Essays on the Political-economy of Large-scale Land Deals in Developing Countries

Anthony Harris

Nuffield College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy in Economics

Michaelmas Term 2014

Abstract

The thesis consists of a short introduction and three self-contained analytical chapters on land policy in developing countries.

Chapter 1 examines the agricultural investment choices of small-scale farmers in Ethiopia whose land will be expropriated to provide space for a large factory. I use data from a survey of households conducted before expropriation occurred, but after the policy was announced. I identify the anticipation effects of land expropriation using variation in whether households own plots located inside or outside the proposed project boundary. Households facing immediate expropriation hedge against future income risk by using less fertilizer on their plots, and and growing less risky crops. These households are more likely to grow sorghum (a safe crop) and less likely to grow wheat (a relatively riskier crop). Households also respond to the threat of expropriation by reducing long-term investments in soil quality. Using two-stage least squares I show that subjective beliefs about the likelihood of expropriation act as a channel through which the threat of expropriation affects investment decisions. The results are robust to a number of other specifications, including some that account for unobservable geographic variation in plot characteristics.

Chapter 2 explores the consequences of land expropriation for small-scale farmers in Ethiopia. Expropriation of farmland is used by all levels of government in Ethiopia as a tool for providing new land for industrial investors, commercial agriculture and expanding cities. Farmers usually receive a cash payment in exchange for their land based on a fixed formula to establish the price of land. I evaluate the impact of such a policy on a group of small-scale farmers and assess the extent to which they make the transition to new livelihoods. On average, households lose 70% of their land and receive compensation payments that are about 5 times the value of their annual consumption expenditure. Using data collected before and after the intervention I examine the impact of expropriation and compensation on household consumption, productive assets, livestock holdings, savings and labour market participation. Households in the treatment group increase their consumption, start more businesses and participate more in non-farm activities than households that do not lose farmland. These households also reallocate livestock portfolios away from oxen and towards small ruminants and cattle, reflecting a shift away from growing crops. However, these shifts to new livelihoods are relatively small compared to the amount of compensation kept as savings: with the exception of a few households, most of the compensation payment is left in commercial banks earning a negative real return.

Chapter 3 focuses on the recent increase in large-scale agricultural land deals across Africa and the nature of the contracts reached by governments and foreign investors. In recent years, multi-national firms and foreign governments have entered into long term contracts with host countries in which large tracts of land are purchased or leased for commercial agricultural production in exchange for promises of infrastructure development, job creation and rural infrastructure improvement. The profitability of these projects is uncertain, especially at a time of increased agricultural commodity price volatility in world markets. Based on stylized facts about land deals I present a theoretical model of land contracts reached by host governments and foreign investors that explains the policy tradeoff between investment timelines, revenue generation and uncertainty. When agricultural projects require fixed infrastructure investment and yield uncertain payoffs, firms benefit from being able to complete the fixed investment in stages. If firms can learn more about payoffs by holding off on investment, they effectively hold an option to abandon the project. The value of this option provides a channel by which uncertainty affects the terms of the land contract. When host governments determine the terms of the contract by setting an income tax, a royalty rate and an investment timeline, the value of this option will affect government's optimal policy choice. In particular, I find that if governments benefit a great deal from investment spillovers the optimal contract will be designed to encourage firms not to abandon a project. But, if governments benefit relatively little from investment spillovers, governments will choose contract parameters to extract the value of the firm's option to abandon the project. I end by examining the effect of increasing uncertainty on the government's optimal policy choice.

Notes for the reader

This thesis contains a brief introduction and three chapters that are each intended to stand on their own. The numbering of sections, figures, tables and equations is specific to each chapter and notation varies across chapters. A list of references and chapter appendices are placed at the end of each chapter.

All three chapters are solely my own work. An earlier version of chapter 3 was submitted as the thesis requirement for the MPhil degree in Economics at the University of Oxford.

Word count: This thesis has approximately 66,192 words excluding references. This is calculated using the number of words on page 2 (336) multiplied by the number of pages (197).

Acknowledgments

My first thanks is to my family. Their love and unwavering support, even from a distance, have sustained my efforts through the last years. The love of learning that they instilled in me is with me today and was instrumental in driving this thesis forward.

I am very grateful to my supervisors, Professors Tony Venables and Stefan Dercon who have encouraged and supported me throughout my time at Oxford. Despite busy schedules and professional obligations outside of Oxford, they found time to engage with my work, guide me as a researcher and provide advice at critical times. Over the years I have learnt much about how to do research from them. This thesis (at least the good parts) is a product of that.

The empirical chapters in this thesis use data from a survey of households in Ethiopia who had their farms taken to provide land for a new factory. While running the survey I had the chance to meet many of the people who lost their land and I have witnessed, often with great sadness, the struggle that many of them now face to find a new future for themselves and their families. It is not lost on me that their hardship and suffering provides the basis for my thesis. I am thus very grateful to the respondents for the time they gave and I only hope that, in some modest way, telling their story can bring about some good.

I want to thank Zewdu, Yeshi, Asmelash and Amanuel and all of the enumerators in Ethiopia who worked tirelessly and with compassion to administer the survey. Their advice, hard work and dedication are what made this project work and their insights have been instrumental in improving the research. I am also very grateful to Rose, Gail, Bet and Lena who provided support and survey supervision from Oxford through various rounds of the survey.

Many other academics and colleagues have offered their time to comment on drafts, provide feedback in seminars and discuss research ideas with me. I am especially grateful to Professor Steve Bond, my College Advisor who has discussed and read many drafts of this thesis and Dr. James Fenske, who met with me on numerous occasions to discuss my work. Simon Quinn, Meg Meyer, Ian Jewitt, and many others also provided valuable comments that have greatly improved the work. CSAE, Nuffield College and the Economics Department provided an invaluable network of colleagues and I am very grateful both for the time they spent discussing work and for their friendship. I owe a very special debt of gratitude to my friend Abhijeet Singh. He provided moral support, intellectual guidance and advice through the last years of my DPhil; he read countless drafts, attended most of my seminars and was always there to pick me up when my confidence faltered. I would not have finished without him.

Finally I want to thank my wife, Samantha, whose patience, support and love has carried me through the process of writing.

Contents

List of Tables

List of Figures

Introduction

This thesis is about land acquisition for industrial and agricultural investment in developing countries, with a particular focus on Africa. Rapid economic development in many African countries has put a premium on rural and urban land, yet formal land markets are often poorly developed and unable to reallocate land to new uses.[1] Governments are therefore frequently involved in land transactions, sometimes directly, by selling government land and sometimes through facilitation, compulsory purchase or expropriation.

This thesis is divided into three substantive, self-contained chapters that focus on different aspects of government involvement in land markets in Africa. Chapters 1 and 2 examine the impact of land expropriation on small-scale farmers who lose their land for new investments, using a novel data-set from Ethiopia collected for this purpose. Chapter 3 shifts the focus to understanding the changing demand for land at a macro level, especially as it relates to large-scale land deals for commercial agriculture. It presents a theoretical model that explains the contract structure of large-scale land deals and seeks understand their recent resurgence.

Expropriation and compulsory purchase are important economic policy tools in many developing countries, enabling governments to consolidate agricultural land from fragmented small-holdings and provide space for new investments and urban expansion. China, for example, has relied heavily on expropriation to meet the demand for land for industry, infrastructure and urban expansion

[1] Work by Adamopoulos and Restuccia (2014) highlights the effect of inefficient land market allocation on cross-country productivity differences.

(Ding, 2007). But, these policies are not without controversy. Across Africa, Asia, and Latin America, land for large-scale commercial farms is often acquired through expropriation (Cotula et al., 2009) and this has been a cause for concern amongst many NGOs and advocacy groups[2]. In West Bengal, the compulsory acquisition of farmland for a Tata factory lead to political unrest and violence, which ultimately undermined the success of the investment project and negatively affected the original land-holders (Banerjee et al., 2007; Ghatak et al., 2012). However, without major changes to land markets, expropriation will continue to play an important part in the process of industrialization and development. Chapters 1 and 2 focus on the impact of land expropriation on small-scale farmers who lose their land for new investments, using data collected in Ethiopia. The data was collected during the course of writing the thesis and is based on two rounds of household surveys conducted with small-scale farmers living in and around the future site of a factory. Land for the factory was acquired through expropriation, with financial compensation paid to anyone who lost land. Households were interviewed before and after the land was taken, allowing for a rigorous identification of the impact of expropriation on small-scale farmers.

Chapter 1 uses the baseline data to explore the effect of anticipating expropriation on the agricultural investment decisions made by small-scale farmers. The survey sample includes both households that lost land for the factory and their neighbours from the same administrative area that did not. I use this variation as well as variation in households' expectations of expropriation to identify the effect of anticipation on agricultural investment. Institutional features of Ethiopia, particularly the absence of rural land sale markets, and the timing of the data collection relative to the announcement of expropriation allows for the anticipation effects of expropriation to be isolated and empirically identified.

[2]Oxfam recently launched a campaign calling for a moratorium on land grabs in part because of the consequences for households displaced by investors (Geary, 2012).

Throughout the chapter I examine two candidate mechanisms by which a household's expectation about expropriation can distort agricultural investment choices. The first mechanism focuses on the link between the income risk imposed by expropriation and the household's crop portfolio and decision to use fertilizer. Previous work shows that rural households respond to income risk by choosing safer crops and reducing the use of expensive inputs.[3] In this context, households facing expropriation will receive financial compensation, but face uncertainty in the immediate term over the size and timing of payments and in the long term over whether they will successfully replace their lost farming income. Households facing expropriation are more likely to grow sorghum, which is a relatively safe crop and less likely to grow wheat, which is higher risk, input intensive crop. The same households are also less likely to use any fertilizer on their plots. The second mechanism focuses on investment incentives, expropriation and the nature of agricultural investments. Agricultural investments that yield a return over the long-term and that are irreversible and unverifiable are often under provided by households that expect their land to be taken in the near future.[4] I find that households facing expropriation are less likely to apply organic fertilizer on their plots, an investment that has long term impacts on soil fertility. These distortions may not be problematic if the government is clear about its plan for expropriation, but the effects described in the paper are a function of expectations, and may be damaging to households if plans for expropriations are delayed or cancelled.

Chapter 2 looks at the consequences of expropriation and the transition to new livelihoods made by households that lose their land. Expropriating farmland

[3]For example, Dercon (1996) finds that households in Tanzania grow safer crops in response to rainfall risk. Fafchamps (1992) finds a similar move towards safer crops when households are faced with price risk. Dercon and Christiaensen (2011) find that households facing greater consumption risk avoid expensive inputs such as fertilizer because of the added downside risk if fertilized crops fail.

[4]There is a long literature that looks at the relationship between tenure security and agricultural investment. See for example Ali et al. (2011); Besley (1995); Deininger and Jin (2006); Fenske (2011); Jacoby et al. (2002).

deprives rural small-holders of one their most important income-generating assets and forces them to find new livelihoods. Governments recognize this, and often provide households with financial compensation. But are these lump-sum payments sufficient to compensate households for the land that is taken? A natural way to assess this is to determine whether a household's permanent wealth has changed as a result of the intervention. Evaluating this requires some understanding of how households adjust their asset and activity portfolio after losing land. I evaluate the extent to which households use financial compensation payments for productive investments, the change in their income-generating activities and the change in levels of consumption following expropriation. I focus specifically on evaluating the effect of the program on a number of key household outcomes, including consumption, asset ownership, livestock holdings, savings and labour market participation.

On average, households in the Ethiopia data lose 70% of their farmland and received compensation payments worth 5,200 USD (the equivalent of 4.7 times total annual consumption expenditure and 9 times the value of livestock). The baseline survey was conducted one year prior to expropriation and the followup survey was conducted eight months afterwards. A group of households that did not lose land were selected from within the same administrative area to serve as a comparison group. The timing of the data collection and the inclusion of a comparison group permit me to evaluate the way in which the program affects households' asset and activity portfolios, their consumption and their ability to cope with the challenge of finding new income generating activities. I find that, on average, the households that lose their land increase consumption, start more businesses and increase their livestock assets. However, with the exception of a few households that make large lumpy investments, the increase in investment in productive assets is dwarfed by massive increases in cash savings. Although households shift assets away from agriculture and towards non-farm work, the change is small, suggesting that investment opportunities are limited.

Chapter 3 seeks to understand the changing demand for land at a more macro level. One important source of demand for land in developing countries, which has received considerable media attention in recent years, is the acquisition of land for large-scale commercial agriculture.[5] International investors and foreign governments have entered into long term contracts with host countries, often in Africa, in which many thousands of hectares of land are sold or leased for large-scale agricultural investment projects. According to a World Bank report published in 2011, 56.6 million ha of land has been included in large-scale land deals, of which 39.6 million ha is in Africa. Median farm size for the proposed projects is 40,000 ha, although only 20% of the reported deals have progressed to implementation (Deininger and Byerlee, 2012). Compensation paid to host countries for the opportunity to use land often comes as promises for rural infrastructure development and job creation, rather than cash transfers, large lease or purchase fees or high levels of taxation (Cotula et al., 2009; Deininger and Byerlee, 2011). What factors drive these deals and how do the determinants affect the way that contracts are structured?

In chapter 3 I present a theoretical model of land deals based on stylized facts drawn from a number of policy reports. I model the strategic interaction between governments and investors negotiating a land deal and draw some conclusions about the relationship between uncertain profits and the key contract parameters observed in many land deals. In particular, I model the trade-off between policies that generate government revenue and those that ensure that firms keep to their investment commitments. I consider how a government might

[5] See for example: "How food and water are driving a 21st-century African land grab," *The Observer*. March 7 2010.; "The scramble for land in Africa," *The Economist*. May 21 2009; "Foreign fields," *Financial Times*. August 20 2008.

choose a timeline for investment and a profit tax rate when the investment is both crucial for the operation of large-scale agricultural projects and of inherent benefit to the host government. Such investments might include, for instance, new transportation, storage or irrigation infrastructure that benefit neighbouring farmers. Large agricultural projects in Africa yield uncertain profit flows both because of increased price volatility in world food markets as well as uncertainty over the profitability of this model of agriculture in many African countries. Uncertain profits affect the prospective value of an agricultural investment project, especially when a large part of the investment has a fixed cost component that can be completed in stages and when a firm is able to prematurely abandon the project. Explicitly modeling large-scale agricultural projects as having fixed costs provides an essential link between price volatility and both the level of agricultural investment and the terms of the land deal.

A key contribution of this chapter is to highlight the central role of the tradeoff between generating revenue and securing investment in determining how land contracts are structured. Greater variability in profits can affect the overall value of an investment project in this framework, and, as such, influence policy outcomes. Increased profits can increase potential government revenue but may do so at a cost to the level of upfront investment. In this way, price volatility or uncertainty about costs affect the decisions of governments through its effect on the value of an investment project where the timing of investments is flexible. The tradeoff between tax rates and the degree of flexibility offered to investors may explain why taxes and transfer fees seem low in large-scale land deals. Many contracts make land use rights contingent on investors completing an investment plan in a timely manner. Mandating investment plans limits the flexibility of large investment decisions and thus reduces the overall expected value of a project to investors. However, making firms commit a large part of the infrastructure investment right away prevents them from abandoning the project prematurely and assures governments that investment will occur, but at the expense of future tax revenue.

References

Adamopoulos, T. and Restuccia, D. (2014). The size distribution of farms and international productivity differences. *American Economic Review*, 104(6):1667–97.

Ali, D. A., Dercon, S., and Gautam, M. (2011). Property rights in a very poor country: Tenure insecurity and investment in ethiopia. *Agricultural Economics*, 42(1):75–86.

Banerjee, A. V., Bardhan, P., Basu, K., Chaudhury, M. D., Ghatak, M., Guha, A. S., Majumdar, M., Mookherjee, D., and Ray, D. (2007). Beyond nandigram: Industrialisation in west bengal. *Economic and Political Weekly*, 42(17):1487–1489.

Besley, T. (1995). Property rights and investment incentives: Theory and evidence from ghana. *journal of Political Economy*, pages 903–937.

Cotula, L., Vermeulen, S., Leonard, R., and Keeley, J. (2009). *Land grab or development opportunity? Agricultural investment and international land deals in Africa*. IIED, FAO and IFAD.

Deininger, K. and Byerlee, D. (2012). The rise of large farms in land abundant countries: Do they have a future? *World Development*, 40(4):701–714.

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

Dercon, S. (1996). Risk, crop choice, and savings: Evidence from tanzania. *Economic Development and Cultural Change*, 44(3):485–513.

Dercon, S. and Christiaensen, L. (2011). Consumption risk, technology adoption and poverty traps: Evidence from ethiopia. *Journal of Development Economics*, 96(2):159–173.

Fafchamps, M. (1992). Cash crop production, food price volatility, and rural market integration in the third world. *American Journal of Agricultural Economics*, 74(1):90–99.

Fenske, J. (2011). Land tenure and investment incentives: Evidence from west africa. *Journal of Development Economics*, 95(2):137 – 156.

Geary, K. (2012). *Our Land, Our Lives': Time out in the global land rush.* Oxfam.

Ghatak, M., Mitra, S., Mookherjee, D., and Nath, A. (2012). Land acquisition and compensation in singur: What really happened? *Working PaperInstitute for Economic Development, Boston University.*

Jacoby, H. G., Li, G., and Rozelle, S. (2002). Hazards of expropriation: Tenure insecurity and investment in rural china. *The American Economic Review*, 92(5):1420–1447.

Chapter 1

Land expropriation and agricultural investment in Ethiopia

Abstract

This paper examines the agricultural investment choices of small-scale farmers in Ethiopia whose land will be expropriated to provide space for a large factory. The data comes from a baseline survey of households conducted before expropriation occurred, but after the policy was announced. I use variation in whether households own plots located inside or outside the proposed project boundary to identify the anticipation effects of land expropriation. Households facing immediate expropriation are less likely to use fertilizer on their plots, and are more likely to grow less risky crops. These households are more likely to grow sorghum (a safe crop) and less likely to grow wheat (a relatively riskier crop). Using two-stage least squares I show that subjective beliefs about the likelihood of expropriation act as a channel through which the threat of expropriation affects investment decisions. The results are robust to a number of other specifications, including some that account for unobservable geographic variation in plot characteristics.

1.1 Introduction

Expropriating land from small-scale farmers is a way for governments to provide land for new investment projects in developing countries, especially in contexts where land markets function poorly. Direct government intervention in land markets is common in China (Ding, 2007) and has sparked controversy in places in India (Banerjee et al., 2007). In Africa, expropriation is becoming more common as pressure on farmland grows with increased industrialization, urbanization, and a renewed interest in commercial agriculture (Cotula et al., 2009). There are many reasons to criticize such liberal use of expropriation, but government intervention in land markets can be a valuable policy tool, particularly when land holdings are small and fractured. Governments can help overcome coordination problems amongst sellers and may mitigate against individual sellers holding out after land has been transferred to an investor. However, even if expropriation is accompanied by financial compensation for households, the anticipation of losing land may distort household behaviour before expropriation has actually occurred. Distortions are problematic if households change their behaviour in anticipation of an expropriation program which ultimately does not occur or if households incorrectly think that they will lose land, as might happen to households living near areas designated for urban expansion or industrialization.

This paper explores the agricultural investment decisions of small-scale farmers facing an imminent risk of expropriation using data collected from the site of an expropriation in Ethiopia. Ethiopia is an interesting context in which to study this question: much of the Ethiopian population still lives in rural areas, which are densely populated and land rights held by small-scale farmers do not extend to the right to sell land so there is no private land market that investors can use. Instead, if investors want land, the government intervenes on their behalf by transferring land from farmers and ensuring that the previous occupants are paid compensation. Urban and industrial expansion in Ethiopia has been rapid in recent years, and governments at all levels have relied on expropriation to provide new land for these purposes. The data in this paper is from a household survey conducted following the announcement of an expropriation, but before any land was taken. The survey was conducted in one administrative area outside of the city of Kombolcha, Ethiopia where 340ha of land was eventually expropriated in order to build a factory. The sample includes both households that lost land for the factory and their neighbours from the same administrative area that did not. I use this variation as well as variation in households' expectations of expropriation to identify the effect of anticipation on agricultural investment. Institutional features of Ethiopia, particularly the absence of rural land sale markets, and the timing of the data collection relative to the announcement of expropriation allows for the anticipation effects of expropriation to be isolated and empirically identified.

Throughout the paper I examine two candidate mechanisms by which a household's expectation about expropriation can distort agricultural investment choices. The first mechanism focuses on the link between the income risk imposed by expropriation and the household's crop portfolio and fertilizer use. Previous work shows that rural households respond to income risk by choosing safer crops and reducing the use of expensive inputs.[1] In this context, households facing expropriation will receive financial compensation, but face uncertainty in the immediate term over the size and timing of payments and in the long term over whether they will successfully replace their lost farming income. Households facing expropriation are more likely to grow sorghum, which is a relatively safe crop and

[1]For example, Dercon (1996) finds that households in Tanzania grow safer crops in response to rainfall risk. Fafchamps (1992) finds a similar move towards safer crops when households are faced with price risk. Dercon and Christiaensen (2011) find that households facing greater consumption risk avoid expensive inputs such as fertilizer because of the added downside risk if fertilized crops fail. There are, of course, many other examples of ex ante risk mitigation strategies, but these are most relevant in my context.

less likely to grow wheat, which is higher risk, input intensive crop. The same households are also less likely to use any fertilizer on their plots. The second mechanism focuses on investment incentives, expropriation and the nature of agricultural investments. Agricultural investments that yield a return over the long-term and that are irreversible and unverifiable are often under provided by households that expect their land to be taken in the near future.[2] I find that households facing expropriation are less likely to apply organic fertilizer on their plots, an investment that has long term impacts on soil fertility.[3] These distortions may not be problematic if the government is clear about its plan for expropriation, but the effects described in the paper are a function of expectations, and may be damaging to households if plans for expropriations are delayed or cancelled.

The paper proceeds with a brief description of the policy context, a conceptual framework for analyzing household decisions and a discussion of related literature. Section 1.3 outlines the main empirical specification and describes the data in more detail. Section 1.4 presents the main results and section 1.5 provides a number of extensions that try to deal with potential endogeneity.

1.2 Policy context and conceptual framework

1.2.1 Land ownership and expropriation in Ethiopia

The Federal government owns all rural land in Ethiopia and can expropriate land from anyone if it is deemed to be in the national interest. User rights over land entitle small-scale farmers to rent out land, bequeath land to their children and

[2]There is a long literature that looks at the relationship between tenure security and agricultural investment. See for example Ali et al. (2011); Besley (1995); Deininger and Jin (2006); Fenske (2011); Jacoby et al. (2002).

[3]Jacoby et al. (2002) analyzes the likelihood that farmland will be redistributed by local administrators in China and finds that households are less likely to apply organic fertilizer to plots that have a higher chance of being taken away.

to receive compensation in cases where the government expropriates their land. But, crucially, individuals lack the right to sell their land(Holden et al., 2009b). Land sale markets are thus very thin and expropriation is the most common way that land is transferred from smallholders to investors. Current land laws state that individuals should be compensated for losing farmland by receiving an equivalent plot of land nearby. When no land is available, as is often the case, individuals are given financial compensation. Compensation for expropriated land is determined by a fixed formula, which in principle reflects the productivity of the parcel lost by the farmer. It is also meant to compensate farmers for fixed investments that have been made on the land, such as irrigation improvements, trees and buildings. The formula for land compensation is usually: ten times the current market value of the average production of the parcel over the previous five years.[4] In practice, it is impossible to assess a value for each parcel of land, so local governments determine a per unit value for land that is paid to all farmers in the project area. Farmers typically receive a higher payment for irrigated land, are compensated for any tree crops and receive replacement land and money to build a new house if their homestead is expropriated. Local heterogeneity in land (or farmer) productivity is not reflected in the per unit value of compensation. This institutional feature allows for a cleaner identification of the effect of expropriation risk on household behaviour, since all households receive the same price for their land and no household can avoid having their land taken.

1.2.2 Project area and timeline

The investment project in this study is located near the town of Kombolcha, Ethiopia. Kombolcha is a town of approximately 60,000 people that lies on the main road going north from Addis Ababa to Ethiopia's most important seaport,

[4] Similar formulae were/are used in China and Vietnam.(Ding, 2007)

Djibouti. It is known for being a transportation and trucking hub, with some existing industry, including a steel works, a brewery and a textile plant, but this has been changing. Kombolcha was recently designated as an industrial city by the Federal government of Ethiopia and has begun to attract a number of large industrial investment and infrastructure projects. Much of the rural land around the town has already been or will be expropriated to make space for these projects. The project in this study was a textile factory cluster, proposed by an Indian firm, to be built in a rural kebele (the smallest administrative unit) that lies immediately to the west of the town. At the time of the baseline survey, the proposed project required 340 hectares out of a total land area in the kebele of 1000 hectares and local officials estimated that 350 out of 1,100 households living in the kebele would be directly affected by the expropriation.

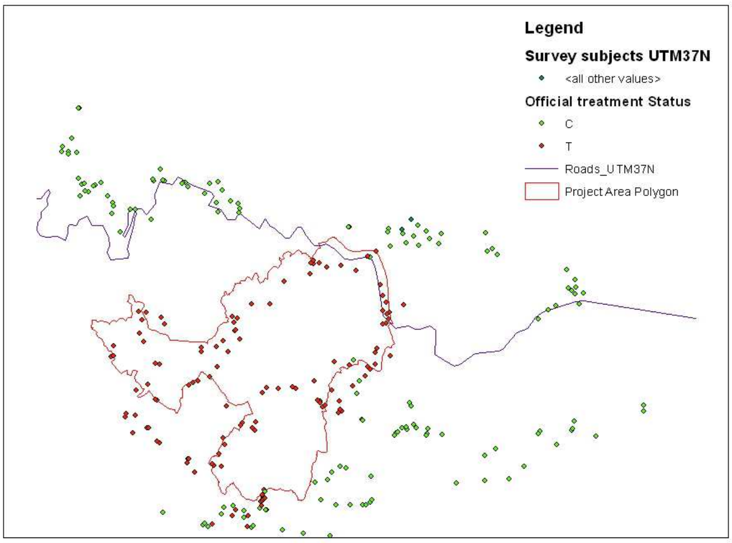
The proposed project was agreed and announced in early 2009. Following an initial agreement, the zonal government (the level of government just below regions) and city administration conducted a cadastral survey of the project area to measure each household's plots and assess their compensation payment. Households were informed that they would lose their land following the end of the main growing season in November. Preliminary research visits to the site were conducted in April 2011, by which point the investment project had been delayed a number of times. Around this time, city officials told households that their land would be remeasured and expropriated in the coming year. Households were told to plant their crops for the main growing season and were assured that their land would only be expropriated following the harvest in November 2011. The baseline survey was conducted in September of that year and focused on various household activities in the preceding 12 months. Following further delays by the investor and a perceived lack of commitment on their part, expropriation was postponed and the investment agreement terminated. The same area of land was later reallocated to the MIDROC investment group who plan to build a steel factory on the site.[5] As of January 2013, the local government began expropriating the land and paying compensation to the farmers. A more detailed description of the compensation procedure can be found in companion paper that analyzes the impact of expropriation and compensation using data from the followup survey. The following timeline sketches out the the key events and decision points for households:

1. June 2009: household's land is mapped for the first time and expropriation is announced.
2. Prior to April 2011: city officials announce that expropriation will occur in November 2011, following the main harvest.
3. April - June 2011: households choose crops to plant for main (meher) season and apply fertilizer. Sorghum is planted during April and requires sufficient rain during the first month of growth to be viable. Wheat is planted in June. Tef is planted in July.
4. September 2011: baseline household survey is conducted.
5. November 2011: households harvest crop, but expropriation is delayed once again.
6. November 2012: land is finally expropriated, following one more full agricultural season.
7. October 2013: followup household survey conducted.

Figure 1 on the following page shows the location and scale of the investment project as well as the geographic distribution of households in the sample. The

[5] MIDROC investment group comprises 41 companies across almost all sectors of the economy and is one of the biggest companies in Ethiopia.

Figure 1: Map of survey area showing HH location - treatment (red), control (green)



area outlined in red illustrates the approximate location and size of the investment project. Red and green markers show the approximate location of households that will, and will not, lose land for the investment project.[6]

1.2.3 Conceptual framework and related literature

This section sketches out a broad theoretical explanation for how expropriation risk might affect investment behaviour and relates it to the literature on investment and tenure security. Households facing imminent expropriation, with the prospect of receiving compensation will adjust their behaviour in anticipation. Some adjustments may not relate to tenure, rather they might be driven by the anticipation of a windfall compensation payment and a reduction in agricultural income. For example, households might borrow money to invest in a new business, retrain members of their household to work in the non-farm sector or insure against

[6]This variable is constructed using the kebele administrators recollection about whether a household will lose land. A detailed discussion follows in section §1.3.2.2 .

a lower future income-stream. Other adjustments to investment relate directly to tenure security and might occur because households fear that they will not recoup their investments (for example, households might skip fallow periods and draw down soil quality). These adjustments depend on at least three things: 1) the timing and likelihood of expropriation, 2) the nature of the investment and 3) the intrinsic characteristics of the households and their land.

There is disagreement in the empirical literature on tenure security and investment in developing countries; across many countries and contexts the same types of investments are not always linked with tenure security. However, for investments with particular features, a consistent link has been found (Fenske, 2011). Agricultural investments affected by tenure security tend to be those that are irreversible, long-term, and unverifiable. These intrinsic features are crucial to understanding how households will respond to the anticipation of expropriation. The most salient features investment in my context are: 1) the time to maturity, 2) the degree to which an investment is verifiable and 3) the reversibility of the investment. If households expect the government to take their land before they can harvest their crops, i.e. before they realize a return on their fertilizer investment, they will be reluctant to invest. Likewise, these households may avoid crops that have a higher risk-return ratio than others. Fertilizer use of any kind is an unverifiable, irreversible investment and its use will not be reflected in compensation payments. However, the returns to chemical fertilizer (as opposed to organic fertilizer) can be recouped at the end of a growing season.[7] Thus, if households facing expropriation placed a higher probability on their land being taken before harvest, they might reduce fertilizer use of any kind. A similar logic can be applied to planting input-intensive crops such as wheat. Investment

[7]Fenske (2011) finds that the link between fertilizer use and tenure security is weak. He suggests that this is because chemical fertilizer is a seasonal investment. The tenure insecurity examined in other surveys tends to be longer term, in contrast to the immediate threat of expropriation examined here. This may explain why I find a link, while others do not.

decisions also depend on the characteristics of households and their land. I address this in more detail in the section on identification where I argue that whether a household faces expropriation is exogenous with respect to their characteristics.

The literature on land tenure security and investment incentives in agriculture spans many contexts, including a number of empirical papers focused on Ethiopia. Ali et al. (2011) find that tenure insecurity, as measured by a low perceived right to transfer land and recurring land reform, reduces investment in cash earning tree crops. Deininger and Jin (2006) find that enhanced transferability rights are correlated with investment in terrace construction (a long term, productive, irreversible investment) and that tenure security and the right to sell encourage investment in trees. There are also a number of papers that evaluate the impact of user-right permits on tenure security, mitigation of land disputes with neighbours and increased farm investment.(Holden et al., 2009a,b) Outside the Ethiopian context, Fenske (2011) finds systematic evidence, in a sample of studies on West Africa, that tenure is linked to investment in fallowing land and planting trees. The evidence for a link with investments such as manure or chemical fertilizer is weaker.

The literature on tenure security and investment often measures tenure using subjective measures of risk, perceived transferability of rights, past experience of conflict with neighbours over land and experience of redistribution. My paper offers something new by focusing on a fairly certain type of tenure insecurity: expropriation. Although households will be compensated for their land, the price is common across farmers and does not reflect any long-term investment, save trees and buildings. The expropriation is also very immediate and so may affect short-run investments (i.e. within one season) more than in other studies. Furthermore, many of the measures typically used as outcome variables may be endogenous, particularly if households make certain types of investments precisely to enhance their tenure security Besley (1995). This paper explicitly identifies the effect of increased tenure insecurity arising from an announced government expropriation. In the next section I argue that this tenure insecurity is exogenous with respect to investment decisions, and is therefore not subject to the same endogeneity problems as in other studies.

The closest study to my paper, in the sense that it specifically examines expropriation and compensation is Ghatak et al. (2012). They examine a case in India where the local government in Singur, West Bengal acquired land from households for a car factory. Rather than evaluating the impact of expropriation on investment, they identify the types of households that were likely to reject the compensation packages offered to them and to determine who lost and gained subsequent to the land transfer. They also analyze what households do with their compensation payments. In a related theoretical paper Ghatak and Mookherjee (2013) argue that compensation payments should be above market value so as not to undermine investment incentives by tenant farmers, and to discourage landlords or local governments from selling too much land to industrialists. This theoretical setup is less applicable in my setting, because households (whether tenant or landlord) have no say in the sale of land to the factory. The type of incentive problems that they identify, i.e. where too much land is sold, may occur on the margins of the project area where plots may be added or subtracted from the proposed area. However, most of the boundaries of the project area are determined by geographical features such as roads or mountains, which leaves little room for manipulation. The biggest scope for corruption in my case comes from influential land-holders working with the local government land measurement committee to have their land classified as irrigable land in order to receive higher compensation payments.[8]

[8] Ghatak et al. (2012) report a great deal of mis-classification of land in their case but they do not attribute this mis-classification to corruption

1.3 Identification and data description

I use plot-level data to test the effect of anticipating expropriation on crop choice portfolios and fertilizer use. Crop choice portfolios are measured using a simple indicator variable for whether the household grows sorghum (safe) or wheat (risky) on a particular plot. Fertilizer use is also measured with a plot-level indicator variable that reflects whether a household used any type of fertilizer on a particular plot.[9] The anticipation effect of expropriation is evaluated in two ways: first, using a household level measure of treatment and second using a plot-level indicator for whether the plot was ultimately expropriated.

1.3.1 Empirical specification

1.3.1.1 Household level treatment effects

The main empirical specification tests whether a household's expectation that they will lose any land influences the crops they choose to grow or the fertilizer use on a given plot. Anticipation effects are modeled using the following specification:

$$y_{ih} = \beta_0 + x'_{ih}\beta + expect_h\gamma + \nu_{ih} \tag{1}$$

where i indexes plots, each belonging to a household indexed by h. y_{ih}, is the binary plot-level outcome of interest, $expect_h$ is the household's subjective assessment of the likelihood of expropriation and x'_{ih} is a vector of controls that vary at the plot level or the household level. Expectations, $expect_h$, are measured

[9] The questionnaire for the baseline survey did not specify the type of fertilizer that was used. Recall data collected as part of the follow up survey asked the same question, but followed up with a question asking the household to specify whether any of this fertilizer was organic fertilizer. The wording of the questions was preserved in order to ensure that answers would be comparable across rounds.

using an indicator for whether the household believes that it is likely or very likely that the government will take any of their land within the next 5 years.

Subjective expectations are likely measured with error, especially if households worry about other reasons for land expropriation unrelated to the project, which would cause attenuation bias. As such, my preferred specification instruments for the household's expectation using the local official's assessment of whether household h will lose their land. Expectations are thus modeled using the following first stage regression:

$$expect_h = \pi_0 + x'_{ih}\pi + T_h\phi + \eta_{ih} \tag{2}$$

where T_h is a an indicator variable for whether the local official thought that household h would lose land for the investment project.[10] I expect that the instrument T_h should to be positive and informative about households' expectations of expropriation, $\phi > 0$. Using 2SLS, the parameter of interest, γ, should be interpreted as the local average intent to treat effect of losing land for the investment project. Thus, γ represents the anticipation effect of expropriation for those households whose expectations were affected by being part of the official treatment group.

1.3.1.2 Plot level treatment effects

In practice, most households in the treatment group only lose a portion of their land. Modeling anticipation effects using a household level measure assumes that each plot is affected by the threat of expropriation in the same way. This is

[10] The official's assessment was collected during the baseline survey. After the follow up data was collected it emerged that in a number of cases the official assessment was inconsistent with realized expropriation. In part this occurred because some land was taken out of the proposed project area and elsewhere in the kebele, additional land was taken to expand part of the town. It isn't clear when these changes were made, so I use the official's treatment status as my preferred measure as it was collected at the same as household expectations were elicited.

a reasonable assumption when households react to the threat of expropriation by adjusting their crop choice portfolio or use of fertilizer, since these actions occur at the household level. However, if households are reacting to the threat of expropriation of a particular plot, they may avoid long term investment in soil fertility only on the affected plot. I use the following specification to test this:

$$y_{ih} = \beta_0 + x'_{ih}\beta + T_{ih}\gamma + \nu_{ih} \tag{3}$$

where instead of using an indicator for treatment that varies at the household level, T_h, I use a plot-specific indicator, T_{ih}. The indicator T_{ih} takes a value of one if household h lost plot i as part of the expropriation. This approach is different than the IV approach used to estimate the household level effect as it does not account for subjective expectations and uses realized outcomes collected during the followup survey rather than the household's treatment status at the time of the baseline survey.[11] This specification does not include an expectations variable because plot specific measures of expropriation risk were not collected.

Finally, I estimate a variant of equation 3:

$$y_{ih} = \beta_0 + x'_{ih}\beta + T_{ih}\gamma + \delta_h + u_{ih} \tag{4}$$

which includes a fixed effect, δ_h for each household. Note that in this case the vector of controls, x_{ih}, only contains plot-level variables. Estimating household level anticipation effects could reflect both investment incentives and risk coping behaviour and it may not be possible to separately identify these mechanisms. The household fixed effects specification eliminates any household level anticipation effects (amongst others), identifying γ using only within-household variation.

[11]The recall data is discussed in more detail in the subsequent data description.

Thus, I use this estimate of γ to isolate the investment incentive mechanism from the household risk coping mechanism.

1.3.2 Identification and data description

1.3.2.1 Identification

In equation 1, which models treatment at the household level, the error term is decomposed as follows:

$$\nu_{ih} = \eta_h + \theta_{ih} + \epsilon_{ih} \tag{5}$$

where, η_h represents unobservable household characteristics, θ_{ih} represents unobservable plot characteristics that are potentially correlated with T_h and ϵ_{ih} is the remaining variation where $E\left[T_h \epsilon_{ih}\right] = 0$. In order to identify γ, unobservable household and plot characteristics must be uncorrelated with a household's treatment status, $E\left[T_h \eta_h\right] = 0$ and $E\left[T_h \theta_{ih}\right] = 0$. In equation 3, which models treatment at the plot level, the error term can be decomposed as above, but the identifying assumptions are slightly different. Instead, I require that a plot's treatment status is uncorrelated with unobservable household and plot characteristics, $E\left[T_{ih} \eta_h\right] = 0$ and $E\left[T_{ih} \theta_{ih}\right] = 0$.

Treatment status is determined at the plot level by whether a plot is located within the investment project area and is not randomly assigned. The selection of plots into the treatment is determined by the needs of the investor, which implies that certain household and plot characteristics will be correlated with either measure of treatment status, T_h or T_{ih}. For example, households that own more land are more likely to be a treated household since there is a higher likelihood that any one of their plots falls inside the project area.[12] Table A.2, which displays household

[12] Households that own no land are an extreme example of this: by definition these households can not be in the treatment group. This specific issue shouldn't affect the analysis, since the treatment and control group only contain households that own land.

level descriptive statistics, shows that treated households own more land and more plots than control households. Similarly, investors chose the project area in part because it was flat, which would imply, all else equal, that treated plots are more likely to be flat. Table A.1, which shows plot level summary statistics, reflects this. Most of these differences that arise because of the selection of land for the project area are observable differences that can be controlled for using plot level and household level control variables and should not bias coefficient estimates.

Unobservable household characteristics, η_h cannot be controlled for. However, a household's treatment status is unlikely to be correlated with any unobservable fixed household characteristics, such as productivity, wealth or credit access that might also affect the outcome variables. There is little reason to believe that the investor targeted specific types of households; the investor chose the project area because it is a flat, contiguous piece of land near to the main road going to Kombolcha. The timing of the expropriation announcement is fairly recent and the nature of the policy intervention means is such that households are limited in their ability to select out of or manipulate treatment. Expropriation is dictated by higher levels of government that leave households with no choice about whether their land will be taken and no ability to negotiate for higher prices. Furthermore, land sales are forbidden in Ethiopia (only 3 out of 1,768 plots in the sample were purchased), which further limits the ability of households to select out of (or in to) treatment by buying and selling land. Finally, the sample is drawn from one administrative region and from a relatively small geographic area, which means that features of the area, such as market access, institutions or access to health and education facilities are common across the sample. One might worry that a plot's treatment status is correlated with unobserved political influence, which would allow households to influence a plot is included or excluded from the project area, but this concern only applies to plots located on the boundary of the project area.[13] The project area is very large relative to the size of each plot and the vast majority of plots are located away from the boundary, which makes this problem relatively minor. All this suggests that unobservable household characteristics will not confound estimates of a household's treatment status or of the treatment status of a given plot.

The project area was selected because it is flat, contiguous and close the the road going to the nearest town. Even conditioning on observable plot characteristics it is possible that plot level unobservable characteristics will correlate with the treatment status of a plot, $E[T_{ih}\theta_{ih}] = 0$ or owning a plot in the treatment area, $E[T_h\theta_{ih}] = 0$. Although flat land is important for growing certain crops, the benefits of this area being contiguous and large do not affect small-scale farmers. Land holdings in this part of Ethiopia are small and fractionalized, which suggests that no particular household would have benefited from selecting into the treatment area. It is also possible that the investors targeted low quality land, but given that the price of non-irrigated land does not vary across the kebele the investors could not have benefited by doing so.[14] Even if investors do not explicitly target unobservable plot characteristics it is still possible that these characteristics are correlated with a plot's treatment status. Since treatment is geographically concentrated in one part of the kebele, unobserved differences in soil quality or plot suitability across treatment and control areas may determine both treatment status and the outcomes of interest. Although the survey area only covers a 4km by 4km area, unobservable plot characteristics can vary greatly even in a small geographic area. Conley and Udry (2010) observe that important growing

[13] There is no reason that $E\left[T_{ih}\eta_h\right] = 0$ should imply $E\left[T_h\eta_h\right] = 0$ or vice versa. The example of political influence is one where uncorrelated household treatment does not imply uncorrelated plot treatment. The opposite might be true if plots were randomly located around the survey area, but some households owned more plots than others; a plot is not more likely to be treated if it is owned by a large land holder, but a large land holder is more likely to be treated.

[14] Table A.1 on page 48 shows that treated plots are less likely to be irrigated. Compensation for irrigated plots is double that for non-irrigated plots, which may have affected the choice to expropriate this land.

conditions for crops in Ghana vary spatially on a scale of hundreds of metres and important plot characteristics are often spatially concentrated within neighbouring plots but vary within villages. Plot-level crop choices and fertilizer use may likewise be a function of spatially correlated unobservable plot characteristics that coincide with the location of the project area.[15] I address this potential endogeneity by running robustness checks that incorporate geographic variation in the location of households.

1.3.2.2 Data description and balance

Baseline data was collected in September 2011. The sample was selected in order to get a representative sample of households living in the kebele at the time of the survey. Sample selection was stratified at the sub-village level, with 16 sub-villages. During the listing process, the kebele administrator identified those households that were losing at least some of their land for the investment project. These households are referred to as "treatment" households, while the rest are referred to as "control" households. Out of a sample of 299 households, 130 were meant to lose land, 167 were not and 2 were missing information. The sample represents approximately one third of the households in the kebele. Total land held by the treatment group accounts for about 200 ha out of 340 ha assigned for the investment project, while the whole sample accounts for 450 ha out of the kebele's 1000 ha. According to a qualitative summary report of the kebele, 20.6% of households are female-headed and own, on average, 1.17 ha of arable land and 0.3 of irrigated land. My sample contains 21% female-headed households and the average household in the sample owns 1.3 ha of arable land and 0.24 ha of irrigated land, which suggests that the sample is fairly representative of

[15] Anecdotal evidence from local agricultural agents and focus group discussions with farmers suggest that agricultural land is not uniform across the kebele. For example, certain parts of the kebele are known to receive more rain and small parts of the treatment area are prone to water-logging.

this particular kebele. A followup survey was conducted in October 2013, eight months after land was expropriated and compensation paid.

Plot level data on 1,758 plots was collected from 299 households in the sample. The analysis uses data from approximately 1,500 plots cultivated by the household during the main growing season.[16] Data for the analysis is drawn from both the baseline survey and a recall module included in the followup survey. The baseline data collected in September 2011 includes detailed plot-level data about the preceding agricultural year (the main harvest is in November). All plot level control variables, x_{ih}, and the main outcome variables, y_{ih} (crop choice and fertilizer use), come from the baseline data. The same plot module was included in the followup survey in October 2013 and asks about plot-level information from the preceding year, which is the first year following expropriation. A shortened recall module was also included in the followup survey to ask about the agricultural season directly preceding expropriation. Data from this season fills the one-year gap between the plot-level data collected at baseline and in the followup survey. In this module, households received a description of their plots using data collected at baseline data and were asked a series of recall questions about each plot. Plot data from the recall section can be directly linked to plot data from the baseline survey.[17] Plot recognition was high; 95% of plots from the baseline survey were recognized by households and 85% of households recognized all of their plots.[18] The recall section included data on whether any of the plot has

[16]Grazing land (132 plots) and land sharecropped out to other households (72 plots) make up the majority of excluded plots. Only 13 plots were left fallow. The final regressions usually use fewer than 1,500 plots, because of missing information.

[17]Prompting households in this way is potentially problematic, since the plot descriptions could bias recall towards the descriptions that households were provided. To limit this, households were provided with the plot name, size, physical characteristics and crop grown. Households were not provided with information about whether they had used fertilizer on the plot.

[18]Households were also asked to link the recall plots to the detailed plot data collected in the followup survey. A successful link allows for a correspondence between plot-level data collected at baseline and plot-data collected in the followup survey. It is possible to check for the quality of the matching by comparing fixed characteristics of plots, such as slope or quality across rounds:

been expropriated, whether the household used any type of fertilizer on the plot, whether the household used any organic fertilizer on the plot, and four yes or no questions about whether the household grew wheat, sorghum, tef or chickpeas on the plot. Plot-specific treatment status, T_{ih}, and organic fertilizer use, y_{ih}, are both used in the main regression specifications. Recall data on crop choice and any fertilizer use are used in robustness checks, but are not used as outcomes in the main specifications.

The regressions outlined in section 1.3.1 use a mix of plot level and household level variables. Summary statistics showing plot-level data split by household treatment status and by plot-specific treatment status can be found in table A.1 on page 48. Summary statistics showing household-level data split by household treatment status can be found in table A.2. Households are considered to be in the treatment group, $T_h = 1$, if the local official believed that they would lose land for the investment project at the time that the baseline survey was conducted. A household expects to lose land, $expect_h = 1$, if they state that it is likely or very likely, that they will lose some land within the next 5 years. Plot-Specific treatment status, T_{ih}, depends on the realized outcome after expropriation occurred. Interestingly, $T_{ih} = 1$ does not always correspond with the baseline beliefs held by local officials: 28% of cultivated land owned by control households ($T_h = 0$) was ultimately expropriated. Nevertheless, official treatment status maps fairly well into various measures of expropriation risk: panel four in table A.2 shows that amongst treated households, 97% had their land mapped by zone officials, 87% believed their land would be taken for the investment project and 84% expected their land to be taken within 5 years compared with 23%, 11% and 38% respectively amongst the control group. Inconsistencies in the official

82% of linked plots are consistent in irrigation status, 81% are consistent in whether the plot is flat and 72% are consistent on whether the plot is considered good quality. Despite some recall bias and problems with matching, households are generally good at recognizing their plots based on descriptions using baseline data.

reports reflect the high degree of uncertainty surrounding the land expropriation in Kombolcha, but do not render T_h an uninformative instrument for household's expectations, $expect_h$. However, care should be taken in interpreting the effect of T_h or T_{ih} on outcomes that are collected in different time periods.

Outcome variables are all measured as plot-level indicator variables for whether the household performs a particular action during the main growing season. Crop choice outcomes are measured by whether the household grows any sorghum or any wheat on a particular plot.[19] Wheat or sorghum is grown by 89% of households in the sample and 40% of households grow both crops. Table A.2 shows that, on average, households dedicate 40% of cultivated land to growing either sorghum or wheat, with treated households dedicating a statistically significant larger share of cultivated land to growing sorghum and less to growing wheat than control households. Fertilizer use is measured by whether the household used any fertilizer on a given plot. The baseline survey did not distinguish between types of fertilizer, instead asking simply: "Did you use fertilizer on this land?" The recall questionnaire included this question with exactly the same wording but added a followup question asking whether households used any organic fertilizer on the plot. At baseline, households report using any fertilizer on 32% of plots. In the recall data, households increase any fertilizer use to 60% of plots and report using organic fertilizer on 41% of plots. It is unlikely that fertilizer use doubled over the course of one year, and is more likely due to changes in the way the question was asked.[20] Nevertheless, treated households use fertilizer on fewer plots than control households across all measures. These differences in means already suggest

[19] Plots tend to be planted with only one type of cereal crop. Plots are subsets of parcels, which are defined as a contiguous areas of land owned by the farmer and unconnected to any other plot owned by the farmer. Households do not have many plots per parcel: there are 1,478 unique household-parcel combinations and 1,760 unique household-plot combinations. This suggests that households plant wheat and sorghum in different locations and reinforces the observations that land ownership is fractured.

[20] Although the wording of the "any fertilizer" question remained unchanged between baseline and recall questionnaires, a consistency check was introduced so that any plot with "organic fertilizer" also had to have "any fertilizer". Neither the organic fertilizer follow-up question nor

that relatively more sorghum is grown by households in the treatment area, but the results presented in the next section include plot level and household level controls.

Reported regressions control for a number of important plot and household characteristics. Plot level controls include: slope, irrigation status, whether the plot can be used in both growing seasons, size and whether the plot was share-cropped. Plot level controls vary systematically with whether a plot has been expropriated, but in ways that would be expected: 90% of treated plots are classified as flat, compared to 80% of control plots and only 10% of treated plots are irrigated, compared to 25% of control plots. These differences reflect the interests of the investor in choosing the project area. Household level controls include the household head's age, literacy, gender and marital status as well as the household's adult labour supply, livestock holding, credit access, consumption and food self-sufficiency. Summary statistics, aggregated at the household level, are in table A.2. The balance between treated and control households is better than at the plot level, although statistically significant differences are present. Treated households report 10% more per capita consumption, have bigger households, are less likely to be headed by females, have the equivalent of half a cow more in livestock and cultivate 0.25 ha more land than control group. These differences are consistent with what one would expect, given that treatment is defined by having at least one plot in the project area: households with greater land holdings are more likely to own at least one plot in the treatment area. However, bias in the treatment effect estimates introduced because of these differences would tend to understate the true effect of the mechanisms I am testing. I am testing whether households choose safer crops and avoid using expensive inputs in anticipation of

---

the consistency check were part of the baseline survey. As such, enumerators could have ignored any plots in the baseline survey that had organic fertilizer but not chemical fertilizer.

losing their land and one might expect households with more wealth, livestock and labour to less worried about income risk associated with expropriation.

1.4 Results

1.4.1 Household treatment results

Results from the IV regression outlined in equation 1 are reported in this section. Equation 1 models the effect of household's subjective expectation of expropriation on the outcomes of interest. In each case, expectations are instrumented using the local official's assessment of whether the household will lose land. The official's assessment is an informative instrument with a coefficient estimate, $\hat{\phi} \approx 0.45$, in the first stage regressions. The local official's assessment of whether a household will be treated is a strong predictor of whether the household believes that it will lose land and the estimate does not vary much when plot and household level controls are included. The full first stage regression results are reported in table A.3. The table shows three regression results: the first includes no control variables, the second includes only plot-level controls and the last includes both plot-level and household-level controls.

Table 1 reports estimates of the effect of expropriation risk on a household's plot-level decision to grow sorghum. Columns 1 through 3 show the results for the main IV regression specification, with the reduced form model (using the instrumental variable directly in the second stage equation instead of expectations) and the un-instrumented regression reported in columns 4 through 9. Each set of 3 columns reports the estimated effect without controls, with plot controls only and with the full set of plot and household controls. Using the IV specification with the full set of controls I find that $\hat{\gamma} = 0.10$. This effect is statistically significant at the 95% level and does not change as controls are taken away. The magnitude of the

Table 1: Effect of expropriation risk on plot level sorghum choice

Dep. var: Sorghum	(1)	(2) IV	(3)	(4)	(5) Reduced form	(6)	(7)	(8) OLS	(9)
Expect expropriation {1,0}	0.15***	0.13**	0.10**				0.046**	0.047**	0.052***
	(0.057)	(0.052)	(0.045)				(0.018)	(0.020)	(0.020)
Official treatment status {1,0}				0.064***	0.056**	0.046**			
				(0.021)	(0.022)	(0.020)			
Observations	1,507	1,506	1,479	1,507	1,506	1,479	1,516	1,515	1,488
R-squared	-0.011	0.079	0.101	0.006	0.089	0.103	0.003	0.088	0.104
Instrument	Y	Y	Y	N	N	N	N	N	N
Controls: Plot level	N	Y	Y	N	Y	Y	N	Y	Y
Controls: Household level	N	N	Y	N	N	Y	N	N	Y
Kleibergen-Paap F-statistic	36.1	36.6	52.1						

Note: This table reports full regression results for the main empirical specification outlined in equation 1. Columns 1 to 3 report results for the IV regression using official treatment status as an instrument for subjective expectations. Columns 4 to 6 report reduced-form results using official treatment status in an OLS specification. Columns 7 to 9 report results from a simple OLS regression using un-instrumented subjective expectations. Samples for each regression are restricted to plots that were cultivated in the main growing season. Clustered standard errors in parentheses are calculated at the sub-village level for 16 clusters. Statistical significance is denoted by the following system: ***p<0.01, **p<0.05, *p<0.1. Plot controls include: slope, irrigation status, plot size and plot size squared, an indicator for whether the plot is used in the main growing season only and an indicator for whether the plot is sharecropped in. Household controls include: head age, female head, marital status, head literacy, household size, livestock, log of per capita consumption, credit access and the number of months households eat mainly from home produced crops.

effect is high: on average, households in the sample grow sorghum on 20% of their cultivated plots. However, some care should be taken in interpreting the meaning of the coefficient: γ represents the change in the outcome variable for those households that had their beliefs about expropriation influenced by being part of the official treatment group. Both reduced form estimates and OLS estimates show a positive relationship between expropriation risk and a household's choice to grow sorghum, although the effect size appears smaller in magnitude than the IV estimate.

Table 2 is set up in the same way as table 1, but reports the estimated effect of expropriation risk on a household's decision to grow wheat. Here the estimated effects are less precisely measured, but are nonetheless statistically significant and robust to the inclusion of plot-level and household-level controls. The negative effects are also present in the reduced form and OLS models. Effect estimates using my preferred specification are reported in column 3 and suggest that households

Table 2: Effect of expropriation risk on plot level wheat choice

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. var: Wheat		IV			Reduced form			OLS	
Expect expropriation {1,0}	-0.088*	-0.098*	-0.11**				-0.022	-0.021	-0.027*
	(0.053)	(0.054)	(0.047)				(0.016)	(0.017)	(0.015)
Official treatment status {1,0}				-0.039*	-0.043**	-0.049**			
				(0.021)	(0.020)	(0.019)			
Observations	1,507	1,506	1,479	1,507	1,506	1,479	1,516	1,515	1,488
R-squared	-0.006	-0.000	0.008	0.003	0.012	0.021	0.001	0.009	0.018
Instrument	Y	Y	Y	N	N	N	N	N	N
Controls: Plot level	N	Y	Y	N	Y	Y	N	Y	Y
Controls: Household level	N	N	Y	N	N	Y	N	N	Y
Kleibergen-Paap F-statistic	36.1	36.6	52.1						

Note: This table reports full regression results for the main empirical specification outlined in equation 1. Columns 1 to 3 report results for the IV regression using official treatment status as an instrument for subjective expectations. Columns 4 to 6 report reduced-form results using official treatment status in an OLS specification. Columns 7 to 9 report results from a simple OLS regression using un-instrumented subjective expectations. Samples for each regression are restricted to plots that were cultivated in the main growing season. Clustered standard errors in parentheses are calculated at the sub-village level for 16 clusters. Statistical significance is denoted by the following system: ***p<0.01, **p<0.05, *p<0.1. Plot controls include: slope, irrigation status, plot size and plot size squared, an indicator for whether the plot is used in the main growing season only and an indicator for whether the plot is sharecropped in. Household controls include: head age, female head, marital status, head literacy, household size, livestock, log of per capita consumption, credit access and the number of months households eat mainly from home produced crops.

that expect to lose land are 11% less likely to grow wheat on any of their plots, an effect that is statistically significant at the 95% level. Overall, 15% of plots are used for growing wheat, which suggests the magnitude of the effect is quite high.

Table 3 presents results showing the effect of expropriation risk on the use of any fertilizer. The effects are not as precisely estimated as in the crop choice variables, but the direction of the effect is consistent with the theoretical predictions. Column 3 shows that households that expect to lose land because of the investment project are 16% less likely to use fertilizer on a plot. This result is statistically significant at the 10% level and the direction of the results is consistent using reduced form estimates. On average, 32% of plots have fertilizer applied. The fertilizer use variable does not explicitly differentiate between chemical and organic fertilizer. Chemical fertilizer is a costly, short term investment, which households can recoup within one growing season, whereas organic fertilizer has long term benefits for soil fertility. It is therefore difficult to know whether households are

Table 3: Effect of expropriation risk on plot level fertilizer use

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. var: Wheat		IV			Reduced form			OLS	
Expect expropriation {1,0}	-0.17	-0.14	-0.16*				0.027	0.041	0.013
	(0.11)	(0.089)	(0.088)				(0.028)	(0.026)	(0.027)
Official treatment status {1,0}				-0.073*	-0.061*	-0.074**			
				(0.041)	(0.034)	(0.036)			
Observations	1,506	1,505	1,478	1,506	1,505	1,478	1,515	1,514	1,487
R-squared	-0.038	0.010	0.045	0.006	0.046	0.080	0.001	0.043	0.073
Instrument	Y	Y	Y	N	N	N	N	N	N
Controls: Plot level	N	Y	Y	N	Y	Y	N	Y	Y
Controls: Household level	N	N	Y	N	N	Y	N	N	Y
Kleibergen-Paap F-statistic	36.1	36.6	52.1						

Note: This table reports full regression results for the main empirical specification outlined in equation 1. Columns 1 to 3 report results for the IV regression using official treatment status as an instrument for subjective expectations. Columns 4 to 6 report reduced-form results using official treatment status in an OLS specification. Columns 7 to 9 report results from a simple OLS regression using un-instrumented subjective expectations. Samples for each regression are restricted to plots that were cultivated in the main growing season. Clustered standard errors in parentheses are calculated at the sub-village level for 16 clusters. Statistical significance is denoted by the following system: ***p<0.01, **p<0.05, *p<0.1. Plot controls include: slope, irrigation status, plot size and plot size squared, an indicator for whether the plot is used in the main growing season only and an indicator for whether the plot is sharecropped in. Household controls include: head age, female head, marital status, head literacy, household size, livestock, log of per capita consumption, credit access and the number of months households eat mainly from home produced crops.

reducing fertilizer use (as it is measured here) because of investment incentives or because of income risk.[21]

Taken together, this behaviour is consistent with the conceptual framework in section 1.2: households in the treatment group are more likely to believe that their land will be taken and react by adjusting their crop choice portfolio and reducing fertilizer use. Sorghum is a safe crop for farmers that does not require inputs. Furthermore, sorghum is drought resistant and, with sufficient rain in the first month of planting (which occurs before any wheat is planted), can provide an assured supply of food. Wheat, on the other hand, benefits greatly from fertilizer application, but has output that can be much more variable. Likewise,

[21]It is likely that this measure reflects chemical fertilizer use. Fertilizer use doubled between the baseline data collection and the recall data collection, despite not changing the wording of the question. In section 1.3.2.2, I argue that adding a follow-up question about organic fertilizer use to the recall questions increased the number of plots reported as having any fertilizer applied. If enumerators only considered chemical fertilizers in the baseline survey, they would have excluded many plots that received only organic fertilizer.

chemical fertilizers can enhance yields, but the downside risk if a crop fails is high. The evidence here is consistent with the idea that households in the treatment group worry about expropriation and the income risks associated with losing land. Although households will receive compensation payments, they may perceive a large degree of uncertainty about when and how much they will be paid.

1.4.2 Plot treatment results

The previous set of results model expropriation risk as something that influences the actions of a household across all plots owned by that household. However, it is possible that households make different decisions depending on whether they expect to lose a particular plot. I do not collect plot-specific expropriation risk, but I can use data on whether plots were ultimately expropriated as a proxy. Table A.7 reports the treatment effect estimates for the regression specified in equation 3, where plot-level outcomes are explained by whether a particular plot is expropriated. Coefficient estimates are reported for regressions on the main outcomes of interest collected at baseline (sorghum, wheat and fertilizer use), their recall analogues and an indicator for organic fertilizer use. Column 1 reports the effect of household level expectations, and is the same as the main results reported in Column 3 in the tables above. Columns 2 to 4 show treatment effects for OLS specifications using plot-level treatment without any controls, with plot-level controls and with the full set of controls. Columns 5 and 6 show treatment effects using plot-level treatment specifications with household level fixed effects, without and with plot-level controls.

Without controlling for plot or household characteristics, households are more likely to grow both sorghum and wheat on treated plots, and less likely to use fertilizer and organic fertilizer. None of these effects are statistically significant, although the effect size is greatest for organic fertilizer use. The raw correlation suggests that treated plots are 9% less likely to have organic fertilizer applied. Introducing plot controls and then household controls reduces the effect size for each of the outcomes. The OLS results suggest that there is not a strong correlation between the treatment status of a plot and the outcomes, with the exception of organic fertilizer use. The non-result for crop choice and fertilizer use is consistent with the idea that expropriation is affecting these choices at the household level: households may choose to shift away from wheat and toward sorghum, but may not care about whether the plot they use will ultimately be expropriated. There may of course be other reasons why treated plots correlate with growing sorghum and wheat: the summary table A.1 showed that treated plots were less likely to be irrigated, more likely to flat and more likely to be used for cereals.

Controlling for household fixed effects strips out any effect of expropriation risk that occurs at the household level, in addition to other fixed household characteristics. This specification is particularly interesting since the effect of plot-level treatment status is identified using only variation within households. Results in column 5 and 6 are striking: within households that do not lose all their land, treated plots are more likely to grow wheat and are less likely to use organic fertilizer. Both effects are statistically significant, and the result for wheat is present both when the baseline data or the recall data are used. The positive effect of treatment on wheat cannot be explained by the conceptual framework in this paper, but nor does it contradict it. Crop choice is considered to be a household level choice that relates to whether the household expects to lose land. The results on organic fertilizer, on the other hand, are consistent with a model of investment incentives and fit well with the conceptual framework. Within households that lose land, treated plots are 13% less likely to have organic fertilizer applied. The magnitude of the effect is large relative to the 40% of plots that have organic fertilizer applied to them. Organic fertilizer has long term benefits for soil fertility and it is reasonable that households apply this type of fertilizer to plots that are not being expropriated.

1.5 Extensions and robustness checks

Identification of the effects discussed above rests on the assumption that households would have made the same crop choice and fertilizer use decisions but for the announcement of the investment project. This section tests the robustness of these results using alternative regression specifications that might account for some sources of potential endogeneity

1.5.1 Running the main specification using recall data

Households do not grow the same crops on the same plot every year, evidence of which can be seen by comparing crop choices on plots across the recall and the baseline data. Out of plots that were used for wheat in either period, 25% of plots had wheat planted in both periods. For plots used for sorghum, 31% had sorghum planted in both years. In contrast, all but 13% of plots that had fertilizer applied in the first period had fertilizer in the second period.[22] If households are indeed making crop portfolio adjustments and fertilizer use choices because of expropriation risk one would expect to see similar results using the recall data. Tables A.8, A.9 and A.10 report regression results that estimate the effect of household level expropriation risk on sorghum, wheat and fertilizer using data collected in the recall module. Estimated effects on sorghum and fertilizer are statistically significant and have the same sign and magnitude as

[22]Fertilizer use reportedly doubled from 30% to 60%, so one would not expect to see many plots without fertilizer that had fertilizer applied at baseline.

the regressions using the baseline data, while the point estimate for the effect on wheat is close to zero. However, tests of equality between the recall estimates and the corresponding baseline point estimates cannot be rejected at the 10% level. These results reinforce the main findings that households in the treatment group adjust their crop portfolios in response to expropriation risk and reduce fertilizer use. However, one should be cautious about these results for a number of reasons. First, it may not be appropriate to use expectations from the baseline to explain data from the recall period; between survey rounds the investment project was once more postponed and the size of the project area was reduced. Second, recall data will be susceptible to the same endogeneity problems as the baseline data if omitted variables are fixed or persistent over time. Any unobservable characteristics such as soil type or plot suitability that determine crop choice or fertilizer use in the baseline data will also drive the results using recall data. Likewise, unobservable fixed household characteristics may be confounding both sets of estimates in the same way.

1.5.2 Ownership and cereal plots

Land laws in Ethiopia make it difficult for households to sell their land and leave the treatment area. However, it is possible that a plot that has been recently acquired by the household was acquired as a result of the project announcement[23]. Furthermore, the threat of expropriation may affect sharecropped plots differently than those owned by a household: households that use sharecropped plots may feel they are at an even greater risk because they will not receive any compensation for investment on plots they cultivate. To account for this possibility, I run the main specification on a sub-sample of plots that are cultivated by the owner and on a sub-sample of plots that were acquired three or more years before the

[23] Although land sale is not possible, land may be acquired from parents or other family members.

baseline data was collected (i.e. prior to any announcement of expropriation). The treatment effect estimates from these regressions are reported in column 2 and 3 of table A.13. Results have the same sign and magnitudes as the main regression results. The results for wheat and sorghum are statistically significant at the 5% level. These checks address the possibility that the treatment group is made up of different types of households than the control group, but the bigger threat to identification is unobserved plot characteristics. Column 4 of table A.13 restricts the sample to plots that are used for growing cereals in the baseline data. The effects remain statistically significant, suggesting that unobservable plot level variables would have to have something to do with the suitability for particular types of cereals.

1.5.3 Spatial statistics

In this study, a household's treatment status is geographically determined by whether they own plots in the project area. Since households mostly live in small hamlets near their plots, treated households are geographically clustered around the project area. This can be seen in the map in figure 1; most of the control households are located relatively far away from the project area, and in some cases on the other side of major obstacles such as roads or hills. Spatial clustering of treatment status poses problems for identification, since important unobservable plot characteristics, may also be geographically clustered. I use GPS data on the location of the household, together with data from a cadastral survey conducted for the investment project to conduct a number of robustness checks that can account for spatial correlation.

Table A.12 presents some basic tests using household location. Observations near the project boundary may be more similar along unobservable dimensions, both at the household level and at the plot level. (The latter is true if household location is a good proxy for the location of a plot). Column 2, 3 and 4 report results from the main regression specification using the sub-sample of houses that live within 500m, 750m and 1000m of the project boundary. The positive effect of expropriation on sorghum goes to zero when the band is extended beyond 750m of the project boundary. The sign of the effect on wheat and fertilizer use is negative for three sub-samples, consistent with the main model. Although I conceive of these robustness checks as tests for unobservable plot characteristics, they may also account for similarities between unobservable household characteristics. The other caveat is that this method does not necessarily identify the treatment effect using neighbouring households; households living on opposite ends of the project area are considered comparable despite the fact that they live far from one another. Column 5 controls for the coordinates of the household's location and column 6 controls for a quadratic expansion of the household's coordinates. The main results are robust to the former, but not the latter. These controls may proxy for unobservable location fixed effects, but are also closely correlated with a household's treatment status, which may explain why the treatment effects tend to zero after controlling for the quadratic expansion of household coordinates.

Ideally, I would like to make comparisons between proximate plots that are owned by a treated and a control household. This comparison would control for any plot level unobservables that are common to that area and would allow for a cleaner identification of the treatment effect. Conley and Udry (2010) and Goldstein and Udry (2008) face similar endogeneity problem and propose a spatial fixed effects method to control for localized plot unobservables.[24] I implement this method, but leave the model description and technical interpretation to section 1.6 in the appendix. Table A.14 reports treatment effect estimates for the spatial fixed effects specifications. Columns 2, 3 and 4 show estimates that implement

[24]Magruder (2012) uses spatial fixed effects to control for unobserved local labour market conditions in neighbouring bargaining councils in South Africa. In this paper he highlights the similarities between spatial fixed effects and regression discontinuities.

the fixed effects transformation using neighbourhoods defined by a 500m, 750m an 1000m radius around each household. The smallest neighbourhoods contain 15 households on average, the next 28 and the largest 45. The estimated coefficients on treatment for most of the specifications are close to zero and are not statistically significant. The exception is the estimated effect on wheat when neighbourhoods have a 500m radius, which is statistically significant and takes the opposite sign as the main regression specification. Although the treatment effects reported in the main specification are not robust to the inclusion of spatial fixed effects, this does not imply that estimated treatment effects are driven by unobservable plot characteristics. The instability and imprecision of the estimated treatment effect (on all three outcome variables) may be a function of the spatial clustering of households by treatment status. The spatial fixed effects model identifies the treatment effect using only information from households whose treatment status differs from at least one of their neighbours. Information on treatment status from any household living in a homogenous neighbourhood is not included in treatment effect estimates. Furthermore, using the location of the household as a proxy for the location of a plot means that this model is also controlling for spatially concentrated characteristics of households. If the threat of expropriation spills over to neighbouring households, this may jointly determine the behaviour of treatment and control households from the same neighbourhood. This would tend to understate the true size of the effect. In either case, these results can be interpreted as evidence that the treatment effect in the main model is being driven by differences between areas more than differences within neighbourhoods.

1.6 Conclusion

Expropriation of rural land has been part of Ethiopia's recent urban and industrial expansion, as it has been in other African countries and around the world. Land laws are in place that entitle farmers to financial compensation when replacement land is not available. However, compensation payments may not be enough to mitigate the distortions brought about by the announcement of expropriation. This paper investigated two mechanisms by which the announcement of an expropriation can distort household behaviour: first, households will employ ex ante coping strategies to mitigate against any income risk they may face in the year following expropriation and second, households will react to investment incentives associated with expropriation risk by reducing investment in immovable, unverifiable investments for which the household will not be compensated. Using a cross-section of households living in and around the site of a proposed factory, I show that the households that anticipate expropriation choose a safer crop portfolio, reduce their use of fertilizer inputs and avoid investing in the long-term fertility of plots that will ultimately be expropriated. Treated households grow more sorghum, which requires fewer inputs and is less risky, and grow less wheat, which is more input intensive. Households also reduce their use of fertilizers, despite the fact that it as an investment which could benefit them in the current growing season.

These findings contribute to the literature on land investment and tenure security, but offer a distinct context. The results presented here are based on short term expectations of a concrete expropriation policy, in contrast to the long term subjective measures of tenure insecurity that often used to evaluate the effects tenure insecurity. Other work focused on program evaluation, such as Field (2007), typically evaluates policy designed to improve to enhance tenure security. Here the policy being evaluated works in the opposite direction: households that happen to own land in the treatment area are having their tenure security eroded. Instead of encouraging investment, the threat of immediate expropriation has the opposite effect.

In some instances, distortions to investment will be of no consequence: drawing down soil quality on land designated for a factory is not problematic. However, delays and inconsistency in implementation and poor communication of expropriation policy can increase insecurity for households thereby distorting their behaviour. The threat of expropriation can exist without it ever being realized; households living near to areas that are expropriated, or households living in areas where a project is cancelled may change their behaviour regardless of the outcome in detrimental ways. The findings in this paper highlight some ways in which these types of households' behaviour is distorted. The mere threat of expropriation seems to be a sufficient reason for households to reduce productive investments and allowing the process of expropriation for investment projects to drag on over many years may have adverse consequences for small-scale farmers.

References

Ali, D. A., Dercon, S., and Gautam, M. (2011). Property rights in a very poor country: Tenure insecurity and investment in ethiopia. *Agricultural Economics*, 42(1):75–86.

Banerjee, A. V., Bardhan, P., Basu, K., Chaudhury, M. D., Ghatak, M., Guha, A. S., Majumdar, M., Mookherjee, D., and Ray, D. (2007). Beyond nandigram: Industrialisation in west bengal. *Economic and Political Weekly*, 42(17):1487–1489.

Besley, T. (1995). Property rights and investment incentives: Theory and evidence from ghana. *journal of Political Economy*, pages 903–937.

Conley, T. G. and Udry, C. R. (2010). Learning about a new technology: Pineapple in ghana. *American Economic Review*, 100(1):35–69.

Cotula, L., Vermeulen, S., Leonard, R., and Keeley, J. (2009). *Land grab or development opportunity? Agricultural investment and international land deals in Africa*. IIED, FAO and IFAD.

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

Dercon, S. (1996). Risk, crop choice, and savings: Evidence from tanzania. *Economic Development and Cultural Change*, 44(3):485–513.

Dercon, S. and Christiaensen, L. (2011). Consumption risk, technology adoption and poverty traps: Evidence from ethiopia. *Journal of Development Economics*, 96(2):159–173.

Ding, C. (2007). Policy and praxis of land acquisition in china. *Land Use Policy*, 24(1):1–13.

Fafchamps, M. (1992). Cash crop production, food price volatility, and rural market integration in the third world. *American Journal of Agricultural Economics*, 74(1):90–99.

Fenske, J. (2011). Land tenure and investment incentives: Evidence from west africa. *Journal of Development Economics*, 95(2):137 – 156.

Field, E. (2007). Entitled to work: Urban property rights and labor supply in peru. *The Quarterly Journal of Economics*, 122(4):1561–1602.

Ghatak, M., Mitra, S., Mookherjee, D., and Nath, A. (2012). Land acquisition and compensation in singur: What really happened? *Working PaperInstitute for Economic Development, Boston University.*

Ghatak, M. and Mookherjee, D. (2013). Land acquisition for industrialization and compensation of displaced farmers. *Journal of Development Economics*, (0):–.

Goldstein, M. and Udry, C. (2008). The profits of power: Land rights and agricultural investment in ghana. *Journal of Political Economy*, 116(6):981–1022.

Holden, S. T., Deininger, K., and Ghebru, H. (2009a). Impacts of low-cost land certification on investment and productivity. *American Journal of Agricultural Economics*, 91(2):359–373.

Holden, S. T., Otsuka, K., and Place, F. M. (2009b). *The Emergence of Land Markets in Africa: Impacts on Poverty, Equity, and Efficiency.* Resources for the Future.

Jacoby, H. G., Li, G., and Rozelle, S. (2002). Hazards of expropriation: Tenure insecurity and investment in rural china. *The American Economic Review*, 92(5):1420–1447.

Magruder, J. R. (2012). High unemployment yet few small firms: The role of centralized bargaining in south africa. *American Economic Journal: Applied Economics*, 4(3):138–166.

Appendices to Chapter 1

A. Tables

Table A.1: Plot-level summary statistics by household treatment status and plot treatment status

	Total		Household level treatment (T_h)				Plot level treatment (T_{ih})			
			Control		Treatment		Control		Treatment	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Baseline outcome variables (y_{ih})										
Sorghum indicator, {1,0}	0.201	0.401	0.169	0.375	0.233***	0.423	0.200	0.400	0.211	0.408
Wheat indicator, {1,0}	0.150	0.357	0.170	0.376	0.132**	0.338	0.138	0.345	0.167	0.373
Any fertilizer indicator, {1,0}	0.324	0.468	0.361	0.481	0.288***	0.453	0.342	0.475	0.296*	0.457
Number of plots	1514		740		767		847		611	
Recall outcome variables (y_{ih})										
Sorghum indicator, {1,0}	0.170	0.376	0.133	0.340	0.205***	0.404	0.151	0.358	0.196**	0.398
Wheat indicator, {1,0}	0.195	0.397	0.199	0.399	0.195	0.396	0.177	0.382	0.221**	0.415
Any fertilizer indicator, {1,0}	0.602	0.490	0.642	0.480	0.563***	0.496	0.624	0.485	0.571**	0.495
Organic fertilizer indicator, {1,0}	0.414	0.493	0.432	0.496	0.401	0.490	0.452	0.498	0.361***	0.481
Number of plots	1473		720		744		853		620	
Plot control variables at baseline (x_{ih})										
Plot irrigated, {1,0}	0.197	0.398	0.242	0.429	0.156***	0.364	0.251	0.434	0.119***	0.325
Used only during main season, {1,0}	0.797	0.402	0.761	0.427	0.832***	0.374	0.733	0.443	0.889***	0.315
Plot sloping, {1,0}	0.160	0.367	0.168	0.374	0.154	0.361	0.190	0.393	0.118***	0.323
Plot steep, {1,0}	0.00859	0.0923	0.00946	0.0969	0.00782	0.0882	0.0153	0.123	0***	0
Plot size (ha)	0.251	0.224	0.260	0.241	0.243	0.209	0.249	0.232	0.256	0.194
Sharecropped or rented in? {1,0}	0.129	0.336	0.150	0.357	0.111**	0.314	0.130	0.336	0.129	0.336
Plot treated? {1,0}	0.419	0.494	0.269	0.444	0.565***	0.496	-	-	-	-
Number of plots	1514		740		767		847		611	

Note: Means and standard deviations across both measures of treatment are calculated above. Statistically significant differences between treatment and control are denoted by stars on the treatment mean, using the following: ***p<0.01, **p<0.05, *p<0.1. Household level treatment is measured using the local official's assessment of whether a household will lose land. Plot level treatment was collected using recall data and reflects the realization of expropriation. Baseline variables were collected during the baseline survey and refer to the agricultural season that ended two years before expropriation. Recall variables were collected during the followup survey and refer to the agricultural season just before expropriation. Summary statistics are calculated only for plots that were cultivated during the main growing season, which accounts for the discrepancy between the number of observations in the recall data and in the baseline data. In addition, recall plots that could not be matched were not included.

Table A.2: Household level summary statistics by official treatment status

	Total		Household level treatment (T_h) Control		Treatment	
	Mean	SD	Mean	SD	Mean	SD
Household control variables ($x_{i,h}$)						
Head age	50.21	15.64	49.97	16.29	50.50	14.86
Female, {1,0}	0.196	0.398	0.234	0.425	0.150*	0.358
Head separated or divorced, {1,0}	0.207	0.406	0.234	0.425	0.173	0.380
Head literate, {1,0}	0.400	0.491	0.367	0.484	0.441	0.498
Household size	5.067	1.937	4.823	1.989	5.370**	1.834
Tropical livestock units	3.436	2.345	3.213	2.481	3.715*	2.142
Log of per capita consumption	8.447	0.498	8.394	0.500	8.513**	0.490
Credit access, {1,0}	0.565	0.497	0.544	0.500	0.591	0.494
Months self-sufficient	9.077	2.897	9.184	2.812	8.944	3.005
Land use, measured in share of cultivated land area (outcomes)						
Sorghum (baseline)	0.253	0.239	0.213	0.244	0.302***	0.225
Wheat (baseline)	0.157	0.192	0.179	0.210	0.131**	0.165
Any fertilizer (baseline)	0.330	0.346	0.357	0.364	0.296	0.321
Sorghum (recall)	0.196	0.229	0.146	0.223	0.255***	0.223
Wheat (recall)	0.212	0.220	0.220	0.235	0.202	0.201
Any fertilizer (recall)	0.614	0.389	0.646	0.389	0.575	0.387
Organic fertilizer (recall)	0.417	0.406	0.427	0.407	0.406	0.406
Land use, measured in share of cultivated land area unless specified (other variables)						
Improved seeds	0.179	0.242	0.196	0.260	0.159	0.216
Irrigated land	0.188	0.348	0.174	0.342	0.206	0.356
Pulses	0.0381	0.103	0.0342	0.105	0.0429	0.1000
Roots	0.0379	0.0945	0.0417	0.101	0.0333	0.0856
Vegetables	0.0610	0.127	0.0568	0.137	0.0663	0.114
Cereals	0.868	0.187	0.842	0.219	0.900***	0.133
Number of plots per household	6.151	3.061	5.639	2.976	6.787***	3.057
Area of land cultivated (ha)	1.375	0.891	1.275	0.916	1.499**	0.847
Measures of expropriation						
Expect land to be taken w/in 5 years, {1,0}	0.589	0.493	0.386	0.488	0.843***	0.366
Share of cultivated land expropriated	0.431	0.418	0.287	0.387	0.602***	0.388
Are you losing land for the investment project? {1,0}	0.446	0.498	0.108	0.311	0.866***	0.342
Have you had your land mapped? {1,0}	0.561	0.497	0.234	0.425	0.969***	0.175
Number of households	285		158		127	

Note: Means and standard deviations across a household treatment status is calculated above. Statistically significant differences between treatment and control are denoted by stars on the treatment mean, using the following: ***p<0.01, **p<0.05, *p<0.1. Household level treatment is measured using the local official's assessment of whether a household will lose land. Baseline variables were collected during the baseline survey and refer to the agricultural season that ended two years before expropriation. Recall variables were collected during the followup survey and refer to the agricultural season just before expropriation. Unless specified, the calculations shown use data from the baseline survey and not the recall data. Summary statistics are calculated only for households that cultivated land during the main growing season. As such, the number of households in the analysis is 285 rather than the 299 surveyed.

Table A.3: First stage regression for baseline results - Full

	(1)	(2)	(3)
	First-stage Regression		
	Expect expropriation indicator?	Expect expropriation indicator?	Expect expropriation indicator?
Official treatment status {1,0}	0.44***	0.44***	0.46***
	(0.073)	(0.072)	(0.063)
Is this plot irrigated? {1,0}		-0.072	-0.051
		(0.053)	(0.051)
Is the plot used only during Meher? {1,0}		-0.064	-0.046
		(0.077)	(0.082)
Is the plot somewhat sloping? {1,0}		-0.0014	-0.0069
		(0.055)	(0.051)
Is the plot steep? {1,0}		0.14	0.15
		(0.18)	(0.17)
Plot size (ha)		-0.093	-0.066
		(0.16)	(0.13)
Plot size (ha)^2		0.016	0.017
		(0.11)	(0.10)
Is the plot sharecropped or rented in? {1,0}		0.046	-0.012
		(0.058)	(0.057)
Head age (years)			-0.0032
			(0.0021)
Female head of household, {1,0}			0.00086
			(0.087)
Head is separated, divorced or widowed, {1,0}			0.040
			(0.11)
Head can read and write simple sentence, {1,0}			0.00056
			(0.067)
Household size			0.022
			(0.017)
Tropical livestock units			-0.0039
			(0.011)
Log of per capita consumption (in Birr)			-0.047
			(0.051)
Household can access 4000 birr in emergency, {1,0}			0.092*
			(0.050)
Number of months self-sufficient			0.023*
			(0.013)
Constant	0.41***	0.49***	0.63
	(0.064)	(0.10)	(0.54)
Observations	1,507	1,506	1,479
R-squared	0.208	0.213	0.272
Controls: Plot level	N	Y	Y
Controls: Household level	N	N	N

Note: This table reports OLS estimates for the first-stage regression specified in the main empirical specification outlined in equation 2. Columns 1 to 3 report OLS regression results first without controls, then with plot controls and finally with plot and household controls. Samples for each regression are restricted to plots that were cultivated in the main growing season. Clustered standard errors in parentheses are calculated at the sub-village level for 16 clusters. Statistical significance is denoted by the following system: ***p<0.01, **p<0.05, *p<0.1.

Table A.4: Plot level sorghum indicator at baseline - Full

	(1)	(2)	(3)	(4)	(5)	(6)
	IV Regression			Reduced Form Regression		
	Sorghum indicator	Sorghum indicator	Sorghum indicator	Sorghum indicator	Sorghum indicator	Sorghum indicator
Expects any plot to be taken w/in 5 years, {1,0}	0.15***	0.13**	0.10**			
	(0.057)	(0.052)	(0.045)			
Official treatment status, {1,0} - Instrument				0.064***	0.056**	0.046**
				(0.021)	(0.022)	(0.020)
Is this plot irrigated? {1,0}		-0.078***	-0.096***		-0.087***	-0.10***
		(0.016)	(0.014)		(0.017)	(0.017)
Is the plot used only during Meher? {1,0}		0.11***	0.096***		0.10***	0.092***
		(0.013)	(0.014)		(0.017)	(0.019)
Is the plot somewhat sloping? {1,0}		0.059**	0.063**		0.059**	0.062***
		(0.029)	(0.027)		(0.025)	(0.024)
Is the plot steep? {1,0}		0.064	0.074		0.082	0.089
		(0.071)	(0.070)		(0.072)	(0.075)
Plot size (ha)		0.56***	0.57***		0.55***	0.57***
		(0.093)	(0.087)		(0.089)	(0.085)
Plot size (ha)^2		-0.18***	-0.19***		-0.18***	-0.19***
		(0.068)	(0.069)		(0.061)	(0.063)
Is the plot sharecropped or rented in? {1,0}		-0.026	-0.011		-0.020	-0.012
		(0.027)	(0.025)		(0.025)	(0.022)
Head age (years)			0.0012***			0.00087**
			(0.00044)			(0.00037)
Female head of household, {1,0}			0.12***			0.12**
			(0.048)			(0.051)
Head is separated, divorced or widowed, {1,0}			-0.14***			-0.13***
			(0.035)			(0.038)
Head can read and write simple sentence, {1,0}			0.013			0.013
			(0.021)			(0.020)
Household size			0.0064			0.0086*
			(0.0063)			(0.0050)
Tropical livestock units			0.0052			0.0048
			(0.0049)			(0.0045)
Log of per capita consumption (in Birr)			0.026*			0.021
			(0.014)			(0.014)
Household can access 4000 birr in emergency, {1,0}			-0.023			-0.014
			(0.028)			(0.028)
Number of months self-sufficient			-0.012***			-0.010***
			(0.0021)			(0.0018)
Constant	0.11**	-0.077*	-0.25**	0.17***	-0.014	-0.19
	(0.044)	(0.043)	(0.11)	(0.017)	(0.025)	(0.11)
Observations	1,507	1,506	1,479	1,507	1,506	1,479
R-squared	-0.011	0.079	0.101	0.006	0.089	0.103
Instrument	IV	IV	IV	Reduced	Reduced	Reduced
Controls: Plot level	N	Y	Y	N	Y	Y
Controls: Household level	N	N	Y	N	N	Y
Kleibergen-Paap F-statistic	36.1	36.6	52.1			

Note: This table reports full regression results for the main empirical specification outlined in equation 1. Columns 1 to 3 report results for the IV regression using official treatment status as an instrument for subjective expectations. Columns 4 to 6 report reduced-form results using official treatment status in an OLS specification. Samples for each regression are restricted to plots that were cultivated in the main growing season. Clustered standard errors in parentheses are calculated at the sub-village level for 16 clusters. Statistical significance is denoted by the following system: ***p<0.01, **p<0.05, *p<0.1.

Table A.5: Plot level wheat indicator at baseline - Full

	(1)	(2)	(3)	(4)	(5)	(6)
	IV Regression			Reduced Form Regression		
	Wheat indicator	Wheat indicator	Wheat indicator	Wheat indicator	Wheat indicator	Wheat indicator
Expects any plot to be taken w/in 5 years, {1,0}	-0.088*	-0.098*	-0.11**			
	(0.053)	(0.054)	(0.047)			
Official treatment status, {1,0} - Instrument				-0.039*	-0.043**	-0.049**
				(0.021)	(0.020)	(0.019)
Is this plot irrigated? {1,0}		0.038	0.037		0.045	0.042
		(0.052)	(0.046)		(0.051)	(0.046)
Is the plot used only during Meher? {1,0}		0.093***	0.094***		0.100***	0.099***
		(0.031)	(0.031)		(0.029)	(0.030)
Is the plot somewhat sloping? {1,0}		-0.021	-0.016		-0.021	-0.015
		(0.020)	(0.020)		(0.019)	(0.020)
Is the plot steep? {1,0}		-0.063	-0.062		-0.077	-0.078
		(0.065)	(0.068)		(0.080)	(0.081)
Plot size (ha)		0.068	0.087		0.077	0.094
		(0.074)	(0.079)		(0.077)	(0.079)
Plot size (ha)^2		-0.066	-0.076*		-0.067*	-0.078**
		(0.042)	(0.040)		(0.039)	(0.037)
Is the plot sharecropped or rented in? {1,0}		-0.012	-0.023		-0.017	-0.021
		(0.015)	(0.015)		(0.016)	(0.019)
Head age (years)			-0.0018***			-0.0014***
			(0.00047)			(0.00053)
Female head of household, {1,0}			-0.011			-0.011
			(0.042)			(0.044)
Head is separated, divorced or widowed, {1,0}			0.070**			0.066*
			(0.035)			(0.039)
Head can read and write simple sentence, {1,0}			-0.019			-0.019
			(0.020)			(0.018)
Household size			0.015**			0.013**
			(0.0063)			(0.0054)
Tropical livestock units			-0.0078*			-0.0073*
			(0.0040)			(0.0039)
Log of per capita consumption (in Birr)			0.039*			0.044**
			(0.023)			(0.019)
Household can access 4000 birr in emergency, {1,0}			0.020			0.010
			(0.016)			(0.015)
Number of months self-sufficient			0.00062			-0.0019
			(0.0025)			(0.0028)
Constant	0.21***	0.13***	-0.19	0.17***	0.079**	-0.26
	(0.039)	(0.048)	(0.23)	(0.014)	(0.032)	(0.19)
Observations	1,507	1,506	1,479	1,507	1,506	1,479
R-squared	-0.006	-0.000	0.008	0.003	0.012	0.021
Instrument	IV	IV	IV	Reduced	Reduced	Reduced
Controls: Plot level	N	Y	Y	N	Y	Y
Controls: Household level	N	N	Y	N	N	Y
Kleibergen-Paap F-statistic	36.1	36.6	52.1			

Note: This table reports full regression results for the main empirical specification outlined in equation 1. Columns 1 to 3 report results for the IV regression using official treatment status as an instrument for subjective expectations. Columns 4 to 6 report reduced-form results using official treatment status in an OLS specification. Samples for each regression are restricted to plots that were cultivated in the main growing season. Clustered standard errors in parentheses are calculated at the sub-village level for 16 clusters. Statistical significance is denoted by the following system: ***p<0.01, **p<0.05, *p<0.1.

Table A.6: Plot level fertilizer use indicator at baseline - Full

	(1)	(2)	(3)	(4)	(5)	(6)
		IV Regression			Reduced Form Regression	
	Fertilizer indicator	Fertilizer indicator	Fertilizer indicator	Fertilizer indicator	Fertilizer indicator	Fertilizer indicator
Expects any plot to be taken w/in 5 years, {1,0}	-0.17	-0.14	-0.16*			
	(0.11)	(0.089)	(0.088)			
Official treatment status, {1,0} - Instrument				-0.073*	-0.061*	-0.074**
				(0.041)	(0.034)	(0.036)
Is this plot irrigated? {1,0}		0.19***	0.19***		0.20***	0.20***
		(0.067)	(0.063)		(0.067)	(0.064)
Is the plot used only during Meher? {1,0}		0.0026	0.0092		0.012	0.017
		(0.059)	(0.056)		(0.058)	(0.055)
Is the plot somewhat sloping? {1,0}		-0.071	-0.073*		-0.071*	-0.072*
		(0.044)	(0.042)		(0.042)	(0.040)
Is the plot steep? {1,0}		-0.30***	-0.30***		-0.32***	-0.33***
		(0.048)	(0.043)		(0.034)	(0.043)
Plot size (ha)		0.093	0.15		0.11	0.16
		(0.15)	(0.13)		(0.14)	(0.13)
Plot size (ha)^2		-0.036	-0.060		-0.038	-0.063
		(0.099)	(0.085)		(0.089)	(0.076)
Is the plot sharecropped or rented in? {1,0}		-0.068**	-0.096***		-0.075***	-0.094***
		(0.030)	(0.030)		(0.029)	(0.029)
Head age (years)			-0.0032***			-0.0027***
			(0.0011)			(0.00081)
Female head of household, {1,0}			-0.015			-0.015
			(0.088)			(0.088)
Head is separated, divorced or widowed, {1,0}			0.15*			0.15
			(0.092)			(0.091)
Head can read and write simple sentence, {1,0}			0.094**			0.094**
			(0.040)			(0.038)
Household size			0.021			0.017
			(0.013)			(0.011)
Tropical livestock units			0.0017			0.0023
			(0.012)			(0.012)
Log of per capita consumption (in Birr)			0.040			0.048
			(0.034)			(0.036)
Household can access 4000 birr in emergency, {1,0}			0.051			0.037
			(0.035)			(0.031)
Number of months self-sufficient			0.0062			0.0024
			(0.0061)			(0.0048)
Constant	0.43***	0.38***	-0.075	0.36***	0.31***	-0.18
	(0.076)	(0.086)	(0.32)	(0.029)	(0.062)	(0.32)
Observations	1,506	1,505	1,478	1,506	1,505	1,478
R-squared	-0.038	0.010	0.045	0.006	0.046	0.080
Instrument	IV	IV	IV	Reduced	Reduced	Reduced
Controls: Plot level	N	Y	Y	N	Y	Y
Controls: Household level	N	N	Y	N	N	Y
Kleibergen-Paap F-statistic	36.1	36.6	52.1			

Note: This table reports full regression results for the main empirical specification outlined in equation 1. Columns 1 to 3 report results for the IV regression using official treatment status as an instrument for subjective expectations. Columns 4 to 6 report reduced-form results using official treatment status in an OLS specification. Samples for each regression are restricted to plots that were cultivated in the main growing season. Clustered standard errors in parentheses are calculated at the sub-village level for 16 clusters. Statistical significance is denoted by the following system: ***p<0.01, **p<0.05, *p<0.1.

Table A.7: Plot-specific expropriation effects with household fixed effects

	(1)	(2)	(3)	(4)	(5)	(6)
	Coefficient on HH level expectation	Coefficient on indicator for whether a plot is treated, {1,0}				
Sorghum plot indicator (Baseline), {1,0}	0.10**	0.012	-0.014	-0.019	-0.071	-0.083
	(0.045)	(0.029)	(0.032)	(0.031)	(0.059)	(0.057)
Wheat plot indicator (Baseline), {1,0}	-0.11**	0.029	0.023	0.021	0.079*	0.073*
	(0.047)	(0.023)	(0.023)	(0.023)	(0.041)	(0.038)
Any fertilizer indicator (Baseline), {1,0}	-0.16*	-0.045	-0.033	-0.027	0.061	0.048
	(0.088)	(0.047)	(0.042)	(0.041)	(0.052)	(0.057)
Sorghum plot indicator (Recall), {1,0}	0.12**	0.045	0.029	0.022	-0.000020	-0.014
	(0.050)	(0.030)	(0.031)	(0.031)	(0.056)	(0.053)
Wheat plot indicator (Recall), {1,0}	-0.016	0.044	0.034	0.035	0.12**	0.12**
	(0.062)	(0.029)	(0.030)	(0.029)	(0.045)	(0.042)
Any fertilizer indicator (Recall), {1,0}	-0.17*	-0.053	-0.044	-0.038	-0.011	-0.0082
	(0.100)	(0.052)	(0.048)	(0.041)	(0.042)	(0.044)
Organic fertilizer indicator (Recall), {1,0}	-0.059	-0.091	-0.073	-0.059	-0.13**	-0.11**
	(0.11)	(0.053)	(0.047)	(0.039)	(0.047)	(0.043)
Maximum number of observations (recall variables in brackets):	1479 (1432)	1458 (1469)	1457 (1468)	1432 (1441)	1458 (1469)	1457 (1468)
Method:	IV	OLS	OLS	OLS	HH-FE	HH-FE
Controls: Plot level	Y	N	Y	Y	N	Y
Controls: Household level	Y	N	N	Y	N	N
Household fixed effect	N	N	N	N	Y	Y

Note: This table shows the coefficient on treatment indicators for various model specifications. Columns corresponds to a different regression specification and rows to a different outcome variable. Column 1 shows the coefficient of interest from the main regression specification; this is a households expectation of expropriation instrumented by the official's assessment of whether the household will lose land. Columns 2 to 6 show the coefficient on an indicator for whether a particular plot was ultimately expropriated. Columns 2, 3 and 4 show OLS estimates and columns 5 and 6 show OLS estimates that absorb any household fixed effects. Samples for each regression are restricted to plots that were cultivated in the relevant growing season. Clustered standard errors in parentheses are calculated at the sub-village level for 16 clusters. Statistical significance is denoted by the following system: ***p<0.01, **p<0.05, *p<0.1. Plot controls include: slope, irrigation status, plot size and plot size squared, an indicator for whether the plot is used in the main growing season only and an indicator for whether the plot is sharecropped in. Household controls include: head age, female head, marital status, head literacy, household size, livestock, log of per capita consumption, credit access and the number of months households eat mainly from home produced crops.

Table A.8: Plot level sorghum indicator (Recall) - Full

	(1)	(2)	(3)	(4)	(5)	(6)
		IV Regression			Reduced Form Regression	
	Sorghum indicator (recall)	Sorghum indicator (recall)	Sorghum indicator (recall)	Sorghum indicator (recall)	Sorghum indicator (recall)	Sorghum indicator (recall)
Expects any plot to be taken w/in 5 years, {1,0}	0.16***	0.15***	0.12**			
	(0.051)	(0.051)	(0.050)			
Official treatment status, {1,0} - Instrument				0.072***	0.068***	0.055**
				(0.019)	(0.019)	(0.022)
Is this plot irrigated? {1,0}		-0.048*	-0.056**		-0.059**	-0.067***
		(0.026)	(0.027)		(0.023)	(0.024)
Is the plot used only during Meher? {1,0}		0.055*	0.059**		0.046*	0.047**
		(0.028)	(0.028)		(0.024)	(0.023)
Is the plot somewhat sloping? {1,0}		-0.0027	0.0015		-0.0055	-0.0012
		(0.034)	(0.033)		(0.033)	(0.033)
Is the plot steep? {1,0}		0.0079	0.0058		0.033	0.026
		(0.092)	(0.087)		(0.087)	(0.086)
Plot size (ha)		0.37***	0.36***		0.36***	0.36***
		(0.12)	(0.11)		(0.12)	(0.11)
Plot size (ha)^2		-0.095	-0.086		-0.11	-0.096
		(0.10)	(0.092)		(0.10)	(0.093)
Is the plot sharecropped or rented in? {1,0}		-0.0021	0.012		0.0080	0.014
		(0.029)	(0.028)		(0.029)	(0.030)
Head age (years)			-0.00053			-0.00084
			(0.00087)			(0.00079)
Female head of household, {1,0}			0.074*			0.073*
			(0.040)			(0.042)
Head is separated, divorced or widowed, {1,0}			-0.054			-0.051
			(0.039)			(0.041)
Head can read and write simple sentence, {1,0}			-0.0029			-0.0063
			(0.022)			(0.023)
Household size			0.014**			0.016***
			(0.0064)			(0.0057)
Tropical livestock units			-0.0022			-0.0028
			(0.0052)			(0.0047)
Log of per capita consumption (in Birr)			0.043			0.038
			(0.030)			(0.031)
Household can access 4000 birr in emergency, {1,0}			-0.055**			-0.045*
			(0.024)			(0.026)
Number of months self-sufficient			-0.0086**			-0.0059*
			(0.0043)			(0.0035)
Constant	0.066*	-0.044	-0.31	0.13***	0.032	-0.23
	(0.040)	(0.062)	(0.22)	(0.014)	(0.035)	(0.25)
Observations	1,460	1,459	1,432	1,460	1,459	1,432
R-squared	-0.023	0.019	0.047	0.009	0.045	0.056
Instrument	IV	IV	IV	Reduced	Reduced	Reduced
Controls: Plot level	N	Y	Y	N	Y	Y
Controls: Household level	N	N	Y	N	N	Y
Kleibergen-Paap F-statistic	29.9	30.4	42.2			

Note: This table reports full regression results for the main empirical specification outlined in equation 1. Columns 1 to 3 report results for the IV regression using official treatment status as an instrument for subjective expectations. Columns 4 to 6 report reduced-form results using official treatment status in an OLS specification. Samples for each regression are restricted to plots that were cultivated in the main growing season. Clustered standard errors in parentheses are calculated at the sub-village level for 16 clusters. Statistical significance is denoted by the following system: ***p<0.01, **p<0.05, *p<0.1.

Table A.9: Plot level wheat indicator (Recall) - Full

	(1)	(2)	(3)	(4)	(5)	(6)
		IV Regression			Reduced Form Regression	
	Wheat indicator (recall)	Wheat indicator (recall)	Wheat indicator (recall)	Wheat indicator (recall)	Wheat indicator (recall)	Wheat indicator (recall)
Expects any plot to be taken w/in 5 years, {1,0}	-0.0090	-0.018	-0.016			
	(0.061)	(0.058)	(0.062)			
Official treatment status, {1,0} - Instrument				-0.0040	-0.0082	-0.0073
				(0.027)	(0.025)	(0.029)
Is this plot irrigated? {1,0}		0.019	0.020		0.020	0.021
		(0.033)	(0.036)		(0.032)	(0.034)
Is the plot used only during Meher? {1,0}		0.081***	0.076**		0.082***	0.077***
		(0.025)	(0.030)		(0.024)	(0.026)
Is the plot somewhat sloping? {1,0}		-0.025	-0.025		-0.025	-0.025
		(0.032)	(0.031)		(0.032)	(0.031)
Is the plot steep? {1,0}		-0.088	-0.082		-0.091	-0.085
		(0.074)	(0.078)		(0.071)	(0.075)
Plot size (ha)		0.31***	0.28***		0.31***	0.28***
		(0.092)	(0.091)		(0.093)	(0.091)
Plot size (ha)^2		-0.19***	-0.16**		-0.19***	-0.16**
		(0.067)	(0.065)		(0.066)	(0.064)
Is the plot sharecropped or rented in? {1,0}		0.010	0.024		0.0092	0.024
		(0.046)	(0.044)		(0.045)	(0.044)
Head age (years)			0.0014*			0.0014*
			(0.00083)			(0.00085)
Female head of household, {1,0}			0.072*			0.073
			(0.043)			(0.045)
Head is separated, divorced or widowed, {1,0}			-0.061			-0.062
			(0.044)			(0.046)
Head can read and write simple sentence, {1,0}			0.015			0.015
			(0.023)			(0.023)
Household size			-0.0018			-0.0021
			(0.0072)			(0.0068)
Tropical livestock units			0.0042			0.0043
			(0.0054)			(0.0055)
Log of per capita consumption (in Birr)			-0.036			-0.036
			(0.026)			(0.026)
Household can access 4000 birr in emergency, {1,0}			0.0065			0.0052
			(0.027)			(0.027)
Number of months self-sufficient			-0.0024			-0.0028
			(0.0020)			(0.0025)
Constant	0.20***	0.089*	0.33	0.20***	0.080***	0.32
	(0.044)	(0.051)	(0.23)	(0.020)	(0.029)	(0.22)
Observations	1,460	1,459	1,432	1,460	1,459	1,432
R-squared	-0.000	0.016	0.021	0.000	0.017	0.022
Instrument	IV	IV	IV	Reduced	Reduced	Reduced
Controls: Plot level	N	Y	Y	N	Y	Y
Controls: Household level	N	N	Y	N	N	Y
Kleibergen-Paap F-statistic	29.9	30.4	42.2			

Note: This table reports full regression results for the main empirical specification outlined in equation 1. Columns 1 to 3 report results for the IV regression using official treatment status as an instrument for subjective expectations. Columns 4 to 6 report reduced-form results using official treatment status in an OLS specification. Samples for each regression are restricted to plots that were cultivated in the main growing season. Clustered standard errors in parentheses are calculated at the sub-village level for 16 clusters. Statistical significance is denoted by the following system: ***p<0.01, **p<0.05, *p<0.1.

Table A.10: Plot level fertilizer use indicator (Recall) - Full

	(1)	(2)	(3)	(4)	(5)	(6)
		IV Regression			Reduced Form Regression	
	Fertilizer indicator (recall)	Fertilizer indicator (recall)	Fertilizer indicator (recall)	Fertilizer indicator (recall)	Fertilizer indicator (recall)	Fertilizer indicator (recall)
Expects any plot to be taken w/in 5 years, {1,0}	-0.18	-0.17	-0.17*			
	(0.12)	(0.10)	(0.100)			
Official treatment status, {1,0} - Instrument				-0.079*	-0.075*	-0.080*
				(0.047)	(0.042)	(0.042)
Is this plot irrigated? {1,0}		0.11**	0.070		0.12***	0.087
		(0.047)	(0.055)		(0.047)	(0.053)
Is the plot used only during Meher? {1,0}		0.022	-0.012		0.032	0.0048
		(0.057)	(0.064)		(0.052)	(0.058)
Is the plot somewhat sloping? {1,0}		-0.043	-0.027		-0.040	-0.024
		(0.072)	(0.067)		(0.067)	(0.063)
Is the plot steep? {1,0}		0.12	0.10		0.090	0.075
		(0.19)	(0.19)		(0.17)	(0.18)
Plot size (ha)		0.076	0.044		0.089	0.047
		(0.16)	(0.17)		(0.17)	(0.17)
Plot size (ha)^2		-0.032	-0.018		-0.021	-0.0034
		(0.11)	(0.11)		(0.11)	(0.11)
Is the plot sharecropped or rented in? {1,0}		-0.083*	-0.095**		-0.094**	-0.097***
		(0.043)	(0.040)		(0.039)	(0.036)
Head age (years)			0.0011			0.0015
			(0.0022)			(0.0021)
Female head of household, {1,0}			0.23*			0.23*
			(0.12)			(0.12)
Head is separated, divorced or widowed, {1,0}			-0.19*			-0.19*
			(0.10)			(0.10)
Head can read and write simple sentence, {1,0}			-0.019			-0.014
			(0.067)			(0.062)
Household size			0.027*			0.023
			(0.016)			(0.015)
Tropical livestock units			0.00030			0.0012
			(0.011)			(0.011)
Log of per capita consumption (in Birr)			0.047			0.054
			(0.058)			(0.060)
Household can access 4000 birr in emergency, {1,0}			0.059			0.045
			(0.058)			(0.058)
Number of months self-sufficient			0.017*			0.014
			(0.0095)			(0.0095)
Constant	0.72***	0.67***	-0.090	0.64***	0.59***	-0.20
	(0.077)	(0.069)	(0.52)	(0.032)	(0.048)	(0.54)
Observations	1,463	1,462	1,435	1,463	1,462	1,435
R-squared	-0.053	-0.038	-0.001	0.007	0.019	0.047
Instrument	IV	IV	IV	Reduced	Reduced	Reduced
Controls: Plot level	N	Y	Y	N	Y	Y
Controls: Household level	N	N	Y	N	N	Y
Kleibergen-Paap F-statistic	29.9	30.4	42.2			

Note: This table reports full regression results for the main empirical specification outlined in equation 1. Columns 1 to 3 report results for the IV regression using official treatment status as an instrument for subjective expectations. Columns 4 to 6 report reduced-form results using official treatment status in an OLS specification. Samples for each regression are restricted to plots that were cultivated in the main growing season. Clustered standard errors in parentheses are calculated at the sub-village level for 16 clusters. Statistical significance is denoted by the following system: ***p<0.01, **p<0.05, *p<0.1.

Table A.11: Plot level organic fertilizer use (Recall) - Full

	(1)	(2)	(3)	(4)	(5)	(6)
		IV Regression			Reduced Form Regression	
	Organic fertilizer indicator (recall)	Organic fertilizer indicator (recall)	Organic fertilizer indicator (recall)	Organic fertilizer indicator (recall)	Organic fertilizer indicator (recall)	Organic fertilizer indicator (recall)
Expects any plot to be taken w/in 5 years, {1,0}	-0.070	-0.060	-0.059			
	(0.13)	(0.12)	(0.11)			
Official treatment status, {1,0} - Instrument				-0.031	-0.026	-0.027
				(0.058)	(0.054)	(0.051)
Is this plot irrigated? {1,0}		0.13**	0.099		0.13**	0.10
		(0.051)	(0.062)		(0.054)	(0.064)
Is the plot used only during Meher? {1,0}		0.013	-0.0019		0.017	0.0039
		(0.056)	(0.065)		(0.055)	(0.064)
Is the plot somewhat sloping? {1,0}		0.029	0.044		0.030	0.046
		(0.068)	(0.064)		(0.068)	(0.064)
Is the plot steep? {1,0}		0.21	0.19		0.20	0.18
		(0.20)	(0.20)		(0.20)	(0.21)
Plot size (ha)		0.0060	-0.026		0.010	-0.026
		(0.18)	(0.19)		(0.19)	(0.19)
Plot size (ha)^2		-0.015	0.0052		-0.011	0.010
		(0.096)	(0.10)		(0.096)	(0.10)
Is the plot sharecropped or rented in? {1,0}		-0.17***	-0.17***		-0.17***	-0.17***
		(0.041)	(0.038)		(0.041)	(0.038)
Head age (years)			0.0019			0.0021
			(0.0018)			(0.0018)
Female head of household, {1,0}			0.18*			0.18*
			(0.096)			(0.098)
Head is separated, divorced or widowed, {1,0}			-0.12			-0.12
			(0.11)			(0.11)
Head can read and write simple sentence, {1,0}			0.022			0.024
			(0.065)			(0.065)
Household size			0.019			0.018
			(0.020)			(0.021)
Tropical livestock units			0.00049			0.00078
			(0.011)			(0.011)
Log of per capita consumption (in Birr)			0.052			0.054
			(0.066)			(0.066)
Household can access 4000 birr in emergency, {1,0}			0.015			0.011
			(0.049)			(0.048)
Number of months self-sufficient			0.016***			0.014**
			(0.0057)			(0.0067)
Constant	0.46***	0.43***	-0.36	0.43***	0.41***	-0.40
	(0.063)	(0.064)	(0.59)	(0.019)	(0.055)	(0.58)
Observations	1,460	1,459	1,432	1,460	1,459	1,432
R-squared	-0.008	0.016	0.036	0.001	0.024	0.043
Instrument	IV	IV	IV	Reduced	Reduced	Reduced
Controls: Plot level	N	Y	Y	N	Y	Y
Controls: Household level	N	N	Y	N	N	Y
Kleibergen-Paap F-statistic	29.9	30.4	42.2			

Note: This table reports full regression results for the main empirical specification outlined in equation 1. Columns 1 to 3 report results for the IV regression using official treatment status as an instrument for subjective expectations. Columns 4 to 6 report reduced-form results using official treatment status in an OLS specification. Samples for each regression are restricted to plots that were cultivated in the main growing season. Clustered standard errors in parentheses are calculated at the sub-village level for 16 clusters. Statistical significance is denoted by the following system: ***p<0.01, **p<0.05, *p<0.1.

Table A.12: Robustness checks using spatial data

	(1)	(2)	(3)	(4)	(5)	(6)
	Main specification	500 metre band around boundary	750 metre band around boundary	1000 metre band around boundary	Control latitude longitude	Control quadratic expansion of coordinates
Sorghum plot indicator (Baseline), {1,0}	0.10**	0.11	0.0069	0.017	0.084**	-0.0078
	(0.045)	(0.12)	(0.049)	(0.049)	(0.037)	(0.067)
Wheat plot indicator (Baseline), {1,0}	-0.11**	-0.057	-0.049	-0.053	-0.098**	0.039
	(0.047)	(0.055)	(0.049)	(0.042)	(0.040)	(0.053)
Any fertilizer indicator (Baseline), {1,0}	-0.16*	-0.15**	-0.13	-0.070	-0.15**	-0.0065
	(0.088)	(0.065)	(0.090)	(0.071)	(0.067)	(0.096)
Maximum number of observations:	1479	832	965	1026	1479	1479
Method:	IV	IV	IV	IV	IV	IV
Controls: Plot level	Y	Y	Y	Y	Y	Y
Controls: Household level	Y	Y	Y	Y	Y	Y
Kleibergen-Paap F-statistic		43.4	64.4	67.5	99.2	107

Note: This table shows the coefficient on treatment indicators for various model specifications. Columns corresponds to a different regression specification and rows to a different outcome variable. Column 1 shows the coefficient of interest from the main regression specification; this is a households expectation of expropriation instrumented by the official's assessment of whether the household will lose land. Columns 2, 3 and 4 run this regression specification on a limited sub-sample of households that live within a specified distance of the project boundary. Column 5 additionally controls for the coordinates of the household's homestead. Column 6 controls for location using a second order polynomial expansion. Clustered standard errors in parentheses are calculated at the sub-village level for 16 clusters in the main regression, but fewer in the regressions that limit the sub-sample. Statistical significance is denoted by the following system: ***p<0.01, **p<0.05, *p<0.1. Plot controls include: slope, irrigation status, plot size and plot size squared, an indicator for whether the plot is used in the main growing season only and an indicator for whether the plot is sharecropped in. Household controls include: head age, female head, marital status, head literacy, household size, livestock, log of per capita consumption, credit access and the number of months households eat mainly from home produced crops.

Table A.13: Main specification run on restricted sub-samples

	(1)	(2)	(3)	(4)
	Main specification	Only plots owned by the household	Only plots acquired prior to announcement	Only plots used for cereal production
Sorghum plot indicator (Baseline), {1,0}	0.10**	0.13**	0.100**	0.085**
	(0.045)	(0.053)	(0.048)	(0.040)
Wheat plot indicator (Baseline), {1,0}	-0.11**	-0.11***	-0.10**	-0.14***
	(0.047)	(0.040)	(0.041)	(0.053)
Any fertilizer indicator (Baseline), {1,0}	-0.16*	-0.18*	-0.12	-0.17
	(0.088)	(0.095)	(0.093)	(0.11)
Maximum number of observations:	1,479	1,278	1,363	1,182
Method:	IV	IV	IV	IV
Controls: Plot level	Y	Y	Y	Y
Controls: Household level	Y	Y	Y	Y
Kleibergen-Paap F-statistic	52.1	49.7	58.9	46.7

Note: This table shows the coefficient on treatment indicators for various model specifications. Columns corresponds to a different regression specification and rows to a different outcome variable. Column 1 shows the coefficient of interest from the main regression specification; this is a households expectation of expropriation instrumented by the official's assessment of whether the household will lose land. Column 2 limits the sample to plots that are owned by the household. Column 3 limits the sample to plots that were acquired 3 or more years before the baseline survey was conducted, implying that these plots were acquired prior to the announcement of expropriation. Column 4 limits the sample to those plots that are used for growing cereals only. Clustered standard errors in parentheses are calculated at the sub-village level for 16 clusters in the main regression, but fewer in the regressions that limit the sub-sample. Statistical significance is denoted by the following system: ***p<0.01, **p<0.05, *p<0.1. Plot controls include: slope, irrigation status, plot size and plot size squared, an indicator for whether the plot is used in the main growing season only and an indicator for whether the plot is sharecropped in. Household controls include: head age, female head, marital status, head literacy, household size, livestock, log of per capita consumption, credit access and the number of months households eat mainly from home produced crops.

Table A.14: Spatial fixed effects estimates

	(1)	(2)	(3)	(4)
	Main specification	Spatial fixed effects specification	Spatial fixed effects specification	Spatial fixed effects specification
Neighbourhood size (radius):		500m	750m	1000m
Sorghum plot indicator (Baseline), {1,0}	0.10**	-0.015	0.012	0.031
	(0.045)	(0.14)	(0.077)	(0.071)
Wheat plot indicator (Baseline), {1,0}	-0.11**	0.092**	0.031	0.0037
	(0.047)	(0.038)	(0.053)	(0.058)
Any fertilizer indicator (Baseline), {1,0}	-0.16*	0.021	0.043	-0.018
	(0.088)	(0.049)	(0.049)	(0.068)
Maximum number of observations:	1479	1479	1479	1479
Average number of neighbourhood:	-	15	29	47
Method:	IV	IV	IV	IV
Controls: Plot level	Y	Y	Y	Y
Controls: Household level	Y	Y	Y	Y
Kleibergen-Paap F-statistic	52.1	154	57.6	70

Note: This table shows the coefficient on treatment indicators for various model specifications. Columns corresponds to a different regression specification and rows to a different outcome variable. Column 1 shows the coefficient of interest from the main regression specification; this is a households expectation of expropriation instrumented by the official's assessment of whether the household will lose land. Column 2 limits the sample to plots that are owned by the household. Column 3 limits the sample to plots that were acquired 3 or more years before the baseline survey was conducted, implying that these plots were acquired prior to the announcement of expropriation. Column 4 limits the sample to those plots that are used for growing cereals only. Clustered standard errors in parentheses are calculated at the sub-village level for 16 clusters in the main regression, but fewer in the regressions that limit the sub-sample. Statistical significance is denoted by the following system: ***p<0.01, **p<0.05, *p<0.1. Plot controls include: slope, irrigation status, plot size and plot size squared, an indicator for whether the plot is used in the main growing season only and an indicator for whether the plot is sharecropped in. Household controls include: head age, female head, marital status, head literacy, household size, livestock, log of per capita consumption, credit access and the number of months households eat mainly from home produced crops.

B. Spatial fixed effects

Define $R(h)$ as the set of plots belonging to the $k \neq h$ households that reside within a given distance of household h and $n_{R(h)}$ as the number of these plots.[25] I estimate the following transformed version of the baseline equations 1 and 2:

$$\begin{aligned}\left(y_{ih} - \frac{1}{n_{R(h)}} \sum_{j,k \in R(h)} y_{jk}\right) &= \left(x'_{ih} - \frac{1}{n_{R(h)}} \sum_{j,k \in R(h)} x'_{jk}\right) \beta \\ &+ \left(expect_h - \frac{1}{n_{R(h)}} \sum_{jk \in R(h)} expect_k\right) \gamma + \left(\nu_{ih} - \frac{1}{n_{R(h)}} \sum_{j,k \in R(h)} \nu_{jk}\right) \quad \text{(B.1)}\end{aligned}$$

$$\begin{aligned}\left(expect_h - \frac{1}{n_{R(h)}} \sum_{j,k \in R(h)} expect_k\right) &= \left(x'_{ih} - \frac{1}{n_{R(h)}} \sum_{j,k \in R(h)} x'_{jk}\right) \pi \\ &+ \left(T_h - \frac{1}{n_{R(h)}} \sum_{j,k \in R(h)} T_k\right) \phi + \left(\eta_{ih} - \frac{1}{n_{R(h)}} \sum_{j,k \in R(h)} \eta_{jk}\right) \quad \text{(B.2)}\end{aligned}$$

There are two related points to make about interpreting the spatial fixed effects model. First, estimated coefficients are identified only where there is variation within neighbourhoods for a given variable. Since treated households are clustered (by definition) in and around the project area, the treatment effect is identified using information only from a subset of households living along the project boundary[26]. In contrast, most of the control variables vary across space and so their estimated coefficients will be identified using information from all households. Second, but related to this, is the importance of the size of the neighbourhood as defined by the critical distance from household h. The neighbourhood size

[25] This complicated structure arises because I only have geographic data on the location of households while my outcomes of interest are measured at the plot-level.

[26] This is only true for treated households that live near enough to control households and vice versa. Examination of the survey map shows that there are some treatment households living near the project boundary who do not live near control households; information from these households would not be used to identify the treatment effect.

affects both how many households are used to identify the treatment effect and the model's ability to control for localized unobservable heterogeneity. A model with very small neighbourhood sizes can control for very localized unobservable heterogeneity, but will estimate the treatment effect imprecisely.[27] A model with very large neighbourhoods only controls for spatially broad fixed effects, but estimates the treatment effect with more precision.[28]

Ideally, this model would be estimated using information on the location of plots rather than using household location as a proxy. This would control for spatially concentrated differences in unobservable plot characteristics, while allowing for variation across households. Spatial fixed effects using household location controls for unobservable neighbourhood characteristics, including household unobservables such as information that is common to the neighbourhood social network or shared beliefs about expropriation. If these are important determinants of crop choice or fertilizer use, the true effect of anticipating expropriation would be absorbed by the neighbourhood fixed effects. Using household location for spatial fixed effects estimates may actually exacerbate the bias introduced by household unobservables. The ideal model would allow for a comparison between plots on either side of the border (which is effectively what spatial fixed effects does), since characteristics of the owners of plots on either side of the boundary would be plausibly exogenous. However, it is possible that comparing a treated households living in a neighbourhood full of controls is actually a poor comparison. A treated household located amongst control households is likely to own more land and to own land further away than their neighbours. These may be the households, within the treatment group, that are least likely to hedge their crop choice portfolio or fertilizer use in anticipation of losing land.

[27]This is analogous to the precision/consistency tradeoff in bandwidth selection for estimating empirical distribution functions.

[28]The extreme case is the model where the neighbourhood includes all households in the survey. This model is not different from the baseline model, except that the constant term is zero.

Chapter 2

Expropriation, compensation and the struggle to find a new livelihood:

Evidence from an expropriation program in Ethiopia

Abstract

Government intervention in land transactions is common in developing countries, especially where land markets function poorly. This is very much the case in Ethiopia, where expropriation of farmland from small-scale farmers has been used by all levels of government as a tool for providing new land for industrial investors, commercial agriculture and expanding cities. Most often, farmers receive a cash payment in exchange for their land, but without land markets, governments rely on a fixed formula to establish the price of land. This paper evaluates the impact of such a policy on a group of small-scale farmers whose land has been taken to make room for a large factory. Baseline data was collected in the year before expropriation and a follow up survey was conducted 8 months after households lost their land and received payment. On average, household lose 70% of their land and receive compensation payments that are about 5 times the value of annual consumption expenditure.

Using a first difference specification, I examine the impact of expropriation and compensation on household consumption, productive assets, livestock holdings, savings and labour market participation. I find that households in the treatment group increase their consumption, start more businesses and participate more in non-farm activities than households that do not lose farmland. These households also reallocate their livestock portfolios away from oxen and towards small ruminants and cattle, reflecting a shift away from growing crops. However, all of these changes are relatively minor compared to the increase in savings: with the exception of a few households, most of the compensation payment is left in the bank.

2.1 Introduction

Government intervention in land transactions is common in developing countries, especially where land markets function poorly and land is held in small, fragmented plots. Expropriation and compulsory purchase of farmland from small-scale farmers are important economic policy tools that enable governments to consolidate agricultural land and provide space for industry and urban expansion.[1] China, for example, has relied heavily on expropriation to meet the demand for land for industry, infrastructure and urban expansion (Ding, 2007). These policies are not without controversy. Across Africa, Asia, and Latin America, land for large-scale commercial farms is often acquired through expropriation (Cotula et al., 2009) and this has been a cause for concern amongst many NGOs and advocacy groups[2]. In West Bengal, the compulsory acquisition of farmland for a Tata factory lead to political unrest and violence (Banerjee et al., 2007; Ghatak et al., 2012). Without major changes to land markets, expropriation will continue to play an important part in the process of industrialization and development.

Expropriating farmland deprives rural small-holders of one their most important income-generating assets and forces them to find new livelihoods. Governments recognize this, and often provide households with compensation, which in some cases takes the form of a lump-sum payment. But are these lump-sum payments sufficient to compensate households for the land that is taken? A natural way to assess this is to determine whether a household's permanent wealth has changed as a result of the intervention. Evaluating this requires some understanding of how households adjust their asset and activity portfolio after losing land.

[1] Work by Adamopoulos and Restuccia (2014) highlights the effect of inefficient land market allocation on cross-country productivity differences.

[2] Oxfam recently launched a campaign calling for a moratorium on land grabs in part because of the consequences for households displaced by investors (Geary, 2012).

It also requires an understanding of what households do with large lump-sum payments; households may use lump-sum payments for productive investments such as livestock or a new business or may look for paid wage employment, choosing to save or consume their payment instead. Changes to permanent wealth may also be reflected in household consumption, with increasing consumption an indication of increased wealth. However, previous liquidity constraints and behavioural biases may render changes in consumption a poor measure of changes in permanent wealth.

Beyond the analysis of the Singur case in West Bengal in Ghatak et al. (2012) and Banerjee et al. (2007), very little is known about what happens to households that lose their land or the way in which they find new income generating activities. In this paper I explore these questions in detail using an impact evaluation of an expropriation and compensation program that occurred in Ethiopia. I evaluate the extent to which households use financial compensation payments for productive investments, the change in their income-generating activities and the change in levels of consumption following expropriation. I focus specifically on evaluating the effect of the program on a number of key household outcomes, including consumption, asset ownership, livestock holdings, savings and labour market participation.

I use data from a survey of rural households in Ethiopia that had their land expropriated to build a factory. Expropriation of land from small-scale farmers is commonly used by the Ethiopian government to provide land for rapidly growing cities and industrial investment projects. Surveyed households received a lump-sum compensation payment for their land, which was intended to help them transition to new income generating activities. On average, these households lost 70% of their farmland and received compensation payments worth 5,200 USD (the equivalent of 4.7 times total annual consumption expenditure and 9 times the value of livestock). A baseline survey was conducted one year prior to expropriation, but following the announcement of the project, and a follow-up survey was conducted eight months afterwards. A group of households that did not lose land were selected from within the same administrative area to serve as a comparison group. The timing of the data collection and the inclusion of a comparison group permit me to evaluate the way in which the program affects households' asset and activity portfolios, their consumption and their ability to cope with the challenge of finding new income generating activities. I find that, on average, the households that lose their land increase their consumption, start more businesses and increase their livestock assets. However, with the exception of a few households that make large lumpy investments, the increase in investment in productive assets is dwarfed by massive increases in cash savings. Households also shift assets away from agriculture uses and spend more time in non-farm work, although the change is not large.

The paper proceeds with a brief review of the literature on expropriation and compensation, and the literature on how rural households use cash transfers. Section 2.2 discusses the policy context and the intervention in more detail. Section 2.3 discusses the data, empirical strategy and identifying assumptions required to estimate the treatment effect. Finally, section 2.4 presents results, robustness checks and a discussion.

2.1.1 Related literature

Land expropriation and cash payments are inextricably linked components of the policy intervention. There is a large literature on both expropriation and cash payments individually, but very little that studies their effect together. The exception to this is Ghatak et al. (2012), who study the aftermath of a land expropriation in India, where the government in Singur, West Bengal acquired farmland from households to build a car factory. In this setting, compensated households experienced slower income growth in the 5 years following expropriation and a majority of households reported spending their payment on house improvements and savings in the bank, rather than investing in new businesses. Households that relied heavily on agriculture for their livelihoods - owner-cultivators and agricultural wage workers - were adversely affected by the policy and to a greater extent than households headed by non-agricultural wage workers.

The large body of research on land expropriation tends to focus on the effect of expropriation risk on ex ante investment incentives and on the impact of improving tenure security in countries where the risk of expropriation is high, rather than on evaluating what happens to households that lose their land. Typically, households that anticipate expropriation adjust their land-based investment, avoiding investments that are immovable or where the value cannot be verified (Ali et al., 2011; Deininger and Jin, 2006; Fenske, 2011; Jacoby et al., 2002). In some cases, households respond in the opposite way, investing in immovable assets precisely to reduce the risk of expropriation (Besley, 1995). Ghatak and Mookherjee (2013) explicitly model land expropriation in a setting where farmland can be expropriated by a landlord or local government and sold on to an industrialist, showing that the size of compensation payment, even if it is not linked to land quality or verifiable investments, affects ex ante investment incentives for farmers.[3] In contrast, this paper illustrates the ex post consequences of expropriation and gives some suggestive evidence about the adequacy of the compensation in the Ethiopian setting.

[3]Ghatak and Mookherjee (2013) argue that households should be over-compensated, in the sense that they should receive a payment that is greater than what the household receives from farming. They argue that households should be given a greater share of the economic surplus that is generated by transforming agricultural land into industry. In a setting like Ethiopia, where there is no established price for farmland, one might think of this value as the expected net change in household income resulting from expropriation.

Cash compensation payments are often provided to help households that lose land make the transition from farming to non-farm livelihoods; the adequacy of compensation depends in part on what households can do with cash transfers. There is a large literature that evaluates how households and small enterprises deal with cash transfers, which finds that, amongst many things, cash payments can assist unemployed youth in transitioning toward trades and formal businesses (Blattman et al., 2013), increase returns for owners of micro-enterprises (de Mel et al., 2008; Fafchamps et al., 2014) and increase consumption, durable assets and investment (Gertler et al., 2012; Haushofer and Shapiro, 2013). These papers fit into a broader context that suggests that rates of return on investment in developing countries are very high (Banerjee et al., 2005). However, for all the benefits that accrue to households from cash transfers, many of these papers also find heterogeneity in the effect of cash payments on investment and household businesses.[4] These papers generally suggest that cash payments can be effectively used by households in developing countries, but the parallels with the intervention studied here should not be taken too far. First, in the literature on cash transfers, the cash transfer is typically the only policy intervention, whereas in my context households lose their land as well. Second, the magnitude of the transfers in these papers is much smaller. Out of the papers discussed, Haushofer and Shapiro (2013) analyzed the effects of the largest cash payment, which was 1,300 USD. This amount is lower than the payment received by 85% of the treated households in my sample. Although cash transfers can be effectively absorbed at a small scale, there is no reason to think that households receiving very large cash payments will be able to use the money effectively.

[4]For example, Fafchamps et al. (2014) and de Mel et al. (2008) find that cash transfers do not yield positive returns for female micro-entrepreneurs. de Mel et al. (2008) also find that returns vary widely with the owner's entrepreneurial ability.

2.2 Policy context, timeline and description of the intervention

2.2.1 Land ownership, expropriation and compensation in Ethiopia

The Federal government owns all rural land in Ethiopia and can expropriate land from anyone if it is deemed in the national interest. Small farmers have user rights over their land, which entitle them to lease out land for short periods of time and to give land to their children. However, individuals do not have the right to sell land, which means that land markets in rural areas are very thin (Ali et al., 2011; Deininger and Jin, 2006). Expropriation of land is therefore the most common way that land is transferred from small-scale farmers to industrial uses. Federal laws on expropriation entitle households to compensation if their land is expropriated. Households can be compensated with equivalent land in another part of the area where they live, or, if there is no land, financial compensation. Officially, households should be paid 10 times the market value of what can be produced on their land. In practice, this figure is impossible to calculate for an individual plot so administrators calculate a common price per square metre for the whole project area instead. For the project evaluated in this paper the compensation price reflects whether the plot is used for grazing, growing rain-fed crops for one season only (Meher land), or growing crops for two seasons (irrigated land). Households are also compensated for any trees or improvements that are on the expropriated plots and for their house. Households that lose their homestead are provided with a plot of land and money to rebuild their house.

2.2.2 Survey area, investment project and timeline

The survey was conducted in a rural administrative area next to the town of Kombolcha, Ethiopia. Kombolcha is a town of 60,000 people that has recently been designated as a target for industrial investment by the Federal Government of Ethiopia. Kombolcha is known for being a major transport hub as it lies on one of the roads the goes from Addis Ababa to Djibouti, and for some existing industry, including a textile factory, a brewery and a steel plant. In recent years, rural land around the town has been expropriated to make space for a new airport and for large industrial projects. The area of land where the survey was conducted was allocated to an Indian textile firm in 2009 as part of an initial investment agreement, however the land was never transferred to these investors.[5] In early 2012, the land concession was transferred to a major Ethiopian company that plans to build a steel factory and in January 2013 land was expropriated. The survey area itself is close to the town of Kombolcha and is well connected by a main road that runs through the area. The closest parts of the survey area border the town, whereas the farthest parts are approximately two hours from the town centre by foot. Given the proximity, most households are well connected to economic activity in the town: most households purchase and sell their produce at the main market, many individuals look for casual labour in town and high-school students attend school in Kombolcha.

The area of land for the investment project covers approximately 340 hectares out of a total land area of 1100 hectares from the survey area and expropriation affected 626 out of 1100 households living in the area. The land was chosen primarily because it is a large contiguous piece of flat land that is well connected to the road. A map of the survey area in figure 1 shows the project area outlined

[5] Further details about the baseline survey can be found in a companion paper by Harris (2014) (other chapter) that studies the effects of anticipating expropriation on agricultural investment decisions.

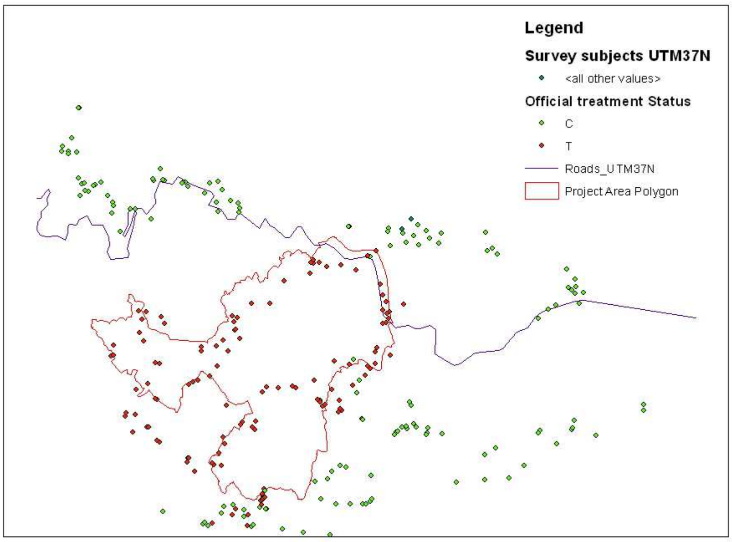


Figure 1: Map of survey area showing HH location - treatment (red), control (green) at baseline

in red with the road to Kombolcha marked in blue. In addition to the project area, this map shows the location of households, with treated households indicated with red markers and control households indicated with green. Many of the households that lose land do not live inside the project boundary, but own at least one plot inside the project boundary. The western edge of Kombolcha town borders the right hand side edge of the map.

Table 1 on the next page summarizes the available administrative data on expropriation and compensation, and breaks down the total household compensation payment for farmland by category.[6] In total, 626 households received compensation for two projects that were happening in the survey area.[7] Households lost 288.3 ha of Meher[8] farmland, 21.6 ha of irrigation land and 25.5

[6]The administrative data does not include information on compensation for lost homesteads.

[7]Land was taken from the centre of the survey area for a steel factory. This area is indicated in red on the map. Land was also taken on the edge of town to make room for a new landfill and to expand urban housing. This occurred at the same time as the expropriation for the factory and so I treat them as the same intervention.

[8]Meher farmland is rain-fed agricultural land used only during the main growing season.

ha of grazing land for which they were paid 14.7 Birr ($0.80), 30.9 Birr ($1.70) and 5.1 Birr ($0.28) per square meter, respectively. Land prices were calculated using 5-year averages of local agricultural yields and market prices for a number of key crops grown on each type of land. Prior to expropriation, local officials conducted a cadastral survey to establish how much land each household owned and to classify each plot as Meher land, irrigated land or grazing land. Basing payment on the land type means that total compensation payments will not correspond one-to-one to the quantity of land a household loses since the type of land affects the unit price of land. Furthermore, households were compensated for any trees on the plot, so payments may vary across plots of the same type. Despite this variation, the quantity of Meher land lost is the most important determinant of compensation payments. Table 1 shows that 96% of treated households received some compensation for Meher land and that, on average, this accounts for 84% of household's compensation.

Table 1: Breakdown of administrative compensation payment data by category

		Mean payment (Birr)	Mean share of total payment	Share of HH receiving any payment
Category of land:	Meher (one season)	68,010	84%	96%
	Irrigation	10,695	8%	19%
	Grazing	2,086	3%	33%
	Trees*	7,275	6%	44%
	Total payment	86,375	-	-

		Mean area taken (ha)	Total area taken (ha)
Category of land:	Meher (one season)	0.46	288.3
	Irrigation	0.035	21.6
	Grazing	0.041	25.5

Note: Data on compensation paid for trees is only available for one of the two projects, which comprises 418 out of 626 affected households. At the time of data collection 1 USD was equal to approximately 19.2 Ethiopian Birr.

The timing of the data collection relative to the timing of the process of expropriation is a critical part of the identification strategy used in this paper. Although baseline data was collected prior to expropriation, households in the survey area were aware that their land had been identified by a potential investor. Investors identified the land in the project area in early 2009 and in June 2009 the government conducted a cadastral survey of the project area to calculate compensation payments for expropriated households. The city government told households that their land would be taken following the main harvest in November, but this did not occur and in the subsequent years the investment timeline was revised and expropriation delayed. In early 2011 the government once again announced that land would be taken in December and once again expropriation was postponed. During September 2011 we conducted the baseline survey. Following further delays, the government transferred the land concession to another investor in January 2012. The land was ultimately expropriated following the main harvest in December 2012 and the follow up survey was conducted in October 2013.

2.2.3 Description of the intervention

Households in the treatment group are formally defined as those that report having had any of their land expropriated between rounds.[9] But what does it mean for households to be in the treatment group? Fundamentally, the intervention has two related components: 1) households lose their farmland and 2) they receive a cash payment that is a function of how much land was taken. Land holdings are fractured and distributed widely across the survey area, so there are many households that lose some, but not all of their land. In addition there is a high degree of uncertainty about how the policy is enacted: households were unsure of when and whether their land would be taken and often have little knowledge of the process of calculating compensation.

[9] Specifically, households were asked: "Did you lose any land as part of the expropriations in this kebele in the last two years?" Although this is a subjective measure of expropriation, it accords fairly well with official records. Appendix A validates these measures using official government records.

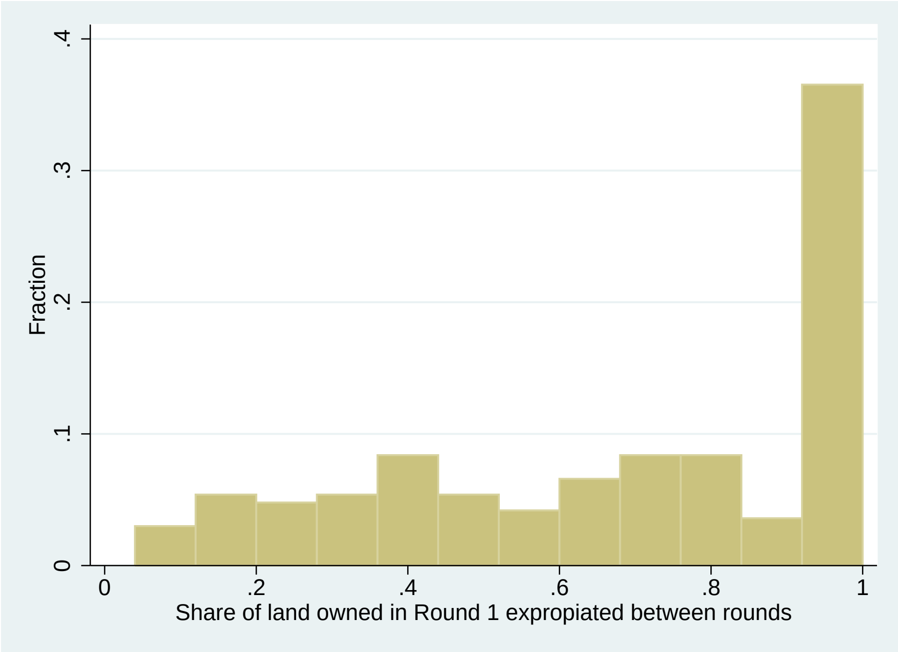


Figure 2: Distribution of share of land expropriated from treated households

Figure 2 plots the distribution of land lost by the treated group calculated as a the share of land owned in round 1 that was expropriated by round 2.[10] Treated households lost, on average, 68% (median 75%) of their initial land holdings and approximately 35% of households in the treated group lost all of their land between rounds. The distribution of compensation payments also varies greatly across the sample of treated households. Figure 3 shows the distribution of reported compensation payments in Birr and indicates the level of median total annual consumption expenditure as a benchmark. The median reported compensation payment was 85,500 Birr (4500 USD) and the average payment was approximately 100,000 Birr (5,200 USD). These payments are very large; on average, compensation payments are 4.7 times higher than total annual household consumption expenditure and for some households as much as 10 times higher. Compared to household assets, compensation payments are also very large: the median household received a payment 8.8 times the value of their livestock and 40

[10]Information about the treatment status of a particular plot comes from recall data that was collected about every plot that households listed during the baseline survey. Details about this data can be found in a companion paper by Harris (2014) (other chapter).

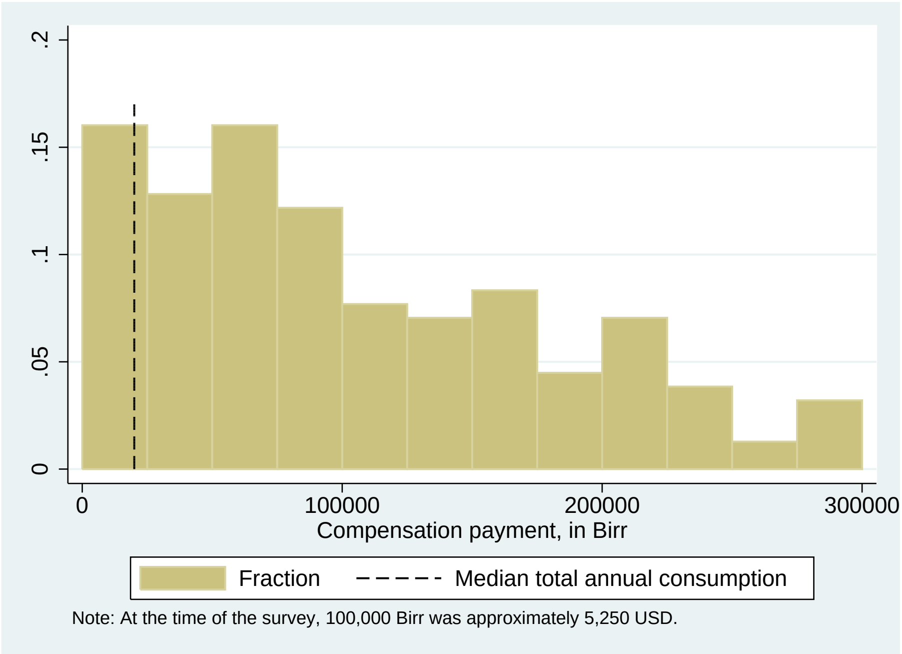


Figure 3: Distribution of compensation payments

times the value of their non-productive assets. This policy intervention represents a major change in life circumstances for treated households, although the portion of land lost and the size of compensation varies across households.

There is uncertainty along a number of dimensions with this policy, both with respect to how the policy was implemented and whether a household would be affected. Half of the households in the treatment group lodged complaints about the measurement of their land and/or the total payment they received, suggesting that they expected different payments. Households also reported receiving their payments at very different times (this may be a function of recall). On average, households received their payment 2 months after they lost their land, with about 10% of treated households facing a delay of more than 5 months. Households have had their compensation payments on average for 8 months, with months since payment ranging from 0 to 16 months. The land was only taken 9 months before the survey, so reports of longer durations may relate to other expropriation

Table 2: Baseline expectations of treatment status compared to realized status

			Share of HHs that actually lost land:
(1)	Do you expect to lose land within 1 year (baseline)?	No (n=169)	37%
		Yes (n=127)	80%
(2)	Local official's assessment of HH treatment status (baseline):	Control (n=164)	34%
		Treatment (n=130)	82%

programs that happened in the area. In rare cases there are households that are still waiting for payment.

Data on expectations of expropriation collected during the baseline survey are sometimes at odds with a household's realized treatment status, suggesting that households faced a high degree of uncertainty about whether they would actually lose land. This is perhaps due to a long history of delays in the implementation of the land expropriation program, and also a lack of communication and information between officials and households.[11] Panel 1 of table 2 shows the discrepancy between household's expectations and realized outcomes; the shares are equivalent to 70% of households in the sample being correct in their assessment of their treatment status.[12] Local officials also had incomplete information about which households would lose their land. Panel 2 of table 2 shows the treatment status of a household reported by the local officials in round 1 compared to the realized outcome: in this case, 72% of households were correctly identified as treatment or control.[13]

The nature of the policy intervention differs greatly depending on how much land a household loses. For some households the policy intervention is very disruptive:

[11] The fact that a number of households were ultimately incorrect about their treatment status could be a very interesting source of variation that would allow one to separate the effect of expected and actual expropriation. Unfortunately, sample sizes are too small to exploit this variation.

[12] Those households that predicted that they would not lose land within 1 year were ultimately proved right as the land expropriation was delayed by one year following the baseline survey.

[13] The final size of the project was reduced from the initial proposed size between rounds, which may explain why households thought they would lose land, but found themselves in the control group. Some households also live near the town of Kombolcha, where land was taken to expand the city at the same time as land was taken for the factory; this ambiguity may explain some of the confusion behind the answers given by local officials and households.

they lose a major income-generating asset and are given a massive cash payment instead. Land is the most important asset for many households in this area; it provides a steady source of food for the household and can be sharecropped out by older households once they are too old to farm the land themselves. As compensation, households receive cash payments that are probably larger than any amount of money that households have held at one time. For households that lose less land and receive smaller payments, the intervention is much less disruptive and may have very different effects.

2.3 Data description, empirical strategy and identification

The baseline survey was conducted in September 2011. We planned to survey 300 households drawn from 19 sub-villages in one Kebele (the smallest administrative unit in Ethiopia) in which the expropriation was to occur. The sample was restricted to households living in the Kebele at the time of the survey. Together with local officials we identified treated and control households from the Kebele administrator's household list. 15 households were randomly sampled from each of 16 sub-villages where the treatment status was common across the village. In two villages which contained a mix of treatment and control households we randomly sampled 30 households. One village of the 19 was excluded from the sample because their land had been taken in previous years. During the analysis of the baseline data it became clear that the official's assessment of the household's treatment status did not coincide with the household's own assessment. At the end-line this discrepancy was confirmed: the treatment status was correctly identified by officials for 72% of the sample. There are three main reasons for this: 1) the size of the factory project area was reduced between rounds, 2) an additional area of land was expropriated to expand the town and 3) any random errors on the part of local officials who did not know exactly which households would lose land.

The end-line survey was conducted in October 2013 and every effort was made to follow households that had moved away and to interview the same respondent as in round 1. Where the household head had died, we interviewed the new household head, if that individual had been part of the household in the previous round. If the household had divided, we interviewed those people remaining in the main home. Out of 299 households interviewed during the baseline survey, 1 household refused to be interviewed and 2 respondents were too old to complete the survey on their own, so others completed partial information on their behalf. Most households still lived in the same place as they had in round 1, except for a small number of households had moved to the nearby towns of Komobolcha (approximately 3km from the centre of the Kebele) or Dessie (approximately 17km from the centre of the Kebele) and one household had moved to the far west of Ethiopia. Out of 298 surveys, 82% were conducted with the same respondent as in round 1 and 95% reported having the same household head, thus attrition is not a major problem.

2.3.1 Conceptual framework

How might expropriation and compensation affect household behaviour? In this analysis, I think of households as having a portfolio of activities chosen with a view to mitigate risk and accumulate assets for future consumption or investment. Income generating strategies employed by the household depend on inherent characteristics of the household, such as business skills or education, household endowments and on the asset portfolio owned by the household. The policy intervention disrupts this balance, essentially forcing households to liquidate their land assets, which had previously been impossible to do. In response to this change, households rebalance their asset portfolios, which now contain a large cash component and choose a new portfolio of income generating activities to replace the flow of agricultural income. As before, households adjust their portfolio to generate income, mitigate risk and accumulate assets. With less land and with large cash reserves, one would expect households to adjust asset portfolios away from agricultural production and towards investment in new businesses, livestock and other assets that yield a return that exceeds what is paid in interest by the bank. One would also expect households to supply labour that is no longer used for agricultural to local labour markets or household businesses.

The degree to which households can react in these ways to expropriation may be complicated by a number of factors including: 1) the slow speed of adjustment to new asset or activity portfolios, and 2) exogenous constraints such as physical limits to livestock herd size or rationing of low-skilled jobs that may limit employment opportunities. The combined adjustment of asset and activity portfolios contribute to a new permanent income for household's which, in turn, affects their consumption decisions. In this sense, changes in consumption due to expropriation could be attributed to changes in permanent wealth. I use this stylized framework to interpret the results of the impact evaluation and ultimately try to say something about the adequacy of compensation payments.

2.3.2 Empirical strategy

I estimate the average treatment effect of losing land and receiving compensation on a number of key outcomes, including, household consumption, savings, asset holdings and off-farm work using the following first difference regression specification:

$$\Delta y_{it} = \rho T_i + x'_{iB}\beta + \Delta\epsilon_{it} \quad (1)$$

where subscript i denotes households and subscript t denotes the time period, taking a value of $t = B$ at baseline and $t = E$ at end-line. Δy_{it} is the change in the outcome of interest for households i, T_i is an indicator for whether the household lost any land to expropriation, and x_{iB} is a set of household controls measured at baseline. The first difference specification eliminates the effect of any fixed household characteristics such as household productivity or risk preference, but does not control for differential trends across various types of households. For this reason I include control variables that enter the regression as differential time trends, based on baseline household characteristics. These characteristics are mostly fixed at baseline and include household size and household head's age, gender, marital status and literacy. I also include the quantity of land owned by the household at baseline, to control for the likelihood that a household will lose any land at all. The object of interest in this regression is ρ, which captures the difference in the outcome of interest that can be attributed to losing some land and receiving a cash compensation payment.

The basic specification simply reflects the average change in the outcome of interest due to losing any land for the investment project. However, as section 2.2 illustrates, the intervention varies greatly across households: some households lose more land than others and as a result receive a larger compensation payment. I test how the size of the compensation payment (and implicitly the loss of land) affects the change in the outcomes of interest by splitting the treatment group into two, where one group is those that receive payments above the median payment and the other is those that receive payments below. I estimate the treatment effect for these groups using the following first difference regression specification:

$$\Delta y_{it} = \rho^A T_i^A + \rho^B T_i^B + x'_{iB}\beta + \Delta\epsilon_{it} \tag{2}$$

where T_i^A is an indicator for households that receive payments above the median payment and T_i^B is an indicator for households that receive payment below the median.

The key outcomes of interest are chosen to reflect elements of the conceptual framework above. They include the most important income generating activities, assets for households and a measure of consumption. I estimate the effect of expropriation on household consumption, non-farm labour supply, asset holdings - specifically, livestock assets, agricultural assets and other assets - and savings. Land holdings are excluded from the outcome variables since it is directly affected by the intervention.

2.3.3 Identification

The average treatment effect can be identified if the household's treatment status is uncorrelated with characteristics of the household at baseline. Differencing eliminates any omitted variables that are fixed across time, but there may be time-varying characteristics that can bias estimates. Furthermore, a household's treatment status will not be random, given the population of households that the sample is drawn from. Treatment status is defined by losing at least some land, which means that a household's treatment status is necessarily related to the quantity of land that they initially own. This implies that households that own no land at all must be in the comparison group[14] and, all else equal, households with more land have a higher likelihood that at least one of their plots will be

[14]Households without farmland could lose their house and would then be included in the treatment group. I abstract from this given the small number of treated households that lose their house.

contained within the project boundary. To account for this, I include baseline land holdings as a control variable in all regression specifications.

However, there are a number of reasons to think that a household's treatment status (conditional on controlling for baseline land holdings) is uncorrelated with omitted variables and that I have a suitable comparison group. First, the sample is drawn from a small 3km by 3km area, so treated households have access to the same markets, schools, health facilities and political institutions as the control group. Second, although the baseline data was collected after the announcement of the project, land law in Ethiopia prevents households from selling their farmland and moving away. This means that households cannot select out of treatment and factors such as human capital, asset holdings and livestock will be predetermined with respect to treatment. Table 3 shows the statistical balance between treatment and control for the included control variables and the outcomes of interest. It is possible that households will adjust some behaviour in anticipation of expropriation, for example by borrowing money against an anticipated cash windfall. (Anticipation effects are explored in more detail in a companion paper by Harris (2014)). Indeed, the third panel of table 3 shows that 60% of households in the treatment group expected their land to be taken within one year of the baseline survey. However, if ex ante behaviour moves in the same direction as behaviour following expropriation, the average treatment effect will be biased downward.[15] Despite this possibility, households actually seem to be quite limited in the short term changes they can make - credit access is limited, so households are not able to borrow in anticipation and labour may have to be kept on the farm while households still own the land.

[15]For example, if households wanted to ease into new off-farm activities, they may choose to do this in anticipation of expropriation. The treatment effect estimate would then understate the true effect of expropriation on occupation choice.

Table 3: Statistical balance between treatment and control households

		Control (n=133)	Treatment (n=164)	t-stat of difference (T-C)
Controls	Household size	4.74	5.18	1.97**
	Female head, {1,0}	0.26	0.18	1.63
	Head age	50.5	50.0	0.42
	Head married, {1,0}	0.71	0.79	-1.56
	Head sep/div/wid, {1,0}	0.27	0.18	1.79*
	Land owned (ha)	1.18	1.29	1.32
Outcomes	Total consumption (Birr)	22,429	24,198	0.97
	Total savings (Birr)	3,411	1,481	-1.98**
	Non-farm work or business, {1,0}	0.48	0.52	0.63
	Agricultural asset value (Birr)	360	391	0.89
	Non-agricultural asset value (Birr)	1,769	2,188	0.94
	Livestock value (Birr)	10,356	11,430	0.93
Other	Expect expropriation w/in 1 year, {1,0}	0.20	0.62	8.26***
	Expect expropriation w/in 5 years, {1,0}	0.36	0.74	6.91***
	Trust land committee, {1,0}	0.90	0.77	3.06***

Notes: t-statistics for the difference in means are calculated assuming unequal variance. Statistical significance is denoted by stars on the coefficient estimates using: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Another threat to identification is that households may manipulate the political process to ensure that their land is or is not included in the project area. However, households' plots tend to be scattered across the project area and so for the majority of households any political manipulation is only likely to affect the magnitude of compensation or the amount of land taken rather than a household's binary treatment status. Finally, we might worry that the project area was selected by the investors because of the characteristics of the households living in it. This is unlikely, since the investors are more concerned with finding a large, flat, contiguous piece of land than targeting particular types of individuals or acquiring land from the most fertile part of the project area. Households across the survey area hold flat land, but their holdings are fractured and small relative to the overall size of the project area, so it would not benefit the investors to target specific households nor would we expect households holding land in the project area to be systematically different from households in other parts of the survey area that were not selected into the project area.

This empirical specification does not account for any spillover effects from treated households to households in the control groups. Treated households receive large cash payments for their land and it is possible that this could lead to improved credit access, informal insurance and local labour hiring conditions for comparison households. Although these may not be first order considerations, they could lead to systematically underestimating the treatment effect (Angelucci and Giorgi, 2009). Similarly, control households are not isolated from the benefits (or costs) that come from building a factory in their neighbourhood and will benefit from any local externalities, such as job creation resulting from the new factory. In this way, control households do not provide an exact counterfactual for losing land and receiving financial compensation. Rather, the control group tells us what the treated households might have done if their neighbours land was taken away rather than their own. Given that large projects such as this factory will often provide different benefits (or costs) for the households that lose their land and for those that remain behind, this comparison can tell us much about how households respond to land expropriation programs.

2.3.4 Inference

Sample selection was stratified at the sub-village level and, as such, I adjust the standard errors on my treatment effect estimates to account for village-level clustering. However, the sample only contains 16 clusters, which can lead to over-rejection of the null hypothesis. Cameron et al. (2008) simulate data with a clustered error structure and show that with 15 clusters an ad hoc adjustment to the t-distribution can adequately correct for the over-rejection of the null hypothesis due to a small number of clusters. They recommend using a T distribution with $(G - k)$ degrees of freedom, where G is the number of clusters and k is the number of included regressors that do not vary within clusters.

Throughout the paper, I calculate p-values using this adjustment, and report statistical significance accordingly.

Furthermore, I am estimating the average treatment effect for a number of key outcome variables. In order to account for running multiple hypothesis tests, I present Bonferroni adjusted p-values alongside the normal p-values. These p-values are almost certainly too conservative since they assume that each outcome, for each household is independently drawn. Considering that many outcomes within categories are correlated I make an ad hoc adjustment to allow for this correlation following Aker et al. (2011). Both adjusted p-values are reported using the following transformation:

$$p_{adj} = 1 - (1 - p_{unadj}(k))^{g(.k)} \tag{3}$$

$$g(.k) = M^{1-r(.k)} \tag{4}$$

where M is the number of outcomes, k, being estimated, $p_{unadj}(k)$ is the unadjusted p-value for outcome k and $r(.k)$ is the average correlation between the $(M-1)$ other outcomes. The Bonferoni adjustment is the limiting case of this transformation where $r(.k) = 0$. The adjusted p-values are reported alongside the the regression results in the main results table below.

2.4 Results

2.4.1 Key indicators

Regression results are presented for each of the key outcome indicators in table 4.[16] Full tables showing regression results for a wider sub-set of outcomes related

[16]Table B.4 in the appendix estimates the treatment effect on each of the key indicators without household demographic controls and without controlling for initial land holdings. The treatment effect estimates are similar to the main results across all specifications.

to each of the key outcomes are in the appendix. I examine each of the indicators in more detail below, but the most striking result, at the outset, is that the most of the money received by households is kept as savings in the bank. Although some households have increased their non-farm assets and started new businesses, this average increase is dwarfed in magnitude by the average effect on household savings. Furthermore, most of the treatment effect on non-farm assets is driven by 9 treated households purchasing expensive motorcycle taxis. There is suggestive evidence that households have increased the value of livestock holdings and their participation in off-farm activities, but these effects are not statistically significant. Even adjusting for running multiple hypothesis tests, the trend is clear: beyond a small effect on consumption, most of the compensation payments are left as savings in the bank.

The following section examines the effect of expropriation on consumption, savings, assets and labour markets in more detail, analyzing a subset of indicators related to and making up the key indicators, and breaking down treatment by the size of payment, to explore the heterogeneity in how households respond to expropriation.

2.4.1.1 Consumption

Total nominal consumption expenditure increased by 6,300 Birr more for households that had land expropriated, an increase that is statistically significant at the 90% level. Measured in round 1 prices, this effect is 20% of the average baseline level of expenditure. What components of the consumption index drive this change? Table B.1 presents regression results using the main empirical specification, where each row corresponds to a regression explaining a component of the total consumption measure. Model 1 refers to the main specification outlined in equation 1 and model 2 refers to the specification that splits the

Table 4: Treatment effect estimates for key outcomes

	(1) Mean of DV at baseline	(2) Treatment effect	(3) N	(4) Unadjusted p-value	(5) Partial adjust- ment	(6) Bonferoni adjust- ment
Total consumption (Birr)	23,225	6,285*	293	0.0643	0.270	0.329
		(3,148)				
Total savings (Birr)	2,199	58,305***	288	0.000	0.000	0.000
		(5,781)				
Participates in off-farm activity, {1,0}	0.502	0.0765	295	0.195	0.642	0.728
		(0.0564)				
Value of agricultural assets (Birr)	376.8	-69.47	295	0.464	0.948	0.976
		(92.37)				
Value of non-agricultural assets (Birr)	1,872	5,404*	295	0.0842	0.341	0.410
		(2,921)				
Value of livestock (Birr)	10,954	2,348	295	0.132	0.489	0.572
		(1,474)				

Notes: Each row corresponds to the first differenced regression in equation 1 explaining the outcome variable specified. All financial variables are reported in Ethiopian Birr and are top-coded at the 99th percentile to limit the influence of outliers. I report only the treatment effect of losing land, but control for differential trends based on household size, head characteristics and baseline land holdings. Standard errors reported in column (2) are clustered at the sub-village level and there are 16 clusters. t-statistics are compared to a T distribution with $(G-1)$ degrees of freedom and statistical significance is denoted by stars on the coefficient estimates using: *** p<0.01, ** p<0.05, * p<0.1. These critical values do not account for multiple hypothesis tests. Column (3) reports the number of observations, which varies depending on whether there is missing data. Columns (4) to (6) calculate p-values with increasingly strict adjustments for multiple hypothesis tests: column (5) reports p-values for a single hypothesis test, column (7) adjusts for the number of hypothesis tests, but allows for correlated outcomes and column (6) implements a full Bonferoni adjustment. The adjustments are explained in equation 3.

treatment group into households that received high and low payments as outlined in equation 2. First, households in the treatment group shifted their consumption expenditure away from own food production towards food consumed or purchased outside of the household. At baseline, households spent 17,600 Birr annually on food items consumed at home, of which 10,400 Birr was spent on food items purchased outside of the household and 6,600 Birr was spent on food produced by the household. Households that lost land increased their consumption of purchased food items by 5,500 Birr relative to the control group and decreased their consumption of home produced goods by 1,500 Birr. The increase in consumption of purchased food is significant at the 90% level and represents a 35% increase relative to baseline levels. The decrease in consumption of home produced goods represents a 20% decrease on baseline levels but is not statistically significant. This pattern doesn't change when the treatment group is split into households that receive payments above and below the median compensation payment. In addition to purchasing more food for home consumption, treated households more than doubled their consumption of food prepared outside of the household, with households that received larger payments increasing their outside food consumption by more than those that received lower payments.

The change in consumption is not driven entirely by changes in food consumption. Although the survey does not collect data on a very many non-food items, treated households increase their spending on clothing and shoes by 1,600 Birr, which represents a 60% increase compared to average baseline spending and is statistically significant at the 90% level. Households receiving larger payments also increased their spending on clothing and shoes by more than those receiving lower payments. This may be explained by the fact that households in the treatment group now having access to cash and are able to afford clothes that they couldn't buy before. Finally, it does not appear that households are squandering their cash payments on alcohol or chat (a popular drug in the area); there is no change in chat or alcohol consumption for households that lost their land nor is there an any change for households that received larger payments.[17] This suggests that at least in this respect, households are not behaving irresponsibly with their payments.

Why might we observe this pattern? One reason to increase consumption in response to receiving compensation is that the household's permanent income has increased. Given that households also lose an unmeasured stream of income from their farmland, it is impossible to test whether this is the case. Another explanation for the increased consumption is that the survey was conducted during the lean season, just before the main harvest; treated households receive a liquid

[17] This finding is in line with Blattman et al. (2013), Haushofer and Shapiro (2013) and a number of other studies of cash transfers that find that households that receive cash windfalls do not increase their spending on alcohol, tobacco and drugs.

asset that can be used to smooth consumption across seasons, whereas households in the comparison group do not.[18] This could explain the relative increase in food consumed at home from purchased sources. Finally, there may be a behavioral explanation: when a household holds a large sum of cash, it is easy to rationalize spending a small amount to supplement household consumption. The change in consumption expenditure observed in the treatment group is large relative to baseline average consumption, however the magnitude of change is relatively small compared to the total payments received. In interpreting these results it is also important to keep in mind the timing of the survey: the survey was conducted within one year expropriation and households were permitted to harvest their land before it was taken from them, which means that treated households may still have had stores remaining.

2.4.1.2 Savings

The biggest effect of the intervention in terms of magnitude is the increase in savings held by households that lost land. Initial savings at baseline were low: 65% of households did not have any savings and 30% had less than 10,000 Birr. These proportions are reflected in the level of savings held by the control group at end-line. In contrast, treated households increased their savings, on average, by 58,000 Birr, a change which is strongly statistically significant. These results are reported in table B.3. Within the treatment group there is a close relationship between the level of savings at end-line and the amount of compensation households received. Figure 4 plots a local polynomial regression of the ratio of end-line savings to total compensation payment against

[18] Note also that the treatment effect for home consumption of purchased goods is similar across households that receive large and small payment, despite the fact that households that receive large payments are more likely to lose all of their land. This suggests that these households want to meet a certain level of consumption, which comparison households are unable to reach.

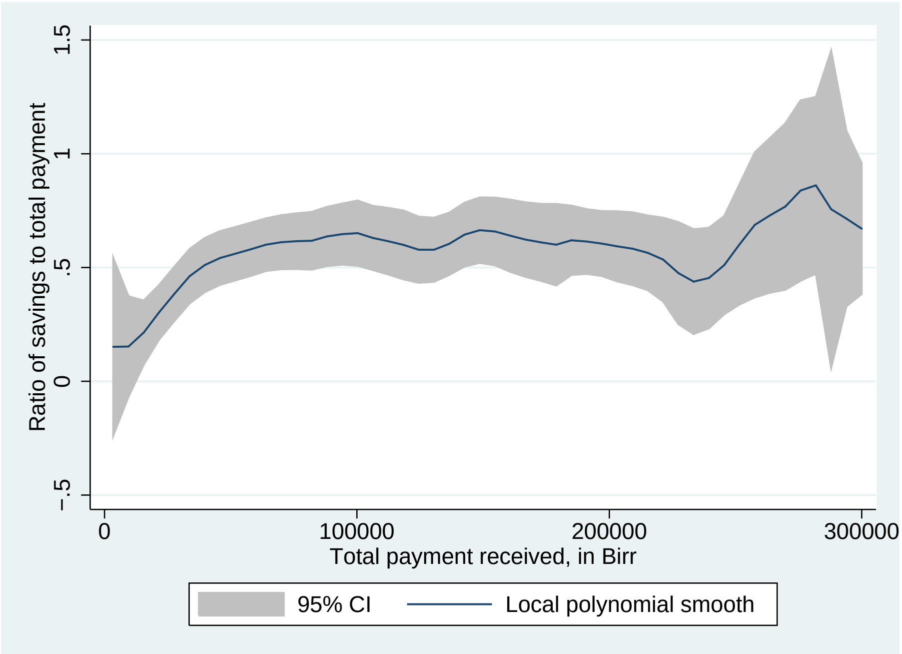


Figure 4: Local polynomial smooth of savings to compensation ratio plotted against total payment

the size of the compensation payment.[19] There is a clear non-linear relationship between compensation payments and end-line savings, with the average savings to compensation ratio increasing below payments of 50,000 Birr, after which households hold, on average, 60% of their compensation as savings. About 15% of households report having no money held in savings, which suggests that they have already spent their compensation.[20] Most treated households with no savings received smaller payments in the first place, with a median payment size of 30,000 Birr compared to a median payment of 85,000 in the treatment group as a whole.

This result is so striking because 8 months after receiving compensation most households keep a majority of their compensation payment in an asset that has a negative real interest rate rather than investing in a new income generating

[19]Households with savings greater than compensation are excluded from the plot, but given the small size of initial savings this does not amount to many households nor does it change the interpretation of this graph.

[20]It is also possible that these households are reluctant to disclose how much money they hold in savings.

activity.[21] Compensation was paid directly into savings accounts at a commercial bank so, in some ways, the amount saved in the bank represents what remains of the lump-sum cash transfer. Households may actively choose to leave their payments in a savings account because they offer a safe, liquid asset. However, saving is costly since annual interest rates on these accounts tend to be below inflation: households reported 0% to 5% annual interest on their savings accounts, but food prices increased by 30% between survey rounds. Why might households choose to keep the money in savings rather than investing in a new business or expanding an existing one? One possibility is that the real return on starting or scaling up a non-farm business is negative and that households prefer to keep their compensation in savings because of a lack of any other option. Most households know how to farm, but may lack the necessary experience or skills to operate a new business at a larger scale. Another possibility is that households are biding their time before making an investment decision because they are waiting for a good opportunity to arise, they may want to learn from what other households do or they may be waiting for construction jobs at the factory. Finally households may simply choose to consume out of their consumption while they make a transition to the local non-farm labour market or migrate elsewhere for work. I return to these questions in the extension sections and try to test some of these possibilities.

2.4.1.3 Other assets: livestock and other assets

Losing land and receiving compensation can affect the asset mix held by households in a number of ways. Where treated households are liquidity constrained, a cash payment enables households to make

[21] Haushofer and Shapiro (2013) present findings on the largest unconditional cash transfer program, by size of payment. In their sample from poor rural Kenya, households that receive 404 USD increase their savings by 10 USD while households that receive 1,500 USD increase their savings by 20 USD. Their follow up surveys were run between 1 and 14 months after the treatment.

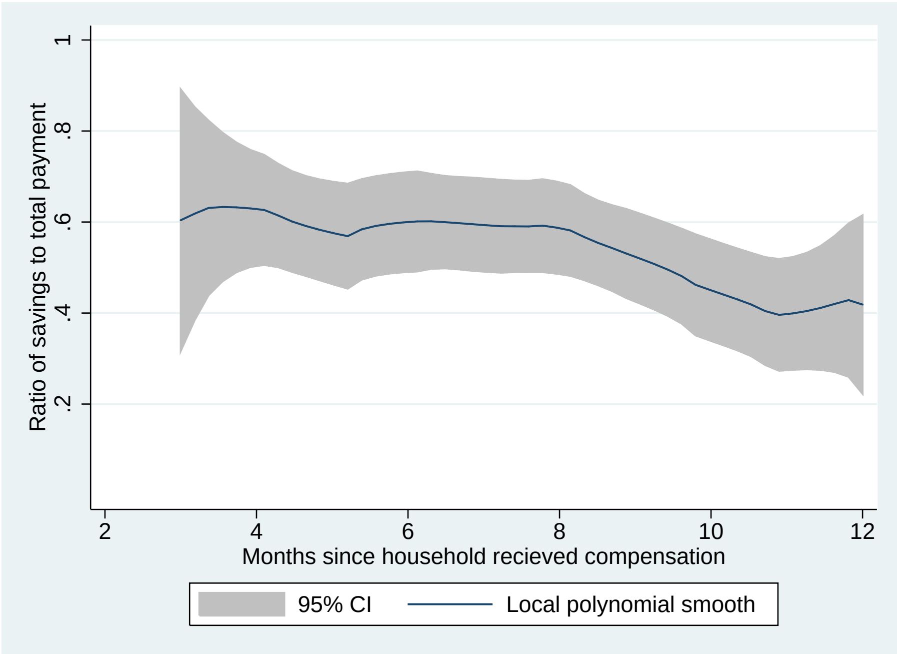


Figure 5: Local polynomial smooth of the savings to compensation ratio over months since payment

lumpy purchases(Fafchamps and Pender, 1997; Rosenzweig and Wolpin, 1993). Likewise, if households lose enough land, they may choose to draw down the value of assets that are directly used in agricultural production. Many of the non-productive assets asked about in the survey could be classed as durable consumption goods and households may increase their investment in these goods as well.[22] Table B.2 breaks down the relative change in key asset indicators for households in the treatment group. These assets fall into three broad categories: productive agricultural assets, non-productive assets and livestock holdings. Many relevant assets are missing from the list due to constraints on the size of the survey, but the analysis of their change gives some insight into how households reallocate their asset portfolio following expropriation.

Table 4 on page 88 shows that treated households increase the total value of their livestock holdings by 2,400 Birr more than the control group, however this effect

[22]Haushofer and Shapiro (2013), in their evaluation of a large unconditional cash transfer program, find that poor households in rural Kenya increase their consumption of durable goods and especially metal roofs.

is not statistically significant above the 90% level. There is also no difference in the size of the treatment effect when the treatment group is divided according to the size of payments. However, a more nuanced picture emerges when the value of livestock holdings is broken down by the type of animal. These results are reported in panel 3 of table B.2. Relative to the control group, treated households increase the value of their holdings of sheep and goats by 1,300 Birr and cattle by 1,100 Birr, while reducing the value of their oxen by 1,300 Birr. The positive effect on sheep and goats and the negative effect on oxen is statistically significant at the 99% and 90% levels respectively. This pattern is even more pronounced when the treatment indicator is split by the size of payment: households that receive payments above the median reduce their holdings of oxen by 2,600 Birr, while increasing the value of their holdings of sheep and goats by 1,500 Birr. The reduction in the value of oxen is statistically significant at the 95% level.[23] There is no change in the value of oxen holdings for households that receive payments below the median payment, but they do increase their holdings of sheep and goats. This pattern is consistent with households that receive large payments losing more land, and therefore having less need of oxen. To test this, figure 6 plots the change in oxen against the area of land that is expropriated. Households that lose more land reduce their head of oxen by more. In contrast to oxen, cattle, sheep and goats represent both a store of real value and a business opportunity for households that lose their farmland, so it is unsurprising to see that treated households have increased their investment in this type of livestock. However, the change in the value of sheep, goats and cattle does not increase dramatically with the size of household's compensation payment, which suggests that herd size may be constrained by other factors such as a lack of grazing land, pens or household labour.

[23]These general patterns persist when I use the number of animals instead of their reported value as the outcome variable.

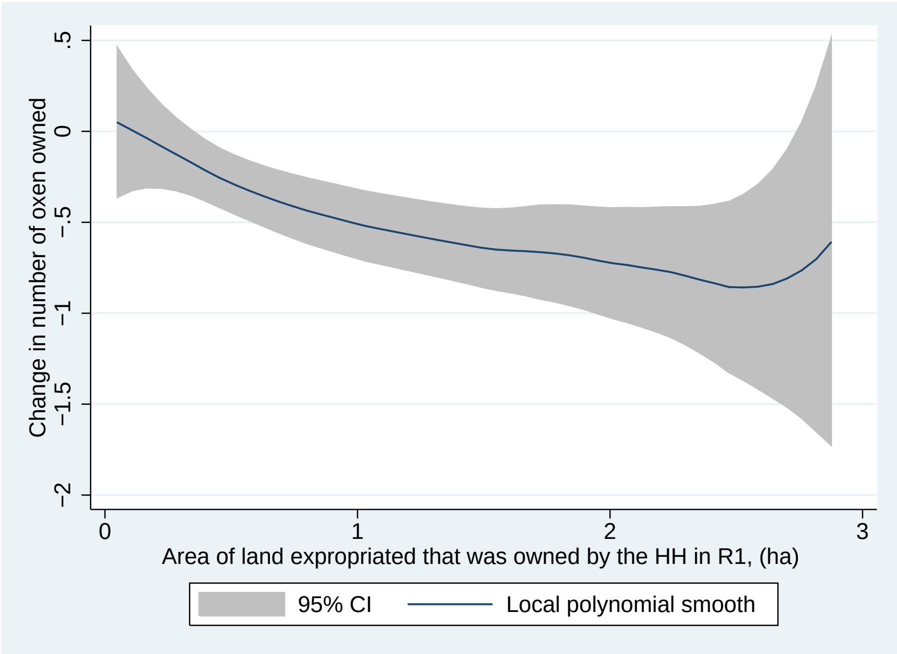


Figure 6: Local polynomial smooth of the change in the number of oxen over the total area of land lost

Panels 1 and 2 in table B.2 show the estimated treatment effect of losing land on productive agricultural assets and set of non-productive assets held by households. Both asset value measures are constructed using self-reported estimates of the cost of purchasing a specific set of items. Productive assets include ploughs, hoes, irrigation pumps, tools and axes and guns, while the set of non-productive assets include furniture (beds, table, sofa, chair), stoves, cellphones, radios, TVs, transport (cars, motorcycles, carts) and jewelry. The treatment effect on productive agricultural assets is small and not statistically different from zero nor is there any difference between treated households that receive large payments and those that receive less. This zero result is somewhat surprising, since one might expect households that retain land and receive some compensation to invest in agricultural production, while those that are forced out of agriculture altogether would liquidate their agricultural assets. To some extent, such a pattern is visible in figure 7 where the change in productive agricultural asset values are plotted against the total land area lost, however this plot excludes

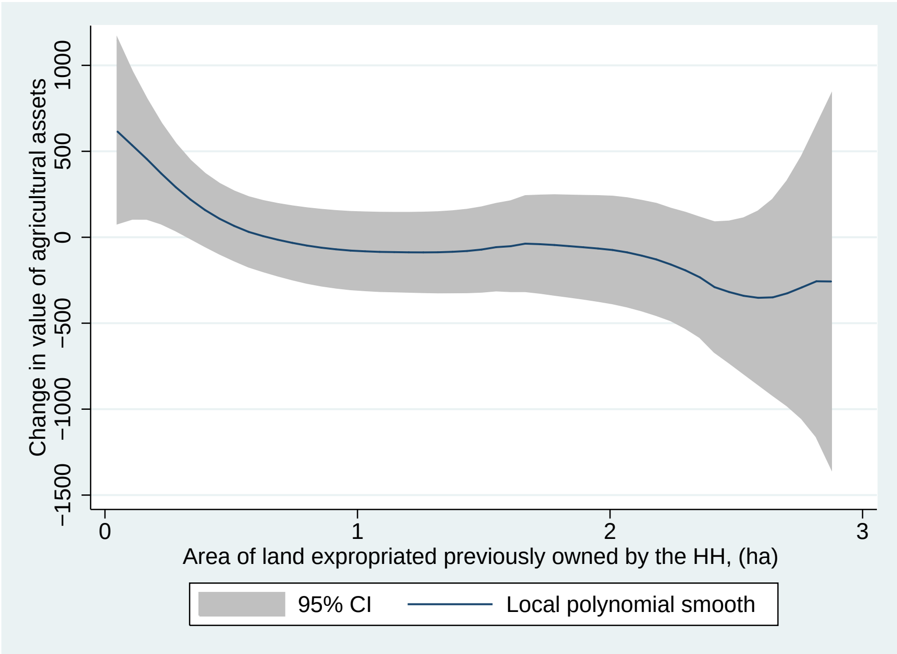


Figure 7: Local polynomial smooth of the change in productive assets and the total area of land lost

10 households that increased the value of their agricultural assets by purchasing irrigation pumps.[24] However, there is also no reason why households would use compensation payments to increase their agricultural assets if they already had the productive assets they need.

Panel 2 in table B.2 reports the treatment effect for a subset of non-productive assets and durable goods. Losing land and receiving compensation increases the value of non-productive assets held by the household by 5,400 Birr, an increase that is statistically significant at the 90% level. This large, positive effect is almost entirely driven by the treatment effect on transport assets, which itself is due to 9 households that purchased Bajaj motorcycle taxis. These households reportedly spent between 50,000 Birr and 231,000 Birr on motorcycle taxis between rounds,

[24]These 10 households increased the value of their productive assets by an average of 5,800 Birr, which is 17 times the mean value of baseline productive assets and 3 times the largest value from round 1. Three of these households are in the control group and seven are from the treatment group, so this might be an example of spillovers from treatment.

while no control households purchased any transport asset. This is the primary example of households making lumpy purchases with their compensation payment.

The average treatment effect on the other assets measured is much smaller in magnitude than the effect on transport assets: treated households increase their cellphone ownership by 280 Birr and reduce the value of jewelry by 530 Birr, effects which are statistically significant at the 95% and 90% level respectively. There is no statistical difference in the change in the value of furniture or televisions owned by the treated and control households. The pattern remains the same even if the treatment indicator is split by payment size, except for the treatment effect on televisions and radios, which is positive and significant for households that receive large payments. The set of assets that make up the non-productive asset indicator is limited, so we might not be measuring other types of investment, however most of the money held by treated households is still in the bank, suggesting that no other large purchases have been made. Overall, this evidence suggests that treated households use their compensation payment to increase their purchases of durable assets, but the size of the effect is not large. The negative effect on jewelry would be consistent with households choosing to substitute jewelry for bank accounts as their primary saving mechanism or with households drawing down assets prior to expropriation (but after baseline) in anticipation of a cash windfall.

2.4.1.4 Off-farm labour, businesses and income shares

Table B.3 evaluates the treatment effect for a variety of indicators that reflect off-farm labour market participation and participation in household businesses. These indicators measure whether a household has any members of the household working for wages in the non-agricultural sector or working for a non-agricultural household business or both. Households also reported the number of weeks each member of the household spent doing non-farm work. The average treatment effect for households that lost land was positive for each of these measures, but not significant at the 90% level. However, if the treatment group is split into those who received payments above the median and those who receive payments below, there are positive and statistically significant effects of losing land on labour market participation for those households that received the bigger payments. These households were 22% more likely to participate in the non-farm labour market and/or a household business unrelated to farming compared to a baseline participation rate of 50%. These households also increased person-weeks worked by 9 weeks over the course of the year, were 18% more likely to participate in wage labour and 15% more likely to participate in a household business. In contrast, households that received smaller payments were no more likely to have increased their participation in non-farm activities than control households. This heterogeneity in response by the size of payment suggests that those households that lost larger amounts of land were being forced to adapt their income generating strategies rather than choosing to work more. Households in the low payment group retain farmland, which may explain why they have chosen not to increase their off-farm participation relative to the control group.[25] This suggests that high payment households are working in business or for wages because their labour is no longer needed on the farm and these jobs provide the household with important cash income.

Losing land and receiving compensation also affects a household's source of cash income. Panel 3 of Table B.3 shows the treatment effect for cash income shares. Households that lose land derive a relatively smaller share of their cash income from selling crops and increase the share of cash income that comes from the sale of livestock, and from household business. At baseline, the average household

[25] Although, note that the size of payment and the amount of land lost are closely correlated, so the effect is not separately identifiable from receiving a large payment. The driving force may actually be the large payments received by households. For example, large payments could enable households to make large fixed investments that are required for starting a business.

acquired 50% of their cash income from the sale of crops, 13% from the sale of livestock and 8% from household businesses. Treated households receive 11% less cash income from crops, 8% more from livestock sales and 6% more from household business, effects which are significant at the 90% level. There is no statistically significant treatment effect on the share of cash income that comes from non-agricultural wages. This pattern suggests that households in the treatment group are adjusting their cash income portfolios as a result of losing land, but that the important source of non-farm income comes from household businesses rather than wage work. It is important in interpreting these results to emphasize that these are shares of cash income and not absolute levels.[26] It is impossible to say anything about what is driving the portfolio readjustment: a shift away from crop income could occur either because of a drop in crop incomes that is not matched by a proportionate drop in other income sources or an increase in income from other sources that is not proportionate to an increase in crop income. Thus, this pattern is both consistent with households reducing their crop income because they have less land and using compensation payments to increase income from other categories.

Households in the treatment group increased their share of cash income from non-farm businesses and increased their labour participation in non-farm businesses. However, the survey only contains basic information on household businesses collected in the end-line survey, which makes it difficult to formally measure how much households invested directly into these business following the intervention. (No detailed information on household enterprises was collected at baseline.) Figure 8 shows the year in which household businesses were started relative to the year of expropriation. At end-line, 23% of all households owned a non-farm business. Of these businesses, 18 were started in the year of or

[26]In the survey, households were asked to allocated 20 tokens representing their yearly cash income across a number of categories of income sources.

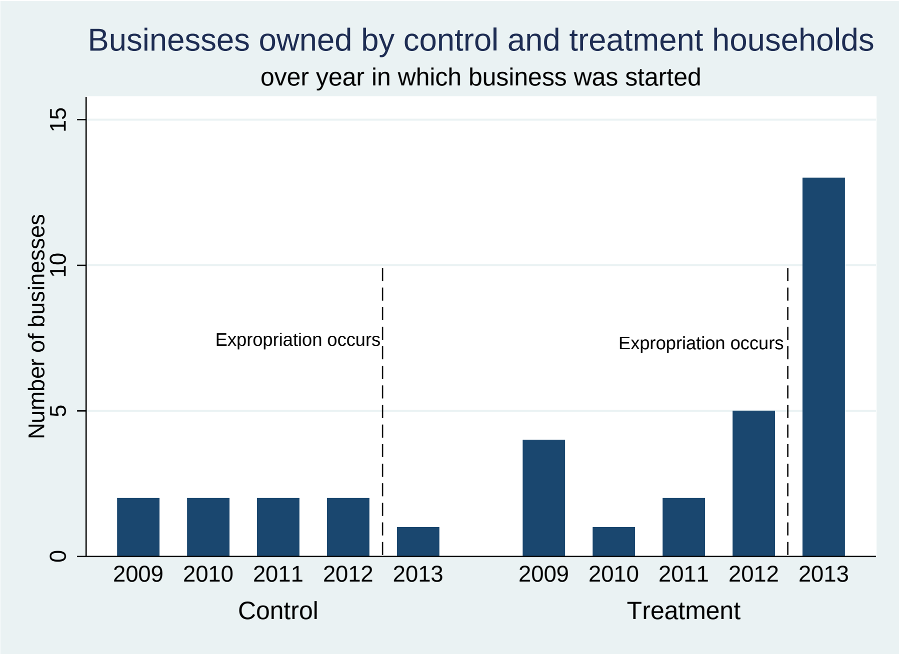


Figure 8: Distribution of year in which household business was started

year following expropriation by treated households and 2 by control households. Amongst the remaining businesses started prior to this time, 22 were owned by control households and 27 by treated households - these numbers are roughly 16% of the total number of households in each group. Although this is evidence is not subject to empirical testing, it strongly suggests that some households that lost their land have responded by starting new businesses. Table 5 on the next page breaks down the businesses by type; most of the businesses, with the exception of those households that have purchased motorcycle taxis, have entered the trading business. These businesses have limits to how much can be invested, and indeed the savings rates of households in the treated group with businesses are not significantly different from those in the treated group without businesses.

Table 5: Breakdown of businesses owned by treated households

Business activity started in year of expropriation	Treated HH (n=164)
Grain trade	2
Livestock trade	1
Other trade	9
Shop/kiosk/retail	2
Transport/cart/bajaj	6
Handicraft	1
Sub-total	**21**
Business established before 2012	**24**
No business reported	**119**

2.4.2 Robustness

Data on the intervention is somewhat messy and one might worry that the treatment effect estimates are biased by this mis-measurement. Table B.5 presents treatment effect estimates for various measures of treatment on the key indicators using the main specification outlined in equation 1. In some cases, households report receiving payments more than 12 months prior to the date of the end-line survey, which should not be possible since expropriation occurred only 8 months before the end-line survey. There is also some inconsistency within household responses. An alternate measure of treatment is derived from the plot roster collected at end-line: households were reminded about the plots that they had listed in the baseline survey and were asked which, if any had been expropriated. Out of 295 households, 22 provided inconsistent answers across these measures. Table B.5 presents treatment effect estimates on the key indicators using various measures of treatment using the main specification outlined in equation 1. Column (2) uses the measure of treatment derived from the plot roster. The results are very similar when this alternate measure is used. Some households also lose their house in addition to losing land and may respond very differently than households that only lose land. Households that lose their homestead are given money to rebuild their home and a new plot of land within the kebele. Column (3) and (4) run the main specification while including a dummy variable for whether the household lost their house in addition to their land. The treatment effect estimates are similar to the main results for households that do not lose their homestead, but households that do lose their house are more likely to divest their productive agricultural assets and less likely to increase the value of their livestock holdings relative to households that only lose land.

The effect of expropriation on some of the outcomes discussed above depends on whether households receive large or small payments. This is intimately related to how much land households lose. Thus, in addition to receiving very different payments, these groups also lose different shares of their initial land holdings. Figure 9 shows the distribution of the share of land holdings lost by whether households receive large or small payments. Households that receive large payments are much more likely to lose all of their land whereas those who receive payments below the median are not. Splitting the treatment group by median payment effectively estimates a non-linear effect of the size of compensation on outcomes. Columns (7) and (8) of table B.5 present alternate specifications of the split treatment variable, running each regression on the total payment size and the log of total payment (giving control households a payment of 1). Both sets of regressions provide estimates that have a similar magnitude and support the estimated coefficients on the indicators for high and low payments.

The final set of robustness checks run the main specification on a set of restricted samples. These results are reported in table B.6. The first set of regressions in columns (2) to (4) run the main specification on a sub-set of households living close to the boundary of the project area. The first regression is run on a sub-sample of households living within 500m of the boundary, followed by those living within 1000m and finally by those within 2500m. Restricting the sample to those households living near to the project boundary limits the influence of unobservable omitted variables that vary by geography. These might include

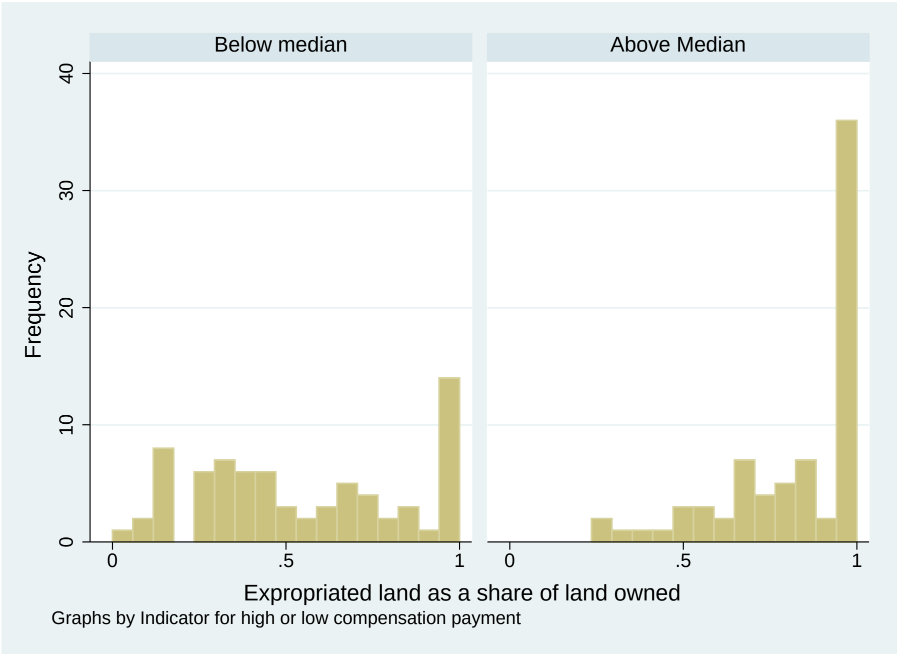


Figure 9: Distribution of share of land lost by large and small payment groups

things like better access to markets or unobserved soil quality. We may also worry that households have been incorrectly identified as treatment or comparison, have been misclassified in some systematic way. This might be the case if households that were identified by local officials as being in the treatment group at baseline had undue political influence and were able to ensure that their land wasn't taken (or comparison households that ensured that their land was taken). Columns (5) to (7) run the main specification on a sub-sample of households for which treatment is consistent across measures. Column (5) restricts the treatment sample to households that report losing their land within the last year, which eliminates any households that lost land in previous expropriations. Column (6) drops any households whose self-reported treatment status was different from what the local official reported prior to the baseline survey. Column (7) restricts the sample to those households whose self-reported treatment status is consistent with whether they report losing any land in the plot recall section described above. The coefficient estimates are broadly similar across these different regressions.

2.4.3 Discussion

Expropriating farmland forces households to find new income generating activities and compensation payments provide a liquid asset that can be invested, consumed or saved. In the 8 months following expropriation, most households have done relatively little with their compensation payment, having chosen to save most of it in the bank where it's real value is eroded by high inflation. If this trend continues, policy makers should worry. First, it suggests that most households are not using their lump-sum payments as seed capital for new businesses. Second, inflation in Ethiopia is high and, if it continues, will erode the real value of saved compensation. Between survey rounds the price of the household's food basket increased on average by 30%, yet nominal interest rates on savings accounts range from 0% to 5% annually.[27]

Evidence on the effect of expropriation and compensation on livestock assets and agricultural assets suggest that households do react to expropriation in ways that might be suggested by the conceptual framework, but the magnitude of their response is dwarfed by the amount of money left in the bank. Treated households do make adjustments to their asset portfolios, reducing their holding of agricultural assets and increasing their livestock holdings, and shift labour into off-farm income generating activities. Treated households also start more new businesses following expropriation, despite only investing small sums of money in these businesses. Where the data permits, exceptions to this trend have been identified: for example 9 households used their compensation payments to make very large capital expenditures in buying motorcycle taxis. But this

[27] Most of the respondents are Muslim and many reported that they did not receive interest on religious grounds. The website for the Commercial Bank of Ethiopia currently advertises annual interest rates of 5% on savings accounts. Although I have no official information on interest rates from the banks in Kombolcha town, 25% of households that had bank accounts reported annual interest rates in the 0% - 5% range. 75% of households that had savings in a commercial bank did not know what interest rate (if any) they were being paid.

group represents a small fraction of treated households and most observed changes in investment are relatively small compared to the size of the household's compensation payment. These findings are consistent with the conceptual framework, but suggest that households face a lack of high-return investment opportunities.

What then explains the high level of savings? Households may already have a diversified asset and activity portfolio and are saving their compensation payment for a 'rainy day'. This behaviour may also be rational if households see future opportunities for work at the factory; if households can find regular employment working in the construction of the factory, they may not need to risk their compensation payment in a new investment project, especially when they don't have the requisite business skills or experience. On the other hand, the household's choice to save their money in the bank may indicate that returns to other activities are lower than the negative real return from keeping the money in the bank. The patterns observed in the first 8 months might be the beginning of a worrying trend: households will continue to slowly consume out of their compensation payment until it is gone, without using the money towards any productive purpose. This would make households worse off as a result of the policy.

Assessing the adequacy of compensation requires us to consider whether the changes that are observed in income generating activities and the investment in productive assets provide households with a level of permanent wealth that is equivalent to what they had before. Treated households increase their consumption expenditure following expropriation along a number of dimensions, increasing their consumption of purchased food, clothing and shoes and durable goods such as cellphones. This finding is consistent with an increase in permanent income, but may just as easily reflect seasonal liquidity constraints faced by agricultural households during the lean season.

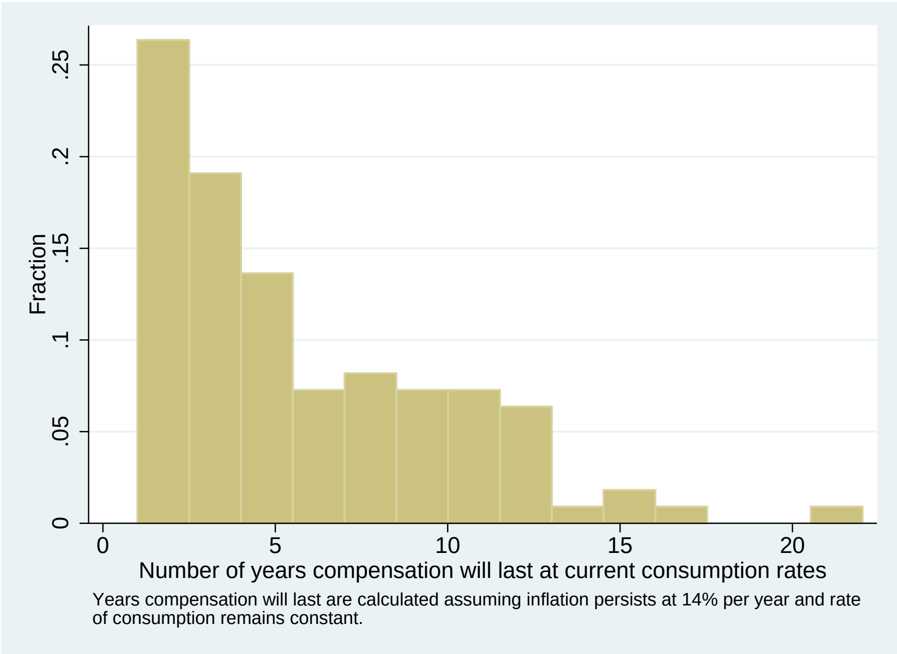


Figure 10: Year in which households will exhaust compensation payment given current rates of consumption

Another way of assessing the adequacy of compensation payments is to consider whether the changes to asset and income portfolios would provide households with a sufficient level of consumption by the time their compensation payment is exhausted. Households were asked to report how much of their compensation they had spent on food and household expenditure since receiving their payment. I use this data in figure 10 to calculate the year in which households would exhaust their payments if the rate of consumption remained unchanged and price inflation stayed constant at 14%[28]. By this measure, most households will deplete their compensation within 5 years. For compensation to be adequate, the adjustments to asset and income portfolios made by households must be sufficient to allow households to consume at the same level as if they still owned all their land.

[28]Inflation was calculated using data on the basket of goods included in the survey's consumption module.

2.5 Conclusions

Households in this area experienced a significant intervention: on average they lost half of their farmland and received the equivalent of 5,200 USD. These households responded by moving into new income-generating activities, investing in new businesses and adjusting their asset portfolios. They also increased their consumption expenditure relative to the comparison group and increased their consumption of durable goods. The magnitude of these effects depends in part on the size of the payment (and the share of land lost): households that lost most of their land also received the largest payments and were more likely to participate in non-farm income generating activities and to sell off oxen. While all of these effects reflect a rational response to losing land and receiving a large compensation payment, by far the biggest and most striking result is that most of the compensation payment is left in the bank. This finding suggests that rates of return, at least in the short term, are very low. While households are able to find a productive use for cash transfers, above a certain value, the additional scope for investing is limited. With the exception of a few households that made very large investments, most households have done very little with their money.

From a normative point of view, households that lose their land should not be made worse off as a result of expropriation and at the very least, should be able to replace the income that they generated with their land. Land in Ethiopia also serves as more than just a productive asset: it serves as insurance and security in old age when it is used for sharecropping. Compensation payments should assist households in making the transition from small-scale agriculture to other income generating activity and yet, in this short time period, it seems that the majority of households are not able to do so.

References

Adamopoulos, T. and Restuccia, D. (2014). The size distribution of farms and international productivity differences. *American Economic Review*, 104(6):1667–97.

Aker, J. C., Boumnijel, R., McClelland, A., and Tierney, N. (2011). Zap it to me: The short-term impacts of a mobile cash transfer program. *Center for Global Development working paper*, 268.

Ali, D. A., Dercon, S., and Gautam, M. (2011). Property rights in a very poor country: Tenure insecurity and investment in ethiopia. *Agricultural Economics*, 42(1):75–86.

Angelucci, M. and Giorgi, G. D. (2009). Indirect effects of an aid program: How do cash transfers affect ineligibles' consumption? *The American Economic Review*, 99(1):486–508.

Banerjee, A. V., Bardhan, P., Basu, K., Chaudhury, M. D., Ghatak, M., Guha, A. S., Majumdar, M., Mookherjee, D., and Ray, D. (2007). Beyond nandigram: Industrialisation in west bengal. *Economic and Political Weekly*, 42(17):1487–1489.

Banerjee, A. V., Duflo, E., Aghion, P., and Durlauf, S. N. (2005). *Chapter 7 Growth Theory through the Lens of Development Economics*, volume Volume 1, Part A, pages 473–552. Elsevier.

Besley, T. (1995). Property rights and investment incentives: Theory and evidence from ghana. *journal of Political Economy*, pages 903–937.

Blattman, C., Fiala, N., and Martinez, S. (2013). Generating skilled self-employment in developing countries: Experimental evidence from uganda*. *The Quarterly Journal of Economics*.

Cameron, A. C., Gelbach, J. B., and Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics*, 90(3):414–427.

Cotula, L., Vermeulen, S., Leonard, R., and Keeley, J. (2009). *Land grab or development opportunity? Agricultural investment and international land deals in Africa*. IIED, FAO and IFAD.

de Mel, S., McKenzie, D., and Woodruff, C. (2008). Returns to capital in microenterprises: Evidence from a field experiment. *The Quarterly Journal of Economics*, 123(4):1329–1372.

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

Ding, C. (2007). Policy and praxis of land acquisition in china. *Land Use Policy*, 24(1):1–13.

Fafchamps, M., McKenzie, D., Quinn, S., and Woodruff, C. (2014). Microenterprise growth and the flypaper effect: Evidence from a randomized experiment in ghana. *Journal of Development Economics*, 106(0):211–226.

Fafchamps, M. and Pender, J. (1997). Precautionary saving, credit constraints, and irreversible investment: Theory and evidence from semiarid india. *Journal of Business 'I&' Economic Statistics*, 15(2):pp. 180–194.

Fenske, J. (2011). Land tenure and investment incentives: Evidence from west africa. *Journal of Development Economics*, 95(2):137 – 156.

Geary, K. (2012). *Our Land, Our Lives': Time out in the global land rush*. Oxfam.

Gertler, P. J., Martinez, S. W., and Rubio-Codina, M. (2012). Investing cash transfers to raise long-term living standards. *American Economic Journal: Applied Economics*, 4(1):164–192.

Ghatak, M., Mitra, S., Mookherjee, D., and Nath, A. (2012). Land acquisition and compensation in singur: What really happened? *Working PaperInstitute for Economic Development, Boston University*.

Ghatak, M. and Mookherjee, D. (2013). Land acquisition for industrialization and compensation of displaced farmers. *Journal of Development Economics*, (0):–.

Haushofer, J. and Shapiro, J. (2013). Welfare effects of unconditional cash transfers: Evidence from a randomized controlled trial in kenya. *Unpublished Working Paper*.

Jacoby, H. G., Li, G., and Rozelle, S. (2002). Hazards of expropriation: Tenure insecurity and investment in rural china. *The American Economic Review*, 92(5):1420–1447.

Rosenzweig, M. R. and Wolpin, K. I. (1993). Credit market constraints, consumption smoothing, and the accumulation of durable production assets in low-income countries: Investments in bullocks in india. *Journal of Political Economy*, 101(2):pp. 223–244.

Appendices to Chapter 2

A. Validation of self-reported treatment using official records

Surveyed households were matched to administrative records for compensation payments using the full name of the household head. However, only 45% of treated households from the survey could be uniquely identified in the administrative data.[29] Although this means that administrative data cannot be used in the analysis, it is possible to use this data to validate the survey questions on the expropriation and compensation process that overlap with the administrative records. In particular, it is possible to check whether households mis-report the compensation payments they receive or the quantity of land taken. Figure A.1 compares the household's reported compensation payment with the record held by the government. There are large discrepancies for some households, but for the most part the self-reported value of compensation corresponds closely to the value held in the administrative records. This suggests that the self-reported compensation payment is a good proxy for the official payment and can be used in the analysis. Figure A.2 compares the self-reported amount of farmland taken (measured as the sum of Meher and irrigation land) with the official records for land taken. In contrast, it shows that households consistently overstate the amount of land taken relative to the government's records. This likely occurs because households were estimating land lost using local units that are the equivalent of 0.25 hectares and thus will tend to round up. An alternative measure of land taken shows a similar pattern: respondents were shown the plot roster from the baseline survey and asked to identify which of their plots had been taken.

[29] There are three main reasons for this. First, many people in this area have the same name and without further information they cannot be distinguished. Second, the administrative data is written using the Ge'ez alphabet, (the Ethiopian script), while the names in the survey are written in the Latin alphabet, which means there will be error in translation. Third, matching was done by hand using a list of names from both surveys, so it is possible that some households were missed.

Summing over the plot size for treated plots also overstates land taken relative to the official record. Households have a tendency to overstate the size of their plots, although they do not seem to over-report compensation payments. This suggests that any regressions using land size should be interpreted carefully.

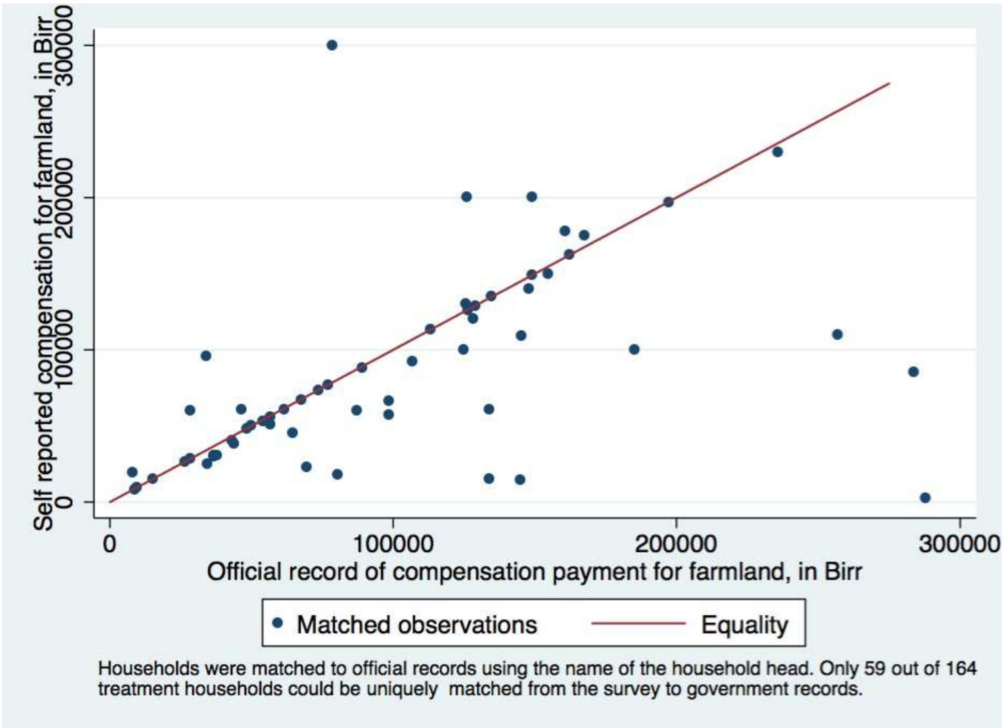


Figure A.1: Self-reported compensation compared to government records

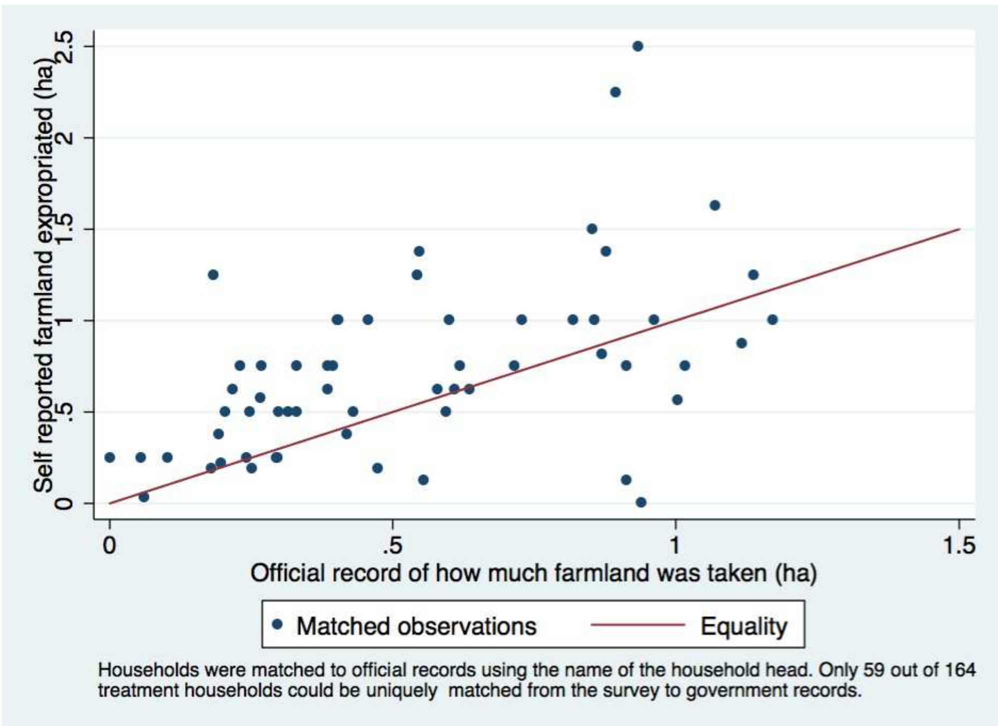


Figure A.2: Self-reported farmland taken compared to government records

B. Main regression tables

Table B.1: Treatment effect estimates: Household consumption indicators

		Model 1	Model 2	
	(1)	(2)	(3)	(4)
Dependent Variable	Mean of DV at Baseline	Treatment indicator {1,0}	Payment above median, {1,0}	Payment below median, {1,0}
Total consumption (Birr)	23,225	6,285*	8,180*	6,438
		(3,148)	(3,896)	(4,270)
Food consumption inside HH (Birr)	17,642	3,370	4,062	3,988
		(2,777)	(3,470)	(3,658)
Food consumption: own production (Birr)	6,580	-1,491	-1,721	-1,365
		(1,135)	(1,129)	(1,499)
Food consumption: purchased (Birr)	10,417	5,470*	5,892	6,452*
		(2,818)	(3,762)	(3,197)
Food consumption: gifts/transfers (Birr)	292.3	-85.11	-165.2	46.92
		(133.0)	(176.4)	(134.1)
Food consumed outside HH (Birr)	656.9	748.8*	1,142**	379.6
		(417.2)	(523.7)	(414.8)
Clothing & Shoes (Birr)	2,649	1,621**	2,554**	1,228
		(590.8)	(966.1)	(704.5)
School expenses, uniforms & textbooks (Birr)	767.8	156.5	188.4	201.2
		(148.8)	(194.7)	(152.8)
Chat expenditure (Birr)	366.5	-160.3	-124.4	-59.79
		(273.7)	(404.0)	(239.1)
Alcohol expenditure (Birr)	12.11	-38.00	6.168	-73.99
		(84.58)	(106.2)	(70.77)
Food consumption, per capita (Birr)	3,857	286.5	414.6	192.6
		(410.0)	(529.1)	(504.0)

Notes: Each row corresponds to a separate outcome variable. All financial variables are reported in Ethiopian Birr and are top-coded at the 99th percentile to limit the influence of outliers. All regressions control for differential trends based on household size, head characteristics and baseline land holdings. Standard errors reported in parentheses are clustered at the sub-village level and there are 16 clusters. Statistical significance is denoted by stars on the coefficient estimates using: *** p<0.01, ** p<0.05, * p<0.1. These critical values do not account for multiple hypothesis tests, but allow for a small number of clusters. Column (1) reports the mean of the outcome variable at baseline across treatment and control households. Column (2) reports the treatment effect for Model 1, which is the first differenced regression in equation 1. Model 2 splits the treatment group into households that received payments above the median compensation payment and below. Columns (3) and (4) show the coefficient estimates on these dummy variables, using the first differenced regression specification outlined in equation 2 on page 82.

Table B.2: Treatment effect estimates: Household assets and livestock

		Model 1	Model 2	
	(1)	(2)	(3)	(4)
Dependent Variable:	Mean of DV at Baseline	Treatment indicator, {1,0}	Payment above median, {1,0}	Payment below median, {1,0}
Value of agricultural assets (Birr)	376.80	-69.47	-202.9	62.81
		(92.37)	(156.7)	(100.6)
Total value of non-agricultural assets (Birr):	1,872	5,404*	11,027*	287.2
		(2,921)	(5,189)	(1,106)
Value of furniture (beds, chairs, sofa, table)	564.20	112.0	44.44	30.85
		(130.7)	(147.5)	(144.6)
Value of cellphones	403.60	276.7**	292.9**	268.6**
		(94.35)	(114.5)	(93.41)
Value of jewelry/gold/watches	574.20	-425.8*	-507.3**	-510.6**
		(202.5)	(213.1)	(226.2)
Value of TV and radio	238.80	127.0	240.3*	53.64
		(77.59)	(126.7)	(95.77)
Value of transport (cart, bajaj, car)	0	5,604*	11,282**	716.3
		(2,928)	(5,135)	(925.5)
Total value of livestock (Birr):	10,954	2,348	2,015	2,846
		(1,474)	(1,330)	(1,936)
Value of oxen	5,656	-1,330*	-2,570**	89.64
		(690.8)	(1,001)	(579.3)
Value of cattle (bulls, cows, calves, heifer)	3,322	1,090*	1,766*	657.1
		(580.7)	(834.8)	(630.4)
Value of sheep and goat	885.10	1,353***	1,504**	821.6*
		(417.7)	(525.0)	(448.3)
Value of transport animals (camel, donkey, horse, mule)	899.40	1,264	1,282*	1,276
		(743.9)	(676.2)	(1,126)
Value of chickens	148.90	49.92	53.96	28.88
		(43.50)	(66.44)	(41.58)
Value of beehives	23.76	-9.675	38.88	-0.704
		(17.22)	(29.22)	(27.65)
Self-reported cost of building house today (Birr)	44,430	1,973	-14,322	13,210
		(7,877)	(11,905)	(8,606)
Sheet roof indicator, {1,0}	0.91	-0.0280	-0.0347*	-0.0190
		(0.0163)	(0.0182)	(0.0175)
Number of sheets on roof	46.51	-0.440	-0.150	2.229
		(2.352)	(2.871)	(2.611)

Notes: Each row corresponds to a separate outcome variable. All financial variables are reported in Ethiopian Birr and are top-coded at the 99th percentile to limit the influence of outliers. All regressions control for differential trends based on household size, head characteristics and baseline land holdings. Standard errors reported in parentheses are clustered at the sub-village level and there are 16 clusters. Statistical significance is denoted by stars on the coefficient estimates using: *** p<0.01, ** p<0.05, * p<0.1. These critical values do not account for multiple hypothesis tests, but allow for a small number of clusters. Column (1) reports the mean of the outcome variable at baseline across treatment and control households. Column (2) reports the treatment effect for Model 1, which is the first differenced regression in equation 1. Model 2 splits the treatment group into households that received payments above the median compensation payment and below. Columns (3) and (4) show the coefficient estimates on these dummy variables, using the first differenced regression specification outlined in equation 2 on page 82.

Table B.3: Treatment effect estimates: Household savings and income sources

		Model 1	Model 2	
	(1)	(2)	(3)	(4)
Dependent Variable:	Mean of DV at Baseline	Treatment indicator, {1,0}	Payment above median, {1,0}	Payment below median, {1,0}
Total savings	2,199	58,305***	98,412***	26,964***
		(5,781)	(7,639)	(4,582)
HH lent money to others, {1,0}	0.0717	-0.0103	0.0353	-0.0369
		(0.0361)	(0.0348)	(0.0501)
HH borrowed money from others, {1,0}	0.0512	-0.139***	-0.135**	-0.132***
		(0.0386)	(0.0519)	(0.0383)
HH participates in off-farm activity, {1,0}	0.500	0.0765	0.217**	-0.0243
		(0.0564)	(0.0823)	(0.0475)
HH participates in non-farm wage labour, {1,0}	0.292	0.0849	0.181**	-0.0188
		(0.0519)	(0.0771)	(0.0492)
HH participates in own business, {1,0}	0.243	0.0773	0.151*	0.0425
		(0.0681)	(0.0770)	(0.0781)
Total number of person-weeks this year	12.21	3.302	8.919**	-0.630
		(2.870)	(3.850)	(3.231)
Share of yearly cash income from:				
Sale of crops	0.501	-0.110**	-0.144*	-0.136**
		(0.0433)	(0.0762)	(0.0558)
Sale of livestock	0.129	0.0817*	0.100**	0.0798
		(0.0418)	(0.0430)	(0.0628)
Sale of livestock products	0.0253	-0.00295	0.00327	-0.00220
		(0.0107)	(0.0156)	(0.00905)
Agricultural wage work	0.00866	-0.00975	-0.000801	0.00198
		(0.0201)	(0.0184)	(0.0225)
Non-agricultural wage work	0.135	-0.00848	0.00347	-0.0178
		(0.0271)	(0.0344)	(0.0355)
Government & NGO transfers	0.0655	0.00545	-0.00853	0.0187
		(0.0217)	(0.0206)	(0.0271)
Gifts and remittances	0.0464	-0.0228	-0.0424**	-0.0138
		(0.0161)	(0.0160)	(0.0246)
Land rental or sharecropping	0.00313	0.00511	0.00108	0.0109
		(0.0114)	(0.0113)	(0.0165)
Household business	0.0764	0.0605*	0.0904***	0.0474
		(0.0303)	(0.0305)	(0.0370)

Notes: Each row corresponds to a separate outcome variable. All financial variables are reported in Ethiopian Birr and are top-coded at the 99th percentile to limit the influence of outliers. All regressions control for differential trends based on household size, head characteristics and baseline land holdings. Standard errors reported in parentheses are clustered at the sub-village level and there are 16 clusters. Statistical significance is denoted by stars on the coefficient estimates using: *** p<0.01, ** p<0.05, * p<0.1. These critical values do not account for multiple hypothesis tests, but allow for a small number of clusters. Column (1) reports the mean of the outcome variable at baseline across treatment and control households. Column (2) reports the treatment effect for Model 1, which is the first differenced regression in equation 1. Model 2 splits the treatment group into households that received payments above the median compensation payment and below. Columns (3) and (4) show the coefficient estimates on these dummy variables, using the first differenced regression specification outlined in equation 2 on page 82.

Table B.4: Treatment effects: Main outcomes with and without controls

	(1)	(2)	(3)	(4)
	Mean of DV at Baseline	Treatment indicator, {1,0}	Treatment indicator, {1,0}	Treatment indicator, {1,0}
Total consumption (Birr)	23,225	7,103**	6,324*	6,285*
		(3,028)	(3,144)	(3,148)
Total savings (Birr)	2,199	58,958***	59,214***	58,305***
		(5,948)	(6,006)	(5,781)
Participates in off-farm activity, {1,0}	0.502	0.0817	0.0824	0.0765
		(0.0542)	(0.0556)	(0.0564)
Value of agricultural assets (Birr)	376.8	-46.41	-67.08	-69.47
		(80.00)	(92.42)	(92.37)
Value of non-agricultural assets (Birr)	1,872	5,139*	5,343*	5,404*
		(2,721)	(2,895)	(2,921)
Value of livestock (Birr)	10,954	3,049**	2,428	2,348
		(1,356)	(1,446)	(1,474)
Controls included:				
Household demographics		N	Y	Y
Baseline land owned		N	N	Y

Notes: Each row in the first panel corresponds to a separate outcome variable. Column (1) reports the mean of the outcome variable at baseline across treatment and control. Columns (2) - (4) report treatment effect estimates from a regression that explains the differenced outcome variable. Column (2) reports the treatment effect for the regression without any controls. Column (3) includes all household demographic controls that are included in the main model specified in equation 1: household size, head age, gender, marital status and literacy. Column (4) adds the amount of land owned at baseline as a control variable. All controls enter as trend effects. Financial variables are reported in Ethiopian Birr and are top-coded at the 99th percentile to limit the influence of outliers. Standard errors reported in parentheses are clustered at the sub-village level and there are 16 clusters. Statistical significance is denoted by stars on the coefficient estimates using: *** p<0.01, ** p<0.05, * p<0.1.

Table B.5: Alternate measures of treatment: Main outcomes

	Model 1	Model 2	Model 3		Model 4		Model 5	Model 6
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Treatment indicator	Plot-based treatment indicator	Treatment indicator	Lose house indicator	Payment above median	Payment below median	Total payment (Birr)	log(payment + 1)
Total consumption (Birr)	6,285*	6,681**	4,576	6,969	8,180*	6,438	0.0312*	606.6**
	(3,148)	(2,990)	(3,276)	(6,448)	(3,896)	(4,270)	(0.0171)	(255.1)
Total savings (Birr)	58,305***	49,176***	49,118***	37,039***	98,412***	26,964***	0.594***	5,975***
	(5,781)	(5,948)	(5,457)	(6,511)	(7,639)	(4,582)	(0.0403)	(578.4)
Participates in off-farm activity, {1,0}	0.0765	0.0412	0.0636	0.0524	0.217**	-0.0243	1.38e-06**	0.00963
	(0.0564)	(0.0560)	(0.0548)	(0.155)	(0.0823)	(0.0475)	(5.05e-07)	(0.00564)
Value of agricultural assets (Birr)	-69.47	-112.6	12.63	-334.2***	-202.9	62.81	-0.00115	-7.951
	(92.37)	(104.6)	(87.92)	(112.8)	(156.7)	(100.6)	(0.000946)	(8.815)
Value of non-agricultural assets (Birr)	5,404*	3,623*	5,369*	140.3	11,027*	287.2	0.0670*	565.8*
	(2,921)	(1,718)	(2,978)	(2,538)	(5,189)	(1,106)	(0.0327)	(310.0)
Value of livestock (Birr)	2,348	1,779	2,718	-1,503	2,015	2,846	0.00747	203.4
	(1,474)	(1,293)	(1,602)	(1,852)	(1,330)	(1,936)	(0.00721)	(123.7)

Notes: Each row in the first panel corresponds to a separate outcome variable. Each model (columns) corresponds to the treatment effect from a regression using the main specification, but with different measures of treatment status. Regressions control for household size, head age, gender, marital status and literacy and land owned at baseline. Column (1) reports the treatment effect from the main specification outlined in equation 1, and is based on the household's self-reported treatment status. Column (2) uses an indicator derived from the plot roster section, in which households were asked about the treatment status of each plot they reported at baseline. Columns (3) and (4) report the coefficients on an indicator for self-reported treatment status and an indicator for whether the household will lose their homestead. Columns (5) and (6) split the self-reported treatment status into an indicator for whether the household receives a payment above and below the median compensation payment. Column (7) uses total compensation payment as a continuous treatment effect, with control households receiving zero. Column (8) uses $log(TotalPay + 1)$ instead. All controls enter as trend effects. Financial variables are reported in Ethiopian Birr and are top-coded at the 99th percentile to limit the influence of outliers. Standard errors reported in parentheses are clustered at the sub-village level and there are 16 clusters. Statistical significance is denoted by stars on the coefficient estimates using: *** p<0.01, ** p<0.05, * p<0.1.

Table B.6: Treatment effect for restricted samples: Main outcomes

		HH lives within distance of project area			Treatment status correctly identified		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Full sample	Within 500m	Within 1000m	Within 2500m	Paid within last year	Paid within last year and official correct	Two measures of treatment consistent
	Treatment indicator	Treatment indicator	Treatment indicator	Treatment indicator	Treatment indicator	Treatment indicator	Treatment indicator
Total consumption (Birr)	6,285*	9,127*	8,231*	6,609*	7,057*	3,991	6,833*
	(3,148)	(4,185)	(3,839)	(3,371)	(3,360)	(4,260)	(3,290)
Total savings (Birr)	58,305***	49,758***	59,126***	56,230***	57,290***	60,039***	58,079***
	(5,781)	(8,227)	(7,872)	(5,698)	(6,031)	(4,533)	(5,813)
Participates in off-farm activity, {1,0}	0.0765	0.0736	0.0657	0.0825	0.0562	0.0205	0.0619
	(0.0564)	(0.0911)	(0.0827)	(0.0554)	(0.0505)	(0.0581)	(0.0603)
Value of agricultural assets	-69.47	-227.9	-131.1	-75.28	-109.6	-47.29	-105.0
(Birr)	(92.37)	(154.3)	(130.8)	(94.16)	(115.3)	(122.5)	(110.9)
Value of non-agricultural assets (Birr)	5,404*	2,709	4,942	5,577*	5,006*	3,193*	4,797*
	(2,921)	(1,944)	(3,112)	(2,995)	(2,453)	(1,774)	(2,441)
Value of livestock (Birr)	2,348	3,354*	3,697***	2,545*	2,349	2,540	2,284
	(1,474)	(1,679)	(1,184)	(1,436)	(1,418)	(1,891)	(1,440)
Sample size (approximate)	295	154	192	290	271	214	271

Notes: Each row in the first panel corresponds to a separate outcome variable. Each column shows the treatment effect estimate from the main regression specification, run on a different sub-sample. Regressions control for household size, head age, gender, marital status and literacy and land owned at baseline. Column (1) reports the treatment effect from the main specification outlined in equation 1. Columns (2) - (4) restrict the sample by the distance between the household's home and the nearest point on the project boundary. Column (5) drops any household that reports receiving compensation more than 12 months prior to the date of the follow-up survey. Column (6) also drops any household whose self-reported treatment status at follow-up contradicts the local official's reported treatment status at baseline. Column (7) restrict the sample to household's whose self-reported treatment status is consistent with whether the household reported losing any plots in the plot recall section of the survey. The actual sample size for each regression is sometimes lower than the approximate sample size indicated because a number of outcome variables are missing data. All controls enter as trend effects. Financial variables are reported in Ethiopian Birr and are top-coded at the 99th percentile to limit the influence of outliers. Standard errors reported in parentheses are clustered at the sub-village level and there are 16 clusters. Statistical significance is denoted by stars on the coefficient estimates using: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Chapter 3

Price volatility, cost uncertainty and the option to abandon investments:

A political-economy model of large-scale land deals

Abstract

Land acquisitions for large-scale agricultural projects in developing countries have increased in recent years. Multi-national firms and foreign governments have entered into long term contracts with host countries in which large tracts of land are purchased or leased for commercial agricultural production in exchange for promises of infrastructure development, job creation and rural infrastructure improvement. The profitability of these projects is uncertain, especially at a time of increased agricultural commodity price volatility in world markets. Based on stylized facts about land deals I present a theoretical model of land contracts reached by host governments and foreign investors that explains the policy tradeoff between investment timelines, revenue generation and uncertainty.

When agricultural projects require fixed infrastructure investment and yield uncertain payoffs, firms benefit from being able to complete the fixed investment in stages. If firms can learn more about payoffs by holding off on investment, they effectively hold an option to abandon the project. The value of this option provides a channel by which uncertainty affects the terms of the land contract. When host governments determine the terms of the contract by setting an income tax, a royalty rate and an investment timeline, the value of this option will affect government's optimal policy choice. In particular, I find that if governments benefit a great deal from investment spillovers the optimal contract will be designed to encourage firms not to abandon a project. But, if governments benefit relatively little from investment spillovers, governments will choose contract parameters to extract the value of the firm's option to abandon the project. I end by examining the effect of increasing uncertainty on the government's optimal policy choice.

1 Introduction

Large-scale land acquisition in developing countries has received considerable media attention in the past years.[1] International firms and foreign governments have entered into long term contracts with host countries, often in Africa, in which many thousands of hectares of land are sold or leased for large-scale agricultural investment projects. According to a World Bank report published in 2011, 56.6 million ha of land has been included in large-scale land deals, of which 39.6 million ha is in Africa. Median farm size for the proposed projects is 40,000 ha, but only 20% of the reported deals have progressed to implementation (Deininger and Byerlee, 2012). News of these deals has stirred a vibrant debate, with some sides viewing the contracts as exploitative and others heralding a new era of agricultural productivity. Compensation paid to host countries for the opportunity to use the land often comes as promises for rural infrastructure development and job creation, rather than cash transfers, large lease or purchase fees or high levels of taxation (Cotula et al., 2009; Deininger and Byerlee, 2011). What factors drive these deals and how do the determinants affect the way that contracts are structured? Based on stylized facts about land deals drawn from a number of policy reports, I model the strategic interaction between governments and investors negotiating a land deal and draw some conclusions about the relationship between uncertain profits and the key contract parameters observed in many land deals. In particular I model the trade-off between policies that generate government revenue and those that ensure that firms keep to their investment commitments.

[1] "How food and water are driving a 21st-century African land grab," *The Observer.* March 7 2010.; "The scramble for land in Africa," *The Economist.* May 21 2009; "Foreign fields," *Financial Times.* August 20 2008. are among some of the many recent new articles on the topic.

I develop a model of investment that links uncertainty in profits to an investment timeline and tax rates agreed by governments and investors. In particular, I consider how a government might choose a timeline for investment and a tax rate when the investment is both crucial for the operation of large-scale agricultural projects and of inherent benefit to the host government. Such investments might include, for instance, new transportation, storage or irrigation infrastructure that benefit neighbouring farmers. Large agricultural projects in Africa yield uncertain profit flows both because of increased price volatility in world food markets as well as uncertainty over the profitability of this model of agriculture in many African countries. Uncertain profits affect the prospective value of an agricultural investment project, especially when a large part of the investment has a fixed cost component that can be completed in stages[2] and when a firm is able to prematurely abandon the project. Explicitly modeling large-scale agricultural projects as having fixed costs provides an essential link between price volatility and both the level of agricultural investment and the terms of the land deal.

1.1 Stylized facts about land deals

Perhaps the best publicized and most widely cited land deal in recent years was the proposed 99-year lease of 1.3 million hectares of farmland in Madagascar by the Korean company Daewoo Logistics, a deal which contributed to the March 2009 coup and ultimately collapsed following the change of government (Cotula et al., 2009). Although atypical, this deal brought attention to other international investors acquiring large tracts of land in developing countries. Many reported deals, although much smaller than the Daewoo-Madagascar deal, still involve land

[2]For example, the timeline might split the fixed cost of investment I into components I_1 and I_2, where I_1 must be completed in the first stage and I_2 in the second stage. I refer to this as the phasing of investment.

transfers in the order of tens of thousands hectares.[3] The international investors involved are often multi-national firms or foreign governments from food-poor countries in Asia and the Middle East.

I limit the scope of the issue by focusing on land deals in Africa. Land deals vary significantly across the world depending on the types of crops that can be grown, the land holding systems and the level of development in the agricultural sector(Deininger and Byerlee, 2012). The agricultural sector is a significant source of income and employment in African countries, with 16.5% of value added in GDP coming from agriculture, (relative to 3% in the rest of the world) and an average of 51% of people employed in the agricultural sector (compared with 30% worldwide) (UNCTAD, 2009). Across many African countries, land rights are derived through a complex mix of customary rights, informal title and state ownership and functioning land markets are often missing. This creates an important role for governments in facilitating land deals and is reflected in the important role played by investment promotion agencies in providing land (Pande and Udry, 2006). Arezki et al. (2013) report that in Africa, 90% of the sellers in verified land deals are governments. I also restrict the focus of this analysis to African countries because of the availability of a preliminary data analysis done by the FAO/IEED/IFAD. The FAO report contains a description of land deals in Mali, Sudan, Ethiopia, Madagascar and Tanzania, which I draw upon to inform my stylized facts.

Land deals tend to involve large tracts of land, transferred to investors at low rental and tax rates for long periods of time. Mali's Office du Niger region actively participates in land deals. It is a region of Mali, along the Niger

[3]The scale of operation that is proposed in these deals is in itself surprising. Many of the countries in which these deals are occurring are land abundant, but the size of the proposed farms is much larger than farms in land abundant developed countries (Deininger and Byerlee, 2012). There is also little evidence that very large farms are efficient, with the exception of plantation crops and large ranches that benefit from high degrees of coordination along the supply chain (Eastwood et al., 2010).

River that is particularly well suited to irrigation improvement. Publicized deals include a 2007 deal with PetrOtech/Agro-Mali for 10,000ha for the production of bio-diesel feedstock, 100,000ha deal, part of a 2009 bilateral investment treaty, with a subsidiary of the Libya Africa Investment Portfolio and an application for 100,000ha from Lonrho, a pan-African conglomerate. Much like many of the land deals reported, land-rents are minimal and taxes are low. In Mali, the fee was between 6$ and 12$ per hectare per year (Cotula et al., 2009). Ethiopian land rents were similar, ranging from 3$ per hectare to 10$ per hectare. A 2008 deal between Tanzania and CAMS Group, a UK energy company, involved a 45,000-hectare lease where no land rents are being paid(Cotula et al., 2009). By comparison, the US Department of Agriculture reported that average land rents for cropland in the USA are about 225$/hectare per year. The low rents being paid on land suggest that governments are not deriving a great deal of revenue from the rents they charge. Taxes are also minimal, with most countries offering tax breaks for the first five years, and some countries charging no tax at all(Cotula et al., 2009). There is also a discrepancy between the reported size of the deal and the actual size of the project once it begins. According to the Tanzanian investment agency, by March 2009, 4 million hectares of land were requested for bio-fuel investments, while only 640,000ha were eventually allocated, of which 100,000ha have received official rights of occupancy(Sulle and Nelson, 2009).

Another evident trend in many land deals is that investment partners agree to improve the land through irrigation, infrastructure projects and other large-scale investments as part of the contract. The contracts often stipulate timelines for these investments that must be adhered to as a condition of maintaining the lease. Mali's Office du Niger requires leaseholders to develop irrigation as part of their lease commitments and stipulates consequences for not adhering to the investment timeline. The PetrOtech/Agro-Mali deal gives the firm 3 years to develop irrigation following the initial feasibility study, after which the project lease can be renewed if at least 50% of the investment outlay has been made.(Cotula et al., 2009) Other deals mandate public infrastructure projects as part of the contract. The 100,000ha Libya-Mali deal includes the construction of a 40km irrigation channel and the building of a 40km road valued at 48 million US dollars (?). Kenya and Qatar signed a deal whereby Qatar leases 40,000ha of land in exchange for financing a 2.3 billion dollar port (Cotula et al., 2009). Similar investment arrangements were reported in a deal between Sudan and Syria where the Syrian government was granted a 50-year lease on 12,600ha provided they develop irrigation for 4,200ha of land outside of the project area. The deal included an investment timeline allowing one year for a feasibility study, three years for the construction of irrigation infrastructure and two years for reaching planned production levels (Cotula et al., 2009). These land deals often contain large investment components and mandated timelines, outlining the timing of the large infrastructure investments.

Host government agencies are an integral part of the land deals in Africa, enabling foreign investors to bypass the intricacies of complex land systems and facilitating deals where formal land markets are underdeveloped. In Ghana, Mali and Madagascar investment promotion agencies facilitate the acquisition of permits, licenses and authorizations. Tanzania's investment agency has a land bank with 2.5 million ha in it that is suitable for investment projects. Host governments, in addition to facilitating foreign investment in land seek to promote it through generous tax holidays, thus making the initial cost of acquiring land even lower for foreign investors (Cotula et al., 2009; UNCTAD, 2009).

Finally, most of these land deals were reached in the last decade, with most deals being announced in the last seven years. Over the years in which these deals were struck world food prices were significantly more volatile than they have been in the decade before (Deininger, 2013). Figure 1 shows the change in monthly FAO food price indices from 1990, illustrating the recent increase in volatility. This suggests that looking at the role of changing volatility may be a useful exercise. Headey and Fan (2008) suggest that this recent increase in volatility as well as the latest spike in prices is attributable to increased demand for biofuel, volatile oil prices and environmental factors, such as climate change. There is also major uncertainty about the profitability of large-scale commercial

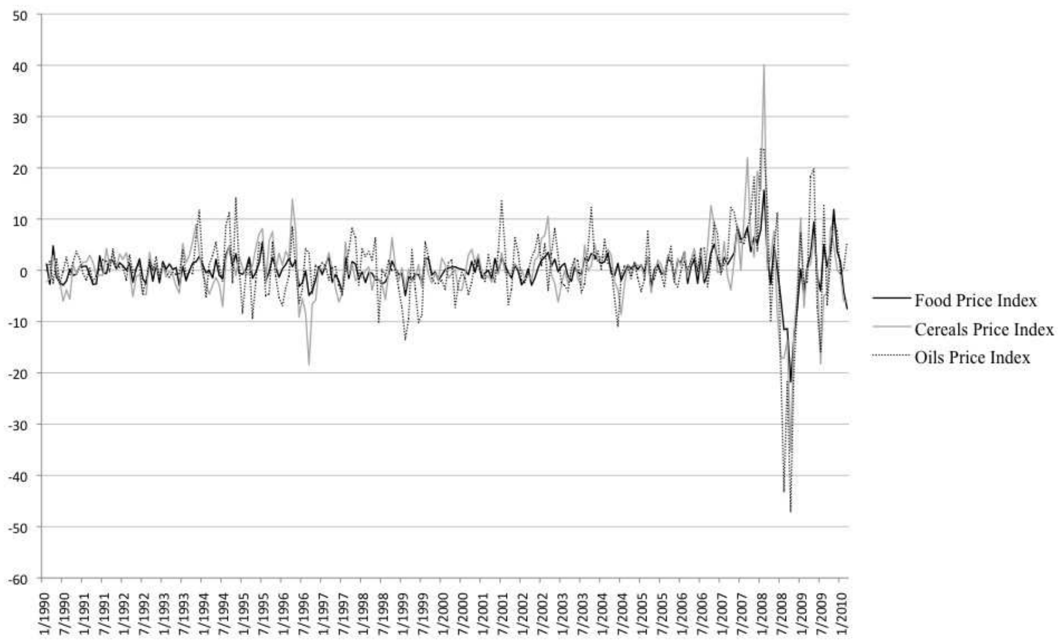


Figure 1: Change in monthly FAO food price index

agriculture in Africa; there is some evidence that large farms have been successful in Latin America (Deininger and Byerlee, 2012), but historical examples from North America[4] and previous experience with large-scale agriculture Africa have

[4]Drache (1964) discusses the boom in bonanza farms in the Red River valley in North Dakota. Land was sold at very cheap prices in the 1870s to investors from the east coast of America who planned to operate large-scale farms, using corporate management structures and the newest machinery. The average farm size was about 2000 ha, with the biggest farms reaching 18,000 ha. Ultimately, the scale of operation was too large and the projects became unprofitable when

not been successful(Binswanger et al., 1995; Deininger and Binswanger, 1995; Deininger and Byerlee, 2012).

What can be drawn from this anecdotal evidence? Over a period of time in which food prices were increasing, but becoming increasingly volatile there was an increase in large-scale land acquisitions. Many of these deals involve large tracts of land that are transferred to international investors at relatively low cost for long periods of time (often in excess of 50 years). Host governments and foreign investors each bring unique resources to the bargaining table that enable these deals to go ahead; governments can navigate complicated land rights issues and are able to facilitate acquisition of large parcels of land while foreign investors bring expertise, capital and access to finance. Although investors are often exempted from taxation there are large infrastructure investments that go along with acquiring the land, which sometimes have indirect benefits for the host government. Contracts contain strict guidelines articulating when and how the large infrastructure investments need to be made. A theoretical model seeking to explain how these deals look should consider these observations and build them into the model.

1.2 Summary of the model

I model large-scale agriculture as an investment opportunity that exists through the cooperation between a government and a foreign investor. I focus on policy tools that raise revenue and regulate investment timelines. The projects are characterized by large fixed infrastructure investments, which are undertaken by the foreign investor rather than the government. International investors may have better access to capital, greater expertise, or may plan to operate a farm at a scale that is large enough to capture enough of the benefits from a public infrastructure

wheat prices collapsed at the end of the 19th century. Eventually the bonanza farms were broken into small, diversified, family run units.

investment. In addition, foreign investors may bring better technology enabling them to extract higher crop yields from the land.[5] Host governments act as both facilitators and regulators of land acquisition. The interaction between investors and host governments, and the resulting structure of land deals, is a function of volatility in profits, government priorities, and the spillover benefits from investment. If governments stand to benefit from infrastructure investments made by international investors, they may find it in their interest to trade off potential future revenue against up-front infrastructure investment, which can commit firms to continue the project. Conversely, government policy that allows flexibility in the phasing of fixed investment increases the value of a deal to investors, allowing for higher taxes. The value of phasing an investment, however, depends on the uncertainty of profits. By explicitly modeling flexibility in the phasing of fixed investments, I give insight into how price volatility in world food markets, or other sources of uncertainty can impact the political economy of large-scale land deals.

Flexibility in the timing of investment provides the link between uncertain payoffs and the investment decisions of international investors, and thus, the outcome of government policy.[6] Investors trade off the cost of making an investment now with the prospect of some unknown payoff in the future. If investors are unable to complete the infrastructure investment in stages, they are required to make their total investment with less knowledge of future returns than if phasing is permitted. Conversely, allowing investors the opportunity to delay part of their investment opens up the possibility that in the interim they learn something new

[5]Collier and Venables (2012) argue that investors might also provide public goods through experimenting with new production techniques or by demonstrating profitable agricultural projects that could benefit firms that follow.

[6]The analysis of investment under uncertainty is a well developed field. Dixit and Pindyck (1994), Abel and Eberly (1996), McDonald and Siegel (1986) and Abel et al. (1996) all examine various models of investment where flexibility is somehow limited. The relationship between uncertainty and investment depends crucially on the nature of the inflexibility. My model assumes that a fixed investment can be completed in stages, but that once any portion of the investment is sunk it cannot be recouped by the firm. This prohibits the possibility that the investor could sell their stake to another investor.

about the future that enables them to make a better investment decision (for example, in some cases, it might be optimal to abandon a large-scale agricultural project). Investors benefit if they can invest after learning that a project will be profitable and avoid investing if they learn that the project is unprofitable.

Increasingly uncertain project profits affect the contracts offered by governments through the policy tools available to them and the way in which governments benefit from these tools. Governments benefit directly from a project by generating tax revenue[7] and indirectly through investment spillovers. With respect to taxation, investors and governments share some incentives. In particular, both governments and investors can benefit from flexibility in the timing of investment; projects that are more profitable in expectation may justify higher taxes, generating greater revenues for governments. However, governments face a policy tradeoff when they are able to affect the timing of the investment project. In particular, if governments benefit from having infrastructure investment occur, regardless of whether the investment project is abandoned and if they can successfully regulate the investment timeline, then they may be willing to forgo some tax revenue in order to ensure that investors commit to completing a project. This incentive to limit flexibility in the phasing of investment is at odds with the profit making incentives of firms and, thus, the revenue generating incentives of government.

A key contribution of this paper is to highlight how the tradeoff between generating revenue and securing investment is a fundamental determinant of how land contracts are structured. Greater variability in profits can affect the overall value of an investment project in this framework, and, as such, influence policy

[7] Governments choose both an income tax instrument and a royalty rate. The inclusion of a royalty is intended to reflect the prevalence of this tax instrument in other resource contracts. I abstract from a number of other potential policy tools that could be used by the government, such as lump sum transfers or land rental prices. I also ignore the possibility that governments can vary the amount of land offered to the investor.

outcomes. Increased profits can increase potential government revenue but may do so at a cost to the level of upfront investment. In this way, price volatility or uncertainty about costs affect the decisions of governments through its effect on the value of an investment project where the timing of investments is flexible. The tradeoff between tax rates and the degree of flexibility offered to investors may explain why taxes and transfer fees seem low in large-scale land deals. Many contracts make land use rights contingent on investors completing an investment plan in a timely manner. Mandating investment plans limits the flexibility of large investment decisions and thus reduces the overall expected value of a project to investors. However, making firms commit a large part of the infrastructure investment right away prevents them from abandoning the project prematurely and assures governments that investment will occur, but at the expense of future tax revenue.

2 Land deals and investment under uncertainty

2.1 Model set-up

I model large-scale agricultural investment projects using a simple real options model of investment adapted from Dixit and Pindyck (1994), in which an investment project requires a fixed investment and yields an uncertain future payoff. Investors must sink a portion of the fixed investment before learning more about the value of payoffs. The land deal, agreed by the government and investor, stipulates an investment timeline characterized by the portion, ρ, of fixed investment that a firm must sink in the initial period and a royalty rate and profit tax that are collected upon completion of the project and commencement of production.[8] Governments benefit both from revenue generation and investment spillovers. International investors are not bound by their investment commitment giving them the option to prematurely abandon the investment project in the event that they learn that it will be unprofitable. Modeling large-scale land deals in this way creates a link between uncertain profits, investment decisions and optimal policy, which would not exist if risk neutral firms simply compared the discounted present value of an investment project with the fixed, up-front, cost of investment. This link creates a tradeoff between policy objectives aimed at revenue generation and those aimed at ensuring investment will be completed. There is an extensive literature that uses more sophisticated real options models in a developing country context to explain the link between levels of foreign direct investment and expropriation risk, foreign investment through joint venture and forestry investment decisions in developing countries, among other things.[9] This paper, in contrast to the others, looks at a setting where the firm holds an option to abandon the project.

Firms and governments share an investment opportunity, which is specific to their relationship since each party brings unique assets to the relationship which cannot be found elsewhere. Firms may have easier access to capital, expertise in operating large scale agriculture projects or capacities for infrastructure improvement. Whereas governments, although less able to invest in the necessary infrastructure, have control over land resources and are in a position to dictate how land is used. The government thus acts as both the facilitator of large scale land acquisition and the agency granting the right to invest. I solve the model under the assumption that many firms are competing for the right to undertake the

[8]There are other policy instruments that governments might use, such as rental fees for land or lump sum transfers. I abstract away from these and consider a royalty and an income tax as the only revenue raising instruments.

[9]Leiblein (2003), Yap (2004), Li and Rugman (2007), Engel and Fischer (2008) and Lukas (2007) are all recent examples of options models applied to developing country contexts. There is a wealth of other options models using complex analytical and numerical techniques to model the link between investment and uncertainty.

investment project, which allows governments to dictate the terms of the contract. This is a simplifying assumption that allows for a tractable analysis.

The rest of the section lays out the model, solves for the optimal behaviour of firms and then establishes the optimal government policy. I end with a discussion and analysis of the impact of changing uncertainty on optimal policy.

2.1.1 Description of the game

The game unfolds over three periods. In the initial period, $t = 0$, the value of the investment project is unknown. The uncertainty is partially resolved in the subsequent period, $t = 1$, when the marginal cost of production is revealed and the uncertainty is fully resolved in the final period, $t = 2$, when price is revealed. The variable profits and total profit of the investment project are described by:

$$\pi_t = p_t - c(1 + \theta_t) \text{ (variable profits)} \tag{2.1}$$

$$V_t = \pi_t - I \text{ (total profit)}$$

p_t is distributed according to a continuous CDF $F(p_t)$

c is a known marginal cost subject to a random error, θ_t

$$\theta_t = \left\{ \begin{array}{ll} 0 & \text{w/ } \Pr(\theta_t = 0) = q \\ \theta & \text{w/ } \Pr(\theta_t = \theta) = (1 - q) \end{array} \right\} \text{ where } \theta > 0$$

I is a known fixed cost of investment

Figure 2 on the next page summarizes the timing of the game and the sequence in which information arrives and decisions are made. Each column corresponds to one time period and the arrival of information and the decisions made in each period occur in the order they are listed.

Actor:	t=0	t=1	t=2
Nature		1.1 Choose marginal cost: Low cost with Prob = q −or− High cost with Prob = (1−q)	3.1 Choose price: Price drawn from continuous distribution: p~F(p)
Government	0.1 Set policy parameters		
Firms	0.2 Given policy parameters: Start fixed investment −or− Not participate	1.2 Given cost: Finish fixed investment −or− Abandon project	2.2 Given price: Produce −or− Not Produce

Figure 2: Timing of the game, with possible actions

Firms make choices at each stage of the game while government only makes choices at $t = 0$. In the initial period government presents firms with a contract, which consists of an income tax, τ_i, a royalty rate, τ_r and an investment plan, ρ, stipulating the portion of fixed investment, I, that must be sunk at $t = 0$. The policy parameters are set before firms make any decisions and I assume that governments can commit to their choices. If firms accept the contract, they must sink the portion ρ of the fixed cost of investment I. In the next period firms discover the marginal cost of production. At this point, firms can either abandon the project losing ρI or continue the investment by sinking the remaining $(1-\rho)I$. In the final period firms learn the price and decide whether to produce.

The final payoffs to firms and government depends on the choices made by each of the actors. If firms choose to produce at $t = 2$, their payoff is:

$$V_{t=2} = [(1 - \tau_r)p_t - c(1 + \theta_t)]\,(1 - \tau_i) - I. \tag{2.2}$$

p_t and θ_t are state dependent variables that are determined by the beginning of period $t = 2$. Payoffs to the government consist of the tax revenue and any spillover benefits from having the fixed investment occur in their country:

$$U_{t=2} = \tau_r p_t + \tau_i \left[(1 - \tau_r)p_t - c(1 + \theta_t)\right] + u\left(\rho I + (1 - \rho)I\right), \tag{2.3}$$

$$\text{where } u'(\cdot) > 0 \text{ and } u''(\cdot) < 0.$$

The final term, $u\left(\rho I + (1 - \rho)I\right)$, captures the benefit from any positive spillovers from the fixed investment and is assumed to display diminishing returns. Firms that choose to abandon a project at any stage lose whatever portion of the fixed investment cost, I, that has been sunk. Importantly, even if firms do not produce, governments still benefit from any fixed investment that has already been made. The final payoffs are illustrated in the decision tree in figure 3. Nature reveals

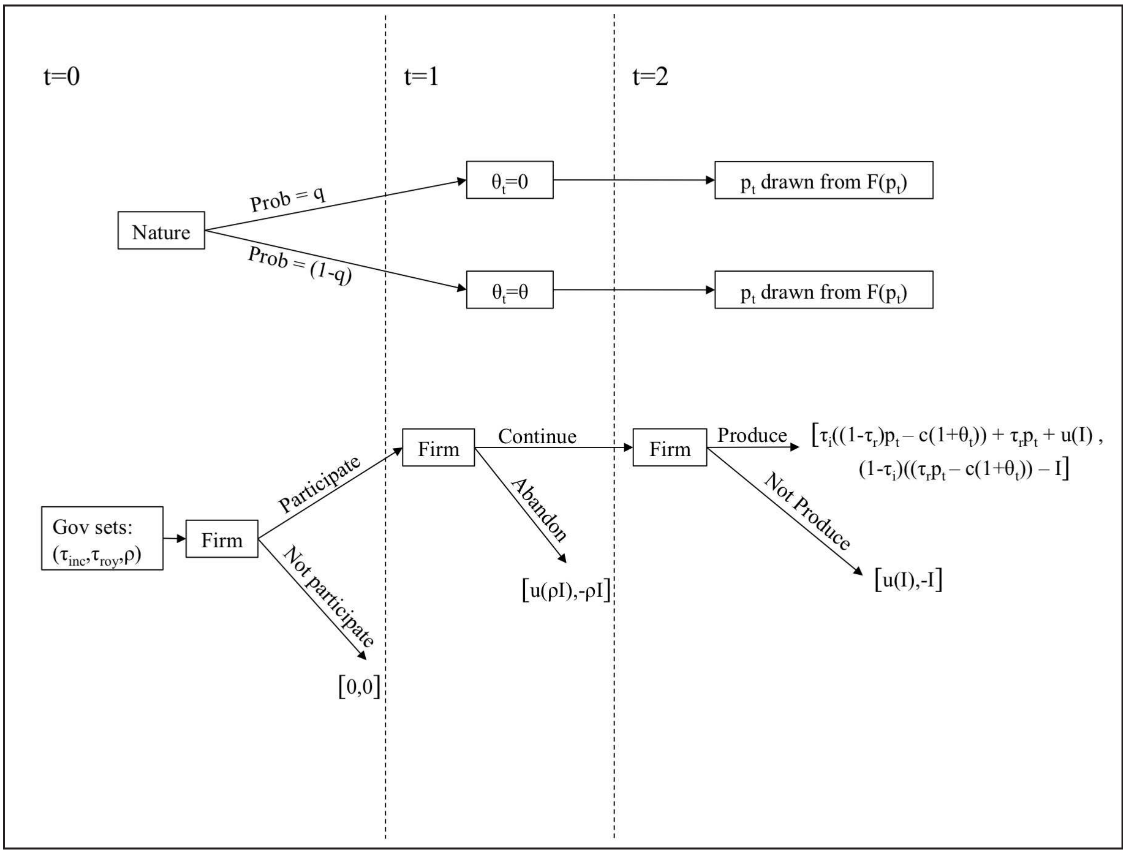


Figure 3: Game tree describing choices, states of the world and payoffs

the random variables at the beginning of each period and the final payoffs at the end of each decision path are denoted by $[U, V]$, where U is the total payoff to government and V is the total payoff to the firm.

2.1.2 Key assumptions

Many key assumptions in the model have been stated explicitly above. The most important are: 1) many firms compete for the opportunity to invest, thus allowing governments to dictate the terms of the land contract without any bargaining over the economic surplus. 2) I assume common knowledge of prices and marginal costs, which rules out the possibility that firms can misreport their profits in order to avoid taxation. 3) I assume that governments can credibly commit to the contract parameters they declare in their contract. 4) I assume that governments always benefit to some degree from the fixed investment, even if it is not complete and that governments are not able to undertake the investment themselves. If governments were able to undertake the infrastructure investment firms could pay a lump sum transfer or user fee to governments to fund the investment; this assumption rules out that possibility.

2.2 Solving for optimal policy

Government sets policy parameters prior to any actions on the part of firms, but do so with the knowledge of the firm's optimal actions at each subsequent stage of the game. These actions are: 1) to participate or not at $t = 0$, 2) to continue or abandon the investment at $t = 1$, and 3) to produce or not at $t = 2$. I solve for the optimal policy, $(\tau_i^*, \tau_r^*, \rho^*)$, using backward induction.

2.2.1 Firm's optimal behaviour at $t = 2$ (production)

At $t = 2$, firms have completed the fixed investment cost, I and must decide whether to produce or not. Firms do so with full knowledge of both price, p_t and the level of marginal cost, determined by θ_t and take (τ_i, τ_r, ρ) as given. Recall that marginal cost can take two possible values. When costs are low, $\theta_t = 0$, firms will produce iff:

$$V(\theta_t = 0 \text{ and } \textit{produce}) \geq V(\theta_t = 0 \text{ and } \textit{not produce}), \tag{2.4}$$
$$[(1-\tau_r)p_t - c]\,(1-\tau_i) - I \geq -I,$$
$$p_t \geq \frac{c}{(1-\tau_r)}.$$

Likewise, when costs are high, $\theta_t = \theta$, firms will produce iff:

$$p_t \geq \frac{c(1+\theta)}{(1-\tau_r)}. \tag{2.5}$$

The firm's decision to produce at $t = 2$ is made with full knowledge of the stochastic parameters and government policy. Since, the fixed investment component, I, has already been sunk, firms will only produce if the price, after paying royalties, is weakly greater than the marginal cost of production. A firm's decision at $t = 2$ is unaffected by the size of the fixed investment, which has already been sunk, or the income tax, which is paid after deducting marginal cost; royalties are the only policy tool that affect firm behaviour at this stage.

2.2.2 Firm's optimal behaviour at $t = 1$ (continuation)

At $t = 1$, firms decide whether to complete or abandon the fixed investment. Firms know θ_t but do not know prices. In order to continue to the production stage, firms must complete the sunk cost investment by paying $(1-\rho)I$. They will do so only if the unconditional expected payoff at $t=1$, given their future optimal production decision, exceeds the certain payoff from abandoning the project, $-\rho I$. Equations (2.4) and (2.5) above give cost-dependent threshold conditions that are used to construct both the probabilities that a firm will produce at $t=2$ and the expected value of producing at $t=2$, conditional on producing. Note that below $V(\cdot)$ refers to firm's final payoffs and may be conditional on completing or abandoning the project, while π refers explicitly to variable profits earned by the firm at $t=2$, before paying income tax, but after paying royalties. Superscript G signifies a good cost realization (where marginal cost is low $\theta_t=0$) and superscript B signifies a bad cost realization (where marginal cost is high, $\theta_t=\theta$). The probability of production at $t=2$, depending on the cost realization, is:

$$\Pr(prod.|\theta_t=0) = 1-F\left(\frac{c}{(1-\tau_r)}\right),$$
$$\Pr(prod.|\theta_t=\theta) = 1-F\left(\frac{c(1+\theta)}{(1-\tau_r)}\right).$$

The unconditional expected payoff from completing the investment when costs are low is[10]:

$$\begin{aligned} E_{t=1}[V^G|C] &= \Pr(prod.|\theta_t=0)\Big((1-\tau_i)E_{t=1}[\pi^G|prod.]-I\Big)+(1-\Pr(prod.|\theta_t=0))(-I) \\ &= \Pr(prod.|\theta_t=0)\Big((1-\tau_i)E_{t=1}[\pi^G|prod.]\Big)-I \\ &= (1-\tau_i)E_{t=1}[\pi^G]-I \end{aligned} \tag{2.6}$$

Firms will complete the investment following a good cost realization if the unconditional expected payoff from completing the investment exceeds the cost of

[10]For ease of notation I work with the unconditional expected variable profits at $t=1$, $E_{t=1}[\pi^G]$. It should be noted that $E_{t=1}[\pi^G]$ is a function of the royalty rate.

abandoning the project:

$$E_{t=1}[V^G|C] \geq -\rho I, \tag{2.7}$$
$$(1-\tau_i)E_{t=1}[\pi^G] - I \geq -\rho I,$$
$$(1-\tau_i)E_{t=1}[\pi^G] \geq (1-\rho)I.$$

The final line illustrates the trade-off faced by firms at this stage of the game; firms will continue only if the cost of completion, $(1-\rho)I$, is lower than the expected profits conditional on producing, weighted by the probability of producing. I refer to this condition as the "continuation constraint" and it summarizes the optimal behaviour of firms at $t = 1$. An analogous continuation constraint exists when high costs are realized:

$$(1-\tau_i)E_{t=1}[\pi^B] \geq (1-\rho)I. \tag{2.8}$$

How do the continuation constraints relate to each other and when is it optimal to abandon an investment? If the continuation constraint is satisfied when costs are high, it must also be satisfied when costs are low. This follows because the expected variable profits are greater when the cost realization is low.[11] Thus, when condition (2.8) is satisfied, it is optimal for firms to complete the investment regardless of whether the cost realization is good or bad. However, if (2.7) is satisfied but (2.8) is not, firms will continue the project when costs are low but abandon the project when costs are high. Depending on the parameters of the model, firms may switch from a strategy of continuing the project regardless of cost, to one of abandoning high-cost projects.

[11]Although prices are not known at $t = 1$, a low marginal costs increases the likelihood of production and the expected variable profit, conditional on production.

The option to abandon a high-cost project is valuable to firms when it enables them to avoid completing a loss-making project. This option has no value when firms find it optimal to continue the project regardless of costs. The investment timeline, ρ, and the level of taxes affect the value of the option to abandon the project and the firm's optimal strategy at $t = 1$. When ρ is low, firms are only required to sink a small investment to reach $t = 1$, where they may choose to abandon the high cost project. When ρ is high, firms will have sunk most of the fixed investment by the time they learn about the profitability of the project at $t = 1$ and are thus more likely to continue regardless of the cost realization. For each pair of tax policy parameters, (τ_i, τ_r), there may be a $\rho = \rho^{switch}$ above which firms will no longer find it optimal to abandon the project following high cost realizations and will instead see the project through to completion. However, firms only arrive at this decision point if they are willing to sink ρI at $t = 0$.

2.2.3 Firm's optimal behaviour at $t = 0$ (participation)

At $t = 0$, firms decide whether to participate at all by investing ρI. This decision depends on the expected value of the project at $t = 0$, conditional on the optimal firm behaviour in subsequent periods. Since marginal cost can take two possible values at $t = 1$, firms must specify a plan of action for each realization, which I denote $[action^G, action^B]$, where G stands for good realizations of cost and B for bad. Firms will choose to sink ρI in two cases: first, if it is optimal to continue regardless of cost, $[C, C]$ and second if it is optimal to continue only for good realizations of cost and abandon otherwise, $[C, A]$.[12] Firms will accept the proposed contract if the unconditional payoff from one or both of these optimal action plans is weakly greater than zero. This condition is referred to as the

[12]Although there are actually four possible strategies at $t = 1$, the remaining two strategies are not on the optimal path. First, if firms abandon the project regardless of costs, $A, A]$, it is never optimal to participate. As shown above, if it is optimal to continue in high-cost states, it must also be optimal to continue in low-cost states. Thus, $[A, C]$ is a dominated strategy.

"participation constraint":

$$E_{t=0}[V] = \max\{E_{t=0}[V|C,A], E_{t=0}[V|C,C]\} \geq 0 \tag{2.9}$$

Where,

$$\begin{aligned} E_{t=0}[V|C,C] &= \Pr(\theta_t = 0)E_{t=1}[V|\theta_t = 0] + \Pr(\theta_t = \theta)E_{t=1}[V|\theta_t = \theta], \qquad (2.10)\\ &= qE_{t=1}\left[V^G|C\right] + (1-q)E_{t=}\left[V^B|C\right], \\ &= (1-\tau_i)\left(qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B]\right) - I, \end{aligned}$$

And

$$\begin{aligned} E_{t=0}[V|C,A] &= \Pr(\theta_t = 0)E_{t=1}[V|\theta_t = 0] + \Pr(\theta_t = \theta)(-\rho I), \qquad (2.11)\\ &= qE_{t=1}\left[V^G|C\right] + (1-q)(-\rho I), \\ &= q\left((1-\tau_i)E_{t=1}[\pi^G] - I\right) + (1-q)(-\rho I). \end{aligned}$$

Increasing τ_i or τ_r reduces firm profits and a sufficient increase will prevent firms from participating in the land deal. The effect of the investment timeline, ρ, on participation is ambiguous and depends on the exogenous parameters of the model (which are discussed in the next section). In some cases, firms will choose not to participate if ρ is too high. In other cases, ρ will not affect the decision to participate at all, rather it will only affect the firm's decision to abandon or continue the project at $t = 1$. However, all three policy parameters affect the real value of holding the option to abandon the investment project following bad realizations of cost. The real option value is defined as the difference between the expected value of the project when the firm has the option to abandon the project and when it does not[13]:

$$ROV = E_{t=0}[V|C,A] - E_{t=0}[V|C,C], \quad (2.12)$$
$$= (1-q)(-\rho I) - (1-q)((1-\tau_i)E_{t=1}[\pi^B(\tau_r)] - I),$$
$$= (1-q)[(1-\rho)I - (1-\tau_i)E_{t=1}[\pi^B(\tau_r)]].$$

The value of the option is decreasing in ρ; the more investment a firm must sink, the lower the potential savings when a firm abandons the project. Similarly, the value of the option is increasing in the income tax, τ_i and the royalty rate, τ_r, since both policy tools reduce the profitability of continuing the project. Both of these policy tools reduce the payoff to the firm, making it less attractive for firms to continue to the production stage if costs are high[14].

2.3 Solution and optimal policy

Governments are risk neutral and can choose any combination of policy parameters prior to the firm's actions at $t = 0$. In states of the world where firms produce, governments collect income tax and royalties and derive benefit from any fixed investment that has been sunk. Subject to the optimal behaviour of firms, governments choose τ_i, τ_r and ρ to maximize their expected payoff at $t = 0$. Government's expected payoff depends on whether firms will continue regardless of cost or abandon bad projects. These payoffs are given by:

$$E_{t=0}[U|C,C] = \tau_r \left(qE^G_{t=1}[p_t] + (1-q)E^B_{t=1}[p_t]\right) + \tau_i \left(qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B]\right) + u(I),$$
$$E_{t=0}[U|C,A] = q\left(\tau_r E^G_{t=1}[p_t] + \tau_i E_{t=1}[\pi^G] + u(I)\right) + (1-q)u(\rho I).$$

[13]Strictly speaking, the value of the option is positive if $E_{t=0}[V|C,A] > E_{t=0}[V|C,C]$ and zero otherwise.

[14]Income taxes and royalties also reduce the firm's expected profits when costs are low at $t = 1$, but because it is not optimal to abandon the project in this state of the world the real option value is not affected.

Case	State dependent profit: $E_{t=1}[\pi^B] - I$	$E_{t=1}[\pi^G] - I$	Expected value if firm must complete investment: $qE_{t=1}[\pi^G]+(1-q)E_{t=1}[\pi^B]-I$	Firm's action plan at $t = 1$:
I	< 0	< 0	< 0	do not participate
II	< 0	> 0	< 0	$[C, A]$
III	< 0	> 0	> 0	$[C, A]$ or $[C, C]$
IV	> 0	> 0	> 0	$[C, C]$

Table 1: Conditions on exogenous variables and resulting outcomes

The production condition at $t = 2$, the continuation constraints at $t = 1$ and the resulting participation constraints at $t = 0$ describe the optimal behaviour of firms and therefore the feasible set of policy parameters available to the government. Governments optimize policy subject to this feasible set and, in some cases, may find it optimal to induce firms to either abandon or continue high-cost projects at $t = 1$. The potential outcomes of the model, and, in particular, the type of strategies that the government will induce firms to pursue at $t = 1$ depend on the exogenous parameters of the model. This section restricts those exogenous parameters to a central case, defines the feasible set of policy parameters and solves for the government's optimal policy.

2.3.1 Exogenous parameters and feasible policy parameter space

The exogenous conditions that affect the outcome of the model and the firm's decision to continue or abandon the project at $t = 1$ are outlined in table 1. There are four possible cases that can arise, which depend on the value of expected variable profits at $t = 1$ and the expected value of the investment project at $t = 0$ if firms were forced to complete the fixed investment. I restrict my analysis to Case III because it is most interesting. Here, the optimal government policy will either induce firms to abandon high-cost projects or continue them to completion, depending on the relative importance of investment spillovers and the degree of uncertainty in the project. In contrast, the other cases never involve a switch in strategy that is both optimal for firms and governments: in case I, firms never participate, in case II, firms cannot be induced to complete bad projects and in case IV, governments never find it optimal to have firms abandon a project. A more detailed discussion of these other three cases is left to appendix A.

Case III requires that the expected value of the investment project, if firms were forced to complete the investment at $t = 0$, is positive:

$$\Big(qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B]\Big) - I > 0, \tag{2.13}$$

and that the high-cost, bad project is loss-making in expectation at $t = 1$:

$$E_{t=1}[\pi^B] - I < 0. \tag{2.14}$$

Under these assumptions, firms will find it optimal to switch their plan of action at $t = 1$ depending on the government's choice of ρ and the level of taxation.

Figure 4 plots the feasible set of (ρ, τ_i) parameters, holding τ_r constant for case III. The boundaries of the continuation constraints for both high-cost and low-cost realizations are represented by the upward-sloping lines plotted in red; these constraints correspond to equations 2.7 and 2.8. For any (ρ, τ_i) pair to the southeast of a boundary, firms will find it optimal to complete the investment project given the cost realization. The area between the two red, upward-sloping lines represents the (ρ, τ_i) pairs for which firms find it optimal to complete the project following good cost realizations and abandon the project following bad ones. As ρ increases or as τ_i decreases firms switch from a strategy of abandoning high-cost projects to one of completing the project regardless of cost. The boundary of the participation constraint is represented by the bold, kinked line and depends on the firms optimal strategy at $t = 1$. The boundary corresponds to the zero profit condition, where expected profits at $t = 0$ are equal to zero. The kink coincides with the point at which firms switch their optimal strategy at $t = 1$.

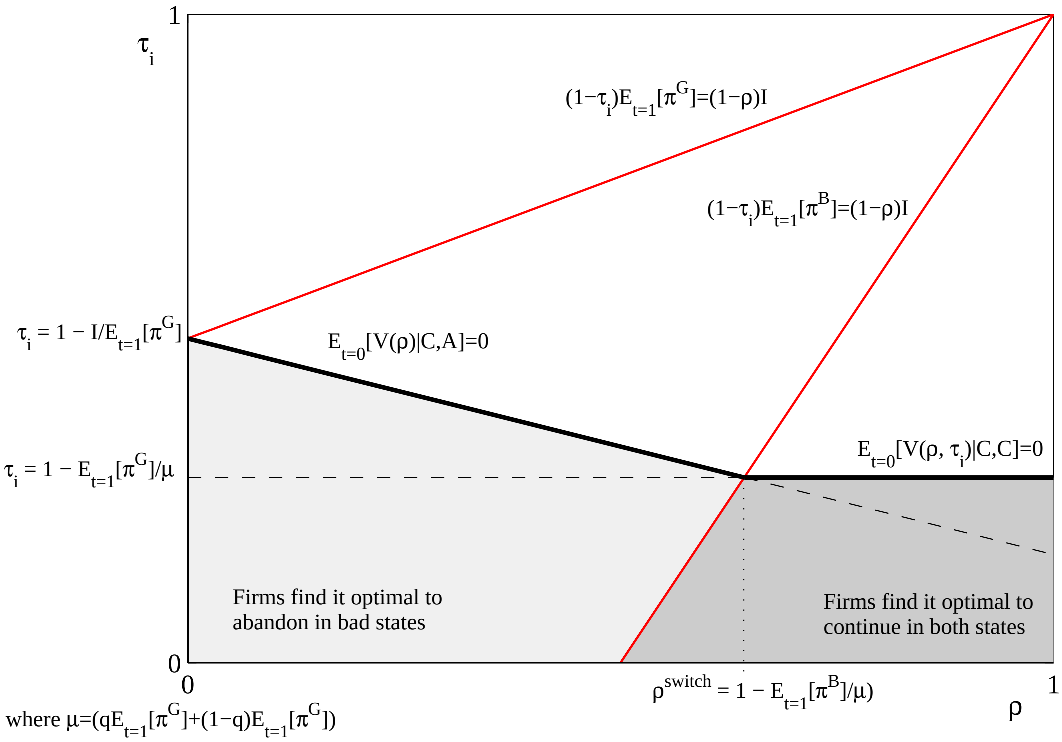


Figure 4: Feasible policy parameter space corresponding to Case III

Any (ρ, τ_i) pair to the southwest of this line satisfies the participation constraints defined by equations 2.10 and 2.11. The dashed lines show the boundaries of these constraints, when the strategy specified for $t = 1$ is sub-optimal. The graphs also includes ρ^{switch}, which is the value of ρ on the boundary of the participation constraint above which firms choose to continue regardless of cost and below which firms decide to abandon high-cost projects. I show below, in (2.21), that $\rho^{switch} \in (0, 1)$ and that the continuation constraint in bad states of the world always intersects the zero profit condition at this point.

2.3.2 Optimal choice of tax instruments, holding investment timeline constant

The income tax and royalty rate only affect the expected government revenue and the expected variable profits for the firm. At each stage in the game, the firm's optimal behaviour is determined by a trade off between the expected variable profit and the cost of proceeding to the next stage. Given any value of ρ there will be a minimum expected revenue that governments must provide to firms in order to induce firms to participate at $t = 0$. Given this constraint, governments choose τ_r and τ_i to maximize their expected revenue at $t = 0$. I show that it is optimal to set the royalty rate equal to zero.

Assume for the moment that firms find it optimal to continue in both states of the world. This requires us to assume that $\rho \geq \rho^{switch}$. Governments will extract the maximum expected revenue by driving the expected payoff of the firm at $t = 0$ to zero, making the participation constraint(2.9) hold with equality. Recall that if $\rho > \rho^{switch}$ and the participation constraint is satisfied, it follows that the continuation constraint is also satisfied. Government's expected payoff at $t = 0$ is increasing in government revenue, thus if the participation constraint for firms did not hold, the government could increase its payoff by increasing revenue. The zero payoff condition for firms simplifies the government's optimal choice of τ_r and τ_i considerably. Governments will choose τ_r and τ_i to maximize their expected payoff, $E_{t=0}[U|C,C]$ such that $E_{t=0}[V|C,C] = 0$. Recall also that the expected variable profit functions, $E_{t=1}[\pi(\tau_r)]$ are functions of the royalty rate.[15]

Proposition 1. *With common knowledge of marginal costs it is always optimal for governments to set the royalty rate, $\tau_r^* = 0$ and raise revenue with an income tax, $\tau_i > 0$*

Proof. See appendix B ☐

The intuition behind this proof is that royalties distort the production decision of firms at $t = 2$ thereby reducing the total expected surplus generated by the

[15]Where $\pi^G(\tau_r) = (1 - \tau_r)p_t - c$ and $\pi^B(\tau_r) = (1 - \tau_r)p_t - c(1 + \theta)$
$E_{t=1}[\pi^G(\tau_r)] = \Pr(p_t \geq \frac{c}{1-\tau_r})E_{t=1}[(1 - \tau_r)p_t - c|p_t \geq \frac{c}{1-\tau_r}]$
$E_{t=1}[\pi^B(\tau_r)] = \Pr(p_t \geq \frac{c(1+\theta)}{1-\tau_r})E_{t=1}[(1 - \tau_r)p_t - c(1 + \theta)|p_t \geq \frac{c(1+\theta)}{1-\tau_r}]$

investment project regardless of project costs. Government can always increase revenue by choosing an income tax instead of a royalty.[16] Increasing royalty rates reduces the probability that firms will produce at $t = 2$ and the profit margins they receive for any given realization of price. Reducing the probability of production and the expected variable profits reduces the unconditional expected surplus of the project at $t = 0$. This result is a variation of the "cut-off grade" problem in resource taxation discussed in Heaps and Helliwell (1985b). The cut-off grade problem occurs when projects that yield lower profits are not under taken because firms have to pay royalties. However, royalties increase the expected price, conditional on production since firms require a higher pre-royalty price in order to produce.

Substituting $\tau_r^* = 0$ into the zero profit condition yields the optimal income tax rate when $\rho > \rho_{switch}$:

$$\tau_i^* = 1 - \frac{I}{(qE_{t=1}[\pi^G(0)] + (1-q)E_{t=1}[\pi^B(0)]},$$

$$\text{where } E_{t=1}[\pi^G(0)] = \Pr(p_t \geq c)E_{t=1}[p_t - c | p_t \geq c],$$

$$\text{and } E_{t=1}[\pi^B(0)] = \Pr(p_t \geq c(1+\theta))E_{t=1}[p_t - c(1+\theta) | p_t \geq c(1+\theta)].$$

The proof that $\tau_r^* = 0$ is essentially identical for cases where firms find it optimal to abandon high cost projects, but instead yields an optimal tax rate that is a function of ρ.

[16]Given that total expected surplus is decreasing in royalties, it would also be optimal both for firms and governments to set royalties to zero in bargaining situations, and for firms to set royalties to zero when they are dictating terms to governments (assuming that some positive surplus must be shared with the government to induce participation). This result relies crucially on assumptions about information; if governments were unable to monitor marginal costs of production while price was public knowledge it may be preferable to employ royalties rather than income taxes. If governments could not monitor marginal costs, low cost firms would report high marginal costs in order to pay less income tax.

2.3.3 Optimal policy choice

The government's optimal policy choice is now a question of the portion of investment that must be sunk and the level of income tax used to generate revenue. The portion, ρ, of fixed investment that must be sunk prior to discovering marginal cost directly affects government revenue only when it is optimal for firms to abandon high-cost projects. When firms plan to abandon high-cost projects the flexibility of being able to wait and see what marginal costs will be at $t = 1$ increases the expected value of the investment project at $t = 0$. This extra option value can be extracted by taxing investors. Forcing firms to sink a larger portion of investment prior to discovering marginal cost can also induce firms to see a project through to completion.

With royalties set to zero, expected variable profits at $t = 1$ are only a function of cost and price. Income tax and the portion of fixed investment that must be sunk affect firm's decisions at $t = 0$ and $t = 1$. Thus, expected variable profits at $t = 1$ for high and low cost projects can be treated as exogenous. The unconditional expected profit at $t = 1$ in each states is:

$$E_{t=1}[\pi^G] = \Pr(p_t > c) \cdot E_{t=1}[p_t - c | p_t > c] \text{ (good cost)},$$
$$E_{t=1}[\pi^B] = \Pr(p_t > c(1+\theta) \cdot E_{t=1}[p_t - c(1+\theta) | p_t > c(1+\theta)] \text{ (bad cost)}.$$

Government can choose τ_i and ρ to induce firms to continue regardless of cost or to abandon high-cost projects. In each situation government solves a different maximization problem to determine the optimal policy parameters given the firm's plan of action at $t = 1$. Whichever pair yield a greater maximized expected utility is chosen by the government. The optimal choice of ρ and τ_i is formally described as:

$$\max_{\rho,\tau_i} \{E_{t=0}[U(\rho^*, \tau_i^*|C, A], E_{t=0}[U(\rho^*, \tau_i^*|C, C]\}$$
$$s.t.\ \rho \in [0,1]\,,\ \tau_i \in [0,1]$$

Where the first argument is:

$$E_{t=0}[U(\rho^*, \tau_{inc}^*|C, A] = \max_{\rho,\tau_i}\{\tau_{inc} q E_{t=1}[\pi^G] + qu(I) + (1-q)u(\rho I)\} \tag{2.15}$$

$$\text{s.t}\quad q(1-\tau_i)E_{t=1}[\pi^G] \geq qI + (1-q)(\rho I) \text{ at } t = 0 \tag{2.16}$$
$$(1-\tau_i)E_{t=1}[\pi^G] \geq (1-\rho)I \text{ at } t = 1$$
$$(1-\tau_i)E_{t=1}[\pi^B] < (1-\rho)I \text{ at } t = 1.$$

And the second argument is:

$$E_{t=0}[U(\rho^*, \tau_i^*|C, C] = \max_{\rho,\tau_i}\{\tau_{inc}(qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B]) + u(I)\} \tag{2.17}$$

$$\text{s.t.}\quad (1-\tau_i)(qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B]) \geq I \text{ at } t = 0 \tag{2.18}$$
$$(1-\tau_i)E_{t=1}[\pi^B] \geq (1-\rho)I \text{ at } t = 1$$
$$(1-\tau_i)E_{t=1}[\pi^G] \geq (1-\rho)I \text{ at } t = 1.$$

Constraints (2.16) and (2.18) describe the feasible set of policy parameters and the optimal behaviour firms, and correspond to the regions illustrated in figure 4 above. The optimal policy may either result in firms abandoning high-cost projects or completing the investment regardless of cost. I proceed by solving each sub-optimization problem and comparing the maximized expected utilities. I determine the conditions under which the government prefers firms to abandon high-cost projects rather than continue a project to completion.

In the central case, expected variable profits are negative for high cost projects, but overall expected profits at $t = 0$ when firms are forced to complete the project are positive. Government is able to dictate any incentive compatible contract to firms and will thus extract all the expected surplus from firms at $t = 0$. This implies that government will choose τ_i and ρ such that the firm's participation constraint, (2.16) or (2.18), holds with equality. When firms find it optimal to continue the project regardless of cost this zero profit condition is:

$$\tau_i = 1 - \frac{I}{(qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B])} \tag{2.19}$$

If, on the other hand, firms find it optimal to abandon the project when costs are high, the zero profit condition is:

$$\tau_i = 1 - \frac{qI + (1-q)(\rho I)}{qE_{t=1}[\pi^G]} \tag{2.20}$$

The zero profit conditions coincide with the bold lines in figure 4.

Setting the zero profit conditions equal to each other, yields the critical value for ρ, above which it is optimal for firms to continue regardless of cost and below which it is optimal for firms to abandon high-cost projects:

$$\begin{aligned} \rho^{switch} &= q\left[\frac{E_{t=1}[\pi^G] - E_{t=1}[\pi^B]}{(qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B])}\right], \\ &= \left[\frac{qE_{t=1}[\pi^G] - qE_{t=1}[\pi^B] + E_{t=1}[\pi^B] - E_{t=1}[\pi^B]}{(qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B])}\right], \\ &= \left[1 - \frac{E_{t=1}[\pi^B]}{(qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B])}\right]. \end{aligned} \tag{2.21}$$

By construction, $E_{t=1}[\pi^B] < E_{t=1}[\pi^G]$, implying that $\rho^{switch} \in (0,1)$. This intersection point must lie along the boundary of the continuation constraint for high cost realizations. This is shown by setting the boundary of (2.8) equal to the zero profit condition (2.18) and solving for ρ:

$$\begin{aligned}
1 - \frac{(1-\rho)I}{E_{t=1}[\pi^B]} &= 1 - \frac{I}{(qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B])}, \\
(1-\rho) &= \frac{E_{t=1}[\pi^B]}{(qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B])}, \\
\rho &= q\left[\frac{E_{t=1}[\pi^G] - E_{t=1}[\pi^B]}{(qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B])}\right].
\end{aligned}$$

This intersection point is shown in figure 4. The optimal policy choice will lie along the zero profit condition, however not all points along this line can be optimal. In particular, for values of ρ just below ρ^{switch}, the government will find it optimal to increase ρ and decrease τ_i in order to prevent firms from abandoning the project.[17]

Returning to the government's optimization problem, I substitute the zero profit condition into the objective function when firms find it optimal to abandon at $t = 1$ and take the first order condition with respect to ρ, which yields the optimal choice of ρ.

$$\begin{aligned}
\max_{\rho,\tau_i} \;\; & \tau_i q E_{t=1}[\pi^G] + qu(I) + (1-q)u(\rho I) \\
\text{s.t.} \quad & \tau_i = 1 - \frac{qI + (1-q)(\rho I)}{qE_{t=1}[\pi^G]} \\
& 0 \leq \tau_i \leq 1 \\
& 0 \leq \rho \leq \rho^{switch}
\end{aligned}$$

[17]To see this, take a point in figure 4 where $\rho < \rho^{switch}$ and the continuation constraint in bad states of the world is slack. Government benefits by increasing τ_i without causing firms to change their behaviour. At the boundary of the continuation constraint, firms are indifferent between completing the investment and abandoning the project whereas governments would strictly prefer the firm to continue, since it receives $u(I)$ in both states of the world. Governments can only make themselves better off by increasing both ρ and τ_i together. This logic suggests that for a range of values of ρ that are just less than ρ^{switch} the government will strictly prefer to set $\rho \geq \rho^{switch}$ and reduce τ_i.

Assuming $\rho < \rho^{switch}$ and $\tau_i \in (0,1)$ then the constrained objective function is:

$$\max_{\rho} qE_{t=1}[\pi^G] + qu(I) - qI - (1-q)\rho I + (1-q)u(\rho I). \tag{2.22}$$

Setting the first derivative with respect to ρ equal to zero yields the following first order condition:

$$\begin{aligned} -(1-q)I + (1-q)\frac{du(\rho I)}{d\rho}|_{\rho=\rho^*} I &= 0 \\ \frac{du(\rho I)}{d\rho}|_{\rho=\rho^*} &= 1. \end{aligned}$$

Government will chose ρ^* optimally so that the marginal benefit of forcing firms to sink more investment just equals the marginal reduction in expected revenue that is required to satisfy the firm's zero profit condition. Government benefits when high-cost projects are abandoned at $t = 1$ from the quantity of abandoned investment, ρI, but for every unit increase in upfront investment that firms will ultimately abandon, they must be compensated with a unit reduction in the expected tax bill.

The optimal choice of τ_i^* when firms find it optimal to continue regardless of cost ($\rho^* > \rho^{switch}$) is simply the zero profit condition:

$$\tau_i^* = 1 - \frac{I}{(qE_{t=1}[\pi^G(0)] + (1-q)E_{t=1}[\pi^B(0)]}. \tag{2.23}$$

If ρ is high enough, firms complete the project regardless of cost realizations at $t = 1$. There is no trade-off between ρ and τ_i since the timing of investment is irrelevant for a firm that plans to continue regardless of cost, although ρ is still a necessary policy tool to commit firms to complete the investment project. The best the government can do, when $\rho > \rho^{switch}$ is set income tax rates at a level that extracts the entire expected surplus at $t = 0$.

2.4 Discussion and comparative statics

2.4.1 Optimal policy using a functional form for investment spillovers

In order to determine the conditions under which it is optimal to induce firms to abandon high-cost projects I solve explicitly for ρ^* using a functional form for $u(\cdot)$ that satisfies the assumptions of strict concavity:

$$u(x) = \alpha \log(x) \text{ where } \alpha > 0 \text{ and } x > 0.$$
$$u(0) = 0$$

The government's optimal policy choice, given that firms abandon high-cost projects is then:

$$\rho^* = \frac{\alpha}{I}, \tag{2.24}$$
$$\tau_i^* = 1 - \left(\frac{I}{E_{t=1}[\pi^G]}\right) - \left(\frac{(1-q)(\alpha)}{qE_{t=1}[\pi^G]}\right).$$

Governments choose ρ^* at a level that maximizes the benefit to government of the sunk investment that is abandoned by the firm. The greater is α, the greater is the relative benefit from investment and the more up-front investment government will ask for at $t = 0$. In order to compensate firms for investment that may be abandoned, governments must reduce the tax rate. This result reflects the quasi-linear structure of the government's benefit function: government chooses a fixed level of $\rho = \frac{\alpha}{I}$, that doesn't change with the expected variable profit $E_{t=1}[\pi^G]$. As the expected variable profits change, governments adjust the tax rate in order to compensate firms for keeping $\rho = \rho^*$. This sub-maximization problem has an interior solution if $\frac{\alpha}{I} < \rho^{switch}$. If $\frac{\alpha}{I} \geq \rho^{switch}$ we will have a corner solution to the sub-maximization problem where firms abandon the high-cost project and it will be optimal to set $\rho^* = \rho^{switch}$. However, at $\rho^* = \rho^{switch}$, firms are indifferent between abandoning high cost projects at $t = 1$ and continuing regardless of cost. Government would strictly prefer firms to continue, and will thus choose a level of $\rho > \rho^{switch}$ so that firms find it optimal to continue regardless of cost. The optimal income tax will be defined by the zero payoff condition (2.18) and the optimal choice of ρ will have to be greater than ρ^{switch}:

$$\rho^* \geq \rho^{switch}, \tag{2.25}$$
$$\tau_i^* = 1 - \left(\frac{I}{(qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B])}\right).$$

Equations (2.24) and (2.25) are the solution to the sub-maximization problems outlined above, however only one of these solutions is the optimal policy choice for governments. The optimal policy choice is the (ρ^*, τ_i^*) pair which yields the greatest maximized utility, conditional on firm behaviour.

2.4.2 Conditions determining optimal policy

Government will choose to make firms abandon high-cost projects iff:

$$E_{t=0}[U(\rho^*, \tau_i^*|C, A] > E_{t=0}[U(\rho^*, \tau_i^*|C, C]. \tag{2.26}$$

These maximized expected payoffs depend crucially on how we parameterize $u(\cdot)$, and on the other exogenous parameters determining the profitability of the project. If $\frac{\alpha}{I} > \rho^{switch}$ then the solution to the sub-maximization problem where firms abandon the project is a corner solution and yields optimal policy parameters, (ρ^*, τ_i^*), that are identical to the sub-maximization problem where firms continue regardless of cost. Comparing the maximized utility functions shows that governments always prefer firms to continue in both states of the world.

Proposition 2. *If* $\frac{\alpha}{I} > \rho^{switch}$ *where* $\rho^{switch} = [1 - \frac{E_{t=1}[\pi^B]}{(q \cdot E_{t=1}[\pi^G] + (1-q) \cdot E_{t=1}[\pi^B])}]$ *then government will always induce firms to continue in both states of the world*

Proof. See appendix. □

This proposition has an intuitive explanation. The government's optimal choice of ρ and τ_i when firms abandon the project is identical to the optimal choice when firms continue regardless of cost. At this policy choice firms are indifferent between abandoning high-cost projects or continuing regardless of cost. However, government always prefers firms to complete the project regardless of cost. A completed investment project is preferable to a partially completed project, because government benefits directly from the extra investment and indirectly from the increased expected tax revenue from the completed high-cost projects.

This result highlights a discontinuity that exists in the problem and has been alluded to earlier. The firm's binary decision to continue a high-cost project, instead of abandoning it results in a discrete jump in the government's expected payoff function. Instead of benefiting from the partially completed project, $u(\rho I)$, government benefit from the completed project, $u(I)$ even if their expected revenue stream is identical. If $\rho = \rho^{switch} - \varepsilon$, where ε is some small number, firms would strictly prefer to abandon the high-cost projects, while government would gain a greater payoff by inducing firm's to continue regardless of cost. With $\rho = \rho^{switch} - \varepsilon$, government can only expects to collect revenues from low-cost projects that are marginally higher than those they could collect if firms were being forced to complete the project regardless of cost. Government also forgoes the expected revenue high-cost projects and the discrete jump in benefits accruing from a completed investment project. Over a range of values for $\rho \leq \rho^{switch}$, government will prefer firms to continue regardless of cost, but will be unable to do so because the firm's incentives are not aligned with their own. ρ^* will have to jump just above ρ^{switch} and τ_i^* will have to drop accordingly.

Proposition 3. *There exists a cutoff value of* α*,* $0 < \alpha < I(1 - \frac{E_{t=1}[\pi^B]}{qE_{t=1}[\pi^G]+(1-q)E_{t=1}[\pi^B]})$ *above which governments will find it optimal to force firms to continue regardless of cost and below which governments will find it optimal to force firms to abandon the high-cost projects. The existence of this cutoff implies that if governments force firms to abandon the project the optimal policy choice must be an interior solution.*

Proof. See appendix. □

As government values the indirect benefits of investment more, it is more likely that they will induce firms to continue regardless of cost. However, when the benefits to investment, α are sufficiently low government derives more benefit from collecting higher taxes from low-cost projects and allowing firms to abandon high-cost projects. However, as spillovers from investment increase in importance, government prefers to induce firms to continue in both states of the world. In the central case that we examined, the expected profits at $t = 1$ from the investment project are assumed to be negative when costs are high. Allowing firms enough flexibility in the timing of investment means that firms will exercise the option to abandon high-cost projects. This adds value to the expected surplus of the investment project which the government can extract through taxation. Limiting flexibility, by increasing ρ, forces firms to commit to the investment project. However, government must compensate firms for undertaking high-cost projects by charging lower taxes. In particular, government must allow firms to receive positive profits in expectation for low-cost projects to compensate firms for undertaking a loss making venture when costs are high. Optimal policy in land deals thus depends on the relative benefit of investment spillovers and expected revenue. However, ignoring the role of uncertainty leaves out a key determinant of optimal policy.

2.4.3 Optimal policy in an increasingly uncertain environment

Greater uncertainty over payoffs can benefit a firm that holds the option to abandon a project. The main theoretical model presents an investment project where marginal cost and price are both initially unknown. I examine the effect of increasing uncertainty in two ways. First, I look at the effects of increasing uncertainty over the expected variable profits at $t = 1$ by examining a mean-preserving spread in expected variable profits at $t = 1$. Second, I look explicitly at the effect of increased uncertainty over prices.

2.4.4 Uncertainty in profits

Marginal cost is a binary outcome that affects the total expected profits of the investment project. It is at $t = 1$, where firms discover marginal cost, that firms benefit from the ability to abandon the investment project; if a firm discovers that it has high costs of production it expects to earn lower profits in the final stage of the game. Thus, if we wanted to analyze the effect of increased uncertainty on policy outcomes we would have to look at a mean-preserving increase in the variance of expected variable profits at $t = 1$. This is equivalent to:

$$\tilde{E}_{t=1}[\pi^G] > E_{t=1}[\pi^G],$$
$$\tilde{E}_{t=1}[\pi^B] < E_{t=1}[\pi^B],$$
$$\text{and } q\tilde{E}_{t=1}[\pi^G] + (1-q)\tilde{E}_{t=1}[\pi^B] = qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B].$$

Greater uncertainty increases the total expected surplus of the investment project when firms intend to abandon high-cost projects. When expected variable profits change in this way, firms expect to earn more from low-cost projects and avoid losing more from high-cost projects. Since government extracts the full surplus from firms, an increase in expected total surplus increases the feasible set of policy parameters available to government. The policy outcome, as in the previous model depends on the spillover benefit government derives from investment. Recall that the optimal policy when government finds it optimal to force firms to abandon high-cost projects is:

$$\rho^* = \frac{\alpha}{I},$$
$$\tau_i^* = 1 - \frac{I}{E_{t=1}[\pi^G]} - \frac{(1-q)(\alpha)}{qE_{t=1}[\pi^G]}.$$

By inspection we see that the optimal tax rate, τ_i^*, increases with increased uncertainty (i.e. when $\tilde{E}_{t=1}[\pi^G] > E_{t=1}[\pi^G]$). Greater uncertainty also increases the level of ρ that is required to induce firms to switch strategies from abandonment to continuation. Thus, when government finds it optimal to force firms to continue, they will have to increase their choice of ρ in order to force firms to commit more investment in the initial period. The optimal policy when government forces firms to continue regardless of cost is:

$$\rho^* > \rho^{switch} = 1 - \frac{E_{t=1}[\pi^B]}{qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B]},$$
$$\tau_i^* = 1 - \frac{I}{(qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B])}.$$

The optimal policy if government wants firms to continue will be unchanged by a mean-preserving spread in the expected variable payoffs at $t = 1$. However, since ρ^{switch} increases with increased uncertainty (because $\tilde{E}_{t=1}[\pi^B] < E_{t=1}[\pi^B]$), government may find it optimal to switch from encouraging firms to continue a project regardless of cost, to abandoning high-cost projects.

The effect on optimal policy of increasing uncertainty in this way depends on whether government initially finds it optimal to induce firms to continue the project regardless of cost or to abandon high-cost projects. If government initially wants firms to abandon high-cost projects, the government will continue to set $\rho^* = \frac{\alpha}{I}$ and extract the added surplus by increasing incomes taxes. Firms will continue to abandon high-cost projects. If, on the other hand, government initially wants firms to continue regardless of cost, it is possible that greater uncertainty will induce the government to encourage firms to switch to a strategy of abandoning high-cost projects.

The effect of increasing uncertainty on optimal policy works through the option to abandon a project. Allowing firms to abandon high-cost projects can give them a valuable option. The real value of the option to abandon the project, in the absence of taxes is:

$$\begin{aligned} ROV &= E_{t=0}[\pi|C,A] - E_{t=0}[\pi|C,C], \\ &= q(E_{t=1}[\pi^G] - I) + (1-q)(-\rho I) - (qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B]) + I, \\ &= (1-q)((1-\rho)I - E_{t=1}[\pi^B]). \end{aligned}$$

As the variance in expected variable payoffs increases, the value of this option increases. Government extracts all the surplus from firms, including the value of the option to abandon the project. Thus, as the option to abandon the project increases in value, government will be forced to set ρ higher in order to induce firms to complete a high-cost project. However, as the value of the option to abandon the project increases, government will extract more surplus for any given level of ρ by increasing taxes.

Analyzing the effect of uncertainty on optimal policy by assuming a mean preserving increase in the variance of expected profits is somewhat limited. Recall that firms can choose not to produce in the final period if marginal cost exceeds price, implying that the lowest expected variable profit a firm might face is zero. This imposes a limit on a mean preserving increase in uncertainty of this type. Furthermore, a mean preserving increase in the expected variable profits at $t = 1$ can be achieved by changing cost uncertainty or by changing price uncertainty, leading to ambiguity about how changes in each stochastic element affect the policy outcome. In the next section I analyze the effect of increased price uncertainty on policy outcomes.

2.4.5 Uncertainty in price

The effect of a change in price uncertainty on the policy outcomes of the model depends crucially on how prices are distributed. In particular, since firms choose to produce if $p_t \geq c$, or $p_t \geq c(1+\theta)$ the unconditional expected variable profits at $t = 1$, $(E_{t=1}[\pi^G], E_{t=1}[\pi^B])$, will be the expectation of a truncated distribution, which can be a complicated object. To allow for an analytical solution, prices are assumed to be uniformly distributed, with marginal costs contained in the support of the distribution:

$$p_t \sim U[a, b],$$
$$c,\, c(1+\theta) \in [a, b]\,.$$

This assumption limits how generalizability of comparative statics since the way in which the expected value of a truncated distribution is affected by the unconditional variance depends crucially on the type of distribution and the location of the truncation.[18] Assuming a functional form for the distribution allows for analytical solutions and comparative statics.

Expected variable profits at $t = 1$ as a function of cost realizations are:

$$\begin{aligned} E_{t=1}[\pi^G] &= \Pr(p_t \geq c)E_{t=1}[p_t - c | p_t \geq c] \\ &= (\frac{b-c}{b-a})(\frac{b+c}{2} - c) \\ &= \frac{(b-c)^2}{2(b-a)}, \end{aligned} \tag{2.27}$$

$$\begin{aligned} E_{t=1}[\pi^B] &= \Pr(p_t \geq c(1+\theta))E_{t=1}[p_t - c(1+\theta) | p_t \geq c(1+\theta)] \\ &= (\frac{b - c(1+\theta)}{b-a})(\frac{b - c(1+\theta)}{2} - c(1+\theta)) \\ &= \frac{(b - c(1+\theta))^2}{2(b-a)}. \end{aligned} \tag{2.28}$$

Analyzing the effect of a mean preserving increase in price volatility implies taking the derivative with respect to b while ensuring that:

$$\begin{aligned} \frac{a+b}{2} &= E[p_t] = \mu \text{ for all } (a, b) \\ a &= 2\mu - b. \end{aligned}$$

A mean preserving increase in the variance of prices increases the bounds of the distribution, in effect permitting firms to realize higher and lower prices. This increase in variance unambiguously increases the expected variable profit at $t = 1$

[18] Furthermore, the change in the unconditional expected value of the payoff from a mean preserving increase in the variance of the price distribution is ambiguous. Increasing the variance of the distribution unambiguously increases the expected value of profits, conditional on producing, but in some circumstances it will reduce the probability of production. If prices distributions are symmetrical about the mean, the probability of production is increasing with variance if the marginal cost cut-off is above the mean and decreasing with variance if the marginal cost cut-off is below the mean.

regardless of costs:

$$\begin{aligned} E_{t=1}[\pi^G] &= \frac{(b-c)^2}{4(b-\mu)} \\ \frac{d(E_{t=1}[\pi^G])}{db} &= \frac{1}{4}[\frac{2(b-c)(b-\mu)-(b-c)^2}{(b-\mu)^2}] \\ &= \frac{1}{4}[\frac{(b-c)(c+b-2\mu)}{(b-\mu)^2}] \\ &= \frac{1}{4}[\frac{(b-c)(c-a)}{(b-\mu)^2}] > 0, \end{aligned}$$

since $c \in [a, b]$. By analogy:

$$\frac{d(E_{t=1}[\pi^B])}{db} = \frac{1}{4}[\frac{(b-c(1+\theta))(c(1+\theta)-a)}{(b-\mu)^2}] > 0, \tag{2.29}$$

since $c(1+\theta) \in [a, b]$.

The expected value of the project regardless of costs is increasing with the variance in prices. It is also the case that the relative profitability of a low cost project is diminishing as the variance increases:

$$E_{t=1}[\pi^G] - E_{t=1}[\pi^B] = \frac{(b-c)^2 - (b-c(1+\theta))^2}{4(b-\mu)}, \tag{2.30}$$

$$\begin{aligned} \frac{d(E_{t=1}[\pi^G] - E_{t=1}[\pi^B])}{db} &= \frac{1}{2}[\frac{c\theta(b-\mu) - bc\theta}{(b-\mu)^2}], \\ &= \frac{1}{2}c\theta\left[\frac{-\mu}{(b-\mu)^2}\right] < 0. \end{aligned} \tag{2.31}$$

For any realization of price, the benefit from being a low cost firm is that marginal costs are lower by $c\theta$. This relative benefit decreases as price volatility increases. Thus, as the variance increases, and the expected price, conditional on producing increases, the relative benefit of being a low cost firm is diminished. This logic extends to any convex combination of the two expected values at $t = 1$, implying that $qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B]$ will be increasing with volatility at a faster rate than $E_{t=1}[\pi^G]$, but at a slower rate than $E_{t=1}[\pi^B]$. This result is key in understanding how increased price variance affects the policy outcomes.

Increasing the variance of prices causes the expected profit functions at $t = 1$ to progress through the four cases outlined in Table 1 (the table outlining when firms do or do not abandon high-cost projects). Beginning at a level of variance where $E_{t=1}[\pi^G] - I < 0$, firms do not participate in the project as it is unprofitable regardless of cost. This coincides with case I. A sufficient increase in the variance of prices makes low-cost projects profitable in expectation at $t = 1$, but the expected value at $t = 0$ if firms were forced to complete the fixed investment is negative. This corresponds to case II where firms will participate, but will always abandon the high-cost project. Increasing the variance yet further can make the project profitable in expectation at $t = 0$, which corresponds to case III. This enables government to induce firms to either complete the investment regardless of cost or abandon high-cost projects. Finally, if the variance in prices increases enough, the investment project will yield positive expected profits at $t = 1$ regardless of cost, corresponding to case IV in which government always finds it optimal to induce firms to complete the project.

In the central case, case III[19], governments either choose to induce firms to abandon high-cost projects or to complete the project regardless of cost, depending on the optimal policy chosen by the government. Recall that the value of $\rho = \rho^{switch}$ along the zero profit condition is the value above which firms switch strategies and continue the project in both states of the world. This threshold is

[19]Case III assumes $E_{t=1}[\pi^G] - I > 0$, $qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B] - I > 0$ and $E_{t=1}[\pi^B] - I < 0$,

decreasing as the price variance increases:

$$\rho^{switch} = 1 - \frac{E_{t=1}[\pi^B]}{qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B]}, \tag{2.32}$$
$$\rho^{switch} = 1 - \frac{(b-c(1+\theta))^2}{q(b-c)^2 + (1-q)(b-c(1+\theta)^2)},$$
$$\frac{d(\rho^{switch})}{db} = -2q\frac{(b-c(1+\theta))(b-c)^2 - (b-c(1+\theta))^2(b-c)}{[q(b-c)^2 + (1-q)(b-c(1+\theta)^2)]^2},$$
$$= -2q\frac{(b-c(1+\theta))[(b-c)^2 - (b-c(1+\theta))(b-c)]}{[q(b-c)^2 + (1-q)(b-c(1+\theta)^2)]^2} < 0.$$

The threshold value of ρ at which firms switch their strategy depends on the ratio of expected profits in bad states of the world to the average expected profits at $t = 0$. Equation (2.30) shows that as the price variance increases, the expected profits from high-cost projects increase at a faster rate than the average of the expected variable profits causing ρ^{switch} to decrease with price variance. Rearranging the expression above and dividing both sides by I, gives clearer intuition for this result:

$$\frac{qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B]}{I} = \frac{E_{t=1}[\pi^B]}{(1-\rho^{switch})I}. \tag{2.33}$$

The right hand side of the expression is the benefit cost ratio of completing the high-cost project. Firms choosing to abandon the project forgo the opportunity to earn $E_{t=1}[\pi^B]$ but avoid paying $(1-\rho^{switch})I$, which is the opportunity cost of abandoning the project. The left hand side is the benefit cost ratio at $t = 0$ of continuing regardless of cost. Firms are indifferent between the two strategies if the benefit of completing the project regardless of cost equals the opportunity cost of abandoning the high-cost project. As price variance increases, $E_{t=1}[\pi^B]$ increases at greater rate than $qE_{t=1}[\pi^G]+(1-q)E_{t=1}[\pi^B]$, causing the opportunity cost of abandonment to grow faster than the benefit of continuation. In order to equalize these two ratios, ρ^{switch} must be at a lower level, effectively increasing the cost of completing the high-cost project at $t = 1$ and thus, reducing the opportunity cost of abandoning the project.

What does this tell us about the optimal choice of governments in the central case? Increased price variance reduces the threshold level of ρ^{switch} required to induce firms to continue the project regardless of cost. As it becomes easier to induce firms to complete the investment regardless of cost government can set higher tax rates, while still encouraging high levels of investment. Recall that if governments find it optimal to induce firms to continue regardless of cost they will set:

$$\rho^* > \rho^{switch} = 1 - \frac{E_{t=1}[\pi^B]}{qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B]},$$
$$\tau_i^* = 1 - \frac{I}{(qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B])}.$$

Equation (2.32) shows that ρ^{switch} is decreasing in price variance meaning that governments can require firms to sink less investment prior to discovering the state. Conversely, income tax is increasing with price variance; as average expected profits increase, governments are able to extract a larger surplus from firms by increasing taxes. When governments find it optimal to induce firms to abandon high-cost projects, the optimal policy choice is:

$$\rho^* = \frac{\alpha}{I},$$
$$\tau_i^* = 1 - \frac{I}{E_{t=1}[\pi^G]} - \frac{(1-q)(\alpha)}{qE_{t=1}[\pi^G]}.$$

Increasing price variance does not affect the level of ρ, but increases the tax rate that government will set. Government sets higher taxes in order to capture the increased surplus produced by firms following an increase in the price variance. This result is a function of the quasi-linear government benefit function.[20]

In the central case, government will choose to induce a particular behaviour in firms depending on how they benefit from investment, determined by α. Government will force firms to abandon the high-cost project iff:

$$1 + \frac{\alpha}{I} \log\left(\frac{\alpha}{I}\right) - \frac{\alpha}{I} \geq \frac{E_{t=1}[\pi^B]}{I}. \tag{2.34}$$

It was shown that the LHS (expected payoff given firms abandon) and the RHS (expected payoff given firms continue) lie between zero and one. However, as we increase the variance in prices, the RHS increases leaving the LHS remains the same. Thus, in the central case, increasing price variance causes governments to switch from a strategy of inducing firms to abandon the project to a strategy of inducing firms to continue regardless of cost. This suggests that when price variance is high, governments are more likely to force firms to complete the investment project.

Another interpretation of this result is that increased price volatility reduces the value of holding the option to abandon the investment project and government derives less benefit from extracting surplus from the option. The real value to the firm of holding the option to abandon a project, in the absence of taxes, is:

$$\begin{aligned} ROV &= E_{t=0}[\pi|C, A] - E_{t=0}[\pi|C, C], \\ &= q(E_{t=1}[\pi^G] - I) + (1-q)(-\rho I) - (qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B]) + I, \\ &= (1-q)((1-\rho)I - E_{t=1}[\pi^B]). \end{aligned}$$

[20]Government benefits linearly from revenue, but with diminishing returns on investment. Governments must compensate firms for any investment that gets abandoned by offering lower taxes on the low-cost projects. Once governments have forced firms to invest the optimal portion of investment in $t = 0$, they will extract any remaining surplus from firms through taxation.

Given that $E_{t=1}[\pi^B]$ is increasing as the price variance increases, the real value of the option to abandon the project is decreasing with price volatility. Government extracts this surplus from firms by inducing firms to abandon the project and setting higher tax rates. As the value of the option decreases, the surplus that governments can extract by inducing firms to abandon the project is diminished and government will prefer firms to complete the project.This result is in contrast to the previous section, where government was more likely to induce firms to abandon the project as the variance in expected variable profits increased. The important distinction between these results is that increasing the price variance reduces the value of the option to abandon, while increasing the variance in expected payoffs increases the value of the option to abandon and is reflected in the way the optimal policy choice changes with different types of variance.

These comparative statics illustrate how the value of the option to abandon a project relates to two different types of change in uncertainty, which in turn affect government's optimal policy. The first type was an increase in the variance of expected variable payoffs. Increasing uncertainty in this way increased the value of the option to abandon the high-cost investment project. This type of increase in uncertainty makes it more likely that government induces firms to abandon high-cost projects in order to capture the surplus from the option value. This result contrasts with the findings when price variance increases. In particular, I find that the value of the option to abandon the project is decreasing as price volatility increases, since increasing price variance diminishes the relative benefit of a low-cost project. As the option to abandon the project decreases in value, government gains less from inducing firms to abandon the project and are more likely to induce firms to complete the investment regardless of cost. The common feature between these two types of changing uncertainty is that optimal government policy subsumes the optimal behaviour of firms and in doing so the effects of increased uncertainty.

3 Conclusion

This paper highlights the role of uncertainty and investment timelines in determining the structure of large-scale land contracts. The terms of the contract, summarized by a tax and an investment timeline arise from a strategic interaction between a foreign firm and a host government. I use a simple model of investment where a fixed investment can be completed in stages, yielding the result that: 1) an increase in the uncertainty over payoffs can increase the value of the option to abandon an unfavourable project and 2) the value of this option is decreasing as firms are required to sink more investment in the initial period. When the value of the option to abandon the project is positive firms will choose to exercise this option to avoid completing high-cost projects. However, if firms are sufficiently committed to an investment project, they will not abandon the project.

The model explains how the terms of a contract are determined when government dictates an income tax, royalty rate, and investment timeline to firms. The optimal contract chosen by governments internalizes the optimal behaviour of firms. In particular, I find: 1) that income taxes are preferred to royalties because royalties distort a firm's production decision and 2) that the tradeoff between income tax and the portion of investment that must be sunk initially is a function of the option value from abandoning a high-cost project. I argue that if government benefits from investment spillovers, it will use the policy tools to manipulate the choices of firms, inducing firms to either abandon or complete a project. I find: 1) that if government benefits a great deal from investment, the optimal contract will be designed to encourage firms not to abandon a project, whereas 2) if government benefits relatively little from investment, government will encourage firms to hold the option to abandon a project and extract its value. In this way, increased uncertainty affects the terms of the contract.

In the final section I examine the effect of a mean preserving increase in the uncertainty of prices and expected profits. I show that increasing uncertainty in prices and expected profits increases the total surplus generated by the project. I find that increasing uncertainty over expected profits results in a contract encouraging firms to abandon the investment while increasing price uncertainty results in a contract encouraging firms to complete the investment project. This result, once again depends on the how the value of the option to abandon a project changes with uncertainty. These results highlight the importance of looking at the value of options when analyzing the decision to invest in large-scale agricultural project. Especially in an environment of increasing volatility, the real value of the option to abandon a project affects the decisions and tradeoffs faced by investors and host governments.

Naturally this is a simple model of investment and there are a number of possible theoretical and empirical refinements. As more data becomes available theoretical predictions can be tested. Currently, many land deals are in the implementation phase, and the success or failure of large-scale agricultural projects is yet unknown. Once data become available, an analysis of the duration of land deals might shed light onto the role that price volatility plays on the investment decisions of firms and the original terms of the contract. A simple and relevant theoretical extension might use this model to look at land speculation in a principle-agent framework. Government, as the principle, may be unaware of whether they are dealing with high cost speculators or low cost producers and may adjust their optimal policy choice to attract certain types of investors. Finally, there is a rich literature in real options theory using more sophisticated dynamic models of investment.

Complicating the model in this way might shed light on how government tax policies and investment regulations affect an investment threshold. This method might suggest a critical price or level of volatility that triggers investment in land deals.

Beyond the academic question of why these deals are occurring and how they are structured, is a number of other pressing issues and areas for research. If the trend in large-scale land deals persists, the economic landscape of rural Africa may change drastically, with serious consequences for the local populations. Large-scale agriculture projects have the potential to put a strain on water resources and may impinge on the customary land rights of rural populations. On the other hand, successful projects might be a source of increased food security, rural infrastructure improvement and jobs. Understanding the causes and consequences of large-scale land deals is a new and exciting area for research and this paper represents an early attempt to explain how increased volatility affects the terms of the contracts.

References

Abel, A. B., Dixit, A. K., Eberly, J. C., and Pindyck, R. S. (1996). Options, the value of capital, and investment. *The Quarterly Journal of Economics*, 111(3):753–777.

Abel, A. B. and Eberly, J. C. (1996). Optimal investment with costly reversibility. *The Review of Economic Studies*, 63(4):581–593.

Arezki, R., Deininger, K., and Selod, H. (2013). What drives the global "land rush"? *The World Bank Economic Review*.

Binswanger, H. P., Deininger, K., and Feder, G. (1995). Chapter 42 power, distortions, revolt and reform in agricultural land relations. volume 3, Part B of *Handbook of Development Economics*, pages 2659 – 2772. Elsevier.

Collier, P. and Venables, A. J. (2012). Land deals in africa: pioneers and speculators. *Journal of Globalization and Development*, 3(1).

Cotula, L., Vermeulen, S., Leonard, R., and Keeley, J. (2009). *Land grab or development opportunity? Agricultural investment and international land deals in Africa*. IIED, FAO and IFAD.

Deininger, K. (2013). Global land investments in the bio-economy: evidence and policy implications. *Agricultural Economics*, 44(s1):115–127.

Deininger, K. and Binswanger, H. P. (1995). Rent seeking and the development of large-scale agriculture in kenya, south africa, and zimbabwe. *Economic Development and Cultural Change*, 43(3):493–522.

Deininger, K. and Byerlee, D. (2012). The rise of large farms in land abundant countries: Do they have a future? *World Development*, 40(4):701–714.

Deininger, K. W. and Byerlee, D. (2011). *Rising global interest in farmland: can it yield sustainable and equitable benefits?* World Bank Publications.

Dixit, A. K. and Pindyck, R. S. (1994). *Investment under Uncertainty*. Princeton University Press.

Drache, H. M. (1964). *The day of the bonanza: a history of bonanza farming in the Red River Valley of the North*. North Dakota Institute for Regional Studies.

Eastwood, R., Lipton, M., Newell, A., Pingali, P., and Evenson, R. (2010). *Chapter 65 Farm Size*, volume Volume 4, pages 3323–3397. Elsevier.

Engel, E. and Fischer, R. (2008). Optimal resource extraction contracts under threat of expropriation. *NBER Working Paper 13742*.

Headey, D. and Fan, S. (2008). Anatomy of a crisis: the causes and consequences of surging food prices. *Agricultural Economics*, 39(s1):375–391.

Heaps, T. and Helliwell, J. F. (1985b). The taxation of natural resources. *Handbook of public economics*, 1:421–472.

Leiblein, M. J. (2003). The choice of organizational governance form and performance: Predictions from transaction cost, resource-based, and real options theories. *Journal of Management*, 29(6):937–961.

Li, J. and Rugman, A. M. (2007). Real options and the theory of foreign direct investment. *International Business Review*, 16(6):687 – 712.

Lukas, E. (2007). Dynamic market entry and the value of flexibility in transitional international joint ventures. *Review of Financial Economics*, 16(1):91 – 110.

McDonald, R. and Siegel, D. (1986). The value of waiting to invest. *The Quarterly Journal of Economics*, 101(4):707–728.

Pande, R. and Udry, C. (2006). *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*, chapter Institutions and development: A view from below. Cambridge University Press.

Sulle, E. and Nelson, F. (2009). *Biofuels, land access and rural livelihoods in Tanzania*. IIED.

UNCTAD (2009). *World Investment Report 2009*. United Nations.

Yap, R. C. (2004). Option valuation of philippine forest plantation leases. *Environment and Development Economics*, 9(03):315–333.

Appendices to Chapter 3

A. Definition of cases I, II and IV

The exogenous conditions that affect the outcome of the model and the firm's decision to continue or abandon the project at $t = 1$ are outlined in table A.1. There are four possible cases that depend on the value of state-dependent expected profits at $t = 1$ and the expected value of the investment project at $t = 0$ if firms were forced to complete the investment.

First, in order for firms to participate at all, we require

$$E_{t=1}[\pi^G] - I \geq 0. \tag{A.1}$$

If this were not the case, firms would never make positive profits on the project and the problem becomes uninteresting. This corresponds to case I. Having made this assumption, there are three possible cases that follow.

Case II assumes that expected variable profits are positive for low-cost projects but that the expected value of the investment project if firms were forced to complete the project at $t = 0$ is negative, implying that the expected value at $t = 1$ of the high-cost project is also negative. In this case firms will always abandon the high-cost project. Solving the optimal policy choice for government is much simpler under this assumption because government need not consider the possibility that firms may change their plan of action at $t = 1$. The feasible parameter space as well as the optimal behaviour of firms is illustrated in figure A.1. The downward sloping bold line represents the zero profit condition subject to the optimal behaviour of firms; as ρ increases governments must compensate firms for the diminished investment flexibility by charging a lower income tax. Because the expected profit of the investment project at $t = 0$ when firms are forced to complete the project is negative there is no $\tau_i \geq 0$ that could induce

Case	State dependent profit: $E_{t=1}[\pi^B] - I$	$E_{t=1}[\pi^G] - I$	Expected value if firm must complete investment: $qE_{t=1}[\pi^G]+(1-q)E_{t=1}[\pi^B]-I$	Firm's action plan at $t = 1$:
I	< 0	< 0	< 0	do not participate
II	< 0	> 0	< 0	$[C, A]$
III	< 0	> 0	> 0	$[C, A]$ or $[C, C]$
IV	> 0	> 0	> 0	$[C, C]$

Table A.1: Conditions on exogenous variables and resulting outcomes

firms to participate if they intend to continue regardless of cost. Furthermore, there is a maximum value of ρ above which firms will not participate. Values of $\rho > \rho^{\max}$ will induce firms not to participate at all; the possible loss associated with abandoning a high-cost project cannot be outweighed by lowering taxes on low-cost projects any further. The upward sloping red lines represent the boundaries of the continuation constraints for high and low cost projects. Any combination of ρ and τ_i to the southeast of these lines will induce firms to complete the investment project given that they arrive at $t = 1$. The grey shaded region illustrates the feasible policy region.

Case III imposes the condition that the expected value of the investment project if firms were forced to complete investment at $t = 0$ is positive:

$$\left(qE_{t=1}[\pi^G] + (1 - q)E_{t=1}[\pi^B]\right) - I > 0, \tag{A.2}$$

And that:

$$E_{t=1}[\pi^B] - I < 0. \tag{A.3}$$

Under these assumptions, firms will find it optimal to switch their plan of action at $t = 1$ depending on the government's choice of ρ and τ_i. Thus, there may be some solutions where government finds it optimal to induce firms to continue regardless of cost and some where governments find it optimal to induce firms to abandon high-cost projects. This will depend crucially on the benefit function $u(\cdot)$. Case III represents the central case since government policy can induce

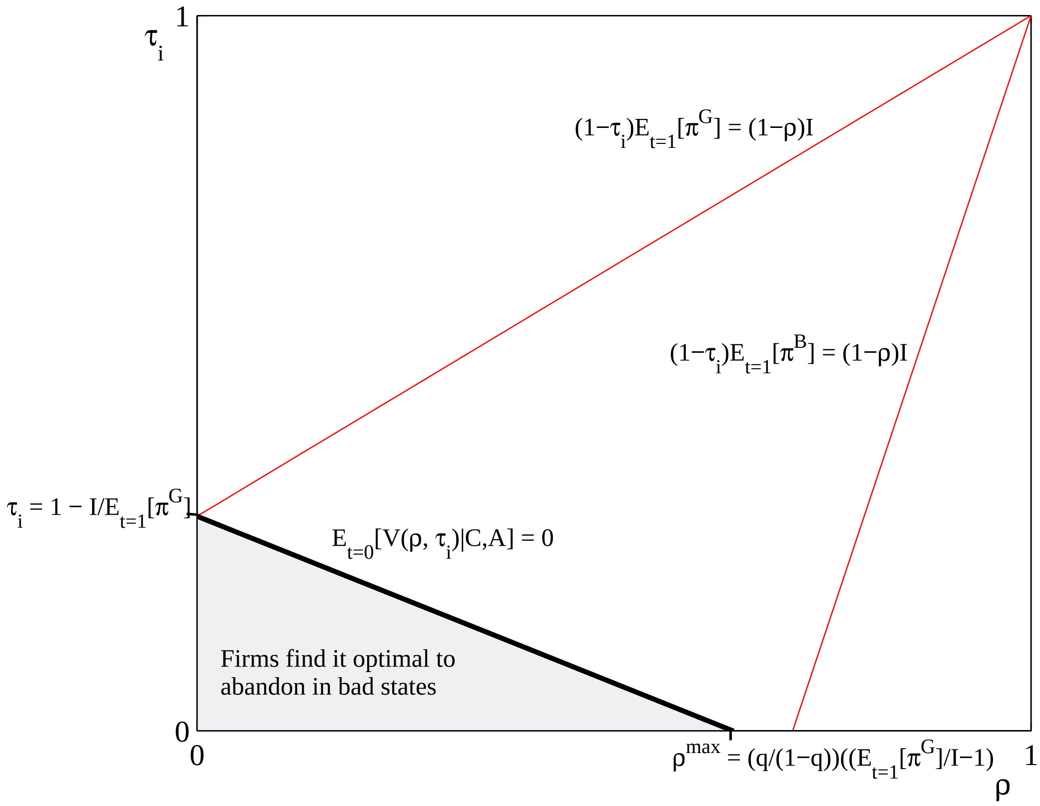


Figure A.1: Feasible policy parameter space corresponding to Case II

firms to either abandon or continue. It is under these conditions that I solve for the optimal policy choice in the main paper. Figure A.2 represents the feasible policy parameters in case III, and contains many of the same features in Figure A.1 including the upward sloping continuation constraints, the downward sloping participation constraint when firms abandon and the bold line illustrating the optimal behaviour of firms along the zero payoff condition. Instead of a maximum value for ρ, $\rho^{\max}$, government must now consider the value ρ^{switch} above which firms choose to continue and below which firms decide to abandon the high-cost project.

The final possible condition on the exogenous parameters described by case IV assumes that the investment project yields positive profits regardless of costs. Under this assumption firms could be induced to continue or to abandon a

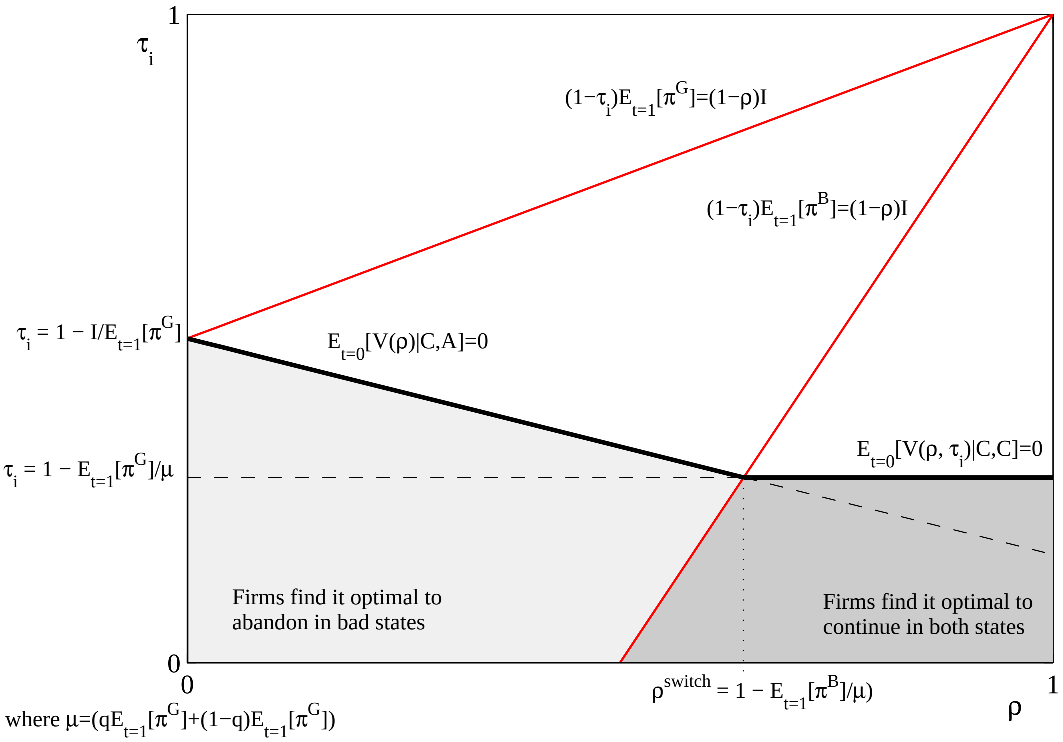


Figure A.2: Feasible policy parameter space corresponding to Case III

high-cost project. However, government will always prefer to induce firms to complete the project, making this case less interesting than case III.

When $E_{t=1}[\pi^B] - I > 0$ and $E_{t=1}[\pi^G] - I > 0$ government always prefers to induce firms to continue the project in good and bad states of the world. Government is able to extract all expected surplus from firms at $t = 0$. This occurs when the participation constraints, subject to the optimal behaviour of firms, are set to zero. When firms find it optimal to continue regardless of cost the zero profit condition yields:

$$\tau_i|C, C = 1 - \frac{I}{(qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B])} \tag{A.4}$$

When firms find it optimal to abandon the high-cost project:

$$\tau_i|C, A = 1 - \frac{qI + (1-q)(\rho I)}{qE_{t=1}[\pi^G]} \tag{A.5}$$

Substituting the zero profit condition into the appropriate government payoff function yields a constrained objective function in terms of ρ:

$$\begin{aligned}
E_{t=0}[U|C, A] &= \tau_i q E_{t=1}[\pi^G] + qu(I) + (1-q)u(\rho I), \\
&= q(E_{t=1}[\pi^G] - I + u(I)) + (1-q)(u(\rho I) - \rho I). \\
E_{t=0}[U|C, C] &= \tau_i(qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B]) + u(I), \\
&= (qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B]) - I + u(I).
\end{aligned}$$

Comparing these constrained objective functions amounts to comparing the expected government payoffs for high-cost realizations. The comparison is

$$\begin{aligned}
E_{t=0}[U|C, C] &\gtrless E_{t=0}[U|C, A], \\
u(I) - I + E[\pi^B] &\gtrless u(\rho I) - \rho I, \\
E_{t=1}[\pi^B] - I &\gtrless u(\rho I) - u(I) - \rho I.
\end{aligned}$$

The LHS is positive by the assumption that $E_{t=1}[\pi^B] - I > 0$. The RHS is negative for all values of $\rho \in [0, 1]$ by the assumption that $u'(\cdot) > 0$. Thus, government would always prefer to induce firms to continue in both states of the world when the investment project yields positive profits in both states of the world.

Intuitively, government internalizes the incentives of firms and makes decisions as if they faced the investment opportunity themselves. If the investment project yields positive profits for high-cost realizations, abandoning the project is never profitable (in the absence of tax) because one forgoes the possibility to earn

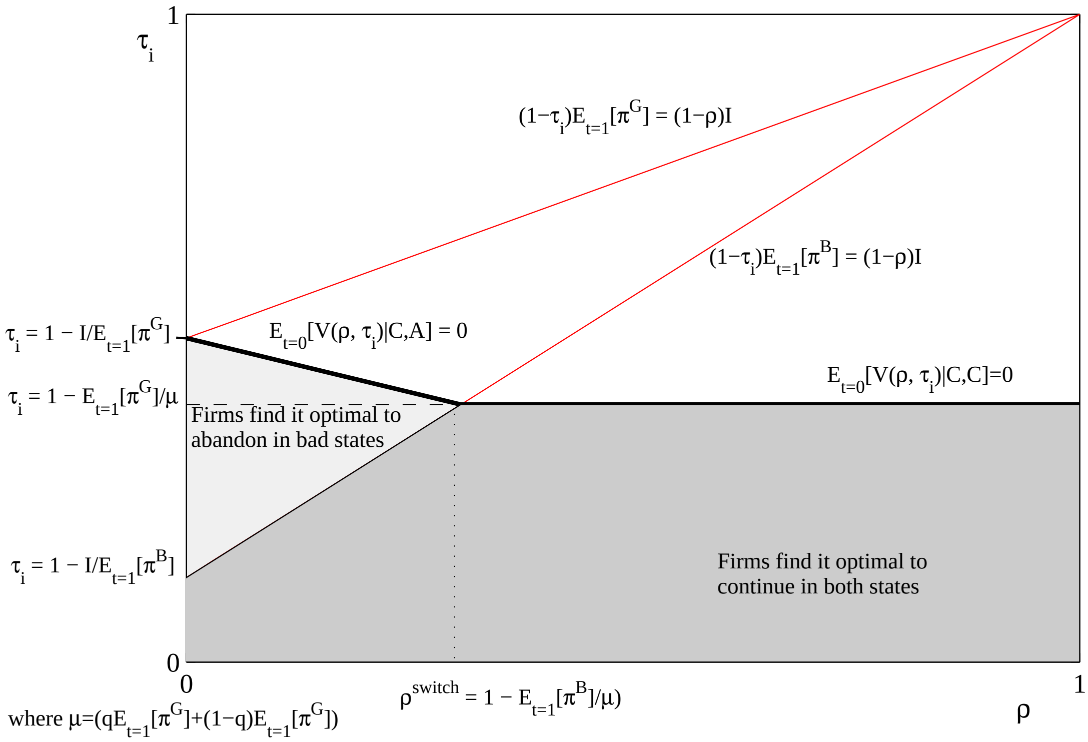


Figure A.3: Feasible policy parameter space corresponding to Case IV

positive profits. In the same way, it is never profitable for the government to induce firms to abandon the high-cost project when the project yields positive profits. Figure A.3 illustrates the feasible policy parameters under these circumstances. This graph contains the same features as figure A.2 except that the intercept for the boundary of the continuation constraint for high-cost projects intersects the τ_i axis at a point above zero. This reflects the notion that for all values of ρ there exists a tax rate that can induce firms to continue in both states of the world and when firms have full flexibility the project still yields a positive expected payoff for high-cost projects at $t = 1$

B. Proofs of propositions

Proposition 1. *With common knowledge of marginal costs it is always optimal for governments to set the royalty rate,* $\tau_r = 0$

Proof. Suppose, at $t = 1$ firms find it optimal to continue in both low and high cost states of the world. In order to induce firms to participate and to maximize total revenue, the participation constraint must be satisfied with equality. Governments choose τ_i and τ_r according to:

$$\max \ E_{t=0}[U|C,C]$$
$$\text{s. t. } E_{t=0}[V|C,C] \geq 0$$

Setting expected firm payoffs at $t = 0$ to zero yields the zero payoff function which constrains the government objective function. It is written as a function of terms in τ_r:

$$\tau_i = 1 - \frac{I}{(qE_{t=1}[\pi^G(\tau_r)] + (1-q)E_{t=1}[\pi^B(\tau_r)]} \tag{B.1}$$

This zero payoff condition is substituted into the government's objective function yielding a constrained object function only in τ_r:

$$= \tau_r(qE_{t=1}[p_t|G] + (1-q)E_{t=1}[p_t|B]) + (qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B]) + u(I) - I \tag{B.2}$$

Recalling that prices are distributed according to a continuous probability density function $f(p_t)$ the constrained objective can be simplified and written:

$$q[\int_{\frac{c}{1-\tau_{roy}}}^{\infty} p_t f(p_t) dp_t - (1 - F\left(\frac{c}{1-\tau_{roy}}\right)(c)]$$
$$+ (1-q)[\int_{\frac{c(1+\theta)}{1-\tau_{roy}}}^{\infty} p_t f(p_t) dp_t - (1 - F\left(\frac{c(1+\theta)}{1-\tau_{roy}}\right)(c(1+\theta)]$$

Note that it is assumed that these arbitrary probability densities contain c and $c(1 + \theta)$ in their support. This need not be the case; if the marginal cost of production were below the set of possible prices, royalties could be increased to a point, without changing the expected price or probability of production. However, for simplicity this special case is ignored. Taking the first order condition of the constrained objective with respect to τ_r yields:

$$\frac{d}{d\tau_r} = q[-\frac{c}{1-\tau_r}f(\frac{c}{1-\tau_r})(\frac{c}{(1-\tau_r)^2}) + cf(\frac{c}{1-\tau_r})\frac{c}{(1-\tau_r)^2}]$$
$$+ (1-q)[-\frac{c(1+\theta)}{1-\tau_r}f(\frac{c(1+\theta)}{1-\tau_r})(\frac{c(1+\theta)}{(1-\tau_r)^2}) + c(1+\theta)f(\frac{c(1+\theta)}{1-\tau_r})\frac{c(1+\theta)}{(1-\tau_r)^2}]$$

Simplifying and rearranging the first derivative shows that for all possible values of $\tau_r \in [0,1]$, except $\tau_r = 0$, the expression is negative:

$$= (\frac{-\tau_r}{1-\tau_r})\{q[f(\frac{c}{1-\tau_r})(\frac{c}{1-\tau_r})^2] + (1-q)[f(\frac{c(1+\theta)}{1-\tau_r})(\frac{c(1+\theta)}{1-\tau_r})^2\} \quad \text{(B.3)}$$

The terms inside the curly brackets are all positive; $f(\cdot) > 0$ as it is a probability density and each term is multiplied by a squared term. The term in front of the curly brackets, $(\frac{-\tau_r}{1-\tau_r})$, is negative for all values of $\tau_r \in (0,1]$. $\tau_r^* = 0$ is the only value of τ_r that satisfies the first order condition. Furthermore, $\tau_r^* = 0$ must be the value of τ_r that maximizes the constrained objective function as the derivative is strictly negative for $0 < \tau_r \leq 1$. □

Proposition 2. *If* $\frac{\alpha}{I} > [1 - \frac{E_{t=1}[\pi^B]}{(q \cdot E_{t=1}[\pi^G] + (1-q)\cdot E_{t=1}[\pi^B])}]$ *then governments will always induce firms to continue in both states of the world*

Proof. If firms abandon the high cost projects, governments will choose the corner solution:

$$\rho^*|C, A = \rho^{switch} = \left[1 - \frac{E_{t=1}[\pi^B]}{(qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B])}\right],$$
$$\tau_i^*|C, A = 1 - \frac{I}{(qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B])}.$$

If firms continue regardless of cost, governments will choose:

$$\rho^*|C, C \geq \rho^{switch},$$
$$\tau_i^*|C, C = 1 - \frac{I}{(q \cdot E_{t=1}[\pi^G] + (1-q) \cdot E_{t=1}[\pi^B])}.$$

Maximized expected payoffs conditional on optimal firm behaviour are:

$$E_{t=1}[U(\rho^*, \tau_{inc}^*)|C, A] = q[E_{t=1}[\pi^G] - I + \alpha \log(I)] + (1-q)[-\rho^{switch} I + \alpha \log(\rho^{switch} I)],$$
$$E_{t=1}[U(\rho^*, \tau_{inc}^*)|C, C] = qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B] - I + \alpha \log(I).$$

Comparing these equations we see that $E_{t=1}[U(\rho^*, \tau_{inc}^*)|C, C] > E_{t=1}[U(\rho^*, \tau_{inc}^*)|C, A]$:

$$(1-q)[-\rho^{switch} I + \alpha \log(\rho^{switch} I)] \gtrless (1-q)[E_{t=1}[\pi^B] - I + \alpha \log(I)],$$
$$-\rho^{switch} I + \alpha \log(\rho^{switch} I) \gtrless E_{t=1}[\pi^B] - I + \alpha \log(I),$$
$$\frac{\alpha}{I} \log(\rho^{switch}) \gtrless \frac{E_{t=1}[\pi^B]}{I} - 1 + \rho^{switch},$$
$$\frac{\alpha}{I} \log(\rho^{switch}) \gtrless \frac{E_{t=1}[\pi^B]}{I} - \frac{E_{t=1}[\pi^B]}{qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B]}.$$

Knowing that $0 < \rho^{switch} < 1$, implies that the LHS will be negative. By assumption, in Case III, $qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B] > I$, implying that the RHS is positive. Thus, $E_{t=1}[U(\rho^*, \tau_{inc}^*)|C, C] > E_{t=1}[U(\rho^*, \tau_{inc}^*)|C, A]$ and governments will always chose (ρ^*, τ_i^*) to induce firms to complete the project regardless of cost. □

Proposition 3. *There exists a cutoff value of* α, $0 < \alpha < I(1 - \frac{E_{t=1}[\pi^B]}{qE_{t=1}[\pi^G]+(1-q)E_{t=1}[\pi^B]})$ *above which governments will find it optimal to force firms to continue in both states of the world and below which governments will find it optimal to force firms to abandon the project in bad states of the world. The* *existence of this cutoff implies that if governments force firms to abandon the project the optimal policy choice must be an interior solution.*

Proof. If firms abandon high cost projects, government's optimal policy, given by (2.24) is:

$$\rho^* = \frac{\alpha}{I} < \rho^{switch},$$
$$\tau_i^* = 1 - \frac{I}{E_{t=1}[\pi^G]} - \frac{(1-q)(\alpha)}{qE_{t=1}[\pi^G]}.$$

If firms continue regardless of cost the optimal policy, given by (2.25) is:

$$\rho^* \geq \rho^{switch},$$
$$\tau_i^* = 1 - \frac{I}{(qE_{t=1}[\pi^G] + (1-q)E_{t=1}[\pi^B])}.$$

Maximized expected payoffs to government conditional on firm behaviour are:

$$E_{t=1}[U(\rho^*, \tau_i^*)|C, A] \gtreqless E_{t=1}[U(\rho^*, \tau_i^*)|C, C]^{21},$$
$$(1-q)[-\alpha + \alpha\log(\alpha)] \gtreqless (1-q)[E_{t=1}[\pi^B] - I + \alpha\log(I)],$$
$$\alpha\log(\alpha) - \alpha\log(I) - \alpha \gtreqless E_{t=1}[\pi^B] - I,$$
$$\frac{\alpha}{I}\log(\frac{\alpha}{I}) - \frac{\alpha}{I} \gtreqless \frac{E_{t=1}[\pi^B]}{I} - 1,$$
$$1 + \rho^*\log(\rho^*) - \rho^* \gtreqless \frac{E_{t=1}[\pi^B]}{I}.$$

If $\frac{\alpha}{I} = \rho^{switch}$ the $LHS < RHS$ and governments prefer to induce firms to continue regardless of cost. If $\frac{\alpha}{I} < \rho^{switch}$, as α approaches zero so too does $\rho^* = \frac{\alpha}{I}$. When

ρ^* approaches 0 from the right, the LHS of the inequality approaches 1.[22] In case III $\frac{E_{t=1}[\pi^B]}{I} < 1$ implying that the $LHS > RHS$ and the governments prefer to induce firms to abandon high cost projects. Finally, the LHS is decreasing in α:

$$\begin{aligned} \frac{d(1 + \frac{\alpha}{I}\log(\frac{\alpha}{I}) - \frac{\alpha}{I})}{d\alpha} &= \frac{1}{I}(\log(\frac{\alpha}{I} - 1) + \frac{\alpha}{I}(\frac{I}{\alpha})\frac{1}{I}, \\ &= \frac{1}{I}\log(\frac{\alpha}{I}). \end{aligned}$$

For all $\alpha \in \left[0, I\rho^{switch}\right]$ this derivative is negative. Thus, as α increases from zero to $I(1 - \frac{E_{t=1}[\pi^B]}{qE_{t=1}[\pi^G]+(1-q)E_{t=1}[\pi^B]})$ government's preferences will switch at some intermediate value from wanting firms to abandon the high cost project to wanting firms to continue regardless of cost. The value of α where $E_{t=1}[U(\rho^*, \tau_i^*)|C, A] = E_{t=1}[U(\rho^*, \tau_i^*)|C, C]$ is defined by the implicit function:

$$\frac{\alpha}{I}\log(\frac{\alpha}{I}) - \frac{\alpha}{I} - \frac{E_{t=1}[\pi^B]}{I} + 1 = 0. \tag{B.4}$$

However, there is no analytical solution for this cutoff value. Governments will prefer to induce firms to abandon the high cost project at some value of α that is strictly less than $I\rho^{switch} = I(1 - \frac{E_{t=1}[\pi^B]}{qE_{t=1}[\pi^G]+(1-q)E_{t=1}[\pi^B]})$. Thus, if governments prefer to induce firms to abandon the project, the solution must be an interior solution to the sub-maximization problem. Recall that any value $\rho^* = \frac{\alpha}{I} < \rho^{switch}$ constitutes an interior solution to the governments optimization problem conditional on firms abandoning the high cost project. □

[22]

write $\rho^* = \frac{1}{x}$, then:

$$\begin{aligned} LHS &= 1 + \lim_{x\to\infty}[\frac{1}{x}\log(\frac{1}{x}) - \frac{1}{x}] \\ &= 1 - \lim_{x\to\infty}\frac{1}{x}[\log(x) + 1] \\ &= 1 \text{ since } \frac{1}{x} \text{ goes to zero, faster than } (\log(x)+1) \text{ goes to } \infty \end{aligned}$$