB World Development Indicators (http://data.worldbank.org) |
| GINI index | United Nations Development Project (UNDP) Human Development Indexes (http://data.un.org) |
| ethnic, religious, linguistic fractionalization | Alesina et al. (2003) |
| country area | WB World Development Indicators (http://data.worldbank.org) |
| highest and lowest geographical points | CIA World Factbook (https://www.cia.gov/library/publications/the-world-factbook) |
| highest and lowest latitude | own observations on political map of the world |

Table B.2: Variables used in regressions in Section 4.4.7: Summary statistics

| | Short definition | N obs | Mean | St. dev | Min | Max | Source |
|---|---|---|---|---|---|---|---|
| **Control variables** | | | | | | | |
| Population | | 181 | 35.7 | 130 | 0.1 | 1303 | WB WDI |
| GDP | | 171 | 210 | 955 | 0 | 10950 | WB WDI |
| GDP per capita | | 174 | 9548 | 14645 | 101 | 81777 | WB WDI |
| openness | (Exp+Imp)/GDP | 179 | 96.5 | 53.7 | 2 | 446 | WB WDI |
| literacy rate | % population | 165 | 81.3 | 20.6 | 16.5 | 100 | WB WDI |
| 1 if in OECD | | 182 | 0.15 | 0.35 | 0 | 1 | |
| **Dependent variables** | | | | | | | |
| Human Development Index | | 168 | 0.7 | 0.2 | 0.3 | 1 | UN |
| Corruption Perception Index | | 169 | 4 | 2.1 | 1.4 | 9.4 | TI |
| Real GDP per Capita Growth, 2000-10 | | 176 | 4.4 | 2.7 | -4.3 | 18.3 | WB WDI |
| General government employment (% total) | | 96 | 2.9 | 2.9 | 0.1 | 15 | WB WDI |
| Procedures to start business | | 136 | 10.2 | 3.4 | 2 | 19 | WB DBI |
| Procedures to enforce contracts | | 167 | 38 | 6.8 | 20 | 58 | WB DBI |

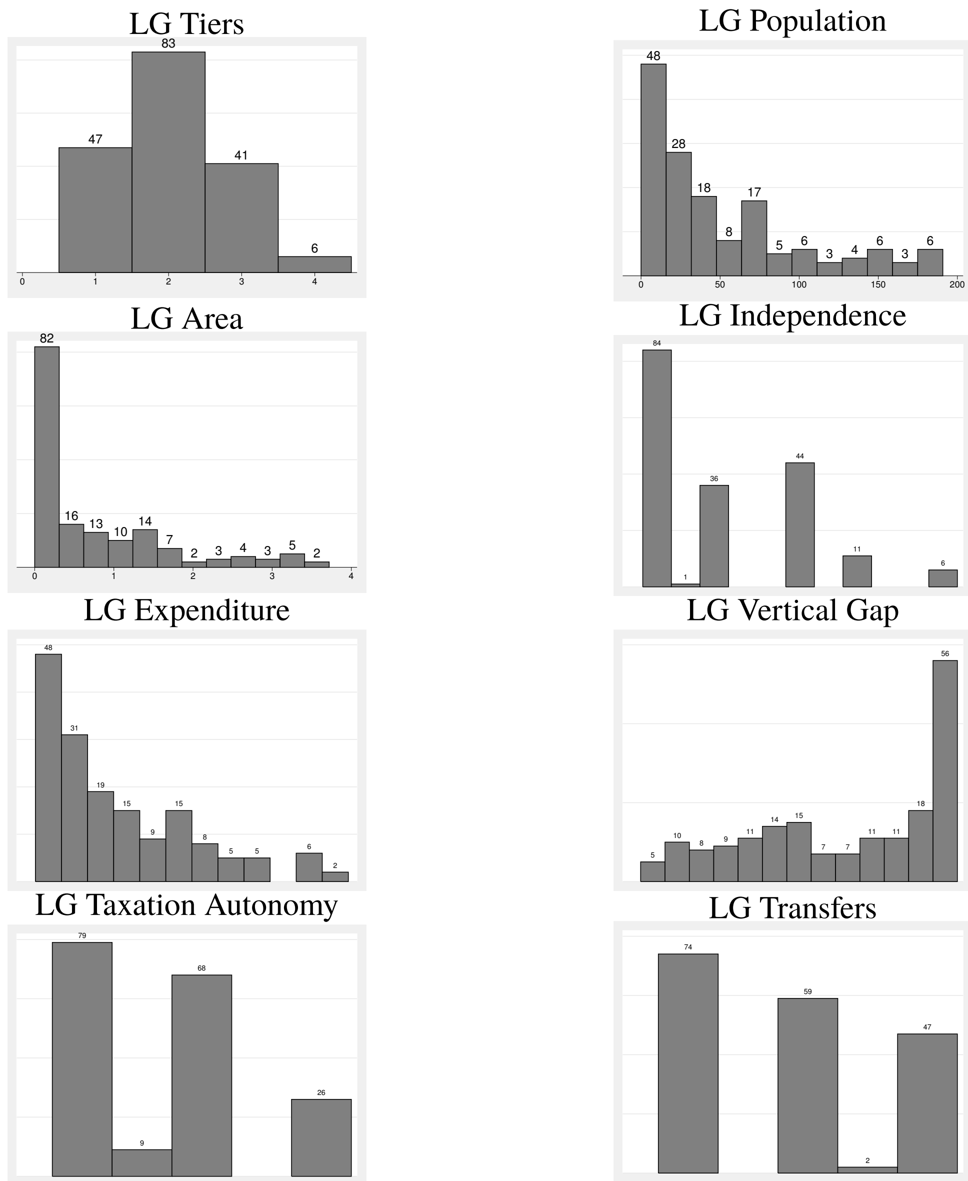


Figure B.1: Decentralization Variables: Frequency Distribution of Countries

*Source:* Authors' calculations based upon data sources reported in Annex Table B.1.
*Note* Distribution of *LG average population* is only for countries with this indicator lower than 200 thousand people (87% of the sample). Distribution of *LG average area* is only for countries with this indicator lower than 4 thousand sq. kilometers (92% of the sample)

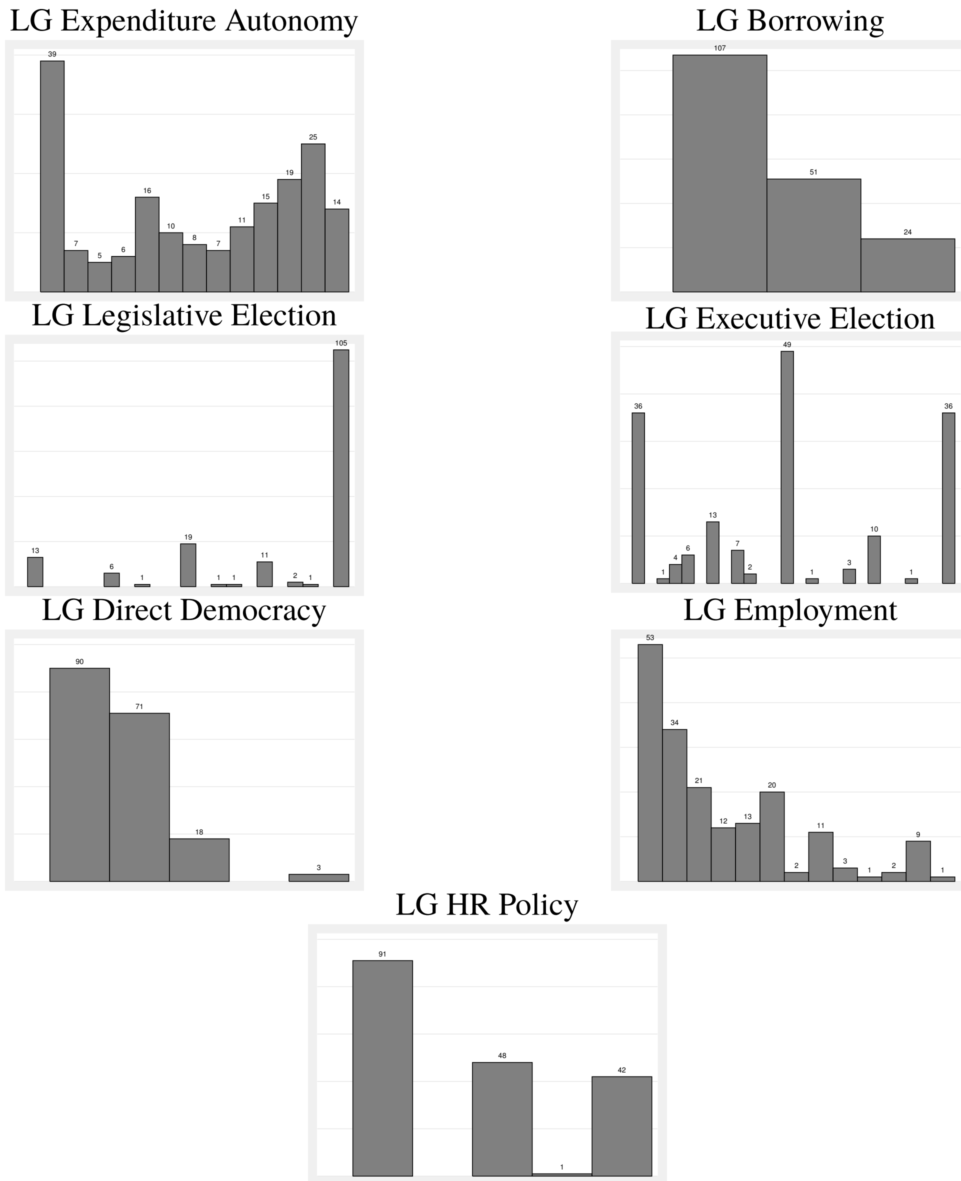


Figure B.2: Decentralization Variables: Frequency Distribution of Countries (continued)

*Source:* Authors' calculations based upon data sources reported in Annex Table B.1.

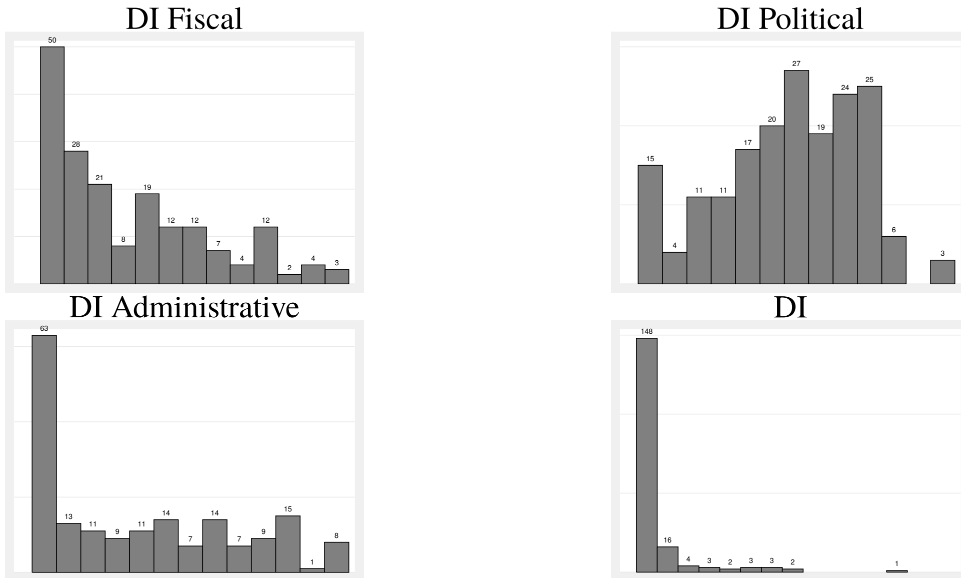


Figure B.3: Decentralization Variables: Frequency Distribution of Countries (continued)

*Source:* Authors' calculations based upon data sources reported in Annex Table B.1.

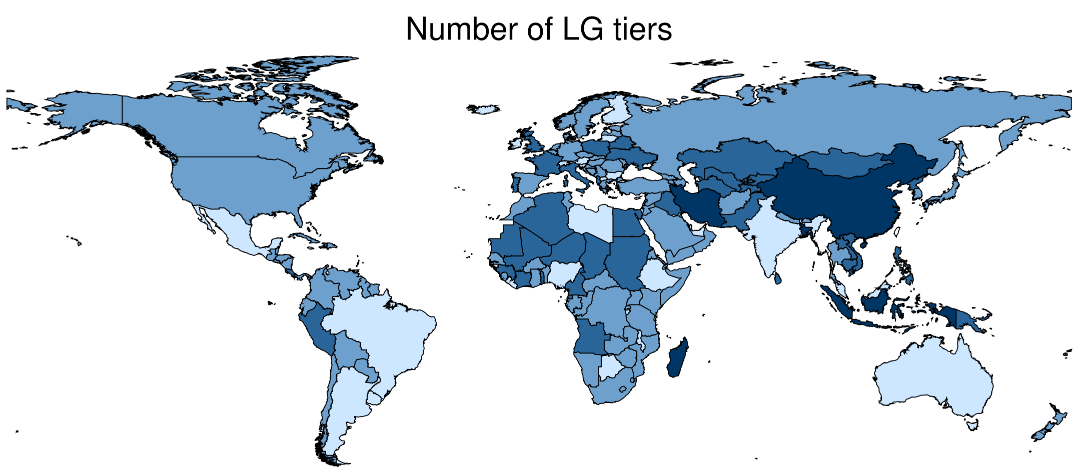


Figure B.4: Number of Tiers of Local Government - World Map

*Source:* Authors' calculations based upon data sources reported in Annex Table B.1.
*Note* Each country is colored according to number of tiers of its local government. The lighter is the shade of color the less tiers are there. Shades of the color correspond to 0-25th, 25-50th, 50-75th, 75-100th percentiles of the corresponding world's distribution.

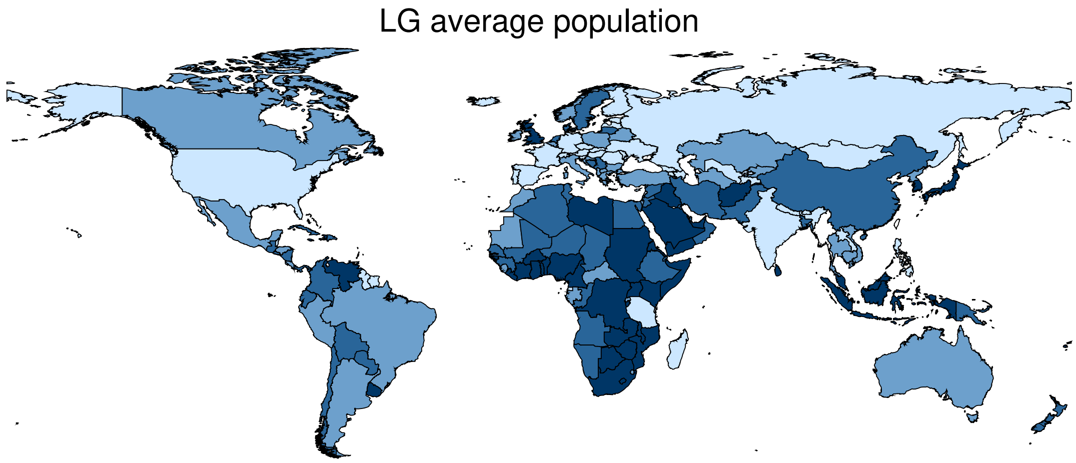


Figure B.5: Population of Local Governments - World Map

*Source:* Authors' calculations based upon data sources reported in Annex Table B.1.
*Note* Each country is colored according to average population of its local government units. The lighter is the shade of color the smaller is the population. Shades of the color correspond to 0-25th, 25-50th, 50-75th, 75-100th percentiles of the corresponding world's distribution.

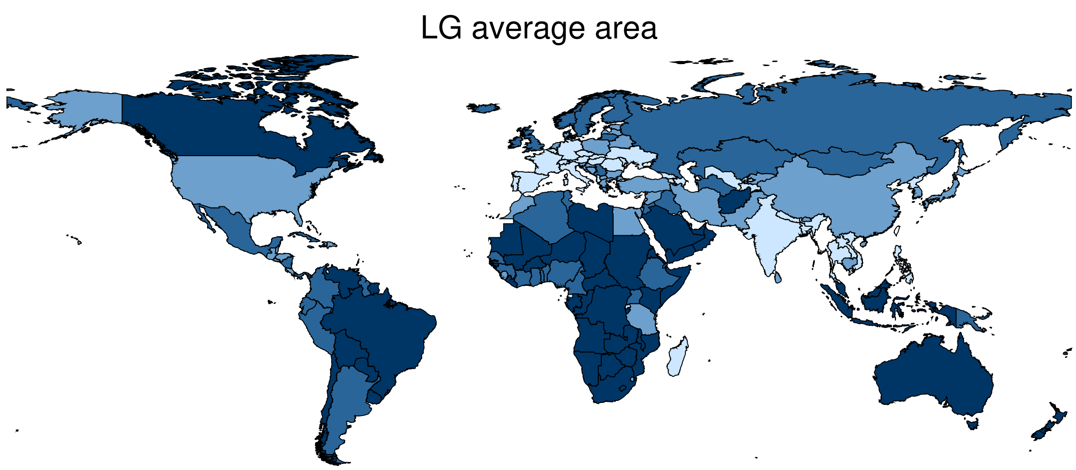


Figure B.6: Area of Local Governments - World Map

*Source:* Authors' calculations based upon data sources reported in Annex Table B.1.
*Note* Each country is colored according to average area of its local government units. The lighter is the shade of color the smaller is the area. Shades of the color correspond to 0-25th, 25-50th, 50-75th, 75-100th percentiles of the corresponding world's distribution.

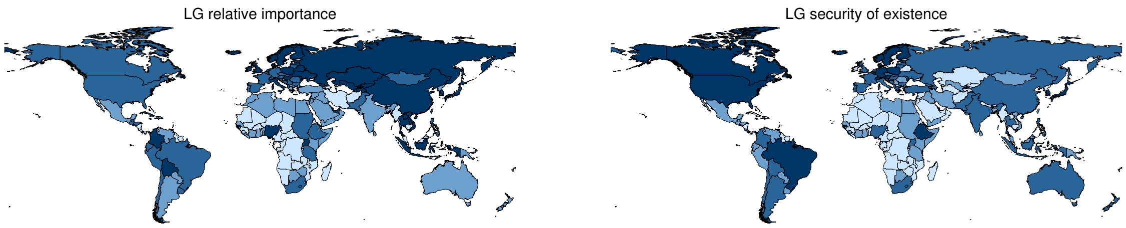


Figure B.7: Relative Importance of Local Governments and Their Independence - World Maps

*Source:* Authors' calculations based upon data sources reported in Annex Table B.1.
*Note* Each country is colored according to value of corresponding decentralization variable (see Table 4.2 for definitions). Shades of the color correspond to 0-25th, 25-50th, 50-75th, 75-100th percentiles of the corresponding world's distribution.

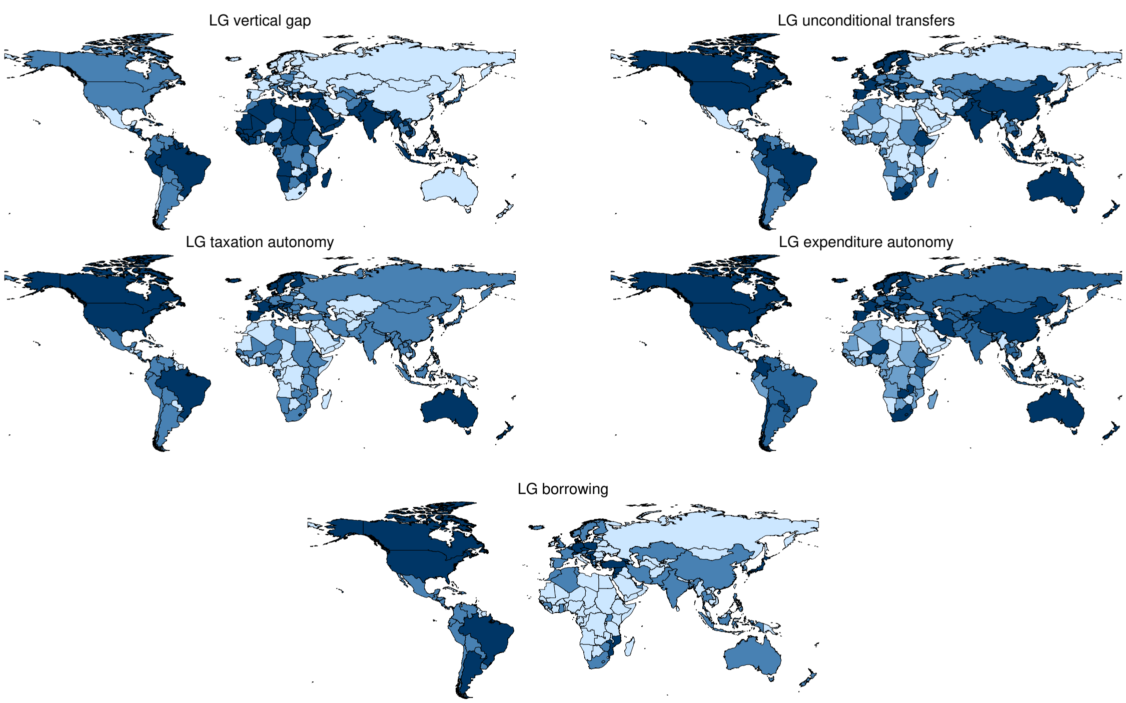


Figure B.8: Fiscal Decentralization Variables - World Maps

*Source:* Authors' calculations based upon data sources reported in Annex Table B.1.
*Note* Each country is colored according to value of corresponding decentralization variable (see Table 4.3 for definitions). The lighter is the shade of color the smaller is the value of a decentralization variable. Shades of the color correspond to 0-25th, 25-50th, 50-75th, 75-100th percentiles of the corresponding world's distribution.

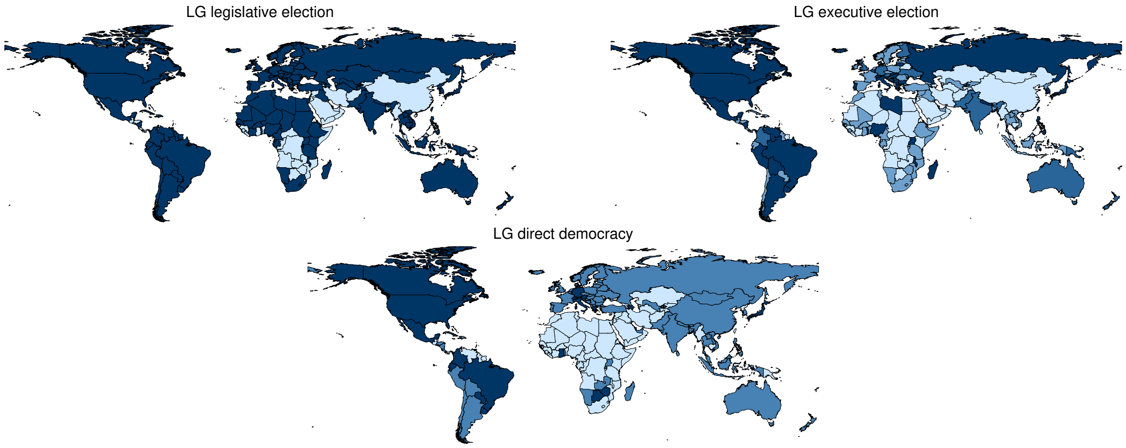


Figure B.9: Political Decentralization Variables - World Maps

*Source:* Authors' calculations based upon data sources reported in Annex Table B.1.
*Note* Each country is colored according to value of corresponding decentralization variable (see Table 4.2 for definitions). The lighter is the shade of color the smaller is the value of a decentralization variable. Shades of the color correspond to 0-25th, 25-50th, 50-75th, 75-100th percentiles of the corresponding world's distribution.

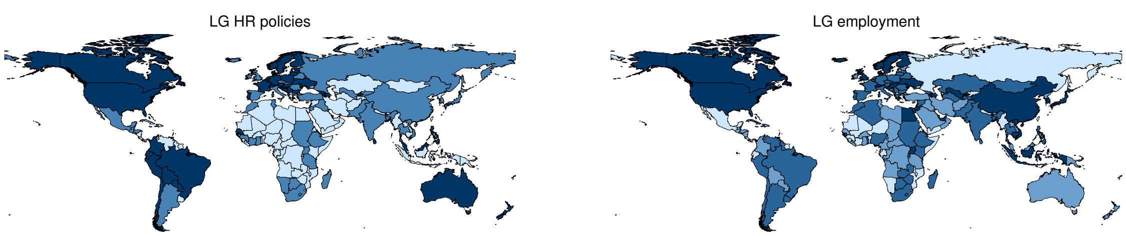


Figure B.10: Administrative Decentralization Variables - World Maps

*Source:* Authors' calculations based upon data sources reported in Annex Table B.1.
*Note* Each country is colored according to value of corresponding decentralization variable (see Table 4.4 for definitions). The lighter is the shade of color the smaller is the value of a decentralization variable. Shades of the color correspond to 0-25th, 25-50th, 50-75th, 75-100th percentiles of the corresponding world's distribution.

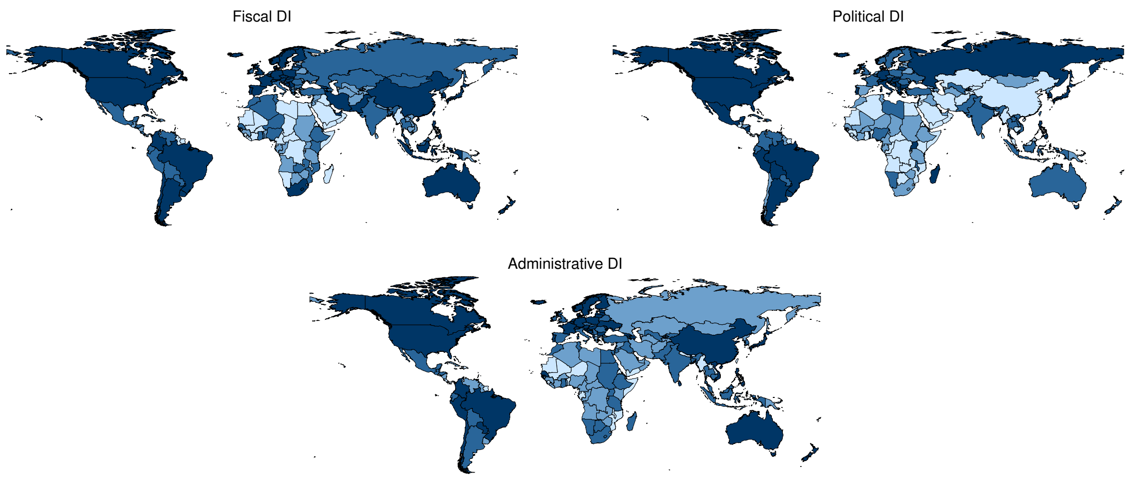


Figure B.11: Fiscal, Political, Administrative Decentralization Indexes - World Maps

*Source:* Authors' calculations based upon data sources reported in Annex Table B.1.
*Note* Each country is colored according to value of corresponding decentralization index. The lighter is the shade of color the smaller is the value of a decentralization index. Shades of the color correspond to 0-25th, 25-50th, 50-75th, 75-100th percentiles of the corresponding world's distribution.

# BIBLIOGRAPHY

## BIBLIOGRAPHY


ActionAid. 2006. *Climate change and smallholder farmers in malawi*. London: ActionAid International.

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat & Romain Wacziarg. 2003. Fractionalization. *Journal of Economic Growth* 8. 155–194.

Arellano, Manuel & Stephen Bond. 1991. Some tests of specification for panel data: Monte carlo evidence and an application to employment equations. *The Review of Economic Studies* 58(2). 277–297.

Boadway, Robin & Anwar Shah. 2009. *Fiscal federalism: Principles and practice of multi-order governance*. New York: Cambridge University Press.

Borlaug, Norman. 2000. Ending world hunger. the promise of biotechnology and the threat of antiscience zealotry. *Plant Physiology* 124(2). 487–490.

Burke, Marshall, John Dykema, David Lobell, Edward Miguel & Shanker Satyanath. 2011. Incorporating climate uncertainty into estimates of climate change impacts, with applications to us and african agriculture. NBER working paper No.17092.

Chirwa, Ephraim & Andrew Dorwards. 2014. The implementation of the 2012/13 farm input subsidy programme. FISP Policy Brief 2014/2.

Deschenes, Olivier & Michael Greenstone. 2007. The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather. *American Economic Review* 97(1). 354–385.

Deschenes, Olivier & Michael Greenstone. 2011. Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the us. *American Economic Journal: Applied Economics* 3(4). 152–185.

Dollery, B. & J. Wallis. 2001. *The political economy of local government*. Cheltenham, U.K.: Edward Elgar.

Dorwards, Andrew & Ephraim Chirwa. 2011. The malawi agricultural input subsidy programme: 2005-6 to 2008-9. *International Journal of Agricultural Sustainability* 9(1). 232–247.

Dorwards, Andrew, Ephraim Chirwa & Tom Jayne. 2010. The malawi agricultural inputs subsidy programme, 2005/06 to 2008/09. Report for Malawian government.

Duflo, Esther, Rema Hanna & Michael Greenstone. 2016. Up in smoke: The influence of household behavior on the long-run impact of improved cooking stoves. *American Economic Journal: Economic Policy* 8(1). 80–114.

Duflo, Esther, Michael Kremer & Jonathan Robinson. 2011. Nudging farmers to use fertilizer: Theory and experimental evidence from kenya. *American Economic Review* 101(6). 2350–2390.

Eckardt, S. & A. Shah. 2008. Decentralized governance in developing and transition countries: A comparative review. In A. Shah (ed.), *Macro federalism and local finance*, 291–322. Washington, DC: World Bank.

FAO. 2008. Fao methodology for the measurement of food deprivation: Updating the minimum dietary energy requirements. Food and Agriculure Organization Statistics Division report.

Fisher, Anthony, Michael Hanemann, Michael Roberts & Wolfram Schlenker. forthcoming. The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather: Comment. *American Economic Review* .

Foster, Andrew & Mark Rosenzweig. 2010. Microeconomics of technology adoption. Yale University Economic Growth Center discussion paper No.984.

Frey, Bruno & R. Eichenberger. 1999. Competition among jurisdictions: The idea of focj. In L. Gerken (ed.), *Competition among jurisdictions*, London: Macmillan.

Gilbert, Robert, Maction Komwa, Todd Benson & Webster Sakala. 2002. A comparison of best-bet soil fertility technologies for maize grown by malawian smallholders. Report by Maize Productivity Task Force.

Griliches, Zvi & Jerry Hausman. 1986. Errors in variables in panel data. *Journal of Econometrics* 31. 93–118.

Huntington, Samuel. 1968. *Political order in changing societies*. New Haven: Yale University Press.

International Monetary Fund. 2012. Macroeconomic policy frameworks for resource-rich developing countries. IMF Policy Paper.

IPCC. 2000. *Special report: Emissions scenarios*. Geneva: International Panel on Climate Change.

IPCC. 2007. *Climate change 2007: Synthesis report*. Geneva: International Panel on Climate Change.

Ivanyna, Maksym, Alex Mourmouras & Peter Rangazas. 2016a. Corruption, public debt and economic growth. Unpublished paper.

Ivanyna, Maksym, Alexandros Mourmouras & Peter Rangazas. 2016b. The culture of corruption, tax evasion, and optimal tax policy. *Economic Inquiry* 54(1). 520–542.

Kerr, Rachel, Peter Berti & Lizzie Shumba. 2011. Effects of a participatory agriculture and nutrition education project on child growth in northern malawi. *Public Health Nutrition* 14(8). 1466–1472.

Kerr, Rachel, Sieglinde Snapp, Marco Chirwa, Lizzie Shumba & Rodgers Msachi. 2007. Participatory research on legume diversification with malawian smallholder farmers for improved human nutrition and soil fertility. *Experimental Agriculture* 43. 437–453.

Liu, Y. & R. Myers. 2009. Model selection in stochastic frontier analysis with an application to maize production in kenya. *Journal of Productivity Analysis* 31(1). 33–46.

Lobell, David, Marianne B anziger, Cosmos Magorokosho & Bindiganavile Vivek. 2011. Nonlinear heat effects on african maize as evidenced by historical yield trials. *Nature Climate Change* 1(6). 42–45.

Mauro, Paolo. 1995. Corruption and growth. *Quarterly Journal of Economics* 110(3). 681–712.

MNVAC. 2005. Malawi baseline livelihood profiles. Prepared by Malawi National Vulnerability Assessment Committee and SADC FANR Vulnerability Assessment Committee, published at http://www.usaid.gov.

Moore, M. 1996. *Creating public value*. Cambridge, MA: Harvard University Press.

Morris, Michael, Valerie Kelly, Ron Kopicki & Derek Byerlee. 2007. *Fertilizer use in african agriculture: Lessons learned and good practice guidelines*. Washington D.C.: World Bank.

Oates, W. 1972. *Fiscal federalism*. New York: Harcourt-Brace-Jovanovich.

Olken, Benjamin & Rohini Pande. 2012. Corruption in developing countries. *Annual Review of Economics* 4. 479–509.

One Acre Fund. 2015. *Maize-legume intercropping. 2014 phase 1 and 2 trial report*. One Acre Fund.

Porta, Rafael La & Andrei Shleifer. 2008. The unofficial economy and economic development. *Brookings Papers on Economic Activity* 39(2). 275–363.

Prud'homme, Remy. 1994. On the dangers of decentralization. World Bank Policy Research Working Paper 1252.

Rickert-Gilbert, J., Tom Jayne & Ephraim Chirwa. 2010. Subsidies and crowding out: a double hurdle model of fertilizer demand in malawi. *American Journal of Agricultural Economics* 93(1). 26–42.

Searle, P., Yuthapong Comudom, D. Shedden & R. Nance. 1981. Effect of maize + legume intercropping systems and fertilizer nitrogen on crop yielding and residual nitrogen. *Field Crops Research* 4. 133–145.

Shah, A. & S. Shah. 2006. The new vision of local governance and the evolving role of local governments. In A. Shah (ed.), *Local governance in developing countries*, 1–46. Washington, DC: World Bank.

Shah, Anwar. 2006. *Local governance in developing countries*. Washington, DC: World Bank.

Shah, Anwar, Theresa Thompson & Hengfu Zou. 2004. The impact of decentralization on service delivery, corruption, fiscal management and growth in developing and emerging market economies: A synthesis of empirical evidence. *Journal for Institutional Comparisons* 2(Spring). 10–14.

Sheahan, Megan, Roy Black & Thomas Jayne. 2013. Are kenyan farmers under-utilizing fertilizer? implications for input intensification strategies and research. *Food Policy* 41. 39–52.

Sileshi, Gudeta, Festus Akinnifesi, Legesse Debusho, Tracy Beedy, Oluyede Ajayi & Simon Mong'omba. 2010. Variation in maize yield gaps with plant nutrient inputs, soil type and climate across sub-saharan africa. *Fiel Crops Research* 116. 1–13.

Snapp, Sieglinde, Malcolm Blackie & C. Donovan. 2003. Realigning research and extension to focus on farmersâĂŹ constraints and opportunities. *Food Policy* 28. 349–363.

Snapp, Sieglinde, Malcolm Blackie, Robert Gilbert, Rachel Bezner-Kerr & George Kanyama-Phiri. 2010. Biodiversity can support greener revolution in africa. *Proceedings of the National Academy of Sciences of the United States of America* 107(48). 20840–20845.

Stigler, G. 1957. The tenable range of functions of local government. In U.S. Congress Joint Economic Committee, Subcommittee on Fiscal Policy (ed.), *Federal expenditure policy for economic growth and stability*, Washington, DC: U.S. Government Printing Office.

Suri, Tavneet. 2011. Selection and comparative advantage in technology adoption. *Econometrica* 79(1). 159–209.

Svensson, Jacob. 2005. Eight questions about corruption. *Journal of Economic Perspectives* 19(5). 19–42.

Tanzi, Vito. 1996. Fiscal federalism aand decentralization: A review of some efficiency and macroeconomic aspects. In Michael Bruno & Boris Pleskovic (eds.), *Annual world bank conference on development economics*, 295–316. Washington, DC: World Bank.

Tilman, David, Christian Balzer, Jason Hill & Belinda Befort. 2011. Global food demand and the sustainable intensification of agriculture. *Proceedings of the National Academy of Sciences* 108(50). 20260–20264.

Treisman, D. 2007. What have we learned about the causes of corruption from ten years of cross-national empirical research. *Annual Review of Political Science* 10. 211–244.

Ugur, Mehmet. 2014. Corruption's direct effects on per-capita income growth: A meta-analysis. *Journal of Economic Surveys* 28(3). 472–490.

UN. 2014. *The millenium development goals report*. United Nations, New-York.

United Cities and Local Governments (UCLG). 2008. *Decentralization and local democracy in the world*. USA.

Wetzel, William, Heather Kharouba, Moria Robinson, Marcel Holyoak & Richard Karban. 2016. Variability in plant nutrients reduces insect herbivore performance. *Nature* 539. 425–427.

White, Roland & Paul Smoke. 2005. East asia decentralizes. In Rolan White & Paul Smoke (eds.), *East asia decentralizes: making local governments work*, 1–25. Washington, DC: World Bank.

Xie, P. & P. Arkin. 1997. Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bulletin of American Meteorological Society* 78. 2539 – 2558.