

Institutional Change and Economic Development

Evidence from Natural and Artefactual
Field Experiments in Ethiopia

Mequanint Biset Melesse

Institutional Change and Economic Development

Mequanint Biset Melesse



Institutional Change and Economic Development

Evidence from Natural and Artefactual Field Experiments in Ethiopia

Mequanint Biset Melesse

Thesis committee

Promotor

Prof. Dr E.H. Bulte

Professor of Development Economics

Wageningen University

Other members

Prof. Dr J. H. H. Wesseler, Wageningen University

Prof. Dr S. T. Holden, Norwegian University of Life Sciences, Ås, Norway

Prof. Dr J.-M. Baland, University of Namur, Belgium

Dr P. S. Dalton, Tilburg University, The Netherlands

This research was conducted under the auspices of the Wageningen School of Social Science (WASS).

Institutional Change and Economic Development

Evidence from Natural and Artefactual Field Experiments in Ethiopia

Mequanint Biset Melesse

Thesis

submitted in fulfillment of the requirements for the degree of doctor

at Wageningen University

by the authority of the Rector Magnificus

Prof. Dr A. P.J. Mol,

in the presence of the

Thesis Committee appointed by the Academic Board

to be defended in public

on Tuesday 15 September 2015

at 4 p.m. in the Aula.

Mequanint Biset Melesse

Institutional Change and Economic Development:

Evidence from Natural and Artefactual Field Experiments in Ethiopia

193 pages.

PhD thesis, Wageningen University, Wageningen, NL (2015)

With references, with summaries in Dutch and English

ISBN 978-94-6257-413-7

To my families.

Contents

Chapter 1	Introduction	1
Chapter 2	Land Certification Programs and Agricultural Productivity	16
Chapter 3	Joint Land Certification Programs and Women's Empowerment	52
Chapter 4	Formal Law and Customary Change	92
Chapter 5	Market Exposure and Risk Aversion	119
Chapter 6	Synthesis	151
	References	165
	Summary (English)	186
	Samenvatting (Dutch)	189
	Acknowledgements	192

Chapter 1

Introduction

1.1 Overview

The post-World War II years have seen striking differences in the distribution of per capita income across countries. Specially, the mid-1970s marked the emergence of a clear bimodal distribution of global income that sharply divided the world into rich and poor countries (Jones, 1997). This was also the period that marked the start of the African growth disaster.¹ Since then, Africa has lagged behind the rest of the world by almost all standards of development. Observing the statistical significance and negative effect of the African dummy in their cross-country growth regressions, Easterly and Levine (1997) have coined the phrase ‘growth tragedy’ to describe the disappointing long-term economic performance of the continent. Evidently, African countries account for about 75% of the lowest quartile of the distribution of the world per capita income.² Average per capita income for Africa is almost less than one-thirteenth that of high income countries. In sum, in the words of Ndulu and O’Connell: “African populations missed out on the unprecedented economic transformation that took place in the rest of the developing world after 1950” (Ndulu and O’Connell, 2008, p. 17).

What has driven this stark cross-country divergence in levels of economic development? The search for explanations to the African growth disaster, broadly the origins of world income divergence, has become one of the most central and challenging themes of research in development economics. This endeavor has prompted a vast body of literature, and several candidate explanations have been advanced over time. Traditional neoclassical growth models attribute differences in per capita income across countries to variations in production factors. Solow (1956) is the landmark to lay the foundation of the neoclassical growth framework. Solow showed that capital accumulation and population growth are the principal determinants of economic growth. Although the model produced simple and empirically testable

¹ Whenever the term Africa is used, it refers to sub-Saharan Africa.

² Our data source is the World Bank’s World Development Indicators 2013. Per capita income refers to per capita gross domestic product, adjusted for differences in purchasing power parity.

hypotheses in predicting the direction of world income differences, it assumed that the factors driving these differences are exogenous. The next generation of growth models endogenized factor accumulation and technical progress, and emphasized that externalities from factor accumulation could induce sustained steady-state growth (Romer, 1986, 1990; Lucas, 1988; Aghion and Howitt, 1998). These models remained squarely within the realm of the neoclassical tradition in explaining differences in growth rates in terms of factor endowments.

While the neoclassical and endogenous growth models have provided useful insights about the mechanics of economic growth, they fall short of providing fundamental explanations for the extremely unequal distribution of global income. Partly, it is because conventional factors of growth alone do not fully account for cross-country differences in levels of economic development, especially those in Africa (Barro 1991, Easterly and Levine, 1997; Rodrik, 1999). Yet, and more importantly, the view that relative income levels differ according to accumulated level of factors of growth begs the obvious question of what are the root causes of differential factor accumulations. Simply stated: why did certain countries fail to save and invest enough, develop and use technologies and have efficient markets? Indeed, North and Thomas (1973) have noted that factor accumulation and productivity are not causes of growth; they are growth. For these reasons, factor accumulation and technical change are considered as proximate causes of growth (Rodrik, 2003). Following this observation, the quest for understanding the root causes of comparative development has raged prominently on the pages of many economic journals and books. While there is no shortage of candidates, two main perspectives have emerged from this rapidly growing literature: *geography* and *institutions*.³

The geography-based perspective on comparative development identifies advantages and disadvantages of both absolute and relative geography as the main driver of long-term economic development. It emphasizes different mechanisms for how geography can affect development, at least in the context of Africa, including depleted ecology, disease burden (malaria), low productivity of available technology and prohibitive transport costs from global 'core markets' (Diamond, 1997; Sachs and Warner, 1997; Gallup et al., 1999; Sachs and Malaney, 2002; Sachs, 2003; Bosker and Garretsen, 2012). Alternatively, the institutional perspective maintains that institutions —broadly understood as the formal and informal rules

³ Alternative explanations for income differences include policies (Rodrik, 2003), culture (Acemoglu et al., 2005) and the colonial past (Acemoglu et al., 2002).

that shape economic and social behavior—are fundamental for understanding the wealth of nations. The institutional view maintains that differences in the quality of institutions are the ultimate cause of differences in economic development (Acemoglu et al., 2005).

Which perspective ‘rules’ has stimulated a hot debate in the literature. While no one can doubt that many African countries face both geographical and institutional challenges and that the two perspectives are not necessarily mutually exclusive causes of economic growth, a large empirical literature has shown that institutions “trump” geography as the dominant cause of long-run (under)development (e.g., Acemoglu et al., 2002; Easterly and Levine, 2003; Rodrik et al., 2004). It has also been increasingly recognized that institutions are, at least partly, endogenous to economic development and cannot be taken for granted (Rodrik, 2000; Dhillon and Rigolini, 2011). Yet, our understanding of their formation and functioning is quite imperfect (Rodrik, 2006). Particularly, much remains to be understood about micro-foundations of institutions and drivers of institutional change. This thesis aims to take us some way in this direction by studying endogenous institutional changes and their outcomes at a micro-level.

1.2 Institutions and economic development

1.2.1 Institutions in the growth process

Institutions are humanly devised incentive frameworks that shape social, political and economic interactions and behavioral outcomes (North, 1990, 1991). They consist of both the “institutional environment” and “institutional arrangements” (Williamson, 2000). The institutional environment comprises both formal (rules, laws and constitutions) and informal (conventions, social norms and beliefs) institutions that set the rules of the game. On the other hand, the institutional arrangements include organizational modes and governance structures that play the game within the purview of the institutional environment. Together, they shape the incentive structure and opportunity sets of economies and determine economic outcomes through their impact on consumption, saving and investment decisions of economic agents. Institutions bring order to economic activities and social relations by affecting information flows, transaction costs, investment risk, expectation formation and the ability of societies to coordinate on collective action in dealing with social dilemmas (Acemoglu et al., 2005).

Apprehension to the notion of institutions dates back to the earliest days of the discipline of economics. Adam Smith (1776) already put forth the famous concept of the “invisible hand”

which disciplines self-interested behavior of individuals in the market. Within this framework, individuals were assumed to be analytically equal in terms of natural endowment of rationality and aspiration. It was the context within which individuals acted that produced different economic outcomes. This context is provided by formal and informal institutions. Despite this initial emphasis on institutions by Adam Smith, classical economists increasingly lost sight of their analysis and importance to economic outcomes. Instead, the demand for modeling gave way to behavioral assumptions in economic analysis, and classical economists had come to take institutions for granted (Buchanan, 1964; Machovec, 1995). In a similar vein, the neoclassical economics theory—the follow-up dominant paradigm in economics thought—emphasized the role of political institutions and private property rights protection for specialization and market exchange, ultimately for economic growth (Hicks, 1969; North, 1991). However, like the classical political economics, the point of emphasis here too was on the functioning of already developed frictionless and static markets, without addressing the ways through which they evolved.

By taking politics and institutions—the necessary building blocks for understanding the process of economic change—as given, the neoclassical economics approach offered limited tools to deal with the subject of economic development. As a result, some of the basic assumptions of mainstream economics have been put into question increasingly. Friedrich Hayek, for example, remarked that: [...] Under the sign of ‘neither individualism nor socialism’ we are in fact rapidly moving from a society of free individuals toward one of a completely collectivist character’ (Hayek, 1948, p. 1). Continuing his critique, Hayek hastened to add that the goal of the Enlightenment writers was not to solve the social dilemma of political and economic order through the assumption of ‘perfect’ individuals, but instead to “find a set of institutions by which man could be induced, by his own choice and from the motives which determined his ordinary conduct, to contribute as much as possible to the need of all others” (Hayek, 1948, p. 13). Only then, the standard emphasis on methodological individualism has been gradually eased in favor of studying institutional structures necessary for economic development. This has reoriented the focus of economic analysis from mere equilibrium theorizing of optimal allocation of resources to a renewed interest in coordination and alternative institutional arrangements, giving way to the emergence of the “new institutional economics” (Williamson, 2000).

Important early contributions to the emergence and development of the new institutional economics include Coase (1960), Demsetz (1967), and North and Thomas (1973). However,

it is in the 1990s that institutions have surfaced to the center stage on the economic development debate. Since then, a growing empirical literature has documented robust evidence that suggests the quality of formal and informal institutions is a fundamental determinant of long-term development (e.g., Hall and Jones, 1999; Acemoglu et al., 2001; Rodrik et al., 2004; Bates, 2006; Angeles, 2011). A follow-up literature has also endeavored to “unbundle” institutions so as to identify dimensions of the institutional framework that matter most for economic development (Acemoglu and Robinson, 2005; Voors and Bulte, 2014).

By and large, that institutions matter for economic development has now become one of the unifying themes in development economics. However, we are still far from a complete understanding of several relevant aspects of institutions. Notably, there has been little work on the mechanisms through which institutions affect development outcomes, and even less on institutional interactions and drivers of institutional change over time. Do a country’s formal and informal institutions coax or clash? How can a country go about improving its institutions? What factors determine the success and impacts of an institutional change? How does an exogenous expansion of the realm of formal institutions affect informal institutions? Do informal institutions respond to exogenous shocks? All of these questions are fundamental, and many are yet to receive formal analysis. The subsequent chapters of this thesis aim to contribute to the literature by addressing (some of) these questions.

1.2.2 Interaction between formal and informal institutions

The institutional grid of economies is made up of differently configured formal and informal institutions. The distinction between formal and informal institutions is typically related to whether or not they are subject to the purview of the state. While formal institutions are centrally designed and enforced, informal ones are subject to self-governance by economic agents (Dixit, 2004; Helmke and Levitsky, 2004).⁴ Though arguably interdependent, research on formal and informal institutions has emerged independently. Particularly, the interaction between formal and informal institutions is not sufficiently understood. In fact, too much emphasis has been put only on the importance of formal institutions (Acemoglu et al., 2005). Not surprisingly, the literature has largely abstracted from studying informal institutions, viewing them often as mere functional substitutes that change incentives to use alternative formal constraints (Zenger et al., 2002).

⁴ The enforcement mechanisms of informal institutions often include exclusion from social groups, ostracism by neighbors and friends, or the loss of reputation (Pejovich, 1999).

Particularly, economic development thinking in the 1980s was motivated by the Washington Consensus, a major initiative aiming to open formal institutions of governance in developing countries to the greater influence of market forces. Reserving a central role for the operation of markets, this initiative stirred many developing countries to engage in extensive macroeconomic reforms, liberalization and privatization programs in order to “get the prices right” (Rodrik, 1990). Yet, economic performance has remained lackluster in these countries, and the initiative has been subject to an intense fire of criticisms for its context-free technocratic reforms and its too little attention to interaction effects (Easterly, 2006; Rodrik, 2006). Now, it is increasingly recognized that context-specific and firmly established informal institutions are not simple appendages to the formal ones (North, 1990; Platteau, 2000). This observation has increased the demand for understanding the nature of the interdependence between formal and informal institutions.

Following this impetus, the interaction between formal and informal institutions has inspired a growing body of literature over recent years (e.g., Boettke et al., 2008; Williamson, 2009). This literature predicts that the interaction between formal and informal institutions is critical to explaining institutional outcomes in communities (Roland, 2004). Typically, formal institutions are often considered as exogenous constraints, while informal institutions are modeled as endogenous self-enforcing rules (Aoki, 2001; Greif and Laitin, 2004; Aldashev et al., 2012b). Within this framework, an exogenous expansion of the realm of formal institutions is expected to have important implications for informal institutions. As result, the interaction between formal and informal institutions may be central to understanding the dynamics of economic development. In this context, Chapter 4 of this thesis studies the interaction between formal law and customary legal institutions.

1.2.3 Institutional change

Today, there are only few other topics on which contemporary economists have come to apparent consensus as readily as on the significance of institutions for development. A wealth of empirical literature has persuasively established the pedigree of both formal and informal institutions for economic performance and distribution. It is now widely accepted that ‘getting institutions right’ is no less important than the ‘getting prices right’ maxim of the neoclassical paradigm. Not surprisingly, some scholars even distill the challenge of economic development to the problem of transforming institutions. North (2005), for example, argues that today’s development challenge lies not in technological solutions to production and distribution, but

in structuring institutions to promote economic exchange and growth. In a similar spirit, Hoff and Stiglitz (2001, p. 389) concluded that “development is no longer seen primarily as a process of capital accumulation but rather as a process of organizational change”.

As a result of this heightened interest in the role of institutions, institutional change has become a subject of first-order relevance in development economics. Unfortunately, our understanding of institutional change is quite far from complete. “Our theories of how fundamental institutional change occurs are underdeveloped” (Evans, 2004, p. 32); “there is much to be learned still about what improving institutional quality means on the ground” (Rodrik et al., 2004, p. 158). In fact, institutional development has been billed as a field with much action and little theory (Dixit, 2009).

Equally, it is known that the ‘technology of institution building’ and its associated challenges are greatly different based on whether the purpose of institution building is to get hold of ‘good’ institutions—in the sense of institutions that promote and sustain growth—where they do not already exist or to transform informal institutions that appear to be incompatible with modern economic growth. Clearly, regarding the former objective, the standard approach has been to transfer or transplant ‘best practice’ institutions that have undergirded the growth of the developed world (Berkowitz et al., 2003; Rodrik, 2006; Couyoumdjian, 2012). Historically, the colonial era and the early years of independence were periods of intensive institutional transfer, albeit involuntary (Brousseau et al., 2011). More recently, the transition of former socialist countries of the Eastern European block has provided another period of intensified institutional transfer (Pejovich, 1999; Grzymala-Busse, 2010). However, this process of wholesale institutional transfer has produced disappointing results in many developing countries (Evans, 2004; Rodrik, 2006).

Institutional transplantation has been subject to a deserved criticism for overlooking local contexts of institutions. Institutions do not function in a vacuum—they require knowledge of ‘place and time’ for their accommodation—but also interact with other institutions (Roland, 2004; Couyoumdjian, 2012). Evidence shows that informal institutions shape the outcome and performance of formal institutions in important and often unexpected ways (Boettke et al., 2008; Williamson, 2009). Accordingly, institutions that perform well in one setting may be inappropriate in other settings, without the supporting norms and complementary institutions. While formal institutions may be the same, the informal norms of behavior, the subjective models of actors, the enforcement mechanisms and the way enforcement occurs are not the

same across countries. Nowhere has this become more apparent than in the diversity of trajectories of post-socialist transforming economies. Douglass North emphasized this point by arguing that institutional forms do not always map into specific economic outcomes: “Economies that adopt the formal rules of another economy will have very different performance characteristics than the first economy because of different informal norms and enforcement. The implication is that transferring the formal political and economic rules of successful Western economies to third-world and Eastern European economies is not a sufficient condition for good economic performance” (North, 1994, p. 366). Indeed, weak state institutions that permeate many developing countries may be a result of ill adapted and incongruent sets of imported institutions.

On the other hand, imported formal institutions may also endure a low level of legitimacy, lack robust complementarity with other formal institutions, or even produce unexpected and undesirable consequences (Grzymala-Busse, 2010). For these reasons, some scholars warn that imposing a set of formal institutions may not be a panacea for the developing world. While denying the existence of “one-size-fits-all” formula for institution building, Rodrik (2000) argues in favor of experimentation and learning to accommodate ‘local knowledge’. Evans (2004) explicitly warned against the danger of ‘institutional monocropping’, i.e. the imposition of standardized institutional blueprints across the board, and called for ‘deliberative development’.

The task of building institutions is even more challenging when it comes to transforming informal institutions that appear to constrain modern development. Informal institutions are often characterized as highly persistent. When change occurs, it is generally expected to be slow, incremental and continuous (Roland, 2004). This is because informal institutions have a more complex genesis and they are firmly grounded in the practices, customs, values and beliefs of indigenous people (North, 1990; Boettke et al., 2008). Institutional economists believe that they emerge spontaneously and endogenously as solutions to coordination problems of communities (Williamson, 2000; Greif and Laitin, 2004). Given their evolutionary origin, informal institutions tend to evolve over time via adapting and responding to all sorts of experiments, new ideas, power plays and external shocks. An equally important aspect of informal institutions is that they cannot be subject to deliberate choice mechanisms. As such, explicit bargaining is less relevant to explaining their evolution, largely due to the fact that their core is tacit (Couyoumdjian, 2012).

Nonetheless, informal institutions do change, and the impetus for their change may come from several sources, albeit with varying degree of success. To begin with the rare window of opportunity for informal institutional change, a sharp break from established procedures can cause informal institutions to move out of path dependence and allow the introduction of a range of new formal institutions with reasonable good chances of adherence (Williamson, 2000; Roland, 2004). For example, Acemoglu et al. (2001) contend that independence from colonial rule presents Botswana such a rare window of change. Alternatively, a sharp break from existing institutions can also be designed on purpose—the radicalism approach to institutional change. This approach is expected to be more comfortable with large-scale changes or departures from existing institutions (Kingston and Caballero, 2009). However, the success of such changes is often dubious—even when applied to end oppressive informal institutions—due to institutional inertia and path-dependence.

Formalization is another important source of informal institutional change. A classic example is formalization of customary land rights. Customary land rights are often formalized on the grounds that their formalization increases transparency and predictability (Deininger, 2003). Arguments supporting formalization of customary land rights often suggest that they are cheap and function well enough for low levels of economic development, but that only formalization provides the certainty that higher levels of investment and development require. However, in many African countries, formalizing customary land rights has produced contentious outcomes. Contrary to the conventional wisdom, the benefits of formalizing customary land rights do not appear to be automatic (e.g., Braselle et al., 2002; Jacoby and Minten, 2007). A host of conditions, including the *de facto* level of tenure security and history of the customary rights, appear to shape the effects of formalizing land rights (Bromley, 2008). In this context, Chapter 2 investigates productivity incentives of formalization of land use rights in Ethiopia, and elaborates how the history of the local institutions of the country might have gauged the outcomes of the process.

Formal institutional change is yet another relevant source of impetus for informal institutional change. This is a direct implication of the ‘interaction’ hypothesis (Pejovich, 1999; Boettke et al., 2008). Formal institutional change may alter the status quo conditions that sustain informal institutions, and thereby trigger them to change. One of the unpleasant features of informal institutions is their propensity to uphold a culture of inequality in communities through entrenching mechanisms that perpetuate local power structures and unequal institutional relations (Platteau, 2000; World Bank, 2002). In such contexts, a change in

formal institutions may reorient the distribution of power and resources within communities, weakening those who benefit from a particular informal institution and empowering those who seek to change it (Helmke and Levitsky, 2004). Chapter 3 explores the effects of introducing joint land certification programs (i.e. titles containing the names of both spouses) on the bargaining position of women in rural households, where male dominant culture is the norm.

Alternatively, inasmuch as informal institutions are endogenous to formal ones, formal institutional change can also alter the costs and benefits of adhering to particular informal institutions (Helmke and Levitsky, 2004), thus serving as a catalyst for informal institutional change. For instance, an exogenous expansion of the realm of the formal law into communities in remote poor rural areas, which were previously excluded from its direct influence, may have implications for both decision outcomes of the customary law and behavior of its adherents. Within this framework, Sandefur and Siddiqi (2013) suggest that improving access to formal law may result in direct gains for those disadvantaged by informal institutions. Yet, if formal legal institutions become accessible and competitive, indirect changes in customary outcomes may be even more salient (Aldashev et al., 2012a, 2012b). Moreover, this penetration of the formal law may also have important consequences on the type of sub-game theoretic equilibria that can be sustained. Caselli (1997) and Dixit (2004) have demonstrated that it may have unforeseen welfare consequences. Chapter 4 looks at the effects of increased ‘competition’ by the formal law on the arbitration outcomes of the customary law and behaviors of its adherents, while commenting on the social welfare implications of such a change.

Finally, understanding institutions themselves is also central to grasping the complexity and dynamics of the relationship between institutional change and economic development. Not always do we observe a one-to-one mapping between forms and functions of institutions (North, 1990). Unforeseen mechanisms may be at work to generate spillover effects from a given institutional change, resulting in ‘functional multiplicity’ of institutions (Chang, 2007). For example, markets, which have been classically assumed as a simple neutral mechanism to allocate resources, have been shown to do more than allocating goods and services (Bowles, 1998). More often than not, empirical studies have documented robust evidence suggesting that markets shape and influence the evolution of rationality and other deeper behaviors of individuals (e.g., List and Millimet, 2008; Henrich et al., 2010; Cecchi and Bulte, 2013). Chapter 5 analyzes whether exposure to market affects risk preferences of poor farm

households, who are largely characterized by self-subsistence and extremely limited integration into markets.

1.3 Objectives and research questions

The contemporary dominant perspective on economic development is that developing countries are poor because their current institutions do not provide appropriate incentives to promote and sustain growth (see Acemoglu et al., 2005). Moreover, there is a high level of agreement in the literature that a good institutional framework is one that is characterized by well-defined and secured property rights, robust legal system, functional political constraints and good governance, and a sense that these elements are interdependent (North, 2005). Yet, from the point of view of developing countries, this is normally no more than a list of good institutions with little understanding of how they could be developed or how exactly they are interrelated.

An earlier debate has focused on whether these ‘best practice’ formal institutions should be transplanted to developing countries. However, such generic institutional solutions have proven time and again ‘unfit’ to the specificities of the very different contexts in these countries (Evans, 2004; Rodrik, 2006). It is now clear that the mere introduction of formal institutions is itself no guarantee of their efficacy. In practice, understanding the process of institutional change requires thinking beyond institutional designs and forms. Institutional change is likely to reflect a wide range of influences and interactions. It is rarely the result of only technocratic, voluntary or pragmatic negotiations. Instead, it is more likely to be the outcome of inevitable power struggles and dynamics (Acemoglu and Robinson, 2006). Equally, the interaction between different institutional domains profoundly matters in shaping the success of an institutional change (Roland, 2004; Boettke et al., 2008). A change in one institutional sphere is likely to affect other institutional spheres to induce a complex dynamics and a given formal institutional blueprint does not generate the same institutional outcome under different conditions. Furthermore, institutional changes may also have unintended effects (Williamson, 2009). As a result, some reconsideration is underway recently to differentiate between institutions and institutional outcomes (see Rodrik et al., 2004).

The overarching objective of this thesis is to improve our understanding of the relationship between institutional change and economic development. It does so by analyzing the endogenous formation of institutions and outcomes of institutional changes on the quality and

sustainability of other institutions and the dynamics of economic development. The main research questions explored in each chapter of this thesis are the following:

- (i) Does formalizing of land rights boost agricultural productivity?
- (ii) Does access to land rights improve the bargaining power of women in the household?
- (iii) How does the threat of competition from the formal law affect arbitration outcomes of the customary law?
- (iv) How does exposure to market influence risk preferences of poor farmers?

1.4 Methodology

An obvious challenge of any research effort that aims to study institutions is coping with measurement problems. This problem is even more laborious when one attempts to gauge the quality of institutions at micro-level, as has been in this thesis. In particular, identifying and objectively measuring informal institutions poses a non-trivial challenge, since they are often related to dimensions of a society's culture, values and norms that are amorphous in their nature. This thesis has attempted to overcome these challenges by combining survey approaches and field experiments in its data collection efforts. Three waves of survey work—two waves in 2012 and the other in the fall of 2013—were conducted in Amhara region, Ethiopia. The survey contained several detailed and specific modules to collect rich data on many aspects of rural households from 31 Kebeles (communities)—the details of which are discussed in each of the chapters.

The survey data were augmented by data collected through field experiments. Field experiments are being increasingly used as a tool in development economics research (see Smith, 1982; Harrison and List, 2004; Gerber and Green, 2012). Field experiments have proven useful not only in assessing the effects of development programs but also in measuring economic concepts that were until recently very difficult to gauge quantitatively (Banerjee and Duflo, 2009). They allow better and objective measure of data in an incentive compatible way by inducing participants to make choices over dilemmas with monetary payoffs. Broadly, field experiments are often classified into three general categories (Harrison and List, 2004). First, artefactual field experiments—also called lab-in-field experiments—mimic lab experiments except for drawing participants from the 'field' of interest. Second, framed field experiments are the same as artefactual field experiments but contextualize the experimental setting with commodities, tasks or information sets used by participants. Third, natural field

experiments occur in the very environment where participants naturally undertake a certain task, but subjects do not know that they are participating in an experiment.

Experiments rely on randomization to facilitate causal attribution. Randomization serves as a convincing instrumental variable by creating variation in treatment among participants (List, 2014). Often, randomized control trials are employed to allow for the most rigorous and clearest causal attribution (Wantchekon and Guardado, 2011). The causal inference problem in simple natural experiments has also been addressed by using other methods that make necessary identification assumptions to create credible counterfactuals—including propensity score matching, difference in difference, regression discontinuity and instrumental variables approaches (Duflo et al., 2008). For the purpose of this thesis, we have specifically used randomized field experiments, propensity score matching and instrumental variables approaches.

The field experiments employed in this thesis consist of several categories, including public goods games, ultimatum games and risk games. Some of the field experiments were completely decontextualized, like the risk games (Chapter 5). Others were framed in the context of real economic environments, such as dispute arbitration by a local customary judge (Chapter 4). In the research conducted in this thesis, these field experiments were used for two main purposes. The first purpose was to measure otherwise murky economic concepts, such as quality of institutions and risk preferences. Particularly, they allowed us to measure internalized informal institutions by eliminating the potential for strategic or reputational considerations, which often characterize non-anonymous interactions of subsistence and hierarchal rural communities in developing countries (Chapters 3, 4 and 5). The second purpose was to engineer exogenous variation in the context of real economic environments to establish credible causal relations and learn about the underlying mechanisms (Chapter 4).

Yet, important practical challenges also pervade efforts in field experiments. Of overriding importance are ethical issues. Often, individuals need to offer their consent to participate in experiments. Rigorously, consent in field experiments will be voluntary only if participants are physically able to provide consent, free from coercion and can comprehend the risks and benefits involved in the experiment (List and Rasul, 2011). In all the field experiments that we carried out in this thesis, we carefully conveyed the risks and practical consequences of participation to all involved individuals. We asked informed consent from participants, granting the freedom to voluntarily terminate the tasks at any time. We usually emphasized

the importance of anonymity both to participants and our experimenters, so as to minimize the risk of any conflict, retaliation or revenge among participants and beyond after the experiments ended.

1.5 Outline

The chapters in this thesis are organized as follows. Chapters 2 through 5 address the four complementary research questions outlined above. Each chapter consists of a stand-alone academic paper, with its own specific contributions to the literature. Chapter 2 investigates the impacts of the Ethiopian land certification and registration program on agricultural productivity, using econometric matching methods. Chapter 3 studies the spillover effects of land reforms in the form of joint land certification programs (i.e., titles containing the names of both spouses) on the bargaining power of women in Ethiopian rural households, by combining econometric matching and field experimental methods. Chapter 4 explores changes in the customary law induced by the increased penetration of the formal law, by examining arbitration decisions made by real customary judges in Ethiopia over lab-in-field disputes. Chapter 5 analyzes how market exposure influences risk preferences, through linking lab-in-field and field behaviors of poor farmers in Ethiopia. Finally, Chapter 6 synthesizes the main findings of the thesis and discusses their implications for policy and future research.

Chapter 2

Land Certification Programs and Agricultural Productivity

Evidence from Ethiopia

Abstract

According to economic theory, tenure security is an important determinant of agricultural investment and productivity. Land certification has been at the center stage of development efforts of many African countries to boost tenure security. We investigate the productivity impacts of the Ethiopian land registration and certification program, employing propensity score matching method in an effort to control for endogeneity. Consistent with theory, we find land registration and certification has robust positive effects on farm productivity, and identify the assurance effect as the most likely channel for impact. Households with land certificates are more likely to adopt soil-fertility management strategies on their plots than households without certificates.

Publication status: Melesse, B. M. & Bulte, E. (2015). Does Land Registration and Certification Boost Farm Productivity? Evidence from Ethiopia. *Agricultural Economics*. In Press

2.1 Introduction

Land remains an important resource on which a large portion of the African population earns its livelihood, and is at the heart of economic, social and political lives in many African economies (Cotula *et al.*, 2004; Toulmin, 2009). Given this vital importance, land reforms have been at the center stage of development efforts of several African countries. Particularly, since millions live and work on land they do not legally own in accordance with enforceable state law, the quest for secure land rights features high on the development agenda of governments and the international development community. This is because tenure security is perceived to be a precondition to leverage agricultural investment and productivity, spurring economic growth (see, e.g., Deininger and Feder, 2009; Holden *et al.*, 2011).

Theoretically, the literature identifies several channels through which tenure security can contribute to economic growth (Besley, 1995; Brasselle *et al.*, 2002; Deininger and Chamorro, 2004; Place, 2009; Toulmin, 2009). First, secure tenure rights can boost land-related investments by providing assurance to landholders that returns on these investments will not be appropriated. This is particularly relevant for investments characterized by longer gestation periods (Jacoby *et al.*, 2002). Second, more secure and transferable land rights allow factor mobility, and provide opportunities for efficiency gains by reallocating land to more efficient users via land sales and rental markets. Third, making land rights more secure and transferable may increase access to formal credit by enabling the use of land as collateral. However, land markets must be liquid enough, and no legal provisions should limit foreclosure to reap this potential benefit (Deininger *et al.*, 2011). Brasselle *et al.* (2002) coin these three major effects of secured land rights as, respectively, the “assurance”, “realizability”, and “collateralizability” effect.⁵

Based on these arguments, many countries have launched land registration and certification programs. However, for the African context it is surprisingly hard to come by conclusive empirical evidence supporting the view that land registration and certification is instrumental to development. While empirical evidence from Latin America and Asia tends to corroborate

⁵ There are other theoretical reasons for commissioning land registration and certification schemes. Secured land rights reduce the level and likelihood of possible land conflicts (Fenske, 2011). In the same vein, tenure security backed by formal mechanisms is more likely to reduce the need to expend private resources to defend land rights and to preclude the need for presence of landholders on a continuous basis in order to safeguard their land claims, thus increasing their participation in off-farm labor markets (Field, 2005; Deininger and Feder, 2009).

this proposition⁶, evidence is rather mixed at best for Africa. While some papers report that stronger land rights are associated with an increased likelihood of investment (e.g., Besley, 1995; Hayes *et al.*, 1997; Deininger and Jin, 2006; Holden *et al.*, 2009; Deininger *et al.*, 2011), others do not (e.g., Place and Hazell, 1993; Place and Migot-Adholla, 1998; Brasselle *et al.*, 2002; Jacoby and Minten, 2007). Empirical work in this domain faces serious challenges in light of well-recognized (self) selection issues, and the notion that investment behavior of households may affect tenure security (e.g., Sjaastad and Bromley 1997, Besley 1995, Brasselle *et al.* 2002, Holden *et al.* 2009, Deininger *et al.* 2011).⁷

Nevertheless, amid such mixed experiences, a new generation of land certification programs has been fueled by several factors; chief among them are new legislation, low-cost certification methods and continuously increasing demand for land (Cotula *et al.*, 2004). The Ethiopian land registration and certification program has been considered as a pilot in Sub Saharan Africa (SSA) for such certification programs. With the goal of issuing every rightful holder of farmland a certificate of usufructs, the program is the largest delivery of non-freehold rights in a short time period (Deininger *et al.*, 2011).

A few careful empirical papers have documented positive effects of the Ethiopian land certification program. Deininger *et al.* (2008) describe the implementation process and provide an evaluation of its early impacts based on subjective assessments by Ethiopian households. Holden *et al.* (2011) find positive effects of land certification on the allocative efficiency of the land rental market in the Tigray region. Likewise, Deininger *et al.* (2011) report that land registration and certification improved tenure security, and positively affected investment and supply of land to the rental market in the Amhara region.

Little work has been done on the program's impact on agricultural productivity. One key exception is Holden *et al.* (2009), who exploit intra- and inter-household variation in certification status to probe the effect of certification on productivity. They found significant positive effects of certification on investment and productivity in the Tigray region. Following Holden *et al.* (2009), our purpose is to investigate the effects of the Ethiopian land registration and certification program on agricultural productivity. Our study complements Holden *et al.* (2009) because we focus on another Ethiopian region with a different registration and

⁶ See, for example, Alston *et al.* (1996) on Brazil; Jacoby *et al.* (2002) on China; Deininger and Chamorro (2004) on Nicaragua; Field (2005) on Peru; Bouquet (2009) on Mexico; and Galiani and Scharfrodsky (2010) on Argentina.

⁷ Empirical work is also complicated because of restrictions on certain forms of land use (e.g. tree planting), or because of public investment programs that might crowd in or crowd out private investments in land quality.

certification program—allowing us to assess the robustness of the earlier findings to another context and program design. For example, while certificates are issued only in the name of the household head in Tigray, the Amhara region requires joint certification with certificates issued in the names of both spouses.⁸ Moreover, the registration process was more intensive and costly than in other regions of the country due to involvement of a survey and land administration team (Deininger et al., 2008). In addition, while the threat of land redistribution never appears far off the agenda in Ethiopia, it was only in the Amhara region that land was actually redistributed as recently as 1997, after the current regime took power. The redistribution was characterized as politically motivated and generated considerable conflict (Ege, 1997).

Unfortunately, we do not have access to panel data. Relying only on cross-section data, we use propensity score matching for causal inference based on comparing certificated plots in certificated Kebeles with plots in Kebeles that were left uncertificated for exogenous reasons (see below). We employ several approaches and quality indicators to test major identifying assumptions behind our method to diagnose its credibility and performance. Nevertheless, as with other (non-randomized experimental) cross-section studies, we realize that reservations may remain about our ability to attribute causal effects to specific treatments (certification). While we seek to minimize concerns about unobserved heterogeneity, we cannot eliminate all potential concerns about endogeneity. Hence, we also provide an important qualification to our findings in the concluding section.

Our main result is consistent with theory and with Holden et al. (2009). Land registration and certification has large and robust positive effects on farm productivity. A simple comparison of the costs and benefits reveals that benefits substantially exceed the cost of implementation. Based on reasoning that eliminates alternative mechanisms, we then proceed by arguing that our results are presumably driven by the so-called assurance effect.

The remaining part of the paper is organized as follows. Section 2 provides a brief account of the Ethiopian tenure system and describes the land registration and certification program in some detail. Section 3 describes the data and outlines the empirical strategy. This is followed by the empirical results in Section 4. Section 5 presents robustness checks and sensitivity

⁸ Note that such joint certification is not a peculiar feature of the Amhara region – the Oromia and the Southern Nations and Nationalities regions also issued joint certificates.

analysis of the results, while Section 6 summarizes the findings of the study and attempts to put them into perspective.

2.2 The Ethiopian tenure system and land certification program

2.2.1 The Ethiopian land tenure system

In Ethiopia, the land tenure system has been at the forefront of policy debates for generations, owing to the prominence of land as a source of livelihood to the majority of the population and as a source of political power (Crewett and Korf, 2008; Bezabih et al., 2011). In the decades prior to and during the imperial era, land was concentrated in the hands of absentee landlords; tenure was highly insecure, and arbitrary evictions posed a serious threat to tenant farmers (Joireman, 2001; Deininger and Jin, 2006).

Soon after overthrowing the imperial regime of Haile Selassie through a military coup, the socialist Derg regime implemented radical reforms that fundamentally altered the agrarian structure and associated access mechanisms to land. Through the land reform proclamations of 1975, it transferred ownership of rural land, and the power to define access, distribution and tenure terms of land use rights, to the state (Rahmato, 2004; Deininger and Jin, 2006; Deininger et al., 2008). The reform abolished exploitative landlord-tenant relations based on land ownership. Transferability of land through lease, sale, exchange or mortgage was prohibited (Holden and Yohannes, 2002). Frequent land redistributions to accommodate new claimants were important features of the tenure system (Holden et al., 2011).

After the fall of the Derg regime in 1991, the current government embarked on liberalizing the economy. However, the set of reforms largely “overlooked” the land problem, and the legacy of the Derg regime continued to define many elements of the current land policy (Crewett and Korf, 2008). Land rights have continued to be vested in the state, which has been enshrined in the Constitution (Rahmato, 2009b). Constitutionally, land redistribution is installed in the land policy to assure the right to land of every Ethiopian who wants to live on agriculture, making generalized tenure insecurity a feature of the land policy. Furthermore, land use rights are made contingent on a host of conditions, like appropriate land and environmental management, and physical residence in a village (Deininger and Jin, 2006; Rahmato, 2009a).

The current government has made several changes to the land policy (Deininger and Jin, 2006; Crewett and Korf, 2008; Fenske, 2011). First, it devolved responsibility for land issues to regions. Second, the frequency of land redistribution has been stalled, but land reform is not completely off the agenda. Third, landholders can transfer their rights to land to their heirs without conditions. Moreover, land rentals are allowed officially, though some regional governments still impose restrictions on the amount of land and terms of rental contracts.

Overall, the Ethiopian land tenure system has been identified as a major culprit for the weak performance of the country's agrarian sector (Bezabih et al., 2011). While the particulars and terms of the discourse are different, the state continues to be a major source of insecurity. The government remains critical of privatizing land holdings, retaining a discourse of social equity and state protection from land concentration in the hands of the few. Some scholars, however, have argued that the government uses land rather as a "carrot and stick" to fulfill political ends (Ege, 1997; Crewett and Korf, 2008). Partly in an effort to respond to these critics, and to bolster farmers' tenure security, the government commissioned a large-scale land registration and certification program in 1998.

2.2.2 The Ethiopian land registration and certification program

The Ethiopian land registration and certification program has delivered the largest non-freehold land rights in a short period in Sub-Saharan Africa (Deininger et al., 2011). The program was implemented in the four major regions of the country: Tigray, Amhara, Oromia and the Southern Nations and Nationalities. It was started in Tigray region in 1998, and had been turned into a comprehensive program in other regions since 2002 (Holden et al., 2009; Rahmato, 2009b). The Amhara region, where our survey was conducted, started a pilot in 2002. Subsequently it started scaling up the program in 2003. Since the onset of the program, more than 20 million parcels of some 6 million rural households have been registered across the country (Deininger et al., 2011). Books and certificates are issued with names of the household head (or names of both spouses), and details related to size, location and land quality of farm plots. They also contain a summary of legal provisions in the form of holders' rights and obligations.

The process of land registration and certification starts with election of an independent Land Use and Administration Committee (LAC)⁹. The LAC assumes responsibility for the main

⁹ The process of the program varies to some extent from region to region. Here the descriptions are mainly based on the experience of the Amhara region, where our survey was conducted.

preparatory tasks of certification and registration, which include plot identification, demarcation and measurement, and recording of personal details of holders and their family members. The responsibility also includes a systematic field-based adjudication of rights through a public process to resolve conflicts in the presence of neighbors and elders.

The registration process is implemented in a highly decentralized, broad-based, high-speed and participatory approach, which allows the program to adapt to local conditions (Deininger et al., 2008; Toulmin, 2009). It has been regarded as an appropriate response to local needs as evidenced by the recipients' readiness to pay to replace lost certificates, high demand for a spatial reference coupled with willingness to pay and overall positive assessment of likely impacts (Deininger et al., 2008; Holden et al., 2011).

Another feature of the program is its cost effectiveness. With estimated average cost of about US\$1 per plot or US\$3.2 per hectare¹⁰ (Deininger et al., 2008),¹¹ the program is said to have set a new standard. This cost is an order of magnitude below what has been reported elsewhere in the literature for similar first-time land registrations, with average costs between US\$20 and US\$60 per parcel, at times even exceeding US\$100 per parcel (Deininger et al., 2011). For instance, Jacoby and Minten (2007) reported an average cost of about US\$350 per parcel for certification upon demand in Madagascar. The Ethiopian land registration costs are low and affordable, owing mainly to the employment of local tools and techniques with strong local participation (Holden et al., 2011). However, as argued by Holden et al. (2009), since low cost may also imply low quality, compromising program effectiveness, it is imperative to test the impact of the program in terms of its stated objectives.

2.3 Data and empirical strategy

2.3.1 Data description

Cross-section data used for this study were collected in Amhara region, which is the second most populous region of the country, in northwestern Ethiopia.¹² The household survey was

¹⁰ US\$1 was approximately equal to 9.6 Ethiopian Birr in 2007/2008.

¹¹ The costs were computed based on data from the Amhara region where the first-time registration process was much more intensive and costly than in other regions due to the involvement of a survey and land administration team in addition to the LAC. It is important to note also that the costs include both public costs of running the program and private household costs for registry books and certificates. Though LAC members work completely voluntarily and free, the project cost estimation accounts for their opportunity costs as well (see Deininger et al., 2008: 1803-05 for a more detailed explanation of the cost estimation).

¹² The administrative structure of Ethiopia is in the following order: federal government, regional government, zone, woreda (district), kebele (the lowest administrative unit).

conducted in the fall of 2012, covering the districts of Womberma, Bure and Jabitehinan. Five certificated and seven uncertificated Kebeles were randomly sampled from these three districts.¹³ The certification program was not implemented across all of Ethiopia simultaneously, but districts and Kebeles are gradually enrolled in the program. We focus on uncertificated Kebeles from a relatively large set of uncertificated Kebeles, making sure to select Kebeles that were uncertificated largely for exogenous reasons (i.e., unrelated to farm management and productivity). Such reasons include human resource and certificate constraints of the districts, border conflicts between Kebeles, and resettlement-related conflicts. We learned from the relevant “natural resources and land use administration” officials that inter- and intra-Kebele conflicts typical involve grazing land claims—a relevant domain from the perspective of rural households, but one that has no bearing on farming (cropping) land. Accordingly, we believe there is no *a priori* reason to assume that certificated and uncertificated households are systematically different (we also test this, see below).

Within each Kebele, we randomly sampled households proportional to the population, resulting in a sample of 325 farm households, of which 148 were certificated. In total, 1678 plots were included in the survey. Identification of average treatment effects (ATE) rests upon a comparison of certificated households in certificated Kebeles, and uncertificated households in uncertificated Kebeles. Within certificated Kebeles, there are also households without certificates, but it is easy to see how including such households could introduce (self) selection bias as such households are likely to be “different” across one or several dimensions.¹⁴ However, we use data from these households to estimate intention to treat (ITE) effects, as a robustness check for our results. For our core specification, we also exclude rented-in and rented-out plots, allowing us to identify the assurance effect more clearly (as opposed to the realizability effect).¹⁵ This yielded 1509 plots for the analysis, of which 588 were certificated and owner-operated and 921 were uncertificated.

¹³ Five Kebeles from Womberma, four Kebeles from Bure and three Kebeles from Jabitehinan.

¹⁴ Households residing in already certificated areas might not be certificated for various reasons, including absence during registration in their villages, failure to collect their certificates considering them as unimportant, household establishment after the registration and changes in household heads.

¹⁵ This also accentuates the incentive problem (Deininger et al., 2013) related to such plots, and rules out the possibility of double counting of plots. Holden et al. (2011) report that land certification has enhanced participation of households in land rental markets. Many of these land rentals were in the form of sharecropping. However, Deininger et al. (2013) find that such rentals lead to lower levels of efficiency because of incentive problems and inefficient contractual arrangements. Tenants would lose incentive to undertake long-term

Table 1 presents an overview of the variables used in the analysis (definitions of variables are given in the appendix). Our explanatory variables include household and plot level variables. Household variables are age and education of the household head, household size, land and livestock owned by the household and market access captured by distance to the nearest daily market. Land size was measured using timads¹⁶, as reported by households for each plot. We then aggregated the plots of the household to reach at the household land size. The interpretation of most variables is straightforward. Average household land size is about 11 timads, which is relatively large by Ethiopian standards with a national average of 6 timads (Future Agricultures, 2006). Plot level variables include biophysical characteristics of plots, like fertility, slope and conflict status of plots captured in dummies, and plot size in timad. While the total household land size affects the allocation decisions of the household, the size of a plot dictates farming activities on that specific plot. Therefore, we include both household land size and plot size in our selection model. Two-sample t tests for equality of means of the covariates indicate that the two groups are not balanced for all relevant covariates (last column of Table 1), hence we control for these differences in our analysis below.

Another potential threat for our identification is that variation across the certified and uncertified Kebeles may be confounded with the treatment. Variations in certified and uncertified Kebeles may stem from two potential sources: agroecological conditions and local social order. While it is unfortunate that we do not have Kebele level data, we do account for them at the household level to mitigate the threat to our attribution effort. First, we selected Kebeles in a sample area with sufficient similarity in economic opportunities and livelihood options to ensure comparability of certificated and uncertificated households. We picked districts based on the notion of “development domains”, defined as geographic locations sharing broadly similar rural development constraints and opportunities (Bernard et al., 2008). This classification is based on agricultural potential, access to markets and population density. The three districts are categorized in the same development domain, characterized as comprising high population density and moisture-reliant highlands with productivity growth and market improvement as priority livelihood strategies (Chamberlin et al., 2006). Moreover, households predominantly produce cereals and operate in similar markets. Second, we control

productivity boosting investments on a land that they rented in for a limited period. We indeed find tentative supportive evidence in our data. A simple two sample t-test for the productivity of rented and owned plots suggests that the productivity of owner-operated plots is significantly ($3.21 > 2.99$; $p\text{-value} = 0.005$) larger than that for rented plots.

¹⁶ One timad is the land area ploughed by a pair of oxen in a day, and approximately equals 0.25 hectare.

for agroecological conditions, like fertility and landscape (slope) of plots,¹⁷ and local social order, as captured by whether plots are conflict affected. Evidently, the two-sample t tests for equality of means indicate that the two groups are not significantly different in these covariates even before matching (last column of Table 1). Third, certifications are initiated by district authorities and we also formally test for unobserved heterogeneity.

Table 1: Descriptive statistics (Standard deviations in parentheses)

Variables	All sample			Certified	Uncertified	t-Test
	Mean	Min	Max	Mean	Mean	
HH-age	44.99 (10.49)	24	71	45.19 (9.68)	44.86 (10.98)	0.552
HH-education	1.85 (2.56)	0	10	2.29 (2.67)	1.57 (2.44)	0.000
HH-size	6.89 (2.13)	2	12	6.99 (2.04)	6.82 (2.19)	0.066
HH-land size	10.69 (4.97)	2	33	10.00 (4.33)	11.13 (5.29)	0.000
Total livestock ¹⁸	7.22 (3.88)	0.08	21.01	7.24 (4.05)	7.21 (3.77)	0.883
Distance to market	2.08 (0.41)	1.15	4.22	2.03 (0.38)	2.12 (0.42)	0.000
Plot size	1.87 (1.21)	0.25	16	1.90 (1.13)	1.85 (1.26)	0.421
Distance to plot	31.72 (35.73)	1	320	26.15 (27.90)	35.27 (39.54)	0.000
Very fertile plot	0.69 (0.46)	0	1	0.70 (0.46)	0.68 (0.47)	0.416
Fertile plot	0.24 (0.43)	0	1	0.26 (0.44)	0.23 (0.42)	0.140
Poor fertile plot	0.07 (0.25)	0	1	0.04 (0.19)	0.09 (0.28)	0.128
Flat plot	0.896 (0.31)	0	1	0.91 (0.29)	0.89 (0.32)	0.215
Gently sloping plot ¹⁹	0.104 (0.305)	0	1	0.09 (0.29)	0.11 (0.32)	0.893
Plot no conflict	0.75 (0.43)	0	1	0.76 (0.42)	0.74 (0.44)	0.291
Plot conflict	0.25 (0.43)	0	1	0.24 (0.43)	0.26 (0.44)	0.855
Value added in log	3.20 (0.43)	1	4.59	3.41 (0.35)	3.07 (0.42)	0.000

¹⁷ We also collected data on agroecological zones of the country (Bereha (hot, hyper-arid lowlands), Kolla (Warm, semi-arid lowlands), Weina Dega (Temperate, cool, sub-humid highlands), Dega (Cool, humid highlands)) to capture the extent of moisture and (possibly) altitude. However, Weina Dega was almost the sole outcome and the data didn't have enough variation. The dummy was then dropped from our selection model.

¹⁸ Tropical livestock unit is a common unit used to quantify a wide range of various livestock species and sizes as a single figure indicating the total amount of livestock owned by a household. Conversion factors vary across geographical regions. We employed a TLU for SSA with conversion factors: matured cow (1.0), ox (1.42), small cattle (0.73), goats/sheep (0.2), horses/mules/donkeys (0.8), and poultry (0.04).

¹⁹ In the survey instrument, we captured the plot slope variable in three constructs of flat, gently and steeply sloping. However, we included steeply sloping plots in the gently sloping plots' category, as there were only 19 observations.

Still another potential source of heterogeneity may also come from district land administrative authorities, who often assume the responsibility for initiating and coordinating the land registration and certification process. Such concerns may stem from either productivity potential of the Kebeles, where Kebeles with high potential productivities received treatment first, or convenience sampling, where nearby Kebeles received certification first. As per the first source, we account for agroecological factors, like plot fertility and landscape, which can help attenuate the productivity differences of the Kebeles (Table 1). Moreover, the three districts are among the high productive districts of West Gojjam zone (Chamberlin et al., 2006). Similarly, controlling for distance to district towns can diminish the threat related to the convenience sampling problem. While we do collect data on distance to district towns at household level, we are not able to include them in our selection model because distance to district towns is highly correlated with distance to permanent (daily) markets, as these markets are located in district towns for many of the households. Figure 1 also shows the locations of certified and uncertified Kebeles relative to district towns.

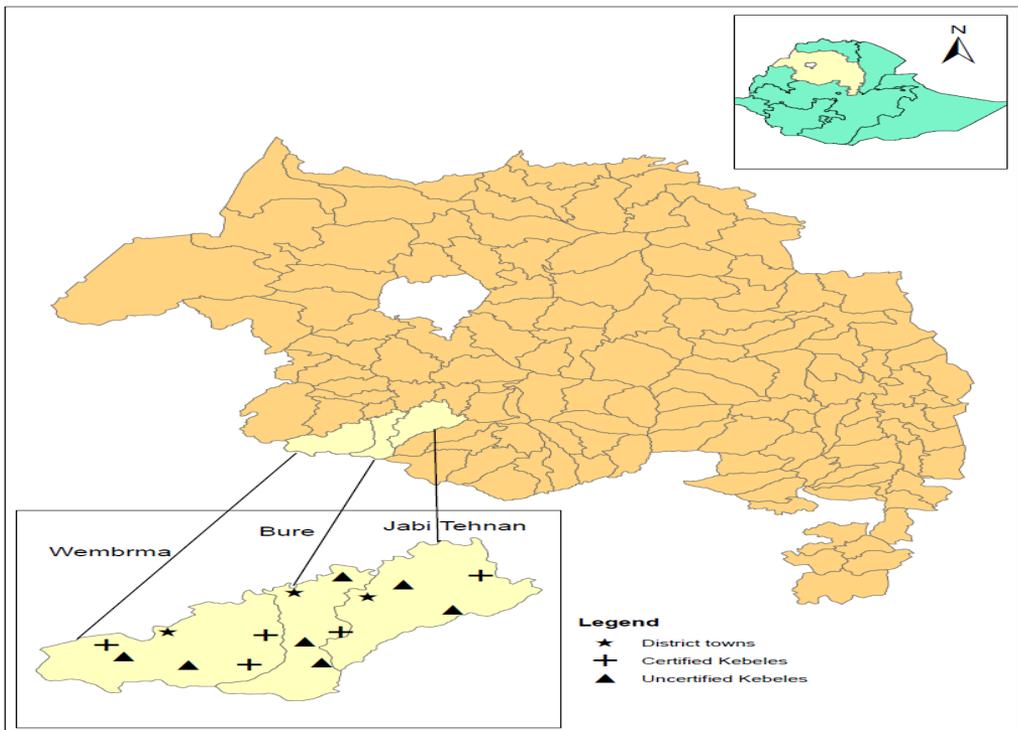


Figure 1: Location of the study areas in Amhara regional state, Ethiopia

Our outcome variable for land productivity is value added per timad, which we transform by taking its logarithm. When computing the value added for plots, we have included opportunity costs of family labor, priced at local wage rates, in addition to agricultural working capital costs. Figure 2 provides the value added distribution for certificated and uncertificated plots. A two-sample t test for equality of the means of the certificated and uncertificated plots shows that the mean for certificated plots is significantly (p -value =0.000) greater than that for uncertificated plots. Moreover, a two-sample Kolmogorov-Smirnov test for equality of the distributions for certificated and uncertificated plots is highly significant (p -value =0.000), revealing that the distributions are significantly different.

a. Value added per plot

b. Log of value added per plot

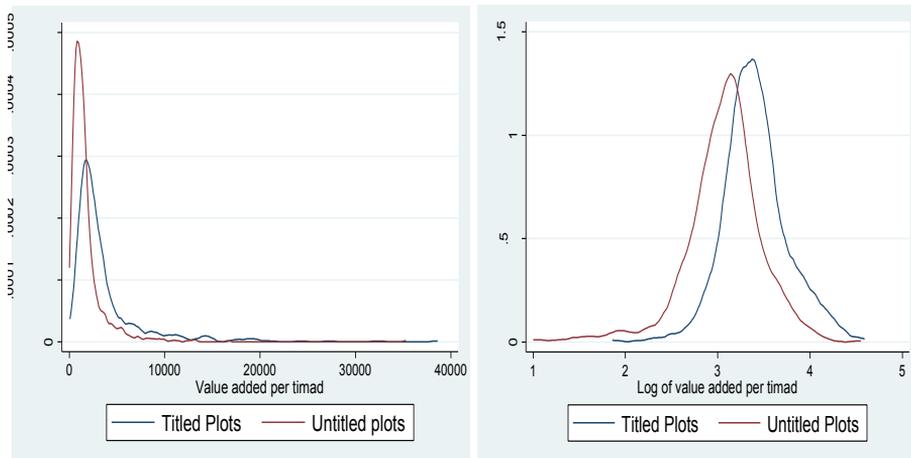


Figure 2: Kernel density distribution of plot-level land productivity for certificated and uncertificated plots

However, such differences between certificated and uncertificated plots should not necessarily be attributed to the certification program. Jumping to causal inference requires controlling for a number of potential sources of biases. Certificated and uncertificated plots may differ in a number of cultivator and plot level characteristics, like cultivator’s resource endowment and plot fertility. To navigate these concerns, we use propensity score matching method to enable attribution of causal effects associated with certification.

2.3.2 Propensity score matching

We employ the propensity score matching (PSM) method to mitigate the problem of selection bias. Let D_i be an indicator of whether plot i is certificated or uncertificated. The potential

productivity outcome of certification, represented by Γ_i , for each plot i can be defined as $\Gamma_i(D_i)$, where $i = 1, \dots, N$, N stands for the total number of plots. The average treatment effect on the treated is computed as:

$$\Delta_{ATT} = E(\Delta | D=1) = E[\Gamma(1) | D=1] - E[\Gamma(0) | D=1] \quad (1)$$

where Δ_{ATT} is the average treatment effect on the treated; $E[\Gamma(1) | D=1]$ is expected productivity for certificated plots and $E[\Gamma(0) | D=1]$ is expected productivity of certificated plots if they had not been certificated. The latter is a counterfactual, and not observed. Computing ATT requires obtaining a credible estimate of this counterfactual, for which we use data from uncertificated plots. However, as certificated and uncertificated plots may differ systematically, simply substituting the mean productivity outcome of uncertificated plots, $E[\Gamma(0) | D=0]$, for the counterfactual will induce selection bias.

Matching helps to construct the counterfactual from uncertificated plots controlling for selection bias due to observed covariates (Heckman et al., 1997). The comparison group must be statistically equivalent to the treated group, and one needs to match all observable covariates. To circumvent the curse of dimensionality, Rosenbaum and Rubin (1983) propose to match on the propensity score, $p(X)$, which is the probability of receiving treatment conditional on all relevant covariates, X .

PSM is not a magic bullet that solves the evaluation problem automatically. It requires imposing two important identifying assumptions (Heckman et al., 1997; Caliendo and Kopeinig, 2008). First is the conditional independence assumption (CIA), which ensures that selection into treatment should solely be based on a set of observable characteristics of the plots determining both the probability of being certificated and the outcome of interest (productivity of plots). Second is the common support condition, which confirms that plots with similar observable covariates have a positive probability of being part of both the treatment and comparison groups. If the CIA and the common support assumptions are adequately met, the PSM estimator for Δ_{ATT} is given as:

$$\Delta_{ATT}^{PSM} = E_{p(X)|D=1} \{E[\Gamma(1) | D=1, p(X)] - E[\Gamma(0) | D=0, p(X)]\} \quad (2)$$

Stated in words, the PSM estimator is simply the mean difference in productivity outcomes, appropriately weighting the comparison units by the propensity score distribution of participants over the common support.

We estimate propensity scores using a logit model. Only variables that simultaneously influence participation and outcomes should be included (Heckman et al., 1997; Smith and Todd, 2005). Accordingly, we include explanatory variables in the model guided by theory, previous empirical studies and the institutional setting. As a matching algorithm, we use the kernel matching method, which uses weighted averages of all subjects in the comparison group to construct the counterfactual outcome (Caliendo and Kopeinig, 2008). However, we probe the model's robustness using alternative matching methods.

Finally, PSM estimators do not account for selection on unobservables. However, we believe selection bias on unobservables has little impact on our results. First, while registration and certification is a voluntary decision of households, the certification intervention is driven by the government, and exogenous to cultivators and communities' unobservables. The compliance rate is very high. Moreover, certificated and uncertificated plots are from restricted geographical areas, sharing similar ecological domains, infrastructure, market development and returns to land-specific investments (Chamberlin et al., 2006). In our base model, we exclude uncertificated plots in certificated areas to mitigate unobservable bias *a priori*. Finally, we carry out robustness analysis of the sensitivity of our results to unobserved heterogeneity using the standard bounding approach test due to Rosenbaum (2002).

2.4 Results and discussion

2.4.1 Estimating propensity scores

Table 2 reports the regression estimates of the marginal effects of household and plot observables used to estimate propensity scores. Both household and plot level observables appear to determine significantly the likelihood of a plot being certificated.²⁰ The predicted propensity scores are used to match certified plots to uncertified plots in the comparison group. Some of the variables included in Table 2 are potentially endogenous to the certification decision. For example, plot fertility may capture (inherent) dimensions of soil quality but may also capture management decisions of the household. Since we collected our data sometime after the certification was finished, some of these management decisions may

²⁰ District dummies do not enter significantly and are not reported.

have occurred in response to the certification status. To attenuate concerns about the endogeneity of some of the variables in column (1), we also estimate a more parsimonious model that only contains variables that are arguably exogenous to certification status. This parsimonious model, presented in column (2), is used for a robustness analysis below.

Table 2: Estimates of the Logit model (standard errors are in parentheses)

Variables	1	2
HH-age head	0.222*** (0.046)	0.240*** (0.046)
HH-age head squared	-0.002*** (0.0005)	-0.002*** (0.0005)
HH-education head	0.167*** (0.024)	0.164*** (0.024)
HH-land size	-0.150*** (0.048)	-0.093** (0.044)
HH-land size squared	0.007 (0.002)	0.001 (0.003)
Total livestock	0.073*** (0.019)	
Distance to market	-0.820*** (0.163)	-0.722*** (0.159)
Plot size	0.461*** (0.149)	0.201*** (0.052)
Distance to plot	-0.009*** (0.002)	-0.009*** (0.002)
Very fertile plot	0.925*** (0.280)	
Fertile plot	0.970*** (0.286)	
Flat plot	-0.443 (0.357)	0.153 (0.185)
Plot no conflict	0.084 (0.133)	
Interaction plot size and flat plot	-0.285* (0.156)	
_cons	-4.453 (1.122)	-4.222 (1.089)
Number of observations	1509	1509
Prob>chi ²	0.000	0.000
Count R ^{2a}	68.00%	66.00%

* Significant at 10%; ** significant at 5%; *** significant at 1%

^a Count R² is computed as the ration of the number of correct predictions to total sample size. It indicates what proportion of the dependent variable is correctly predicted by the model (Green and Hensher, 2010).

2.4.2 Effect of land certification on plot productivity

Once propensity scores are estimated, the next step is to match certificated and uncertificated plots based on ‘closeness’ in terms of propensity scores. We report and discuss the results

using the kernel matching algorithm,²¹ which lowers the variance as it uses more information in the matching process (Heckman et al., 1998). Matching has been done only on plots that are on the common support. The results of the match are summarized in Tables 3 and 4.

Table 3 presents estimates of the average effect of land certification on certificated plots. The productivity of matched certificated plots is 35.4 percentage points higher than matched uncertificated plots. This difference is statistically robust ($p=0.000$). The number of certificated plots used for the estimation of ATT after matching is 569 (Table 4), which is an important diagnostic indicator of success in matching power. The information loss is relatively low since only 19 certificated plots are off the support region.

Table 3: Average treatment effects: Kernel matching^a

	Certificated (treated)	Uncertificated (control)	Difference
Unmatched	3.406	3.066	0.340
Matched(ATT)	3.410	3.056	0.354*** (0.023)

* Significant at 10%; ** significant at 5%; *** significant at 1%

We use the default option for kernel matching (kernel type = epanechnikov kernel; bandwidth =0.06).

^aThe matching is performed by using the psmatch2 program in STATA software (Leuven and Sianesi, 2003).

Standard errors in parentheses. The standard error for the ATT is bootstrapped standard error of 100 replications.

Table 4: Number of plots with the Kernel matching on the common support

	On support	Off support	Total
Uncertificated	921	0	921
Certificated	569	19	588
Total	1490	19	1509

Given this statistically significant and large impact of certification on productivity, what is its economic relevance? The increase in productivity due to certification of plots is equivalent to an annual gain of 1281 Birr (US\$75.40)²² in value added per plot.²³ This increment in value added per plot exceeds the program's one-off social costs (about US\$1 per plot) by a wide margin.²⁴ Net benefits are likely to remain positive even if the land administration is comprehensively restructured to include a cadastral index map, common property resources,

²¹The results remain robust for other well-known matching algorithms (see section 5 for robustness checks).

²² US\$1 was approximately equal to 17 Ethiopian Birr in 2011/2012.

²³ The increase in value added per plot in Birr is computed based on the average value added per timad of unmatched certificated plots, which is equal to 3619.081 Birr. The same value for the matched certificated plots is slightly greater than this value.

²⁴ US\$1 was approximately equal to 9.6 Ethiopian Birr in 2007/2008 when the program costs had been valued.

land used for residence and rural service facilities, and a mechanism to keep records up-to-date. Of course, one may question the necessity of such a “quality upgrade” in light of the impressive performance of the simple, low-cost technology.

What channel explains this productivity gain? As discussed, Ethiopian land is state property and land can neither be sold nor mortgaged. This has important implications for the relative significance of potential channels. For example, land is unlikely to be used as collateral for accessing formal credit services and collateralizability effects are not anticipated from the program. Further, we have excluded all plots that were rented-in (or out) from the analysis, which most likely minimizes the observable role of the realizability effect in driving our results. However, it is important to note that excluding rented-in (or out) plots from our analysis does not completely attain the realizability channel from contributing to our results. This is because farmers may undertake productivity-enhancing investments on a plot they own today, hoping to rent-out (or bequeath to their heirs) this plot to a more efficient farmer in the future. If so, the productivity gain from such investments will appear in the data even if the time of the transaction has not come yet. In view of these considerations, the assurance effect appears to be the most likely candidate to explain the difference between certificated and uncertificated plots, but we acknowledge that the anticipation of realizability effects may also be relevant.

Can we substantiate the importance of the assurance effect? We asked farmers about their subjective assessment of the impact of the certification program. About 96% of the households with land certificates believed that certification enhances their tenure security. Similarly, 91% of these households stated that owning a land certificate increases the incentive to plant trees (see also Holden et al. 2009, 2011; Deininger et al. 2011). In addition, 83% of the households indicated that land certification reduces land border conflicts, and provides protection against encroachment by neighbors. 74% also thought that certification increases the probability of receiving compensation in the event of land takings for public purpose. These statements are supported by actual land management practices “on the ground”. Table 5a shows that households with certificates are more likely to adopt soil-fertility management strategies on their plots than households without certificates. However, the household level analysis does not control for the endogeneity of the treatment. For this reason, we probe further if these differences in the land management decisions between certificated and uncertificated households still exist once endogeneity has been properly

accounted for. The two-sample *t*-test results (Table 5b) show that this is indeed the case.²⁵ Differences in land management practices between certificated and uncertificated plots are significant after matching. This provides tentative evidence that the land management decisions are consequences of certification.

Table 5: Land management practices of certificated and uncertificated households
5a. Household level analysis

Land management practices	Certificated HHs		Uncertificated HHs		<i>t</i> -Test
	Mean	Std. Error	Mean	Std. Error	
Dummy for household plants trees on its farm	0.81	0.032	0.58	0.037	***
Dummy for household uses organic manure on its farm	0.86	0.029	0.60	0.037	***
Dummy for household constructs stone terraces on its farm	0.76	0.035	0.67	0.036	**

** Significant at 5%; *** significant at 1%

5b. Plot level analysis

Variable	Sample	Mean Certificated (treated)	Mean Uncertificated (control)	% reduction (bias)	t-test p> t
Dummy for household plants trees on its farm	Unmatched	0.845	0.477		0.000
	Matched	0.852	0.491	2.0	0.000
Dummy for household uses organic manure on its farm	Unmatched	0.838	0.470		0.000
	Matched	0.844	0.495	5.5	0.000
Dummy for household constructs stone terraces on its farm	Unmatched	0.810	0.503		0.000
	Matched	0.807	0.583	27.0	0.000

2.4.3 Assessing matching quality

We need to assess the matching quality to verify the performance of PSM in terms of eliminating differences in observables between certificated and uncertificated plots. The credibility of PSM hinges on two identifying assumptions: CIA and the common support condition. Regarding CIA, we ask whether the propensity score adequately balances the distribution of relevant variables in matched certificated and uncertificated groups of plots. We employ a two-sample *t*-test for equality of means on each observable of certificated and uncertificated plots before and after matching and a chi-square test for joint significance of all variables in the logit model before and after the match. Test results are provided in Tables 6 and 7, respectively.

²⁵ The tests for the differences in land management practices are done at plot level analysis. Following this, one can observe a difference between the means of the land management practices at the household level analysis (Table 5a) and their respective means for the unmatched plot level data (Table 5b).

The results in Table 6 indicate that, while not all variables are balanced before matching, all explanatory variables are balanced after matching. Matching has reduced differences between certificated and uncertificated plots considerably as indicated by the percentage of reductions in bias.²⁶ An exception is the *HH-age head squared* variable. However, the difference between the two groups for this variable was already rather small before matching.

Table 6: T-tests for equality of means for each variable before and after the match

Variable	Sample	Mean Certificated (treated)	Mean Uncertificated (control)	% reduction (bias)	t-test p> t
HH-age head	Unmatched	45.194	44.864		0.552
	Matched	45.195	44.907	12.5	0.616
HH-age head squared	Unmatched	2136.1	2133.2		0.956
	Matched	2137.8	2109.2	-907.7	0.594
HH-education head	Unmatched	2.2857	1.5722		0.000
	Matched	2.1336	2.0799	92.5	0.730
HH-land size	Unmatched	10.001	11.127		0.000
	Matched	10.056	9.9418	89.9	0.663
HH-land size squared	Unmatched	118.74	151.79		0.000
	Matched	119.98	118.86	96.6	0.870
Total livestock	Unmatched	7.2379	7.2077		0.883
	Matched	7.1808	7.1698	63.7	0.962
Distance to market	Unmatched	2.0285	2.1173		0.000
	Matched	2.0333	2.0343	98.9	0.962
Plot size	Unmatched	1.9001	1.8485		0.421
	Matched	1.8818	1.8716	80.2	0.887
Distance to plot	Unmatched	26.151	35.268		0.000
	Matched	26.153	27.955	80.2	0.288
Very fertile plot	Unmatched	0.70068	0.68078		0.416
	Matched	0.69772	0.71313	22.5	0.569
Fertile plot	Unmatched	0.26361	0.23018		0.140
	Matched	0.26538	0.25059	55.8	0.569
Flat plot	Unmatched	0.90816	0.88817		0.215
	Matched	0.91213	0.9104	91.4	0.918
Plot no conflict	Unmatched	0.76361	0.73941		0.291
	Matched	0.76098	0.77662	35.4	0.532
Interaction plot size and flat plot	Unmatched	1.7079	1.6522		0.411
	Matched	1.7096	1.6912	67.0	0.806

Table 7: Chi-square test for joint significance of all variables before and after the match

Sample	Pseudo R ²	p > chi2
Unmatched	0.089	0.000
Matched	0.002	0.998

The chi-square test indicates that all variables in the logit model are not jointly significant after the match (prob > χ^2 = 0.998) (Table 7). In contrast, the same test is rejected before the

²⁶ The bias is calculated as the difference of the mean values of the certificated group and the unmatched/matched uncertificated group divided by the square root of the average sample variance in the certificated group and the unmatched uncertificated group (Rosenbaum and Rubin, 1985).

match ($\text{prob} > \chi^2 = 0.000$). This is corroborated by the pseudo- R^2 values of the model before and after matching. Since there are no systematic differences in distribution of covariates between certificated and uncertificated plots after matching, the pseudo- R^2 is fairly low compared with its value before matching.

Likewise, we check whether the common support condition is met as we wish to avoid comparing ‘incomparables’ when estimating the counterfactual (Dehejia and Wahba, 1999). We thus check the presence of enough overlap between certificated and uncertificated groups. The distribution of the propensity scores for the whole sample ranges between 0.012 and 0.924. The predicted propensity scores for certificated (treated) and uncertificated (control) groups range from 0.029 to 0.924 and 0.012 to 0.741, respectively. Consequently, the effective region of common support ranges from 0.029 to 0.741, which suggests adequate overlap between the groups. Observations outside this region are considered outliers, and dropped from the analysis.

Visual inspection of the propensity scores distributions for certificated and uncertificated plots also gives us a reading of the extent of overlap in propensity scores of the two groups. We present the density distribution graphs and histograms of propensity scores for certificated and uncertificated groups both before and after matching in Figures 3 and 4, respectively. Both graphs reveal the existence of a clear overlap in distributions after matching.

a. Before matching

b. After matching

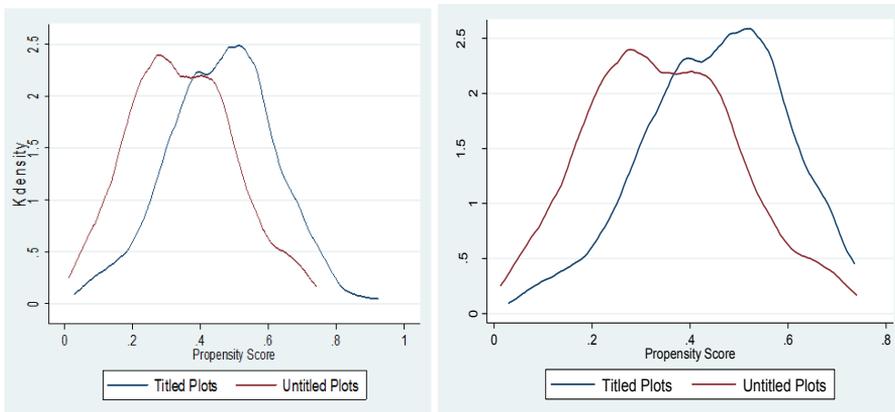
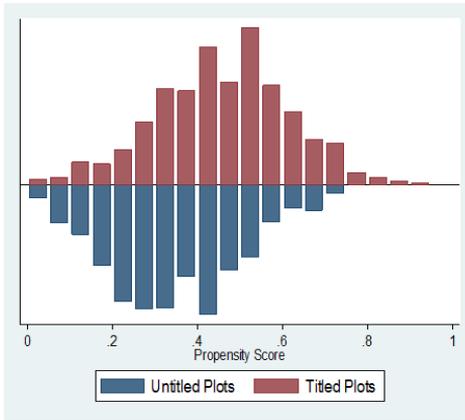


Figure 3: Kernel density distribution of propensity scores for certificated and uncertificated plots

a. Before matching



b. After matching

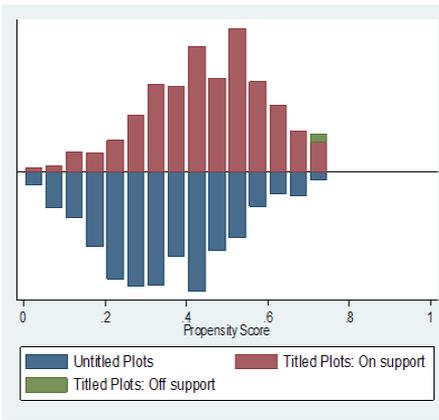


Figure 4: Histogram distribution of propensity scores for certificated and uncertificated plots

Overall, the matching quality tests indicate that we are not comparing incomparables. The matching procedure has generated samples of uncertificated plots that can reasonably be considered similar enough to certificated plots to construct the counterfactual for ATT estimation.

2.5 Sensitivity analysis and extension

2.5.1 Robustness of the ATT estimate to different matching algorithms

Some scholars argue that the choice of matching algorithm may significantly influence estimated results. Baser (2006), for example, proposes that sensitivity analysis of matching algorithms is important since none is *a priori* superior to others. We report estimates of ATT based on different matching algorithms in Table 8. Tests for matching quality are satisfied for all algorithms. Across estimators, all explanatory variables are balanced after matching, and chi-square tests show that variables are not jointly significant after the match. Pseudo- R^2 values are also substantially lower after matching than before matching. Generally, the ATT remains robust not only to different matching algorithms, but also to the meanest specifications of each algorithm.

Land Certification Programs and Agricultural Productivity

Table 8: Robustness of the ATT estimates for different algorithm estimators

Matching algorithms	Obs. on the common support		ATT ^a	t-test for distributions of each variables	Joint significance chi square test			
	Certificated	Uncertificated			Pseudo R ²		P > chi2	
					Unmatched	Matched	Unmatched	Matched
Kernel matching								
Bandwidth =0.001	485	921	0.344*** (0.029)	Satisfied	0.089	0.009	0.000	0.641
Badwidth =0.01	569	921	0.351*** (0.024)	Satisfied	0.089	0.003	0.000	0.985
Bandwidth =0.05	569	921	0.354*** (0.023)	Satisfied	0.089	0.002	0.000	0.998
Nearest neighbor²⁷								
NN=1 (no replacement)	569	921	0.347*** (0.023)	Satisfied	0.089	0.009	0.000	0.399
NN=1 (with replacement)	569	921	0.352*** (0.034)	Satisfied	0.089	0.014	0.000	0.278
NN=2	569	921	0.340*** (0.029)	Satisfied	0.089	0.010	0.000	0.338
NN=3	569	921	0.358*** (0.027)	Satisfied	0.089	0.010	0.000	0.311
Radius matching								
Caliper = 0.001 (no replacement)	365	921	0.345*** (0.029)	Satisfied	0.089	0.010	0.000	0.777
Caliper =0.001 (with replacement)	485	921	0.347*** (0.034)	Satisfied	0.089	0.010	0.000	0.495
Caliper = 0.01 (no replacement)	473	921	0.334*** (0.025)	Satisfied	0.089	0.008	0.000	0.765
Caliper =0.01(with replacement)	569	921	0.352*** (0.034)	Satisfied	0.089	0.014	0.000	0.277
Mahalanobis matching ²⁸	588	921	0.370*** (0.036)	Satisfied	0.089	0.010	0.000	0.264

* Significant at 10%; ** significant at 5%; *** significant at 1%

^a Standard errors are in parentheses.

2.5.2 Exogenous estimation of the propensity scores

Proper estimation of propensity scores requires employment of baseline data – data that are not available for the case analyzed here. Some variables included in the logit model for estimating propensity scores may be endogenous as we use cross-sectional data collected some time after the introduction of the land certification program. We therefore also estimate propensity scores including only exogenous explanatory variables (column (2) of Table 2). The new matching results are contained in Tables 9 and 10. The ATT effect is 0.354, which is exactly comparable to the original estimate (Table 3). The results for matching quality indicators are reported in Tables A1 and A2 in the appendix. Matching was done on the common support region and all quality indicators reveal that quality of the match is good.

²⁷ Matching without replacement is implemented only with 1-to-1 propensity score matching.

²⁸ This algorithm requires specifying a number of options at the same time. We just use the default option, which is matching on a quadratic metric with the generated weighting matrix.

Table 9: Average treatment effects: Kernel matching

	Certificated (treated)	Uncertificated (control)	Difference
Unmatched	3.406	3.065	0.341
Matched(ATT)	3.410	3.056	0.354*** (0.023)

We use the Epanechnikov kernel functional form with bandwidth =0.01.

* Significant at 10%; ** significant at 5%; *** significant at 1%

Std. error for ATT in parentheses. This standard error is bootstrapped standard error of 100 replications.

Table 10: Number of plots with the Kernel matching on the common support

	On support	Off support	Total
Uncertificated	921	0	921
Certificated	570	18	588
Total	1491	18	1563

2.5.3 The Intention to treat effect

We have also estimated intention to treat (ITT) effect. ITT gauges the effects of an intervention by comparing groups based on whether or not the group received the treatment, regardless of the treatment status of individuals (Khandker et al., 2010). It provides a more conservative estimate of treatment effect as it includes so-called noncompliers (i.e., households in certificated kebeles who chose not to certify their land). To estimate the ITT effect, we include the uncertified households residing in “certificated locations”. This procedure has resulted in 642 plots in treated villages, with 54 extra plots.

Table B1 in the appendix provides estimates of marginal effects of household and plot level observables used to estimate propensity scores. Matching results are reported in Tables 11 and 12. The ITT effect is 0.349, so plots in certified locations have on average 34.9 percentage points higher productivity than uncertified plots in comparison locations. This effect is comparable with the ATT effect (ATT = 0.351) for the same specification of the kernel matching algorithm (Table 8). This is not unexpected as compliance is high (about 92% of the households in the certificated Kebeles are actually certificated in our sample). Table 12 shows the number of certified and uncertified plots on and off the common support region. Results for matching quality indicators are reported in Tables B2 and B3 in the appendix and, generally, matching quality tests suggest the matching algorithm has performed well in terms of eliminating systematic differences in the distribution of covariates between the two groups.

Table 11: Intention to treat effects: Kernel matching

	Certificated (treated)	Uncertificated (control)	Difference
Unmatched	3.399	3.066	0.333
Matched(ATT)	3.400	3.052	0.349*** (0.023)

We use the Epanechnikov kernel functional form with bandwidth =0.01.

* Significant at 10%; ** significant at 5%; *** significant at 1%

Std. error for ITT in parentheses. This standard error is bootstrapped standard error of 100 replications.

Table 12: Number of plots with the Kernel matching on the common support

	On support	Off support	Total
Uncertificated	921	0	921
Certificated	620	22	642
Total	1541	22	1563

2.5.4 Heterogeneous treatment effects

The effect of a treatment may vary across households, depending on certain characteristics. The Ethiopian land tenure system is characterized by a land policy where land redistributions result in a dynamic process of leveling down households' land size, as households with more than average land are more likely to be among the losers in future land redistributions. In other words, households with relatively large land size tend to be more tenure insecure, and may gain more from certification. We therefore hypothesize the ATT effect is larger for households with large land size. The literature uses subsample analysis or interaction terms to test for heterogeneous treatment effects (Angrist, 2004). We follow the subsample analysis approach, and divide our dataset into two groups at the median household land size.²⁹ This procedure results in 754 farm plots of “small land size” group and 755 farm plots of “large land size” group.

Table C1 in the appendix contains regression estimates of the marginal effects used to estimate propensity scores, and matching results are reported in Tables 13 and 14 (results for matching quality indicators are reported in Tables C2 through C4 in the appendix). The ATT effects are 0.337 and 0.348 for the small and large land groups, respectively. Both effects are statistically significant and quantitatively similar, so tenure security appears equally important for both groups. While periodic land redistributions are more likely a concern for households

²⁹ The median household land size is 10.25 timads while the mean household land size is 10.69 timads.

with large land size, other features of the land policy may create uncertainties bearing on tenure security for both groups alike. Major sources of uncertainty in the Ethiopian tenure system are the absence of a clear justice system for resolving land conflicts, frequent revisions of the law, authority given to different government agencies to intervene in land affairs coupled with limited legislative awareness from farmers and local officials, and rapid urban expansion and land grants to investors (e.g., Holden and Yohannes, 2002; Rahmato, 2009a; Deininger et al., 2011). Also, competing claims from community members, like land border conflicts and encroachment by neighbors, might be more important sources of tenure insecurity for households with small land size and weak power.

Table 13: Average treatment effects: Kernel matching

	Certificated (treated)		Uncertificated (control)		Difference	
	Small land	Large land	Small land	Large land	Small land	Large land
Unmatched	3.402	3.412	3.055	3.075	0.347	0.337
Matched(ATT)	3.405	3.410	3.067	3.062	0.337 *** (0.0367)	0.348 *** (0.0398)

We use the Epanechnikov kernel functional form with bandwidth =0.01.

* Significant at 10%; ** significant at 5%; *** significant at 1%

Std. error for ATT is in parentheses. The standard error is bootstrapped standard error of 100 replications.

Table 14: Number of plots with the Kernel matching on the common support

	On support		Off support		Total	
	Small land	Large land	Small land	Large land	Small land	Large land
Uncertificated	420	501	0	0	420	501
Certificated	324	236	10	18	334	254
Total	744	737	10	18	754	755

2.5.5 Sensitivity of ATT estimate to unobserved heterogeneity

An important identifying assumption of PSM is the CIA, implying that selection into treatment is based on only observed variables. Matching estimators are not robust to hidden bias resulting from assignment into treatment by unobserved variables (Caliendo and Kopeinig, 2008). One threat to identification in our study is that unobserved heterogeneity explains which kebeles receive the certification treatment and also correlates with investments in agricultural productivity. For example, one reason why some kebeles receive treatment earlier than others has to do with conflicts about demarcation of grazing lands (typically between

kebeles). One might be concerned that this indicates a lack of well-recognized authority, be it an official or informal order, which could guarantee tenure security (affecting investments in agricultural productivity). While such an interpretation would not diminish our argument about the importance of tenure security (for agricultural productivity), it would compromise our ability to attribute the differences between certificated and uncertificated households to the certification process.

However, several features of the Ethiopian certification process attenuate this concern. First, we argue the role of heterogeneity is modest. Customary land laws do not prevail in Ethiopia. Competing claims based on conflicting rules, a relevant source of local heterogeneity elsewhere, are not a major source of tenure insecurity in Ethiopia. The Amhara region has a well-established conflict resolution system known as the *Shimagelle* system (e.g., see Gowok, 2008; Baker, 2013), which is operational throughout the region. Both certificated and uncertificated communities have a long tradition of solving land-related and other types of conflicts – maintaining societal order. While border conflicts between Kebeles regarding the demarcation of grazing lands may invite the postponement of certification, such conflicts do not signal weak local governance. Unobserved heterogeneity is also reduced because, as mentioned, certificated and uncertificated plots are from restricted geographical areas, sharing similar ecological domains, infrastructure, market development and returns to land-specific investments (Chamberlin et al., 2006).

Second, and minimizing the impact of remaining heterogeneity, the responsibility for the land registration and certification process has been bestowed on Woredas (districts), not on Kebeles or villages. Certification is not initiated at the local level (and compliance rates exceeding 90% imply that self-selection into certification at the household level is not likely to be relevant).³⁰ Certificate and human resource constraints of these districts—another leading explanation for uneven progress in certification across kebeles—are typically determined by factors beyond developments at specific Kebeles. These constraints should therefore not be viewed as indicative of unobserved heterogeneity at local levels.

³⁰ Moreover, the source of tenure insecurity in Ethiopia has been the State itself through its land redistribution efforts. The decision to redistribute land is not made locally, but at the level of the regional parliament (“house of regional representatives”). Therefore, when it is decided to redistribute land, the decision will be effected throughout the region. Land proclamations do not leave discretion for making distribution decisions to zonal administrations, let alone to lower level units. This also considerably reduces the possibility of heterogeneity between villages and Kebeles on the source of the tenure insecurity.

Third, we also seek to formally check the robustness of our results to departures from the strong CIA assumption, and employ the bounding approach due to Rosenbaum (2002). This allows us to determine how strong an unobservable variable must be to influence the selection process so as to undermine our results. Results of Rosenbaum bound tests are reported in Table 15. We find that the estimates of ATT are robust to unobservable covariates, even to the extent that would triple ($\gamma = 3$) the log odds of differential assignment to treatment (which seems unlikely in light of observable heterogeneity in our sample).

Table 15: Robustness of the ATT estimates to unobserved heterogeneity (rbound tests)

Gamma	sig+	sig-	t-hat+	t-hat-	CI+	CI-
1	0	0	0.3422	0.3422	0.3159	0.3696
1.1	0	0	0.3293	0.3559	0.3024	0.3837
1.2	0	0	0.3171	0.3683	0.2905	0.3968
1.3	0	0	0.3059	0.3803	0.2795	0.4086
1.4	0	0	0.2956	0.3911	0.2693	0.4204
1.5	0	0	0.2865	0.4011	0.2596	0.4312
1.6	0	0	0.2777	0.4105	0.2505	0.4417
1.7	0	0	0.2696	0.4201	0.2420	0.4514
1.8	0	0	0.2619	0.4286	0.2340	0.4603
1.9	0	0	0.2542	0.4372	0.2270	0.4690
2	0	0	0.2474	0.4453	0.2203	0.4775
2.1	0	0	0.2408	0.4527	0.2137	0.4858
2.2	0	0	0.2345	0.4597	0.2073	0.4937
2.3	0	0	0.2289	0.4667	0.2014	0.5014
2.4	0	0	0.2236	0.4733	0.1954	0.5084
2.5	0	0	0.2183	0.4799	0.1893	0.5150
2.6	0	0	0.2133	0.4862	0.1840	0.5218
2.7	0	0	0.2083	0.4923	0.1786	0.5282
2.8	0	0	0.2037	0.4981	0.1736	0.5344
2.9	0	0	0.1994	0.5036	0.1688	0.5407
3	0	0	0.1947	0.5091	0.1640	0.5470
3.1	0	0	0.1900	0.5141	0.1596	0.5529

Gamma: log odds of differential assignment due to unobserved factors

Sig+ : upper bound significance level (assumption: overestimation of treatment effect)

Sig- : lower bound significance level (assumption: underestimation of treatment effect)

t-hat+ : upper bound Hodges-Lehmann point estimate

t-hat- : lower bound Hodges-Lehman point estimate

CI+ : upper bound confidence interval ($\alpha = 0.95$ and assumption: overestimation of treatment effect)

CI- : lower bound confidence interval ($\alpha = 0.95$ and assumption: underestimation of treatment effect)

2.6 Conclusions and discussion

Theory predicts that land certification is a catalyst for economic growth by providing incentives to boost agricultural investment and productivity. Yet, empirical evidence for Africa is mixed. A review of the empirical evidence shows that land reforms have not lived up to theoretical expectations, particularly in terms of productivity gains. Interestingly, a body of work is emerging that challenges this dismal outcome. Several studies of the Ethiopian land

registration and certification program produce more favorable assessments of impact (e.g., Deininger et al., 2008; Holden et al., 2009; Deininger et al., 2011; Holden et al., 2011).

Consistent with the gist of these studies, we find that land registration and certification has produced significant positive effects on farm productivity. We use propensity score matching as the basis of our empirical approach, and verify whether the so-called conditional independence assumption and common support condition are satisfied by our data. We perform various checks and tests supporting the validity of PSM for our analysis. Our results also survive several sensitivity analyses. Nevertheless, we acknowledge that attribution remains difficult for analyses based on non-randomized experimental cross-section data. While we demonstrate that unobserved heterogeneity is unlikely to drive our results, we cannot rule out that tenure security aspects beyond formal certificates may explain part of the differences we find between certified and uncertified households.

Notwithstanding this caveat, we believe certification matters and, moreover, argue that the assurance effect is mainly responsible for the effect of certification on agricultural productivity (even if we cannot rule out a potential realization effect by certificated households anticipating to rent out plots in the future). The magnitude of this effect exceeds the program's implementation costs by a wide margin. The program managed to reduce tenure insecurity due to frequent land redistributions, and was tailored to fit local conditions with highly decentralized, transparent and cost-effective procedures.

Why are outcomes different for Ethiopia from other African countries? We believe the reason can be found in the local institutions history of Ethiopia. The impact of land registration and certification critically relies on the level of security that holders enjoy prior to the program. Assessing *de facto* and subjective tenure security prior to reform is difficult, as land tenure in Africa is complex, encompassing widely varying customary, religious and statutory arrangements (Cotula et al., 2004). In many African countries, these complex customary tenure systems provide sufficient levels of *de facto* tenure security to induce farmers to invest in productivity-enhancing measures (Migot-Adholla et al., 1991; Brasselle et al., 2002; Jacoby and Minten, 2007; Bromley, 2008; Toulmin, 2009; Fenske, 2011). The variation in pre-formal arrangements is often only crudely mirrored in national land laws. Simplistic and blanket calls for formalization of land rights fail to account for the complexity and history of existing institutional arrangements, and may therefore not be appropriate. Such reforms can make matters worse if they create a parallel system that even increases insecurity (Atwood, 1990;

Bromley, 2008; Deininger and Feder, 2009; Toulmin, 2009; Fenske, 2011; Obeng-Odoom, 2012). Land reforms are not a simple technical matter, but represent a complex and deep social intervention in a multidimensional construct (Lemel, 1988; Galiani and Scharngrodsky, 2010).

The Ethiopian case is different. Customary land laws do not exist in Ethiopia, owing to the discontinuous evolution of its tenure system (Deininger et al., 2008). Before the certification program, the main source of tenure insecurity for Ethiopian farmers has been the state itself through its policy of land redistribution. Deininger and Jin (2006) characterize Ethiopia as more tenure insecure than other African countries, which explains the demand from Ethiopian farmers for land registration. Streamlining state-led redistribution initiatives had, therefore, tremendous effect on tenure security. However, the results of this study, and other studies of the Ethiopian certification program, should not be simply transferred to other African contexts—the external validity may be limited. In general, land certification reforms should not dismiss the role of competing interests and indigenous tenure forms, as these are fundamental to provide basic levels of tenure security where they exist. Neglecting such evidence through one-size-fits-all prescriptions may reduce rather than enhance *de facto* tenure security (Place, 2009).

Appendix

Appendix A: Exogenous estimation of the propensity scores

Table A1: T-tests for equality of means for each variable before and after the match

Variable	Sample	Mean Certificated (treated)	Mean Uncertificated (control)	% reduction (bias)	t-test p> t
HH-age head	Unmatched	45.194	44.864		0.552
	Matched	45.132	44.874	21.9	0.655
HH-age head squared	Unmatched	2136.1	2133.2		0.956
	Matched	2131.7	2106.8	-777.4	0.643
HH-education head	Unmatched	2.2857	1.5722		0.000
	Matched	2.1684	2.2471	89.0	0.620
HH-land size	Unmatched	10.001	11.127		0.000
	Matched	10.109	10.048	94.5	0.817
HH-land size squared	Unmatched	118.74	151.79		0.000
	Matched	120.9	122.55	95	0.814
Distance to market	Unmatched	2.0285	2.1173		0.000
	Matched	2.0313	2.0383	92.1	0.739
Plot size	Unmatched	1.9001	1.8485		0.421
	Matched	1.8776	1.9057	45.7	0.697
Distance to plot	Unmatched	26.151	35.268		0.000
	Matched	26.454	28.22	80.6	0.306
Flat plot	Unmatched	0.90816	0.88817		0.215
	Matched	0.90877	0.90826	97.4	0.976

Table A2: Chi-square test for the joint significance of all variables before and after the match

Sample	Pseudo R ²	p > chi2
Unmatched	0.074	0.000
Matched	0.003	0.906

Institutional Change and Economic Development

Appendix B: The Intention to Treat Effect

Table B1: Estimates of the Logit model (standard errors in parentheses)

Variable	Coefficient
HH-age head	0.193*** (0.044)
HH-age head squared	-0.002*** (0.0005)
HH-education head	0.159*** (0.023)
HH-land size	-0.156** (0.045)
HH-land size squared	0.001 (0.002)
Total livestock	0.076*** (0.019)
Distance to market	-0.832*** (0.158)
Plot size	0.200*** (0.051)
Distance to plot	-0.009*** (0.002)
Very fertile plot	0.622* (0.249)
Fertile plot	0.631* (0.255)
Flat plot	-0.004 (0.201)
Plot no conflict	0.124 (0.129)
_cons	-3.343 (1.122)
Number of observations	1563
Prob>chi ²	0.000
Count R ²	67.00%

* Significant at 10%; ** significant at 5%; *** significant at 1%

Land Certification Programs and Agricultural Productivity

Table B2: T-tests for equality of means for each variable before and after the match

Variable	Sample	Mean Certificated	Mean Uncertificated	% reduction (bias)	t-test p> t
HH-age head	Unmatched	44.734	44.864		0.809
	Matched	44.819	44.609	-61.2	0.705
HH-age head squared	Unmatched	2096.6	2133.2		0.467
	Matched	2104.5	2085.5	48	0.712
HH-education head	Unmatched	2.3209	1.5722		0.000
	Matched	2.1435	2.1115	95.7	0.831
HH-land size	Unmatched	10.134	11.127		0.000
	Matched	10.219	10.091	87.1	0.628
HH-land size squared	Unmatched	122.84	151.79		0.000
	Matched	124.75	124.83	99.7	0.991
Total livestock	Unmatched	7.3279	7.2077		0.550
	Matched	7.2511	7.2946	63.9	0.844
Distance to market	Unmatched	2.0247	2.1167		0.000
	Matched	2.0309	2.0256	94.3	0.794
Plot size	Unmatched	1.9108	1.8485		0.317
	Matched	1.9012	1.8709	51.4	0.666
Distance to plot	Unmatched	26.424	35.268		0.000
	Matched	26.529	28.092	82.3	0.341
Very fertile plot	Unmatched	0.70561	0.68078		0.296
	Matched	0.69839	0.71432	35.8	0.538
Fertile plot	Unmatched	0.25078	0.23018		0.348
	Matched	0.25645	0.23957	18.0	0.492
Flat plot	Unmatched	0.91277	0.88817		0.113
	Matched	0.9129	0.91222	97.2	0.966
Plot no conflict	Unmatched	0.77103	0.73941		0.154
	Matched	0.76613	0.78228	48.9	0.497

Table B3: Chi-square test for the joint significance of all variables before and after the match

Sample	Pseudo R ²	p > chi2
Unmatched	0.080	0.000
Matched	0.004	0.940

Institutional Change and Economic Development

Appendix C: Heterogeneous Treatment Effects

Table C1: Estimates of the Logit model (standard errors in parentheses)

Variable	Small land	Large land
HH-age head	0.113* (0.059)	0.342*** (0.086)
HH-age head squared	-0.001 (0.0006)	-0.004*** (0.0008)
HH-education head	0.122*** (0.034)	0.201*** (0.037)
HH-land size	-0.026 (0.260)	-0.097 (0.145)
HH-land size squared	-0.0001 (0.019)	-0.0015 (0.004)
Total livestock	-0.025 (0.028)	0.183*** (0.029)
Distance to market	-0.557** (0.219)	-1.212*** (0.251)
Plot size	0.296*** (0.102)	0.151** (0.063)
Distance to plot	-0.0178*** (0.003)	-0.0024 (0.0025)
Flat plot	0.155 (0.255)	0.0266 (0.298)
Plot no conflict	-0.035 (0.183)	0.511** (0.210)
_cons	-2.037 (1.598)	-7.359** (2.478)
Number of observations	754	755
Prob>chi ²	0.000	0.000
Count R ²	64.00%	72.00%

* Significant at 10%; ** significant at 5%; *** significant at 1%

Land Certification Programs and Agricultural Productivity

Table C2: T-tests for equality of means for each variable before and after the match (small land group)

Variable	Sample	Mean Certificated	Mean Uncertificated	% reduction (bias)	t-test p> t
HH-age head	Unmatched	44.626	41.4		0.000
	Matched	44.429	43.391	67.8	0.196
HH-age head squared	Unmatched	2100	1823.5		0.000
	Matched	2079.2	1985.4	66.1	0.206
HH-education head	Unmatched	1.9012	1.5405		0.038
	Matched	1.8333	2.1892	1.4	0.071
HH-land size	Unmatched	7.0413	6.8427		0.180
	Matched	7.0759	7.043	83.4	0.832
HH-land size squared	Unmatched	53.585	50.917		0.177
	Matched	54.054	53.425	76.4	0.765
Total livestock	Unmatched	5.1355	5.1783		0.847
	Matched	5.1454	5.0344	-159.5	0.639
Distance to market	Unmatched	2.0759	2.1418		0.017
	Matched	2.0795	2.0573	66.4	0.403
Plot size	Unmatched	1.616	1.4968		0.047
	Matched	1.581	1.6132	73.0	0.625
Distance to plot	Unmatched	23.144	35.519		0.000
	Matched	22.488	24.783	81.4	0.218
Flat plot	Unmatched	0.90419	0.8881		0.474
	Matched	0.90432	0.90271	90.0	0.945
Plot no conflict	Unmatched	0.73653	0.76905		0.303
	Matched	0.74383	0.72334	37.0	0.556

Table C3: T-tests for equality of means for each variable before and after the match (large land group)

Variable	Sample	Mean Certificated (treated)	Mean Uncertificated (control)	% reduction (bias)	t-test p> t
HH-age head	Unmatched	45.941	47.768		0.017
	Matched	46.174	46.832	64.0	0.416
HH-age head squared	Unmatched	2183.5	2392.9		0.004
	Matched	2208.8	2269.8	70.9	0.428
HH-education head	Unmatched	2.7913	1.5988		0.000
	Matched	2.6314	2.1443	59.2	0.057
HH-land size	Unmatched	13.892	14.718		0.009
	Matched	13.938	14.203	67.9	0.447
HH-land size squared	Unmatched	204.42	50.917		0.007
	Matched	205.91	236.35	60.0	0.311
Total livestock	Unmatched	10.003	8.9089		0.000
	Matched	9.7098	10.328	43.5	0.034
Distance to market	Unmatched	1.9661	2.0967		0.000
	Matched	1.9706	1.982	91.2	0.734
Plot size	Unmatched	2.2736	2.1433		0.240
	Matched	2.2564	2.361	19.7	0.432
Distance to plot	Unmatched	30.106	35.058		0.114
	Matched	30.305	33.179	42.0	0.354
Flat plot	Unmatched	0.91339	0.88822		0.284
	Matched	0.91949	0.9292	61.4	0.691
Plot no conflict	Unmatched	0.79921	0.71457		0.012
	Matched	0.78814	0.83047	50.0	0.243

Table C4: Chi-square test for the joint significance of all variables before and after the match

Sample	Pseudo R ²		p > chi2	
	Small land	Large land	Small land	Large land
Unmatched	0.070	0.143	0.000	0.000
Matched	0.008	0.021	0.762	0.237

Variable definitions

HH-age: Age of the household head in years, rounded down to the last birthday

HH-education: The number of completed years of education of the household head

HH-size: The number of people sharing the same roof and sharing the same pot

HH-land size: Total household land holding measured in timad (local measurement of land)

Total livestock: Total livestock of the household measured in tropical livestock units (tlu)

Distance to market: Distance to a daily market center from the homestead of the household in hours on foot

Plot size: A variable indicating the size of plot *i* measured in timad

Distance to plot: Distance to plot *i* from the homestead of the household in minutes on foot

Very fertile plot: A dummy variable taking a value of 1 if plot *i* is very fertile, 0 otherwise

Fertile plot: A dummy variable taking a value of 1 if plot *i* is fertile, 0 otherwise

Poor fertile plot: A dummy variable taking a value of 1 if plot *i* is poor fertile, 0 otherwise

Flat plot: A dummy variable taking a value of 1 if plot *i* is flat, 0 otherwise

Gently sloping plot: A dummy variable taking a value of 1 if plot *i* is either gently or steeply sloping, 0 otherwise

Plot no conflict: A dummy variable taking a value of 1 if plot *i* is not affected by any type of conflict, 0 otherwise

Plot conflict: A dummy variable taking a value of 1 if plot *i* is affected by any type of conflict, 0 otherwise

Value added in log: A variable indicating the logarithm of value added per timad

Chapter 3

Joint Land Certification Programs and Women's Empowerment

*Evidence from Ethiopia*³¹

Abstract

This paper connects two important development policy issues: women's empowerment and land certification. We employ propensity score matching to study the impact of the Ethiopian joint land registration and certification program on women's empowerment. We use data collected using surveys and a field experiment to construct different indices for empowerment as our dependent variables. Our main result is that joint land certification has significant effects on women's empowerment, particularly on dimensions that indicate women's participation and roles outside the home. This result is robust to various sensitivity checks and alternative model specifications.

Publication status: Melesse, B. M., Dabissa, A. & Bulte, E. (2013). Joint Land Certification Programs and Women's Empowerment: Evidence from Ethiopia. Under review at *Economic Development and Cultural Change*.

³¹ This paper has won the Uma Lele Prize for Best Paper Award on Gender Issues of the IAAE, 2015.

3.1 Introduction

Gender based discrimination is common in many developing countries. Many women have limited economic opportunities outside their home, and spend most of their time indoors. Gender asset gaps are also substantial (Agarwal, 1994; Deere and Doss, 2006). In an effort to reverse these realities, gender equality has received a great deal of attention in the recent development agenda. This is evidenced, for example, by its standalone inclusion in the Millennium Development Goals –empowering women and “gender mainstreaming” have become integral components of project interventions and policy reform efforts. However, how to empower women has continued to be a long-standing debate in the literature. Often, this debate centers on opportunities for women's participation in economic activities and access to resources. For example, the provision of land rights to women has been regarded as a key to their empowerment (Agarwal, 1997; Yngstrom, 2002; Deere and Leon, 2003).

In many developing countries, women do not have equal land rights to men. They are often viewed as secondary claimants, through their male relatives – husband, sons or father. Their access to and rights over land tends to be limited by customary practices (Yngstrom, 2002, Deere and Leon, 2003). Reasons accounting for the gender gap are varied, and include male preference in inheritance, male privilege in marriage, male bias in land distribution programs, and bias against women in land markets. The lack of autonomous access to property rights has led to uncertainty and instability in the lives of many women – especially in contexts of widowhood, divorce, or polygamy (Joireman, 2008).

Women's rights to land are at the forefront of development debates, both as an important policy goal in its own right as well as a policy instrument to improve household welfare. In its own right, land rights are expected to empower women by expanding their opportunity sets (World Bank, 2006). Moreover, empowering women is expected to have multiplier effects on household welfare via its impact on household resource allocation (Agarwal, 1994). Evidence suggests resources controlled by women are more likely to benefit a wide swath of household members, particularly children (e.g., Thomas, 1990; Duflo, 2003; Quisumbing and Maluccio, 2003; Park, 2007; Martinez, 2013).

In theory, joint certifications to land rights can empower women for various reasons. It may improve their outside opportunities since they will now have the right to claim their lawful share of household land, improving the “threat point” in intra-household bargaining problems and hence increasing the share of the surplus (Manser and Brown, 1980; McElroy and

Horney, 1981). Furthermore, when women contribute more to the common good of the household, their voice increases through the general norms of “influence according to contribution” (Sen, 1990). Land holding can also capacitate women to push boundaries on detrimental deep-seated gender stereotypes, norms and perceptions (Agarwal, 1997).

However, the empirical literature documenting impacts of land rights on women’s empowerment is thin. A handful of studies attempt to explore the issue empirically. For example, Allendorf (2007) found that women who own land are more likely to have the final say in household decisions in Nepal. Similarly, Wiig (2013) showed that Peruvian women in certified communities tend to have considerable participation in household decision-making. A few papers have focused on Ethiopia. Holden and Tefera (2008) provide a broad and qualitative evaluation of the early gender impacts of the land certification program, which is also the subject of our study. They generally reported favorable subjective assessments by both women and men of land certification on the participation of women in land-related decisions and on the perception of women’s rights. Finally, and somewhat distinct, two studies assess the impact of the Ethiopian certification program on economic outcomes for female-headed households. Holden et al. (2011) found a positive effect on the participation of female-headed households in the land rental market. Similarly, Bezabih et al. (2012) reported modest gains in terms of productivity for female-headed households.

The aim of this paper is to analyze the effects of the Ethiopian joint land certification program on women’s empowerment within the household. This low-cost land registration and certification program conveys several elements of private property rights that particularly strengthen the rights of women. The recent national and regional land proclamations, which provide the content for the land certification efforts, strongly reiterate the concept of gender equality and non-discrimination based on gender. They state that women have equal user land rights, and that husband and wife shall reach consensus in relevant decisions regarding their land. Equally, the recent reform in the country’s Family Code has emphasized gender equality (Kumar and Quisumbing, 2012; Hallward-Driemeier and Gajigo 2013; Kumar and Quisumbing, 2015). Upon divorce, women are entitled to get half of the household land. Widows are eligible to keep the household land after their husbands have died, and girls have the same privileges as their brothers to inherit land from their parents. As a result of these

efforts, (most) women in Ethiopia³² now enjoy the same *de jure* rights as men in key aspects of life, including marriage, property rights and inheritance.

From a research perspective, two key challenges stand out. The first one is credible causal inference, which may be compromised by the potential endogenous placement of land certification efforts (the sequence of implementation). We use propensity score matching and propensity score weighted regression for causal inference, taking advantage of the sequential introduction of the land certification program to treat land formalization as a natural experiment to create credible counterfactuals. The second challenge is how to measure empowerment. For this purpose, we combine survey-based and experimental approaches.

Our study contributes to the existing literature in two ways. First, while the investment and productivity impacts of the Ethiopian land reform have received considerable scrutiny,³³ the impact on the bargaining position of women in the household remains unclear. Second, we adopt a nuanced approach to measuring women's empowerment, which we regard as a multidimensional construct (e.g., Mason, 1986; Mahmud et al., 2012). Empirical research to date has emphasized economic indicators of empowerment, oftentimes ignoring other dimensions of power. In an effort to address this concern, we employ several indicators, capturing a wide variety of issues associated with the autonomy of women in the household.

Our main result is that joint land certification empowers women, but not equally across all dimensions. Specifically, we find strong significant differences between certified and uncertified women in community participation, knowledge of land rights and legal provisions, the level of tenure security and willingness to protect rights, as well as in our measure of institutional inclusion. In contrast, we find only modest effects on indicators of physical mobility and autonomy. Aggregating across dimensions, we construct an overall index of empowerment, and find that joint land certification improves the state of women's empowerment. While we discuss and test the assumptions that allow causal inference with our model to the extent possible, we provide some qualifications on the interpretation of our results in the conclusion.

³² The Tigray region is the only region that issued certificates only in the names of household heads, who are traditionally men.

³³ A few papers document positive effects of the land certification program, particularly on agricultural investment and productivity, and the workings of the land rental market (e.g., Holden et al., 2009; Deininger et al., 2011; Holden et al., 2011).

The remaining part of the paper is organized as follows. Section 2 briefly describes the Ethiopian land registration and certification program. Section 3 provides an account of women's empowerment, and discusses a basic theory to rationalize why the formalization of land rights affects empowerment. Section 4 discusses our data and identification approach. This is followed by the results and their discussion in Section 5. Section 6 presents a sensitivity analysis and probes the robustness of the results. Section 7 gives concluding thoughts.

3.2 The Ethiopian land registration and certification program

As is true for much of Africa, land in Ethiopia is a major source of livelihood, income and employment for the majority of the population. Land is more than just a production factor – it remains a source of political power, a foundation for complex relations of alliance and reciprocity, and an integral element of social identity (Crewett and Korf, 2008). It is no surprise, therefore, that land rights tend to occupy a central place in policy debates (Deininger et al., 2008; Bezabih et al., 2011). During the imperial regime of Haile Selassie, land was controlled by the elite, and exploitative landlord-tenant relations were prominent. In 1975, the socialist *Derg* regime came to power and nationalized all rural land and natural resources, and started to redistribute land to the tiller. The current government, which came to power in 1991, retained state ownership of land and natural resources, giving only usufruct rights to land lords excluding the right to mortgage or sell land. While maintaining land redistribution as an important policy tool for achieving egalitarian land distribution, the current government introduced additional features in the land rights proclamation. This included decentralization of land policy responsibility to regional governments in an effort to secure tenure in major regions of the country (e.g., Deininger and Jin, 2006; Crewett and Korf, 2008).

To improve tenure security, a rapid and low-cost land registration and certification program was implemented. The motivation for the land certification program was in line with the standard economic argument that formalizing land rights enhances tenure security, which promotes sustainable and profitable use of the land through boosting land-related investments. Moreover, certification may reduce land conflicts (Deininger et al., 2008). As mentioned above, and the key focus of this paper, the certification program had a gender component, aiming to strengthen women's rights to land through joint certification.

With the goal of issuing every rightful holder of farmland a certificate of usufruct, the Ethiopian land certification program has delivered the largest non-freehold land rights in a

short period in Sub-Saharan Africa. Since the onset of the program, more than 20 million parcels of some 6 million rural households have been registered across the country (Deininger et al., 2011). Books and certificates are issued with the names of both spouses (except in the Tigray region), and they contain details related to size, location and quality of farm plots as well as a summary of legal provisions in the form of holders' rights and obligations.

For this study, we focus on Amhara region (Ethiopia), which provides an interesting case for a number of reasons. First, substantial gender-based disparities in virtually all dimensions of household decision-making are prevalent in this region. For example, Fafchamps and Quisumbing (2002) show that productive resources are usually controlled by the household head –a role culturally identified with men. Prior to the reform, few women had ownership rights to land, and even fewer had control over it in terms of production and management decisions. For example, data from the FAO gender and land rights database show that only about 19 percent of women possessed land on average, while 81 percent of men did (FAO, 2002). Second, the customary land tenure system was traditionally characterized by patrilineal inheritance and virilocal residence, ignoring women's land rights. In the gist of the custom, girls move to the home of their husband upon marriage, and inherit no land from their parents. Moreover, harmful traditional practices such as early marriage, kidnapping and female genital circumcision are prevalent in the region.³⁴

The Amhara region started a pilot in 2002, and scaled up the joint certification program in 2003, emphasizing equity and joint ownership between genders. The program was implemented in a decentralized, broad-based, high-speed and participatory approach, allowing the program to adapt to local conditions (Deininger et al., 2008; Toulmin, 2009). Moreover, the program costs were low and affordable, owing largely to the employment of local tools and techniques with strong local participation (Holden et al., 2011). It has also been regarded as “wealth neutral” since poor households had similar probabilities of receiving land certificates as richer households (Deininger et al., 2008). The process of land certification starts with a series of public meetings to create a common understanding and awareness about land registration.³⁵ Following that, an independent Land Use and Administration Committee is elected at the kebele level. The committee is versed with the responsibility for major

³⁴ Traditionally, women resort to various strategies to access resources within the household. Among others, excelling in domestic tasks, giving birth to a preferred sex, hiding their income from off-farm work and mediation through their parents and relatives are the chief strategies for women to have access to household resources.

³⁵ The process of the program varies to some extent from region to region. Here the descriptions are mainly based on the experience of the Amhara region, where our survey was conducted.

preparatory tasks of certification and registration, which include plot identification, demarcation and measurement, and recording of personal details of holders and their family members. It also includes a systematic field-based adjudication of rights through a public process to resolve conflicts in the presence of neighbors and elders. Land certificates and books are issued with the names and photos of both spouses. This makes selling or renting out land more difficult for husbands if they do not have their wives' consent.

3.3 Theory: Land rights and women's empowerment

The concept of empowerment reflects one's control over resources, which affects the outcomes of bargaining processes. Different scholars have conceptualized empowerment in different ways. For example, Narayan (2005) adopted an institutional approach to empowerment, defining it as the expansion of assets and capabilities of poor people to participate in, negotiate with, influence, control, and hold accountable institutions that affect their lives. Kabeer (1999) proposed an "ability approach" to empowerment, conceptualizing it as the expansion in people's ability to make strategic life choices. Empowerment then incorporates three inter-related constructs: (a) resources, (b) agency (the ability to engage in actions), and (c) achievement (whether the desired outcome is really in one's interest). Hence, women's empowerment is typically related to the expansion of asset bases and capabilities to make strategic life choices.

Why may joint certification empower women? This would not be evident from the unitary household model, pioneered by Becker (1991). According to this model, a benevolent dictator or representative agent pools household resources and allocates them to maximize a household utility function. Conflicting preferences among household members are ignored, and changes in control over resources would have no effect on outcomes. This unitary approach has been criticized because it falls short of methodological individualism, and leaves the household itself as a black box (Bourguignon and Chiappori, 1992). It is unable to explain such real-life phenomena as household formation and dissolution, and intra-household allocation decisions. Further, empirical evidence speaks against theoretical predictions of the unitary model (e.g., Attanasio and Lechene, 2002; Quisumbing and Maluccio, 2003; Morozumi, 2012; Martinez, 2013). In recent years, emphasis has shifted to collective models of household behavior.

Several (cooperative) bargaining models demonstrate that securing access to assets is a means to empower women. Manser and Brown (1980) demonstrate that both the Nash bargaining

rule and the Kalai-Smorodinsky bargaining solution predict that improving the “threat point” of agents – the outcome that is realized outside the marriage – increases their bargaining power, so that collective choices are more closely aligned with that agent's preferences. It is obvious why joint certification improves the threat point of women – in case of divorce they no longer need to return, asset-less, to their parental households, but now have the choice to eke out an existence based on the land at their disposal (or live off rental income). We present a formal exposition of this argument as follows.

Assume a household consisting of two people: male (m) and female (f). Each individual i ($i = m, f$) possesses a neoclassical utility function U^i defined over a consumption of a public good (X_0), a private good (X_i) and own leisure (L_i). Before marriage, the utility functions of m and f are represented by $U_0^m(X_0, X_m, L_m)$ and $U_0^f(X_0, X_f, L_f)$, respectively. If not married, m and f each maximize their separate utilities subject to their individual full income constraints. This gives rise to their respective indirect utility functions $V_0^m(P_0, P_m, W_m, Y_m; \phi_m)$ and $V_0^f(P_0, P_f, W_f, Y_f; \phi_f)$, where P_0, P_i and W_i are prices of the public good, the private good and own leisure of individual i , respectively. Y_m and Y_f are their respective nonwage incomes.³⁶ The vectors ϕ_m and ϕ_f are “extra-household environmental parameters” (McElroy, 1990) and represent any variables that shift the maximum value of utility attainable by individuals outside of marriage.

If m and f are married, they jointly allocate resources according to a two-person, cooperative Nash game. Key features of this cooperative bargaining model are utility gains from marriage and threat points. Individuals m and f would choose to get married because of positive Nash surplus within marriage, as compared to their utility as a single. Utility gains from marriage are possible because of joint consumption of household public goods (sharing a house), and caring. Once married, the utility of each individual depends not only on the household public good, the private good and own leisure, but also on the welfare of his or her partner. After marriage, the individual utility functions are given by $U^m = U^m(x)$ and $U^f = U^f(x)$ where x is a vector. Accordingly, the gains from marriage for the spouses equal $U^m(x) - V_0^m(P_0, P_m, W_m, Y_m; \phi_m)$ and $U^f(x) - V_0^f(P_0, P_f, W_f, Y_f; \phi_f)$. Marriage is worthwhile when $U^m(x) \geq V_0^m(P_0, P_m, W_m, Y_m; \phi_m)$ and $U^f(x) \geq V_0^f(P_0, P_f, W_f, Y_f; \phi_f)$.

³⁶ Generally, Y_i can be interpreted as some fixed endowment, such as land and other assets.

The expressions $V_0^m(P_0, P_m, W_m, Y_m; \phi_m)$ and $V_0^f(P_0, P_f, W_f, Y_f; \phi_f)$ are the threat points for each spouse, and represent the reservation utility that (s)he would receive in case of divorce. It follows immediately that the utility of a spouse in the Nash bargaining solution is an increasing function of that spouse's threat point utility, which is a function of access to household assets (including land) after the marriage has dissolved.

The couple solves the utility-gain product function to get the Nash solution. This implies maximizing

$$N_{\max X} = [U^m(x) - V_0^m(P_0, P_m, W_m, Y_m; \phi_m)] [U^f(x) - V_0^f(P_0, P_f, W_f, Y_f; \phi_f)] \quad (1)$$

subject to the full income constraint.³⁷ The relevant household budget constraint is $\Sigma px = (W_m + W_f)T + Y_m + Y_f$, where T is the number of hours allocated to work by each spouse and p is a vector of the relevant prices. Changes in extra-household environmental parameters have no effect on the budget constraint, and affect only the utility-gain product function through the threat points (so long as spouses remain as a couple). The solution to this system gives demand functions of the form $x_i = x_i(p, Y_m, Y_f, \phi_m, \phi_f)$.

To investigate how joint land certification empowers women, we derive the comparative statics of the Nash model. McElroy and Horney (1981) define the family rate of substitution of x_i for x_j as minus the slope of the iso-gain product curve:

$$FRS_{ij} = -\frac{dx_j}{dx_i} = \frac{N_i}{N_j}, \quad (2)$$

They then show that:

$$\begin{aligned} \frac{\partial FRS_{ij}}{\partial \phi_f} &= \frac{U_j^f U_j^m}{N_j^2} \left[\frac{U_i^f}{U_j^f} - \frac{U_i^m}{U_j^m} \right] \left[\frac{\partial V_0^f}{\partial \phi_f} (U^m - V_0^m) - \frac{\partial V_0^m}{\partial \phi_f} (U^f - V_0^f) \right] \\ &= \frac{U_j^f U_j^m}{N_j^2} [\Delta MRS_{ij}] Q. \end{aligned} \quad (3)$$

³⁷ The Nash solution is known to be characterized by Pareto optimal outcomes, independence of irrelevant alternatives, symmetry and invariance with respect to positive linear transformation of the utility functions.

An increase in ϕ_f changes the family's marginal rate of substitution, which causes a reallocation of family resources towards goods and services preferred by women. The first and third terms in this expression are positive. The second term, or the difference in the spouses' individual marginal rates of substitution, is also positive as long as women are "sufficiently selfish". An increase in the bargaining power of women moves the household's consumption basket in favor of her preferences.

Finally, taking a different perspective (but reaching similar outcomes), Lundberg and Pollack (1993) proposed the so-called separate spheres bargaining model—treating non-cooperative behavior within marriage as the relevant threat point. Control over own work effort and other assets is an important determinant of the power balance within the household. In Lundberg and Pollak approach, joint land certification implies greater command over land and land-related products for rural women. This could reshape traditional roles, and relax the constraints associated with the spheres that women would have says in decisions.

3.4 Data and empirical approach

3.4.1 Data description

In this section, we introduce our data and identification strategy. We try to adopt a nuanced approach to measuring women's empowerment, accounting explicitly for its multidimensional nature. We combine data collected using survey and field experiment approaches to generate a rich and multi-faceted dataset for our analysis. Our survey data give a real world context to the outcome measures and our analysis. The field experiment allows us to collect data in an incentive-compatible way with real payoffs at stake, which would enhance the objectivity of our outcome measures, and increase the potential to obtain a general picture.

3.4.1.1 Survey data

We collected cross-sectional data in Amhara region, the second most populous region of Ethiopia.³⁸ The household survey covered the districts of Womberma, Bure and Jabitehinan, and was conducted in November and December 2012. Overall, 325 households were randomly sampled from five certified and seven uncertified kebeles from the three districts,³⁹ and 141 households were jointly certified (see below). The survey contained different modules on household issues, including general information, marriage history, household

³⁸ The administrative structure of Ethiopia is in the following order: federal government, regional government, zone, woreda (district), kebele (the lowest administrative unit).

³⁹ The distribution of the kebeles was as follows: five from Womberma, four from Bure and three from Jabitehinan.

assets, and modules on the different empowerment indicators. We interviewed men and women separately.

Table 1 summarizes our explanatory variables (definitions are relegated to the appendix). The explanatory variables aim to capture the setting and resource conditions of women. They include the demographic situation (e.g., age of the woman, presence of children above 15 years), economic situation (e.g., assets brought to marriage), social status (e.g., woman literacy and husband literacy), and information about the marriage (e.g., years of marriage, age difference of the couple, distance from parents). We also collected data on community conditions (e.g., distance from main roads, NGO presence). On average, wives are 8 years younger than their husbands, indicating the common practice of early marriage of girls in the study areas. Women are largely illiterate and upon marriage tend to move out of the village of their parents (it takes about 3 hours on walk for an average woman to visit her parents). While 84% of the women in our sample have a written marriage contract with their current husband, only 19% of them reported that they knew their husbands before marriage. This reflects the traditional practice that parents and relatives arrange most marriages in rural Ethiopia (Fafchamps and Quisumbing, 2005a). In Table 2, we split the sample, and distinguish between certified and uncertified women. A series of two-sample *t-tests* demonstrates that some covariates are differently distributed between the two groups. This suggests the need to control for these variables in our analysis (see below).

Table 1: Summary statistics of explanatory variables

Variables	Obs.	Mean	Std. dev.	Min	Max
Age of woman	325	36.08	8.88	18	57
Years of marriage	325	19.70	9.75	1	50
Couple difference	325	-8.06	6.47	-30	14
Woman literacy	325	0.26	0.44	0	1
Husband literacy	325	0.766	0.42	0	1
Children above 15 years old	325	0.615	0.49	0	1
NGO presence	325	0.05	0.21	0	1
Land brought to marriage ⁴⁰	325	0.56	1.32	0	12
Livestock brought to marriage ⁴¹	325	0.46	0.86	0	9.1
Distance from parents	325	210.42	269.09	0	1440
Distance from main road	325	44.66	42.43	0	180
Marriage contracta	325	0.84	0.36	0	1
Know husbanda	325	0.19	0.39	0	1

^aThese two variables are among three additional variables that are assumed to affect only women's empowerment and are included only in the Tobit regression model.

⁴⁰ The land brought to marriage variable includes both land brought to marriage and land inherited after marriage by women. Land at marriage and land inherited after marriage may be subject to different processes (e.g., Fafchamps and Quisumbing, 2005a; 2005b). However, we aggregated them as there were only 10 observations for land inherited after marriage.

⁴¹ To construct tropical livestock units we use the Sub-Saharan conversion rates: mature cow (1.0), ox (1.42), small cattle (0.73), goats/sheep (0.2), horses/mules/donkeys (0.8), and poultry (0.04) (Njuki et al., 2011).

Joint Land Certification Programs and Women's Empowerment

Table 2: T-test for explanatory variables between certified and uncertified groups

Variables	Certified women	Uncertified women	Difference
Age of woman	36.489	35.761	0.728
Years of marriage	20.752	18.886	1.866*
Couple difference	-8.199	-7.946	-0.253
Woman literacy	0.355	0.190	0.165***
Husband literacy	0.858	0.696	0.162**
Children above 15 years old	0.660	0.582	0.078
NGO presence	0.014	0.071	-0.057**
Land brought to marriage	0.604	0.533	0.071
Livestock brought to marriage in tlu	0.337	0.546	-0.209**
Distance from parents	111.227	286.440	-175.213***
Distance from main road	42.156	46.571	-4.415
Observations	141	184	

* Significant at 10%; ** significant at 5%; *** significant at 1%

Next, we turn to our outcome variables, or our measures of women's empowerment. Empowerment is complex and has multiple dimensions (e.g., Mason, 1986; Kabeer, 1999; Mahmud et al., 2012). While the empirical literature emphasizes economic empowerment indicators, it is well understood that institutions and other social factors, such as kinship networks, social norms and culture, also play central roles in defining perimeters within which women and men interact and negotiate (Agarwal, 1997). The employment of multiple indicators enables us to capture a range of dimensions. In total, 54 survey items were used to capture specific components of women's empowerment.⁴² Appendix A shows the descriptive statistics of each underlying variable included in the construction of the empowerment indicators, and certified women tend to respond differently compared with uncertified women. We aggregated the responses to these 54 questions into six complementary indices: (i) women's participation in household decision making, (ii) participation in community activities, (iii) knowledge of land rights and land related provisions, (iv) level of tenure security and dare to protect their land rights, (v) freedom of physical mobility and autonomy, and (vi) institutional inclusion.

The variable '*participation in household decisions*' is based on responses to 18 survey questions, covering various aspects of household decision making. The questions used to construct this index include land-related and other decisions within a household. The index '*participation in community activities*' captures responses to seven questions designed to measure the extent of women's participation in community activities. The questions include frequency of participation in village meeting, voting, public information meeting and

⁴² Most of these items were binary response questions with 'yes' and 'no' answer options. Those items which had more than two response options were recoded to binary dummies to facilitate construction of indices.

participation on various community development activities. Our '*knowledge of land rights and related legal provisions*' index is based on responses to 17 questions on key aspects of the land law. Our intent was not to examine women's legal knowledge in the deeper sense of the term but to assess about their legal literacy at a basic level. Knowledge of land rights and related legal provisions is as important as the letter of the law when it comes to the effects of land reforms. Knowledge of land rights and relevant legal provisions is not only a *sine qua non* for enforcement of rights, but empowers owners to defend their rights and promote their interests (Deininger et al., 2008). While the land law affects everybody, the land certification program makes women aware of their rights as certification was preceded by series of public information meetings. Moreover, issued land books and certificates summarize legal provisions (emphasizing rights and obligations for holders).

The index '*women's level of tenure security and dare to protect their land rights*' aims to measure the confidence and expectation of women in state enforcement to protect their land rights. The questions used to construct this indicator focus on women's subjective level of tenure security and their confidence in the legal system in protecting their land rights, whether they dare to protect their land rights, whether they expect equal land share upon divorce, and expect their rightful land share upon death of their husbands. We hypothesize that women in certified households should enjoy higher level of tenure security and confidence in formal institutions than uncertified women. The indicator '*physical mobility and autonomy*' gauges the extent of women's independence within the household in taking certain decisions. This indicator is drawn from women's responses to such questions as whether she can go to the market alone, decide by her own to travel and visit family and friends, take her children to a health clinic alone, and whether she needs permission from her husband to attend meetings. Our final indicator for empowerment is '*women's institutional inclusion*', which captures women's participation and inclusion in local institutions. We used women's membership in women's associations, *Eqqub*⁴³ and village development committee to capture the extent of their institutional inclusion.⁴⁴ Women's institutional inclusion can provide them with opportunities to voice and express their interests (Weldon, 2002). Finally, we construct an overall women's empowerment index by aggregating all 54 items.

⁴³ Eqqub is a kind of rotating savings and credit association (ROSCA). It is often organized by members themselves and can measure women's capacity to organize themselves in groups for mutual beneficial activities.

⁴⁴ The survey included other questions meant to measure women's institutional inclusion at local level, like membership in *Iddir* (funeral group), school committee or parent-teacher associations, political party or movement, village NGO or civic group, and health and sport groups. However, we excluded these questions from index construction since almost all interviewed women gave the same answer (the answer for all individuals was 'no').

Index outcomes and mean differences between certified and uncertified women are provided in Table 3. A two-sample *t-test* for mean differences of empowerment indicators reveals that indicators are significantly different. A tentative conclusion from this simple comparison of means is that joint certification is associated with the empowerment of women. However, causal inference requires controlling for a number of potential sources of biases (see Table 2), so in what follows we employ propensity score matching method for making causal inference (see below).

Table 3: Women’s empowerment indicators

Variables	Joint certified women	Uncertified women	Difference
Overall participation in HH decisions	0.9196	0.8584	0.0612 **
Participation in community activities	0.6429	0.4570	0.1859 ***
Women’s knowledge of their land rights and land related legal provisions	0.8990	0.7132	0.1858 ***
Level of tenure security and dare to protect their land rights	0.9560	0.8804	0.0756***
Physical mobility and autonomy	0.5585	0.4647	0.0938 ***
Women’s institutional inclusion	0.3830	0.1993	0.1837***
Overall women’s empowerment index	0.8230	0.6951	0.1279 ***
Experimental empowerment index ^a	0.6090	0.3278	0.2812***
Observations	141	184	

* Significant at 10%; ** significant at 5%; *** significant at 1%

^aThe mean difference for this indicator is calculated only for 244 observations since the remaining are set to missing due to the index construction method (see below).

3.4.1.2 Experimental data

We also aim to augment our survey data through incentive compatible and objective experimental data. Following a rapidly growing literature on household behavior using experiments (e.g., Bateman and Munro, 2005; Ashraf, 2009; Carlsson et al., 2012), we use an incentivized field experiment to elicit individual and joint preferences of spouses. Spouses played simple allocation games, first separately and then jointly. Appendix B provides detailed instructions. We kept wife and husband in separate rooms, providing each with a separate endowment of 15 Ethiopian Birr⁴⁵ (ETB)—around half of the average daily income in the area. We requested them to allocate the 15 ETB between a private account and a common pool account of their respective groups. The amount allocated to the private account was their money to take right away. The amount contributed to the common pool by each player was doubled, and later divided equally among all group members (i.e., a standard

⁴⁵ We set deliberately the value of the endowment to be odd number so as to avoid simple equal allocations. 1USD was about 17 ETB during the survey period.

public good game). Each group consisted of four randomly selected and anonymous fellow villagers. Upon completion of the private game, spouses were brought together to play the same game but jointly, and each was again given an endowment of 15 ETB to allocate between a private and common pool account. However, they now first had to settle on an agreed-upon allocation of endowments. New groups of four people (from the same community) were created for the joint decision-making stage, and spouses were again allocated to different groups. At the end of the experiments, the total payment to players was the sum of money they chose to keep (private stage) and their share of the common pool (joint stage).⁴⁶ On average, each individual earned approximately 50 ETB.

Care was taken to ensure that each spouse understood the rules of the game. Players were not cognizant that a joint stage would follow their private stage of the game. Results from private stages of the experiment were not revealed until after spouses played the joint stage to reduce possible learning effects. Subjects were also told their spouses would not be a member of their ‘public good group’ in any stage of the game. Moreover, all experiments in a single village were completed on a single day to minimize potential learning effects.

We use joint decisions relative to individual decisions to construct a game-specific measure of relative bargaining power. Following Carlsson et al. (2012), we assume joint decisions to be a function of individual preferences revealed in the private stage of the game, and the relative bargaining power. We construct our experimental empowerment index as the difference between the joint and the husband decisions, relative to the difference between the wife and the husband decisions. Given the limited choice sets, some joint decisions are expected to lie outside the interval of private decisions.⁴⁷ If so, we interpret this as evidence that the spouse whose private decision lies closest to the joint decision has full bargaining power. Following Wiig et al. (2011) and Carlsson et al. (2013), our empowerment index is calculated based on the following formula:

⁴⁶ If the amount allocated to the common pool is C_i , then the amount an individual retain for himself or herself is $15 - C_i$. The payoff function (P_i) for each individual is given as $P_i = 15 - C_i + \frac{1}{4}(2) \sum_{j=1}^4 C_i$. This incentive structure was maintained in both parts of the game.

⁴⁷ Similar reasoning is reported in the domain of risk taking decisions of spouses. There is a significant trend of reporting towards either the higher or the lower private contribution in the literature. For example, Bateman and Munro (2005) found that the contribution in joint decision making in risk games tends to be lower than the average of individual choices, i.e. agree closer to the most risk averse alternative.

Experimental Empowerment Index

$$= \begin{cases} 1 & \text{if joint decision outside private interval on women's side} \\ \text{abs} \left(\frac{\text{joint decision} - \text{male decision}}{\text{female decision} - \text{male decision}} \right) & \text{if joint decision in the interval between the two private decisions} \\ 0 & \text{if joint decision outside private interval on man's side} \end{cases}$$

The experimental empowerment index is continuous and is censored at 0 and at 1.⁴⁸ An empowerment index score with value of 0.75 implies a joint decision inside the interval between the two private decisions, and unequal to both of the private decisions, but closer to the wife's private decision. Likewise, values of 1 and 0 indicate full bargaining power of wife and husband, respectively. Couples who made similar private allocation decisions were dropped because we could not derive any information about bargaining power for such couples.⁴⁹ This procedure has yielded useful experimental data for a subsample of 244 couples, 97 of which are certified. A two-sample *t-test* for the mean differences of the experimental empowerment indices between joint certified and uncertified women is significantly different (see Table 3).

Decisions made together might differ from those made privately for various reasons that may not be related to preferences and bargaining power (e.g., Ashraf, 2009; Luhan et al., 2009), particularly in case of sensitive decisions where people may be unwilling to reveal their preferences to their spouse. Other reasons include learning effects from discussions, or that joint choices may reflect outcomes for the whole household (rather than private interests). We investigate the joint decision process to probe for this possibility. We control for differences between decisions of women in the private and joint games by including a dummy to capture whether the private allocation of the wife exceeded the private allocation of her husband at the private game (in our sensitivity analysis).

Table 4 provides summary statistics of the games (definitions of variables are given in the appendix). On average, wives and husbands contributed to the common pool 4.2 and 4.3 ETB, respectively. The difference between them is not significant at conventional significance levels⁵⁰. However, wives had lower expectations of average contributions by other group

⁴⁸ We use this nature of the data to employ Tobit model with additional covariates in our robustness analysis.

⁴⁹ Dropping couples who make similar private allocation decisions should be random and should not influence to whom this variable is defined for. We test this hypothesis using a simple two-sample *t-test* and Appendix C shows the balance test statistics for non-missing and missing couples. As expected, no significant difference was observed for key variables, indicating that the mechanism for missing was random. Alternatively, we construct another 'experimental empowerment index' for the entire sample by assigning 0.5 to both couples when private contributions are equal, and the results remain largely similar (results available on request).

⁵⁰ P-value for mean comparison of paired data is 0.8697.

members than did husbands (4.9 and 5.6 ETB, respectively).⁵¹ Moreover, spouses contributed on average 4.2 ETB from their endowments to the common pool in the joint stage of the game.

Table 4: Summary statistics of the public goods games

Variables	Obs.	Mean	Std. dev.	Min.	Max.
Private kept wife	325	10.8246	2.8279	0	15
Private allocated wife	325	4.1785	2.8282	0	15
Expected average contribution by wife	325	4.9308	2.9217	0	15
Private kept husband	325	10.746	2.8298	0	15
Private allocated husband	325	4.3138	2.8241	0	15
Expected average contribution by husband	325	5.5538	3.8331	0	15
Joint kept	325	10.7554	2.5548	0	15
Joint allocated	325	4.2446	2.5548	0	15
Trend	244	0.4898	0.5009	0	1
Experimental empowerment index	244	0.4518	0.4638	0	1

3.4.2 Identification and empirical approach

In empirical work probing the determinants of bargaining power, analysts need to address potential endogeneity concerns. For example, using current assets or wage income as proxies for bargaining power is inappropriate, as these variables may be a consequence of the prevailing power structure – rather than being its cause. The same is true for uptake. Another threat to identification related to our study is non-random placement of the certification program. We now deal with these two issues.

Consider the challenge of non-random program placement first. The land certification program was not introduced at once across Ethiopia. Instead, districts and kebeles were gradually enrolled in the program (see also Bezabih et al., 2012 for details on how districts sequentially enrolled into the program). Identification of average treatment effects on the treated (ATT) relies on comparing certified households in certified Kebeles and uncertified households in uncertified Kebeles. However, if certified kebeles were different from uncertified kebeles prior to the certification program, and if these differences were correlated with bargaining power of women, then correlations between women’s empowerment and land certification could be spurious. Cognizant of this threat, we constructed a sample frame comprising kebeles that were uncertified for ‘exogenous’ reasons (i.e., unrelated to intra-household organization). These reasons included human resource and certificate constraints at the district level, border conflicts between kebeles, and resettlement-related conflicts (see also

⁵¹ P-value for mean comparison of paired data is 0.0154.

Chapter 2 for details).⁵² Our uncertified kebeles were then randomly chosen from this set. We also obtained key information from relevant “natural resources and land use administration” officials that the conflicts typically involve competing claims on grazing lands that are obviously relevant from the perspective of rural livelihoods, but have no bearing on distribution of power within households in settled farming communities. We argue certified and uncertified kebeles were not systematically different *ex ante* in terms of intra-household power distribution.

The second challenge to identification is non-random uptake of joint land certification (in the sub-sample of certified kebeles). That is, in kebeles selected for the land certification program, a non-random subsample of the households may decide to take the effort to obtain a certificate. Insofar as households with empowered women are more likely to seek formal titles, a simple comparison of certified and uncertified households in terms of women's empowerment would be misleading. The data summarized in Table 2 are consistent with the assumption of non-random uptake, and therefore necessitate controlling for the relevant covariates.

We use propensity score matching (PSM) approach to tackle the problem of selection bias due to observable covariates (Heckman et al, 1997). Let D_i be a binary treatment indicator that equals 1 if woman i receives a joint land certificate, and 0 otherwise. The women empowering impact of certification is defined as $Y_i(D_i)$ for each individual i , where $i = 1, \dots, N$ and N denotes the total population of women in the study. The average treatment effect on the treated (ATT) is computed as:

$$T_{ATT} = E(T | D=1) = E[Y(1) | D=1] - E[Y(0) | D=1] \quad (4)$$

where T_{ATT} is the average treatment effect on the treated; $E[Y(1) | D=1]$ is the expected empowerment level for women who received joint land certificates and $E[Y(0) | D=1]$ is the expected level of empowerment for joint certified women if they had not been certified. The latter is a counterfactual, which is not observed. PSM uses data from uncertified women who share similar observable characteristics with treated women to credibly estimate this counterfactual.

⁵² All the three districts in our study belong to the high-productive and moisture-reliant highlands ‘development domain’ of the country (Chamberlin et al., 2006). The resettlements typically involve households coming to some kebeles in these districts, where extra land is expected for redistribution. We excluded resettled households from our analysis to mitigate some of the unobservable bias *a priori*.

PSM involves two identifying assumptions: common support condition and conditional independence assumption (CIA). The common support condition requires the presence of sufficient overlap in propensity scores across treated and non-treated women samples (Heckman et al., 1998). The CIA implies that all relevant observable covariates that affect the outcome and the treatment are properly controlled for (Caliendo and Kopeinig, 2008). However, conditioning on all relevant observables is limited, particularly in the case of high dimensional covariates vector X . To deal with the dimensionality problem, Rosenbaum and Rubin (1983) showed that balancing on all relevant covariates X is equivalent to balancing on a propensity score $p(X)$, which is the probability of receiving treatment conditional on all relevant covariates X . Given that the common support and CIA assumptions hold, the PSM estimator for T_{ATT} is given by:

$$T_{ATT}^{PSM} = E_{(p(X)|D=1)} \{E[Y(1)|D=1, p(X)] - E[Y(0)|D=0, p(X)]\} \quad (5)$$

We use a standard logit model to estimate propensity scores. Only variables that simultaneously influence participation and outcomes should be included in the model (Heckman et al., 1997; Smith and Todd, 2005). We consulted theory and previous empirical studies to select appropriate explanatory variables (e.g., Fafchamps and Quisumbing, 2002; 2005b). We use kernel matching algorithm to match treatment and control women. Kernel matching performs better in reducing standardized mean bias between treated and control groups compared to other matching algorithms (Heckman et al., 1998). To probe the model's robustness, we also estimate ATT using alternative matching algorithms. We also carry out several matching quality assessment tests.

Finally, one potential remaining concern in our causal identification effort is that selection due to unobservables may play a role in program uptake decisions. While PSM estimators do not control for such selection bias, we believe our results are unlikely to suffer from selection bias due to unobservables for several reasons. First, while registration and certification is a voluntary decision of households, the certification intervention is driven by the government, and exogenous to household unobservables. The responsibility for initiating land registration and certification is solely bestowed on Woredas (districts), not on Kebeles or villages, let alone on individual households. This in turn lessens the concern related to the role of local institutional heterogeneity in program uptake. Moreover, very low and affordable program

participation costs and high compliance rates (exceeding 90%) in certified kebeles⁵³ imply that self-selection into the program at household level is not likely to be relevant.⁵⁴ Second, certified and uncertified households are from districts that belong to the same 'development domains' with sufficient similarity in economic opportunities and livelihood options (Chamberlin et al., 2006). The categorization into development domains is made dependent on agricultural potential, access to markets and population density. The three districts are included in the same domain, characterized by high population density and moisture-reliant highlands with productivity growth and market improvement as their priority livelihood strategies. Finally, we formally check the robustness of our results for selection bias due to unobserved heterogeneity using the standard bounding approach test, following Rosenbaum (2002).

3.5 Results and discussion

3.5.1 *Effects of joint land certification on women's empowerment*

The first step in propensity score matching is to estimate the propensity scores for matching certified and uncertified women from the pool of subjects in the comparison kebeles. Table 5 reports standard logit model estimates of marginal effects for the various observables. From the Table, several covariates have a significant effect on the likelihood of program uptake. All observables but 'NGO presence' have the expected sign.⁵⁵ The likelihood of receiving a joint land certificate for a household increases with both the wife and husband's literacy level. Women with smaller social networks and fewer kinship relations in the vicinity of their residence (captured by distance to her parents) are less likely to obtain a joint certificate. The results of the logit model for program selection support the use of PSM to account for the

⁵³ For example, see Deininger et al. (2011).

⁵⁴ The source of tenure insecurity (i.e., driver of certification) in Ethiopia has been the state itself through its frequent land redistribution policies (Deininger and Jin, 2006). Moreover, customary land laws do not exist in Ethiopia, owing to the discontinuous evolution of its tenure system (Deininger et al., 2008). Land proclamations do not leave discretion for making land distribution decisions beyond regional state administrations. All these features contribute to minimizing potential roles of local level heterogeneity in certification uptake. Moreover, all sample households belong to the Amhara ethnic group and Orthodox Christianity.

⁵⁵ NGO presence seems to impair women from receiving joint land certificates. We conjecture two potential reasons as possible explanations for this contrary effect. First, gender related NGOs might increase tension in the household though it enhances women's outside participation and decision making ability. This is possible because people may perceive it as a direct violation to the culture of men dominance in the household. Second, our study districts are known for deep-seated early marriage traditions. NGOs aiming to empower women are often directed to villages where women are more marginalized and less empowered due to such traditional practices and gender stereotypes (Orloff, 1993). However, means-tested safety net programs, like food for work, are not common in our study districts, as they are among the productive and self-sufficient areas of the country.

significant covariates to purge systematic differences between certified and uncertified women.

Table 5: Program selection model^a

Variable	Marginal effects
Age of woman	-0.003 (0.0283)
Years of marriage	0.0602 (0.0569)
Years of marriage squared	-0.0008 (0.0012)
Couple difference	0.0466 (0.0887)
Interaction of age and couple difference	-0.0017 (0.0023)
Woman literacy	0.6362** (0.2949)
Husband literacy	0.8784*** (0.3284)
Children above 15 yrs.	0.0030 (0.3669)
NGO presence	-1.6127* (0.8553)
Land size brought to marriage	0.0764 (0.1101)
Livestock size brought to marriage	-0.2502 (0.1891)
Distance from main road	-0.00004 (0.0031)
Distance from parents	-0.0036*** (0.0008)
Constant	-1.2162 (1.1739)
Pseudo-R ²	0.1575
Observation	325

* Significant at 10%; ** significant at 5%; *** significant at 1%

^a Standard errors are in parentheses.

After estimating propensity scores, the next step is matching certified and uncertified women based on ‘closeness’ in terms of propensity scores, and then estimating the ATT. We use kernel matching algorithm, since this method uses more information and has a lower variance (Heckman et al., 1998). Since we have eight empowerment indicators, including the experimental index, we estimate eight different ATTs. The results of the matching are summarized in Table 6. Matching is restricted only on the common support region. The fifth and sixth columns show the respective numbers of joint certified and uncertified women on the common support who provide useful information for computing the relevant ATTs. Only 9 certified women are off the common support region, implying relatively small information loss.⁵⁶ This is a reasonably good diagnostic indicator of success in matching power. The

⁵⁶ 15 women are off the common support region for the experimental empowerment index.

fourth column of Table 6 gives the average effects of joint land certification on the different indicators of women's empowerment.

First, certified women are more likely to participate in community activities than do uncertified women with similar characteristics. This difference is statistically significant ($p=0.000$). Participation in community activities provides women the gateway to ensure their representation and voice in community-wide affairs. Second, certified women are more informed about their land rights and land-related legal provisions, compared with uncertified women. Knowledge of land rights and related legal provisions is a vital prerequisite for holders to enforce and defend their rights, and it also empowers title holders to promote their interests. Third, certified women are more likely to have greater levels of tenure security, confidence in the state apparatus in enforcing and defending their rights, and are more willing to protect their land rights.⁵⁷ Fourth, women with joint land rights appear to have more freedom in physical mobility and autonomy relative to uncertified women. Fifth, certified women have greater institutional participation than do comparable uncertified women. The result shows that certified women are more likely to be a member of local organizations, and hence benefit from enhanced opportunities to voice and express their interests and perspectives.

Sixth, and in contrast to the earlier results, the likelihood of women's participation in household decisions does not increase with joint certification. Certified and uncertified women are not significantly different with regard to participation in household decisions. Their respective index scores are 92% and 88%, indicating already high level of participation. This is perhaps not unexpected, as specialization in gender roles is common in Ethiopia. This results in a tacit division of responsibilities within separate spheres of the household (Lundberg and Pollak, 1993). For example, while they are also expected to participate in agriculture, women assume the sole responsibility for doing home chores and services.

Lastly, the overall survey-based empowerment index reveals that certified women score significantly higher. Results from the experimental index confirm the results from the survey data. Certified women have greater bargaining power relative to their spouses than do

⁵⁷ Many women respondents in our fieldwork confirmed this statistically robust difference. One of the members of the 'land use and administration committee' at the kebele level in Jabitehnan district said that "*after the land registration and certification program, joint certified women tend to claim their rights confidently and dare to take any possible and affordable legal action when their husbands treat them unfairly and endanger their land rights*".

uncertified women. The respective ATTs for the overall women's empowerment and experimental indices are highly statistically significant ($p=0.000$).

The effects are also substantial in magnitude. Particularly, the ATT effects are relatively large for such indices as '*knowledge of land rights and land related legal provisions*', '*participation in community activities*' and '*women's institutional inclusion*'. Generally, the ATT effects range from 7-17 in percentage points. Land certification has the greatest effect on women's empowerment as measured by the '*knowledge of land rights and related legal provisions*' index, and the smallest effect for the '*physical mobility and autonomy*' indicator. The ATT effects are even larger for the survey-based '*overall women's empowerment index*' and the '*experimental empowerment index*'. These increments amount to about 11 and 26 percentage points, respectively.

Finally, the last column of Table 6 provides some diagnostic results to assess the statistical precision of the different empowerment estimators based on the minimum detectable effect (MDE) approach. Intuitively, MDE is the minimum threshold at which we can be reasonably confident to have enough power to detect true program effects (Bloom, 1995). The reported minimum detectable effects are computed for a one-sided hypothesis test at 90% statistical power and 5% level of significance. They can then be read as the smallest true effects of the respective estimators that have 90% chance of being detected with the 5% level of statistical significance. The reported results demonstrate two important details related to our different ATT estimates. First, true effects smaller than MDE have less than 90% chance of being detected. This is the case for our index measuring participation in household decisions. The MDE (0.0911) is greater than its actual ATT estimate (0.0409), which is indeed not statistically significant at conventional levels. Second, true effects larger than MDE have more than 90% chance of being reliably detected by estimators. The MDEs for all other our estimators are smaller than their respective actual ATT estimates. Accordingly, we can be reasonably confident that these differences between certified and uncertified women are due to the certification program effects.

Overall, results show that the joint land certification program has increased women participation and representation, both at the household and community level. This enables them in decision making in various aspects of their lives. Joint certification has enhanced women's fallback options through giving them the right to access to land – a crucial asset of agrarian communities, and strengthening their agency.

Table 6: Average treatment effect estimates: Kernel matching

Matched	Certified women	Uncertified women	ATT ^a	Certified women on common support	Uncertified women on common support	MDE
Overall participation in HH decisions	0.9154	0.8745	0.0409 (0.0311)	132	184	0.0911
Participation in community activities	0.6367	0.4784	0.1583*** (0.0290)	132	184	0.0850
Women's knowledge of their land rights and land related legal provisions	0.8997	0.7341	0.1656*** (0.0189)	132	184	0.0554
Level of tenure security and dare to protect their land rights	0.9561	0.8803	0.0757*** (0.0193)	132	184	0.0565
Physical mobility and autonomy	0.5511	0.4888	0.0623* (0.0124)	132	184	0.0363
Women's institutional inclusion	0.3788	0.2327	0.1461*** (0.0338)	132	184	0.0990
Overall women's empowerment index	0.8201	0.7134	0.1067*** (0.0167)	132	184	0.0489
Experimental empowerment index ^b	0.6018	0.3519	0.2499*** (0.0450)	92	137	0.1319

* Significant at 10%; ** significant at 5%; *** significant at 1%

Matching was done using epanechnikov kernel matching method using the psmatch2 Stata program of Leuven and Sianesi (2003).

^a Standard errors are bootstrapped 100 times. Standard errors are given in parenthesis.

^b The difference for this indicator is calculated based on data only for 244 households since the remaining are set to missing due to the index construction procedure.

3.5.2 Matching quality assessments

The PSM method is credible when matching eliminates differences in observables between certified and uncertified women. Different approaches are available to assess the matching quality of PSM estimators. Matching should sufficiently balance the distribution of relevant explanatory variables across certified and uncertified women. A two-sample *t*-test for equality of means on each observable, and a chi-square test for joint significance of all variables in the logit model before and after matching are two commonly used criteria for balancing test. Tables 7 and 8 present the results for these tests, respectively.

Table 7 contains covariate balancing test results for two-sample *t*-test of mean differences between certified and uncertified women, and the standardized bias of these mean differences. We report both before and after matching results. In all cases, sample differences for the unmatched sample significantly exceed those for the matched sample. Importantly, all covariates are balanced after matching, since mean differences between the certified and uncertified samples are not significant. Moreover, the standardized bias⁵⁸ for almost all

⁵⁸ The bias is calculated as the difference of the mean values of the certified group and the unmatched/matched uncertified group divided by the square root of the average sample variance in the certified group and the unmatched uncertified group (Rosenbaum and Rubin, 1985).

covariates after matching is below 5%, which is the benchmark for sufficient covariate balancing (Rosenbaum and Rubin, 1985).

Table 7: Covariance balancing *t*-tests before and after matching for Kernel matching

Variable	Unmatched Matched	Mean		%bias	%reduction bias	p> t
		Treated	Control			
Age of woman	Unmatched	36.489	35.761	8.3		0.464
	Matched	36.227	36.167	0.7	91.7	0.955
Years of marriage	Unmatched	20.752	18.886	19.2		0.087
	Matched	20.091	19.852	2.5	87.2	0.840
Years of marriage squared	Unmatched	523.460	451.460	16.7		0.136
	Matched	492.880	488.750	1.0	94.3	0.937
Couple difference	Unmatched	-8.199	-7.946	-3.9		0.728
	Matched	-8.220	-8.712	7.6	-94.6	0.550
Interaction of age and couple difference	Unmatched	-298.070	-279.860	-7.1		0.526
	Matched	-297.820	-317.500	7.7	-8.1	0.549
Woman literacy	Unmatched	0.355	0.190	37.5		0.001
	Matched	0.311	0.297	3.2	91.6	0.807
Husband literacy	Unmatched	0.858	0.696	39.7		0.001
	Matched	0.849	0.821	6.7	83.1	0.551
Children above 15 yrs.	Unmatched	0.660	0.582	16.1		0.153
	Matched	0.636	0.631	1.2	92.8	0.925
NGO presence	Unmatched	0.014	0.071	-28.2		0.016
	Matched	0.015	0.020	-2.6	90.8	0.751
Land brought to marriage	Unmatched	0.604	0.533	5.4		0.631
	Matched	0.545	0.563	-1.4	74.4	0.915
Livestock brought to marriage in tlu	Unmatched	0.337	0.546	-25.1		0.03
	Matched	0.349	0.327	2.7	89.3	0.781
Distance from main road	Unmatched	42.156	46.571	-10.4		0.353
	Matched	42.167	40.792	3.2	68.9	0.789
Distance from parents	Unmatched	111.230	286.440	-71.9		0.000
	Matched	117.140	111.630	2.3	96.9	0.736

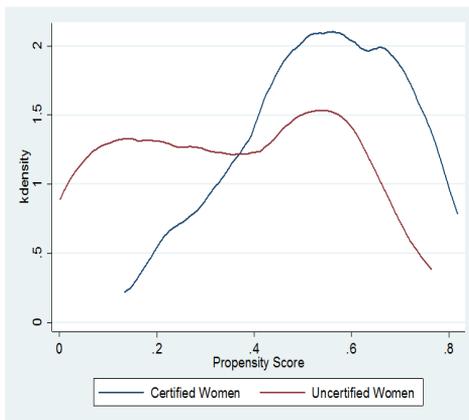
Table 8 shows the results for the joint significance of all covariates, before and after matching. The likelihood ratio test for the joint significance of all covariates for propensity score estimation reveals that all variables in the logit model are not jointly significant after matching ($p > \chi^2 = 1.000$), while the same test for joint significance is rejected before matching ($p > \chi^2 = 0.000$). The pseudo- R^2 values from the propensity score estimation logit model before and after matching echoed similar results for the balancing tests. There should be no systematic differences in the distribution of covariates between joint certified and uncertified women after matching. As expected, the pseudo- R^2 , which indicates how well covariates jointly explain the probability of participation in the certification program, is very low compared with its value before matching. Finally, the respective standardized mean differences for the model before and after matching are 22.3% and 3.3%, indicating a substantial bias reduction after matching and ascertaining once again the balancing power of our estimation.

Table 8: Chi-square test for joint significance of all covariates before and after matching

Sample	Pseudo R ²	p > chi2	Mean standardized bias
Unmatched	0.160	0.000	22.300
Matched	0.004	1.000	3.300

Another important condition for the PSM approach to create a plausible counterfactual is the presence of adequate common support for the treated and control groups. While we impose this condition in our matching, we seek to formally inspect if this key identifying assumption is met in our data. This requires checking the presence of enough overlap between the propensity score distributions of certified and uncertified women groups. The distribution of the predicted propensity scores for the entire sample ranges from 0.002 to 0.818. The propensity scores distributions of the certified and uncertified groups range, respectively, from 0.134 to 0.818, and 0.002 to 0.765. Following the ‘minima and maxima criterion’ of Caliendo and Kopeinig (2008), the effective common support is defined over the interval ranging from 0.134 to 0.765. Observations which lie outside this region are considered as outliers and discarded from the analysis. Alternatively, the density distribution graph of the propensity scores of the two groups can give a better visual reading of the extent of the common support. We present the density distribution graphs of the propensity scores before and after matching for the certified and uncertified groups in Figure 1. The figure shows the presence of clear common support in propensity score distributions of the two groups, especially after matching.

a) Before matching



b) After matching

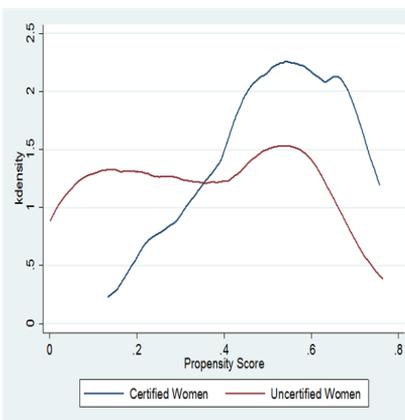


Figure 1: Density distributions of propensity scores for certified and uncertified women

Overall, the matching quality tests indicate that the matching procedure has created a high degree of covariate balance between joint certified and uncertified women.

3.6 Sensitivity analysis and extension

3.6.1 Robustness of the ATT estimates to different matching algorithms

It is worth checking the robustness of our ATT estimates to different matching algorithms, since no specific matching algorithm is *a priori* superior to others (Baser, 2006). We report ATT estimates only for overall women's empowerment for illustrative purpose in Table 9. ATT estimates for this women's empowerment index remain robust across different matching algorithms.⁵⁹ The different matching methods also perform well in covariate balancing tests. The likelihood ratio tests show that covariates are not jointly significant after matching for the different estimators. Pseudo- R^2 values are also substantially lower after matching than before matching across estimators. Finally, all matching algorithms succeed in reducing the standardized mean difference considerably. Similar results are obtained for the other empowerment indices.

Table 9: Robustness of results to matching methods and matching quality

Matching Method	NN ^b matching	NN matching	NN + calliper = 0.01	5-NN + calliper = 0.01	Radius with calliper = 0.01	Kernel matching
Replacement	Yes	No	Yes	Yes	Yes	Yes
Common support	Yes	Yes	Yes	Yes	Yes	Yes
No. of treated	132	132	77	77	126	132
No. of control	184	184	184	184	184	184
Overall women's empowerment (ATT) ^a	0.0706*** (0.0208)	0.1055*** (0.0160)	0.0673*** (0.0253)	0.0763*** (0.0279)	0.1051*** (0.0177)	0.1067*** (0.0167)
Pseudo R ² before matching	0.160	0.160	0.160	0.160	0.160	0.160
Pseudo R ² after matching	0.036	0.029	0.058	0.045	0.015	0.004
$P > \chi^2$ before matching	0.000	0.000	0.000	0.000	0.000	0.000
$P > \chi^2$ after matching	0.431	0.642	0.490	0.724	0.971	1.000
Mean standardized bias before matching	22.300	22.300	22.300	22.300	22.300	22.300
Mean standardized bias after matching	11.600	7.300	12.100	8.500	6.500	3.300

* Significant at 10%; ** significant at 5%; *** significant at 1%; ^a Standard errors are bootstrapped 100 times. Standard errors are given in parenthesis; ^b Denotes nearest neighbor matching.

⁵⁹ We also checked the robustness of other women's empowerment indicators to different matching algorithms and ATT estimates remained largely robust for the indicators (results available on request).

3.6.2 Sensitivity of ATT estimate to unobserved heterogeneity

PSM estimators are not robust to unobserved heterogeneity (DiPrete and Gangl, 2004; Caliendo and Kopeinig, 2008). They are consistent estimators based on the assumption that selection into treatment is only due to observed covariates. We use the Rosenbaum (2002) bounds sensitivity analysis to check the robustness of our results to departures from this strong identifying assumption. The results for the Rosenbaum bounds sensitivity analysis to hidden bias are contained in Table 10. The critical levels of gamma (Γ), which specify how strongly an unobservable variable must influence the selection process to question the causal inference of significant impact of land certification, are shown in the last column for the different empowerment indicators. For example, the critical level for overall women's empowerment index is 5.35-5.40, which implies that individuals that have similar covariates must differ in their odds of joint land certification by a factor of 5.40 to undermine the positive significant effect of certification on empowerment, which is less likely.

Generally, ATT estimates for the different indicators of empowerment are robust to unobservable covariates. Nevertheless, the extent of robustness to hidden bias varies considerably across the eight outcome variables. The results are least robust to possible presence of hidden bias for 'physical mobility and autonomy' and 'women's institutional inclusion'. However, this need not present a critical concern as the Rosenbaum bounds represent the most conservative scenario (DiPrete and Gangl, 2004).

Table 10: Robustness of the ATT estimates to unobserved heterogeneity

Matched	Joint certified women	Uncertified women	ATT ^a	Critical value of gamma(Γ) ^b
Overall participation in HH decisions	0.9154	0.8745	0.0409 (0.0311)	2.00-2.05
Participation in community activities	0.6367	0.4784	0.1583*** (0.0290)	2.90-2.95
Women's knowledge of their land rights and land related legal provisions	0.8997	0.7341	0.1656*** (0.0189)	9.95-10.00
Level of tenure security and dare to protect their land rights	0.9561	0.8803	0.0757*** (0.0193)	4.00-4.05
Physical mobility and autonomy	0.5511	0.4888	0.0623* (0.0324)	1.30-1.35
Women's institutional inclusion	0.3788	0.2327	0.1461*** (0.0438)	1.45-1.50
Overall women's empowerment index	0.8201	0.7134	0.1067*** (0.0167)	5.35-5.40
Experimental empowerment index ^c	0.6018	0.3519	0.2499*** (0.0750)	2.55-2.60

* Significant at 10%; ** significant at 5%; *** significant at 1%

^a Standard errors are bootstrapped 100 times. Standard errors are given in parenthesis.

^b Computed using rbounds program of Gangl (2004). The bounds given are only for positive self-selection case (Sig+).

^c The difference for this indicator is calculated based on data only for 244 households since the remaining are set to missing due to the index construction procedure.

3.6.3 Exogenous estimation of the propensity scores

Ideally, we would use baseline data to estimate propensity scores. However, we collected our data after the certification process was completed (in program kebeles). Accordingly, some of the variables we use in the matching stage may be endogenous – affected by the certification program itself. For example, empowered women may have access to local women’s associations to facilitate and encourage NGO presence in their village. To mitigate such concerns, we also estimate our ATTs using propensity scores estimated by a more parsimonious model containing only variables that are exogenous to certification status. Examples include land and livestock brought to current marriage, distance from main road and distance from parents.

The new matching results are reported in Table 11. ATT estimates of the different indicators remain statistically robust. Indeed, the ATT estimate for the indicator ‘*physical mobility and autonomy*’ now becomes significant at 5% significance level. Matching is done on the common support region and balancing tests for matching quality (not reported) are well satisfied. The ATT estimates are also robust to unobserved heterogeneity (last column).

Table 11: Average treatment effects and their robustness to unobserved heterogeneity

Matched	Certified women	Uncertified women	ATT ^a	Treated on common support	Control on common support	Critical value of gamma(Γ) ^b
Overall participation in HH decisions	0.9152	0.8798	0.0354 (0.0258)	131	184	2.00-2.05
Participation in community activities	0.6372	0.4807	0.1565*** (0.0307)	131	184	2.95-3.00
Women’s knowledge of their land rights and land related provisions	0.8990	0.7320	0.1670*** (0.0211)	131	184	10.10-10.20
Level of tenure security and dare to protect their land rights	0.9557	0.8807	0.0750*** (0.0172)	131	184	3.75-3.80
Physical mobility and autonomy	0.5534	0.4936	0.0598** (0.0293)	131	184	1.25-1.30
Women’s institutional inclusion	0.3791	0.2315	0.1476*** (0.0431)	131	184	1.50-1.55
Overall women’s empowerment index	0.8200	0.7151	0.1049*** (0.0181)	131	184	5.25-5.30
Experimental empowerment index ^c	0.6070	0.3239	0.2831*** (0.0587)	98	137	2.75-2.80

* Significant at 10%; ** significant at 5%; *** significant at 1%

Matching was done using epanechnikov kernel matching method using the psmatch2 Stata program of Leuven and Sianesi (2003).

^aStandard errors are bootstrapped 100 times. Standard errors are given in parenthesis.

^bComputed using rbounds program of Gangli (2004). The bounds given are only for positive self-selection case (Sig+).

^cThe difference for this indicator is calculated based on data only for 244 households since the remaining are set to missing due to the index construction procedure.

3.6.4 Propensity score weighted regression

PSM has been criticized for imposing the common support condition for credible counterfactual estimation, implying that individuals falling outside the region of the common support have to be disregarded—leading to loss of information. This information loss may have its own cost, especially if treatment effects are heterogeneous (Caliendo and Kopeinig, 2008). Another critique is that testing the statistical significance of treatment effects and computing their standard errors is not straightforward. The problem is that the estimated variance of the treatment effect includes the variance due to the estimation of the propensity score. That is why we rely on bootstrapping to obtain correct standard errors.

The recent literature suggests that propensity score weighted regressions can offer additional benefits as an alternative to matching. For example, Hirano et al. (2003) show that weighting with the inverse of the estimates of the propensity score leads to an efficient estimation of average treatment effects. The basic idea is that, although adjusting for differences in the propensity score removes all bias, matching on the propensity score need not be as efficient as adjusting for differences in all covariates. Propensity score weighted regression helps to address this concern, since it allows for controlling for all covariates in the estimation of treatment effects. By using the observations of the entire sample, it has also the benefit of attenuating the information loss due to the common support restriction. The weighting approach uses the estimated propensity scores, λ , to weight each control group observation (women without certificates). While each woman with a certificate is weighted with 1, each control group woman is weighted with the odds $\lambda/(1-\lambda)$.

In effect, this method assigns more weight to control group women that are more similar to certified women in terms of observable characteristics that are potentially correlated with both participation in the certification program and key outcomes, and assigns less weight to those that are less similar. With these weights applied, the result is that all observable characteristics are better balanced across the certified and control women groups, boosting the credibility of the counterfactual outcomes. A regression of outcomes on a treatment indicator using these weights then gives an estimate of the average treatment effect on the treated. We report treatment effects based on propensity score weighted regressions in Table 12. Essentially, the results remain robust and consistent with our results from the PSM.

Table 12: Propensity score weighted regression estimates

Matched	Treatment effects ^a	Observations
Overall participation in HH decisions	0.0422* (0.0231)	325
Participation in community activities	0.1564*** (0.0276)	325
Women's knowledge of their land rights and land related legal provisions	0.1565*** (0.0175)	325
Level of tenure security and dare to protect their land rights	0.0671*** (0.0162)	325
Physical mobility and autonomy	0.0598** (0.0285)	325
Women's institutional inclusion	0.1361*** (0.0363)	325
Overall women's empowerment index	0.1025*** (0.0142)	325
Experimental empowerment index ^b	0.2635*** (0.0577)	244

* Significant at 10%, ** significant at 5%, *** significant at 1%

^a Standard errors are in parentheses.

^bThe difference for this indicator is calculated based on data only for 244 households since the remaining are set to missing due to the index construction procedure.

3.6.5 Tobit model results

In this section, for illustrative purposes, we extend our analysis to a parametric econometric analysis. To economize on space, we use only the experimental empowerment index as our dependent variable for this part, which is continuous, and censored at zero and at one. This data structure is appropriate for Tobit model estimation, so we estimate the following standard Tobit model specification:

$$y_i^* = x_i' \beta + \varepsilon_i \quad (3)$$

$$y_i = \begin{cases} 0 & \text{if } y_i^* \leq 0 \\ y_i^* & \text{if } 1 < y_i^* < 1 \\ 1 & \text{if } y_i^* \geq 1 \end{cases} \quad i = 1, 2, 3, \dots, N$$

where N is the number of observations, y_i^* is a latent variable of censorship, x_i is a vector of explanatory variables, β is a vector of unknown coefficients, and ε_i is an error term assumed to have a normal distribution with zero mean and constant variance σ^2 .

In addition to being a robustness check to the PSM approach results, the Tobit regression serves two other purposes. First, it gives us an opportunity to expand the set of our covariates by including covariates that arguably affect only women bargaining power but not the likelihood of certification. We assess whether the inclusion of such covariates impairs the

positive effects of land certification on women's empowerment. We include two new explanatory variables: whether a woman has a marital contract, and whether she knew her husband before she married him. Second, decisions made together might differ from those made privately for reasons not directly related to preferences and relative bargaining power. The Tobit model permits us to test whether outcomes at the joint stage of the public goods game may be affected by other reasons (i.e., unrelated to preferences and bargaining power). We include a trend dummy indicating whether the private choice of the wife is higher than the private choice of her husband in the private parts of the game.

Table 13 presents the Tobit model regression results, which are broadly consistent with the outcomes of the non-parametric propensity score matching approach. Land certification status significantly affects women's empowerment as measured by the experimental index. Moreover, land and livestock brought to current marriage significantly empower women. Women who knew their husbands before they got married are also more likely to have greater bargaining power in household decisions. This is reasonable as women who directly participate in the marriage market have better opportunity to carefully select their partners and actively engage in household decision making than those women whose marriage is arranged by parents. Similarly, if a woman is older than her husband, she is more likely to influence household decisions. Note the trend dummy is not significant, indicating that outcomes of the resource allocation games reflect women's preferences and bargaining power.

Table 13: Tobit regression results

Explanatory Variables	Women's empowerment index ^a
Land certification status	1.5982*** (0.4180)
Age of woman	-0.0253 (0.0291)
Years of marriage	0.0031 (0.0275)
Couple age difference	0.0465* (0.0255)
Woman literacy	0.2030 (0.4058)
Husband literacy	-0.0593 (0.4106)
Children above 15 yrs.	0.5658 (0.4431)
Marriage contract	0.4433 (0.4805)
Know husband before marriage	1.2809*** (0.4758)
NGO presence	-0.1370 (0.8046)
Land brought to marriage	0.3619** (0.1642)
Livestock brought to marriage in tlu	0.5058** (0.2383)
Distance from main road	-0.0024 (0.0039)
Distance from parents	-0.0012 (0.0007)
Trend	0.1464 (0.3282)
Constant	0.0401 (1.1157)
Sigma	2.051 (0.2989)
Pseudo R ²	0.1018
Observations	244

* Significant at 10%; ** significant at 5%; *** significant at 1%

Observation summary: 112 left-censored observations at empowerment index ≤ 0 ; 40 uncensored observations and 92 right-censored observations at empowerment index ≥ 1 .

^a Standard errors are given in parenthesis.

3.7 Conclusions and discussion

In this paper, we bring together two key components of the current development policy agenda: female empowerment and joint land certification. We demonstrate these issues are logically and robustly inter-related. Our main result suggests that joint land certification has a positive and statistically significant impact on various (but not all) dimensions of women's empowerment. Specifically, we find that certified women are more likely to participate in community activities, be aware of their land rights and land related provisions, have a higher perceived level of tenure security, and are more willing to protect their land rights. We also document positive impacts in terms of the freedom of physical mobility, autonomy, and

institutional inclusion. Our 'overall women's empowerment' index, based on survey data, shows that certified women are more empowered than uncertified women. This is consistent with the results from a field experiment designed to measure bargaining power. We also document robust results using a parametric Tobit regression approach. Our results are largely robust to different robustness checks and sensitivity analyses, and consistent with theory. Of course, we acknowledge that attribution concerns may remain, as with other studies based on (non-randomized experimental) cross-section data. While we demonstrate to the extent possible that unobserved heterogeneity is unlikely to influence our results, we cannot completely rule out the possibility that some unobserved aspects of household organizations may account for part of the differences we find between certified and uncertified women.

The implications of the effects of joint land certification on women's empowerment can be far-reaching and welfare enhancing due to various multiplier effects within and beyond the household (e.g., Quisumbing and Maluccio, 2003; Allendorf, 2007; Bezabih et al., 2012; Martinez, 2013). Women tend to be more concerned with the well-being of other household members (Agarwal, 1997). Particularly, mothers tend to more strongly internalize children nutrition and education benefits. We will seek to test this issue empirically in future work. In the meanwhile, we note that earlier studies revealed strong effects of the certification program on investments in land productivity (e.g. Holden et al., 2009; Melesse and Bulte, 2015) and the functioning of land (rental) markets (Holden et al., 2011). We therefore tentatively conclude that joint land certification has both important equity and efficiency gains. Moreover, since the costs of implementation are modest, the Ethiopian case could be a useful example for other African countries.

Finally, it is important to note that close scrutiny of our empowerment indices suggests a continued cause for concern. Though statistically different, the actual figures of most empowerment indices remain very low (for both certified and uncertified women). Particularly, the levels of participation in community activities, physical mobility and autonomy, and institutional inclusion of both certified and uncertified women are low. This implies a lot remains to be done to fully empower women. Presumably complementary measures aiming at expanding the opportunity sets of women, like education, employment, health, credit access and market participation are important as well.

Appendix

Appendix A: Women's empowerment indicator categories

Variables	Joint certified women	Uncertified women	Difference
1. Overall participation in HH decisions	0.9196	0.8584	0.0612**
Crop choice	0.9078	0.8967	0.0111
Land allocation	0.9149	0.8696	0.0453
Land inheritance	0.9007	0.9022	-0.0015
Land rent-out	0.8511	0.8207	.0304
Land rent-in	0.8652	0.8641	0.0011
Plot responsible person	0.9078	0.8696	0.0382
Land investments and management	0.9220	0.8370	0.0850**
Construction of house on the land	0.9362	0.8804	0.0557*
Whether to use agricultural technology	0.9574	0.8804	0.0770**
Whether to save HH income	0.9220	0.8750	0.0470
Expenditure in large items	0.9362	0.8478	0.0883**
Expenditure in small items	0.9645	0.8804	0.0841***
Number of offspring to be born	0.9220	0.7663	0.1557***
Whether to send children to school	0.9787	0.9348	0.0439*
Whether to use contraceptive	0.9149	0.7717	0.1432***
Whether a sick HH member has to go to health center	0.9433	0.8967	0.0465
Whether to sell or buy animals	0.9078	0.8370	0.0708*
Labor supply decision	0.9007	0.8207	0.0801**
2. Participation in community activities	0.6429	0.4570	0.1858***
Frequency of participation in village meeting	0.7305	0.4783	0.2522***
Participation in voting	0.6099	0.3804	0.2295***
Participation in public information meeting	0.3404	0.1467	0.1937***
Participation in soil and water conservation	0.8582	0.5652	0.2929***
Participation in construction of public center	0.6667	0.5109	0.1558***
Participation in road maintaining	0.3830	0.2989	0.0841***
Participation in school, mosque or church building	.8511	.7228	0.1282***
3. Women's knowledge of their land rights and land related provisions	0.8990	0.7132	0.1858***
A person >18 and whose livelihood is in agriculture has the right to access land	0.7730	0.5489	0.2241***
Equal joint land certification	0.9504	0.7500	0.2004***
Equal rights upon divorce	0.9858	0.9239	0.0619**
Equal division of harvest from field upon divorce	0.8440	0.7446	0.0994**
The right to transfer land through inheritance	0.9716	0.8696	0.1021***
Women's equal land inheritance right	0.9645	0.8370	0.1276***

Joint Land Certification Programs and Women's Empowerment

Appendix A: continued			
Variables	Joint certified women	Uncertified women	Difference
Use right to land has no time limit	0.8440	0.4185	0.4255***
Illegal to sell land	0.9362	0.8370	0.0992***
HH head shall get the agreement of the family to lease-out land	0.9787	0.8315	0.1472***
A wife can deny her husband to rent out part of the HH land	0.9787	0.9239	0.0548**
Land rental contract should be reported to kebele	0.9787	0.7609	0.2179***
Land renter is responsible for rented land management	0.9574	0.8152	0.1422***
It is illegal to plant eucalyptus trees on farmland	0.7872	0.6304	0.1568***
Household shall lose its land if not conserved	0.8936	0.6793	0.2143***
Procedure how to litigate land conflicts at courts	0.5461	0.2663	0.2798***
A complaint has a right to appeal to higher courts	0.9574	0.7120	0.2455***
Land can be expropriated with full compensation	0.9362	0.5761	0.3601***
4. Level of tenure security and dare to protect their land rights	0.9560	0.8804	0.0756***
Confidence in the legal system to protect their land rights	1.0000	0.9130	0.0870***
Tenure security women	0.9787	0.8804	0.0983***
Dare to protect their land rights	0.8865	0.7717	0.1148***
Expect equal land share upon divorce	0.9291	0.8859	0.0432
Expect equal land share upon death of husband	0.9858	0.9511	0.0347*
5. Physical mobility and autonomy	0.5585	0.4647	0.0938***
Can decide by her own to travel and visit family and friends	0.3404	0.2772	0.0633
Can take her children to health clinic alone	0.8369	0.6685	0.1684***
Whether she need permission from her husband to attend meetings	0.2766	0.3315	-0.0549
Can go to market alone	0.7801	0.5815	0.1986***
6. Women's institutional inclusion	0.3830	0.1993	0.1837***
Member of local women's associations	0.4043	0.1957	0.2086***
Member of Iqqub (ROSCA)	0.3972	0.2717	0.1254**
Member of village development committee	0.3475	0.1304	0.2171***
7. Overall women's empowerment	0.8230	0.6951	0.1278***
Observations	141	184	

* Significant at 10%; ** significant at 5%; *** significant at 1%

Appendix B: Experimental instructions

Stage 1: Private part of the game

1.1 The wife and the husband are brought to a separate and private location within the premises of the experimental field by enumerators. Enumerators then carefully explain the public goods game to their participant as follows.

1.2 You are a part of a group that consists of four randomly chosen people including you. As each of you is guaranteed anonymity, none of you will know who the other people in your group are. Husband and wife are never in the same group.

1.3 Each group member will be given 15 Ethiopian Birr by us. Then, you can choose how much of the 15 Birrs you would like to keep for yourself and how much of it you would like to contribute to your group. The amount of money you choose to keep is simply yours to take right away. But we double the amount of money that you give to the group. It means that for each Birr that you give to the group, the group earns 2 Birrs. We collect the money earned by the group in a “common pool” bag.

1.4 All other group members make the same decision. All of you make the decision anonymously and we will not reveal your decision to anyone, including your spouse.

1.5 The sum that all group members contribute to the common pool will be divided equally amongst the four group members upon the completion of the task. Everybody will get the same amount from the common pool, no matter how much each contributes to the common pool. But nobody, including your spouse, will know the amount receive at the end of the game. Your total earnings will thus consist of two parts: the amount of money you choose to keep and your share from the common pool.

1.6 Then a number of trial rounds are performed until the participant clearly understands the rules of the game. After the trial rounds, the respondent is asked to explain the outcome of the game. If the explanation is correct, the actual game takes place, otherwise the enumerator explains the rules again until complete understanding.

1.7 After the private game is completed, enumerators record the information of the private the game including payments to group members. However, payments were not made until the second stages of the game were played.

Stage 2: Joint part of the game

2.1 Upon the completion of the private game, spouses are told that they are going to play the same game jointly and brought together.

2.2 New groups of 4 participants will be drawn, while the spouses will still not be in the same groups.

2.3 Each is again given a new endowment of 15 Birr to allocate between the private and common pool account of their respective groups. The only difference from the private part of the game is that now you have to decide upon the same contribution level and make the same contribution level to your respective groups.

2.4 All the other steps of the private part are then repeated and in the end the spouses negotiate to settle on an agreed-upon allocation of their endowment.

Joint Land Certification Programs and Women's Empowerment

2.5 After the completion of the second stage of the game, payments of the common pool (from both the private and joint parts) were made secretly to each participant.

Appendix C: T-test for explanatory variables between non-missing and missing groups

Variables	Non-missing	Missing	Difference
Age of woman	36.351	35.238	1.114
Years of marriage	19.988	18.800	1.188
Couple difference	-8.306	-7.288	-1.018
Woman literacy	.241	.325	-.084
Husband literacy	.759	.788	-.029
Children above 15 years old	.637	.550	.087*
NGO presence	.045	.050	0.005
Land brought to marriage	.518	.703	-.185
Livestock brought to marriage in thu	.456	.452	.004
Distance from parents	205.327	226.038	-20.711
Distance from main road	45.318	42.625	2.693

* Significant at 10%; ** significant at 5%; *** significant at 1%

Variable definitions

Variable	Definition
<i>Age of woman</i>	Age of a woman in years, rounded down to the last birthday
<i>Years of marriage</i>	The number of years since the spouses married
<i>Couple difference</i>	The difference between the age of the wife and husband
<i>Woman literacy</i>	Dummy for education of woman (1= read and write, otherwise 0)
<i>Husband literacy</i>	Dummy for education of husband (1= read and write, otherwise 0)
<i>Children above 15 years old</i>	Dummy for presence of child (son or daughter or both) above 15 years (1= if there is any, otherwise 0)
<i>NGO presence</i>	Dummy for NGO presence at village level (1=if exist, otherwise 0)
<i>Land brought to marriage</i>	Total hectares of land (either inherited or got as gift) brought to current marriage by the woman
<i>Livestock brought to marriage</i>	Total livestock measured in tropical livestock unit (tlu) brought to current marriage by the woman
<i>Distance from parents</i>	Distance from a woman's parents measured in minutes on her feet
<i>Distance from main road</i>	Distance from main road measured in minutes on her feet
<i>Marriage contracta</i>	Marriage contract of a woman with her current husband (1=written, 0=verbal or no contract)
<i>Know husbanda</i>	Whether the woman knew her husband before her marriage (1=if she knew him, 0=if not)
<i>Private kept wife</i>	Amount kept by the wife in the private game
<i>Private allocated wife</i>	Amount contributed by the wife to the common pool in the private game
<i>Expected average contribution by wife</i>	Wife's expectation of average contributions of other group members in the private game
<i>Private kept husband</i>	Amount kept by the husband in the private game
<i>Private allocated husband</i>	Amount contributed by the husband to the common pool in the private game
<i>Expected average contribution by husband</i>	Husband's expectation of average contributions of other group members in the private game
<i>Joint kept</i>	Amount kept in the joint game
<i>Joint allocated</i>	Amount contributed to the common pool in joint game
<i>Trend</i>	A dummy capturing the joint decision process and takes value 1 when the private choice of the wife is higher than the private choice of her husband, otherwise 0.
<i>Experimental empowerment index</i>	Women's bargaining power constructed based on the experiment data

Chapter 4

Formal Law and Customary Change

a Lab-in-Field Experiment in Ethiopia

Abstract

Do customary courts strategically adapt arbitration outcomes if they face increased competition by the formal law? Through a lab-in-field experiment with villagers and real customary judges in rural Ethiopia, we show that post-arbitration payouts to agents disfavored by the customary system are downwardly biased. Introducing a costly formal law reduces these biases and draws the decisions of customary judges significantly closer to the law. Contrarily, agents advantaged by the law do not exploit their increased bargaining power. Instead, they make offers that are less advantageous to themselves. In equilibrium, only a fraction of them make direct use of the formal law, which is coherent with the presence of social sanctioning against norm-deviating behavior. Our results suggest that local customary dispute resolution institutions may have a role to play in shifting preexisting customs towards a desired outcome. In areas where formal legal institutions have limited outreach and the custom is strong, most effects of increased competition between formal law and customary legal institutions may arise from changes in the latter, rather than from plaintiffs seeking justice under the rule of law.

Publication status: Cecchi, F. & Melesse, B. M. (2015). Formal Law and Customary Change: A Lab-in-Field Experiment in Ethiopia. Revised and Resubmitted at *Journal of Economic Behavior & Organization*.

4.1 Introduction

Formal laws play a marginal role in governing the lives of many African citizens, particularly those residing in rural areas (Chirayath et al., 2006). Instead, customary legal systems provide prompt, accessible and culturally coherent justice services (Wojkowska, 2006). Customary courts oversee and enforce customs and informal rules of behavior, typically taking into account local egalitarian and redistributive norms (Platteau, 2000). Their adherence to minimum standards of justice and human rights remains nevertheless disputed. Customary courts may persistently discriminate against the underprivileged, entrenching mechanisms that perpetuate local power structures (e.g. Ordioni, 2005; Asfaw and Satterfield, 2010; Pimentel, 2010). Local gender biases, for instance, may affect the distributional decisions of customary dispute resolution institutions (Asfaw and Satterfield, 2010). Understanding the effects of increased competitiveness of formal law in predominantly customary institutional environments is, therefore, central to achieving fair and functional legal systems—a primary driver of economic development (Acemoglu et al., 2001; Rodrik et al., 2004). Yet, data on extrajudicial and customary disputes are rarely available (Landeo et al., 2007), and the interaction between customary legal institutions and formal law has been subjected to little rigorous empirical analysis so far.

The work of Sandefur and Siddiqi (2013) in Liberia is a notable exception. They find that the demand for mediation by paralegals trained in formal law are greater for plaintiffs disadvantaged by the customary system, and that direct access to the formal law results in strong socioeconomic gains for the underprivileged. Increased competition of formal law may nonetheless also foster indirect changes to the norms enforced by customary dispute resolution institutions. These indirect effects are particularly salient if agents face strong disincentives to appeal to formal legal institutions—e.g., if customary norms are backed by credible social sanctioning against defection and appeal to alternative forums. If those disadvantaged by the customary system fear that overruling its decisions may be costly—thus continue complying with preexisting customs even when the law is individually preferred—much of the direct socioeconomic gains generated by the increased competition of formal law may be dissipated. However, indirect effects may be at play too, through changes in customary arbitration outcomes. Aldashev et al. (2012a, 2012b) provide clear theoretical predictions on the evolution of customary legal outcomes induced by the introduction or empowerment of formal laws. If customary authorities fear jurisdictional and

reputational erosion⁶⁰, they may strategically adapt arbitration outcomes in response to the introduction of a competing formal law.

In this paper, we empirically investigate the hypothesis by Aldashev et al. (2012a, 2012b). Through a lab-in-field experiment in rural Ethiopia—where controversies are habitually settled through customary courts—we study the effects of introducing a costly legal fallback, on the arbitration decisions of local customary judges and the behavior of plaintiffs. In West Gojjam, we randomly select 60 customary judges, known among the local Amhara people as *Shimagelle*, to rule over controversies born from an ultimatum game with outside option, played by 532 villagers. For a random subsample, we allow participants to further appeal the arbitration through a costly fixed law. While some studies have looked at the influence of extraneous factors on formal judicial rulings (e.g. Danziger et al., 2011), the relative scarcity of naturally occurring data on customary rulings has limited their analysis. By bringing the lab into the field (see List, 2007), this work is the first to bridge this gap, studying the arbitration decisions of real customary judges in Ethiopia.

In line with previous literature, we find evidence of significant arbitration bias against female participants, and in favor of plaintiffs known by the customary judge and advantaged by the custom. Our main finding is that introducing a legal fallback reduces such biases, and that customary arbitration outcomes are drawn significantly closer to the formal law. Contrarily, we find that agents favored by the law do not take direct advantage of the increased bargaining power offered by the legal fallback. Instead, they make offers that are less advantageous to themselves. In equilibrium, only a fraction of them make direct use of the formal law, and women are less likely to apply the law than men. These findings are coherent with the presence of social sanctioning against norm-deviating behavior, and complement the work of Sandefur and Siddiqi (2013) by highlighting the importance of indirect customary responses to the increased competitiveness of the formal law. The “threat of law” may induce significant gains for those disfavored by the custom, even if they do not actively seek justice under the rule of law.

The rest of the paper is organized as follows. Section 2 discusses the literature on formal law and customary institutions. Section 3 provides a brief account of the Ethiopian institutional context

⁶⁰ Jurisdiction erodes as plaintiffs begin to use formal courts instead of customary forums. In so far as customary judges face a positive utility in ruling over a controversy, jurisdictional erosion will reduce their utility. Moreover, customary judges may face an intrinsic disutility in seeing their decision overruled. Reputation thus erodes when plaintiffs reject the customary arbitration decision and appeal to formal legislation.

and legal system. Section 4 outlines the experimental design. Section 5 discusses the empirical strategy. Section 6 illustrates the results, and Section 7 concludes.

4.2 Customs, legal institutions, and the law

Legal institutions encompass both formal and informal structures, and are central to enforcing the “rules of the game” that govern everyday life (North, 1990). Formal legal institutions typically preside over written constraints, such as formal laws and constitutions—prescribed and enforced by exogenous legislative authorities. Customary legal institutions, instead, oversee the ‘codes of conduct, norms of behavior and conventions’ that take form in a particular social setting (North, 1990: 36). Yet, formal and informal legal systems are not necessarily mutually exclusive, and often coexist and overlap to a considerable extent. Legal pluralism is thus prevalent in numerous countries and regions worldwide, including large portions of sub-Saharan Africa (Merry, 1988; Bennett, 2006; Tamanaha, 2008).

The norms upheld by customary legal institutions typically represent fairness standards intended at maintaining peace and social cohesion, but may also result in systematic discrimination against certain disadvantaged categories. In either case, they may have substantial consequences on investment decisions and long-run growth (Platteau, 2009; Baland et al., 2011).⁶¹ Replacing undesired customary norms with the rule of law has proven to be a complex and daunting exercise (e.g. Andre and Platteau, 1998; Kuyu, 2005; Sacco, 2008). Formal legal institutions will not successfully replace incompatible or unwanted customary norms, unless they become a “focal point” of convergence in the expectations of agents (Basu, 2000; Aoki, 2001; Greif, 2006). In other words, in the presence of preexisting customary focal points, the fear of social punishment may inhibit the consolidation of formal laws that contradict such customs. In fact, norms influence behavior not only through internal incentives (e.g., guilt aversion, or a taste for moral virtuousness), but also through external ones (Polinsky and Shavell, 2007). People may be willing to punish non-compliers even if the punishment is costly and doesn’t yield direct private benefits (Fehr and Gächter, 2000). As a result, agents find it harder to deviate from norm-compliant behavior even when the law is individually preferred. This is especially true for rural communities, where social pressure and sanctioning are more pervasive, and the cost of social exclusion is greater (e.g. Crook, 2004; Gedzi, 2012).

⁶¹ Customary norms may also add to the uncertainty over property rights, in turn affecting the investment decisions of individuals. Goldstein and Udry (2008), for instance, find that competing claims and higher insecurity of tenure over specific plots cultivated by a given individual correspond to lower intensity of investments on those plots.

While certainly a source of institutional “stickiness”, customary legal institutions need not necessarily discourage institutional change (see Boettke et al., 2008). Instead, they are often crucial to the functioning of formal institutions—such as the legal system (Platteau, 2000; Aoki, 2001). Their interaction with statutory law has thus surfaced as a pivotal issue to effectively reach the objectives sought by legislators (Richman, 2012). Several theoretical studies have examined this interaction (Greif and Laitin, 2004; Helmke and Levitsky, 2004; Dixit, 2004; Dhillon and Rigolini, 2011; Aldashev et al., 2012a, 2012b). Typically, these studies focus on the behavior of agents, not institutions. Formal laws alter the net benefit of adhesion to prevailing norms— affecting the bargaining power and fallback position of economic agents—such that established patterns of behavior may evolve. Studies empirically investigating the empowerment of formal laws are somewhat less prevalent. Banerjee et al. (2002), for instance, study the effects of a land tenancy reform in West Bengal. They find that the increased bargaining power that the tenants acquire, once a legal fallback is introduced, has positive effects on productivity, whilst arbitrary evictions by landlords all but disappear.⁶²

Sandefur and Siddiqi (2013) propose a “forum shopping” model in which plaintiffs choose between the customary and formal systems based on rational tradeoffs. If plaintiffs face high entry costs to the formal legal system, they will bring their cases to customary forums even when these are systematically biased against them. After providing evidence of such barriers to access in Liberia, they investigate the outcomes of a randomized intervention that increases the competition between formal and customary law by offering complimentary mediation and advocacy services through community paralegals trained in the formal law. They find that female and ethnic minority plaintiffs—facing poor odds in the customary system—are more likely to adopt and to be satisfied with the paralegal service, and that the program increases average household wellbeing. Nonetheless, less than 10% of the recorded cases were brought to the attention of paralegals. The relatively low pick-up of complimentary paralegal services, even among self-selected paralegal clients⁶³, may be an indication of social sanctioning dynamics that create strong disincentives to appeal to any authority alternative to customary ones. Under such constraints much of the socioeconomic gains of increased outreach may be dissipated, unless increased competition of formal law produces indirect changes in customary arbitrations.

⁶² The theoretical model underpinning these findings, however, assumes that the legal innovation replaces a sort of institutional vacuum, in which “landlords wielded a lot of power within the village and were therefore able to intimidate tenants” (Banerjee et al., 2002: p. 242).

⁶³ Their sample stems from a sub-category of villagers that had self-selected into contacting the paralegals before the baseline. It represents a “snapshot of potential paralegal clients, and therefore not representative of Liberians as a whole” (Sandefur and Siddiqi, 2013: p. 28).

Aldashev et al. (2012a, 2012b) provide clear theoretical predictions on the evolution of customary legal outcomes induced by the introduction or empowerment of formal laws. Central to their argument is that not only agents but also customary institutions respond to incentive structures. In fact, if customary judges insist on imposing custom-compliant outcomes once the legal fallback is introduced, they do not only increase the likelihood of appeal to the formal law by unsatisfied plaintiffs but also the distance between the expected outcome and the custom. Provided the formal law is not too radical—i.e., it is not excessively costly for customary judges to deviate from the preferred arbitration outcome—increasing its competitiveness will thus shift the conflicting custom in the direction intended by the legislator. This can occur even without the use of the formal law in equilibrium—it need only be a threat. The law, therefore, not only provides a direct alternative to customary dispute resolution institutions; it also changes the custom, indirectly improving the welfare of the disadvantaged sections of the population (Aldashev et al., 2012a, 2012b). In this paper, we empirically investigate this hypothesis by observing the arbitration outcomes stemming from real customary judges in rural Ethiopia, half of which face the risk of seeing their decisions overruled by a fixed law. To this end, we formulate the following research questions:

RQ1: *Are post-arbitration payouts to agents disfavored by custom downwardly biased?*

RQ2: *Are biases against agents disfavored by the custom reduced by the formal law?*

RQ3: *Does increased competition by the formal law draw customary decisions closer to the law?*

RQ4: *Do disfavored agents take advantage of the increased bargaining power created by the law?*

4.3 Legal institutions in Ethiopia

Ethiopia is home to more than sixty customary legal systems (Donovan and Assefa, 2003). Statutory law was first introduced in the 1950s, and customary institutions remain very vibrant. The formal legal system is far from penetrating and, since the mid-nineties, the government is committed to recognize and preserve local customary dispute resolution authorities. Most Ethiopian ethnic groups have their own customary systems for dispute settlement and conflict resolution. Famous examples include the *Shimagelle* system of the Amhara, the *Gadaa* system of the Oromo and *Sharia* courts of Muslim communities. In this paper, we investigate Amhara's *Shimagelle* system, which contains elements of customary law practices that are very much in use across the whole country. This traditional institution can best be viewed in line with principles of

arbitration, where the arbitrators are mostly religious leaders and village elders who review existing evidence and arguments from both sides and issue a verdict to settle the case based on customary norms. The fact that customary judges are local elders and religious leaders ensures that arbitration outcomes are strongly embedded into community dynamics. This in turn favors the enforcement of deliberations, but also provides strong disincentives to use alternative forums, such as formal courts.

Typically, the adjudication proceeds focusing on narrowing of differences through negotiations, rather than through adversarial procedures. The verdict may vary depending on the nature and gravity of the dispute, ranging from a simple apology for petty disputes to blood money for homicide crimes. At the end of the dispute resolution, the restoration of prior relationships is marked through customary rituals or ceremonies to which both parties take part. Once the arbitrators have held their verdict, they closely monitor its enforcement. Nominally, arbitrators lack the coercive powers of the formal law to ensure compliance; however, they rely effectively on the presence of social pressure and sanctions to enforce their decisions. A party failing to abide by the outcomes will be considered as insulting the arbitrators and will be shunned by the community. Social sanctions—including ostracism by neighbors and friends—and loss of reputation soon follow.

The Shimagelles are expected to provide their services without an explicit fee. Yet, there are several individualistic as well as altruistic reasons for the engagement of customary judges in dispute settlement. First, they are motivated by social recognition. Among the Amhara, there is a tradition of holding a special funeral ceremony for those who are believed to have played an important role for their community. Elders often serve their community hoping that they will be mourned accordingly. Second, in such strictly hierarchal communities, arbitration is an age-ascribed role, that provides a mix of status and responsibility. Third, arbitrators are expected to ensure the welfare of both parties and, through that, the wellbeing of the community. Fourth, the vast majority of *Shimagelles* serve some sort of religious function too, and being recognized as a *Shimagelle* is a signal of piety and righteousness. While accounting for religious scriptures, local norms and customs, deliberations will, therefore, reflect the need of customary judges to maintain their reputation, reaffirm their social and moral standing, and minimize potential jurisdictional erosion.

The official status of Ethiopia's customary courts has not been unequivocal in the legal history of the country. Both the Imperial and *Derg* regimes opted for a centralized legal approach that did

not embrace legal pluralism. Only since 1995 does the state recognize customary legal systems (Gopal and Salim, 1998).⁶⁴ The jurisdiction of customary courts is accepted as long as both parties to the dispute give their consent to be heard at customary forums, and the verdicts thereof are in conformity with human rights provisions. The Constitution preserves the mandate to adjudicate criminal matters solely to the formal law, but customary courts enjoy *de facto* wider jurisdictions—spanning from petty offences, land tenure and inheritance issues to violent crimes and homicide. In many respects, Ethiopian customary institutions stand out as more functional and powerful than the formal law. The political stalemate that followed the country's 2005 general election, for example, was ultimately settled through the arbitration of selected *Shimagelles*.

The relative dominance of Ethiopian customary institutions is in fact related to the problems encountered by the formal legal system. First instance courts are available in every Woreda (district), averaging approximately 100,000 people per tribunal (Guttman et al., 2004). However, as is true for much of Africa, the formal legal system in Ethiopia is regularly reproached as dysfunctional and inaccessible to ordinary people (Gowok, 2008; Asfaw and Satterfield, 2010). Moreover, the formal legal system is limited in its capacity by lack of physical resources, personnel, infrastructure and inadequate information systems (Baker, 2013): access to formal justice can be a daunting task for the average Ethiopian. Formal law is often perceived as costly to access and punitive—rather than conciliatory—and regularly fails to deliver proper redress to aggrieved parties (Sandefur and Siddiqi, 2013). In contrast, customary courts provide prompt and accessible dispute resolution mechanisms. However, their capacity to produce unbiased sentencing is contested. Asfaw and Satterfield (2010) study land and property dispute settlements by *Shimagelles* in the Zeghie Peninsula, in the Amhara region. They find that the formal justice system is both inaccessible and dysfunctional, while customary arbitration outcomes are largely unfavorable to women. This in turn reinforces gender inequality and entrenches local power relations.

4.4 Experimental design

4.4.1 Sample and setting

We develop a multi-stage laboratory experiment involving 532 villagers and 60 real local *Shimagelles*. Participants belong to 18 *Kebele* (municipalities) of West Gojjam, in the Amhara region of Ethiopia. At each of the fifteen game-sessions, 4 customary judges were randomly

⁶⁴ The provisions of the Constitution embracing legal pluralism are provided under Articles 34(5) and 78(5).

selected from a pool of well-established local *Shimagelles*. The villagers belong to a list of randomly selected farmers responding to an agriculture-productivity related survey, administered in the same area in the previous year. Respondents are the main income earners (household heads), 92% are male, they average 43 years of age, and have slightly more than two completed years of formal education. Households are comprised of about 6 family members, and almost the entire sample is orthodox Christian. 82% of respondents take part in at least one informal safety net mechanism. Table 1 reports balance statistics for individual and experimental characteristics by treatment. No variable differs significantly at the 5% level, indicating a successful treatment randomization. The slight difference in sample size across treatments is due to the randomization at game-session level. This ensured that villagers and judges participating in the same game-session were all instructed on the same treatment, minimizing potential spillovers and confusion.

The experiment was comprised of four stages: (1) private investment; (2) joint venture; (3) ultimatum game with arbitration by a customary judge as outside option; and (4) the application of the formal law treatment—the only difference between the two treatment settings.⁶⁵

Table 1: Balance and summary statistics for the experimental samples

Variable	Customary only (N=252)	Customary + Law (N=280)	Diff.	Std. Err.
Male (%)	93.7	89.6	4.1	(2.40)*
Age	43.1	43.9	-0.8	(0.97)
Married (%)	92.5	91.1	1.4	(2.40)
Orthodox (%)	98.8	97.1	1.7	(1.24)
Education	2.6	2.3	0.3	(0.16)
Household size	6.4	6.2	0.2	(0.18)
Non-farm income	0.2	0.2	0.0	(0.03)
Informal safety nets (%)	81.3	84.3	3.0	(3.33)
First risk game (selected ball)	26.4	24.0	2.4	(2.31)
Second risk game (invested %)	46.4	46.6	0.2	(0.22)
Private Endowment	118.9	125.8	-6.7	(4.88)
Joint venture endowment	276.5	289.8	-13.3	(10.22)

Notes: Group means; t-test. Confidence: *** ↔ 99%, ** ↔ 95%, * ↔ 90%.

4.4.2 Private investments and joint venture

The first two stages only involved the villagers (from here onwards called agents), not the judges. In our setting, ultimatum offer rejections are a necessary premise to arbitration by customary judges. We make use of previous findings in economic literature to increase the likelihood of litigation in each anonymous pair, without biasing the randomness of relative endowments. Agents were randomly allotted an endowment of either 80 or 120 tokens⁶⁶, and made two

⁶⁵ Detailed procedures for the different stages are given in Table A1 (see Appendix).

⁶⁶ Each token was worth 0.5 Birr, resulting in endowments of USD 2.2 and 3.3 respectively, or around twice the average daily income in the area.

individual risky investment decisions that could increase or decrease the endowment. By having each agent make private investment choices, we made use of the so-called “earned endowment effect”: people exhibit more self-interested behavior in bargaining and sharing games when relative wealth is earned in some way rather than obtained through a pure windfall gain (e.g., Gantner et al., 2001; Cherry et al., 2002; Frohlich et al., 2004; Oxoby and Spraggon, 2008). The first risk game combined the design of Eckel and Grossman (2002) with that of Holt and Laury (2002).⁶⁷ The resulting endowment was then used in the second game, which followed Gneezy and Potters (1997). Agents decided how much of their endowment to invest in a lottery with 50% probability of doubling and 50% probability of halving the invested amount. By randomly assigning different initial endowments, and by playing two different investment games, we create sufficient exogenous variation to ensure that post-game private endowments are not correlated with risk preferences. This is especially important in the next stage, in which each participant is paired with another participant with a different private endowment.⁶⁸ To further strengthen a sense of ownership with respect to the private endowment (see Kahneman et al., 1990), agents were given a sleek and colorful endowment card reporting private earnings at the end of this stage.

The second stage of the experiment involved paired anonymous decisions, and took place on the following day. Participants to the previous stage were ranked according to their private endowments and the sample was split into half at the median. Agents were then assigned to a randomly selected anonymous partner with higher or lower private endowment. Each pair merged their individual endowments and jointly decided about the same risky investments mentioned above.⁶⁹ On average “higher investors” (i.e., participants contributing a higher share in the joint venture) contributed two thirds of the joint capital, while “lower investors” contributed the remaining third.⁷⁰ In other words, one agent typically invested twice as many stakes as their partner into the joint venture. We expect such endowment heterogeneity to lower ultimatum offers in the next stage (Cherry et al., 2005). Particularly, we expect high endowment agents to prefer

⁶⁷ Agents selected one out of eight balls, with exponentially increasing value from 1 to 128, as their winning prize. They then drew one ball from a bag containing all eight balls: if the extracted ball was worth at least as much as the selected ball, they won the value on the selected ball; if the extracted ball was worth less than the selected ball, they would lose the amount specified on the extracted ball (see Appendix Table A2).

⁶⁸ In other words, by design the risk propensity of each agent should not significantly influence the likelihood of being paired with a higher or a lower endowed partner (because of the random assignment of initial endowments). We test this through a Pearson’s Chi-squared test and find no evidence of significant correlation, with $p=0.611$ and $p=0.303$ respectively for the first and second game, with respect to being paired with someone with a higher or lower endowment. Individual risk preferences are thus not correlated with the relative size of the investment in the joint venture.

⁶⁹ The rationale of repeating choices in a paired setting follows that of intra-household bargaining literature (e.g. Bateman and Munro, 2005; Carlsson et al., 2012; He et al., 2012).

⁷⁰ “Higher Investors” into the joint ventures had an average private endowment of 165 tokens, while “Lower Investors” had on average 80 tokens.

offers proportional to investment shares, and low endowment individuals to prefer egalitarian redistributions (Rutström and Williams, 2000). Pairs could bargain on their investment choices for up to eighteen rounds, through oral messages collected and delivered by experimenters. On average, 4.5 counteroffers were made across the two games, with a maximum of fourteen bargaining rounds before reaching an investment agreement. Once the joint investments were completed, the next task was to split the final outcome through an ultimatum game—the source of our experimentally induced litigations.

4.4.3 *Ultimatum and arbitration*

In the third stage, pairs play an ultimatum game with outside option (see Güth et al., 1982; Schmitt, 2004). One agent was randomly selected to become an ultimatum sender, the other one becoming the receiver. The sender was asked to make an offer about how to split the joint endowment, which the receiver could either accept or reject. In case the offer was rejected, a randomly assigned local *Shimagelle* would mediate a resolution to the litigation. The customary judge independently studied the game history of each player and made an independent arbitration verdict. The verdict overruled the ultimatum offer, and imposed a new division of the joint venture capital.⁷¹ Importantly, experimenters exposed the name of the assigned customary judge before the ultimatum decision, and stressed that in case of arbitration the judge would be informed about the players' names. While joint venture partners remained anonymous to each other throughout the game, the arbitration process was thus non-anonymous (although strictly confidential). Because of this non-anonymity, we expect the experiment to reproduce disincentives to deviate from norm-compliant behavior, closely related to those faced by agents in their daily life.⁷²

Customary judges were paid a fixed fee for their participation in the game. They, therefore, did not have clear interests in any particular outcome of the arbitration. However, the non-anonymity of arbitration ensured that decisions of agents and customary judges were rooted into local reputational dynamics. We expect systematic biases in arbitration decisions to reflect the customary favoritisms present in our setting (e.g., biases against women and in favor of well-known plaintiffs). On top of this, we expect customary judges to enforce local egalitarian and redistributive norms. In this sense, random relative endowment heterogeneity offers an

⁷¹ Throughout the first and second stage of the experiment, customary judges were trained on understanding the game process and implications. The training only halted once each customary judge was able to individually explain the game procedure without external assistance.

⁷² Henrich et al. (2006) show that ultimatum bargaining behavior mirrors local egalitarian and redistributive norms. They observe that laboratory behavior is consistent with economic patterns of everyday life in several small-scale societies, and that community characteristics explain experimental patterns better than individual level variation.

experimentally generated source of “exogenous discrimination”: in the absence of a legal fallback, we expect arbitration decisions to disfavor higher inventors to the joint capital, in favor of lower investors (i.e. egalitarian redistribution).

4.4.4 The formal law

For about half of the game sessions, the game ended with the decision of the customary judge. The remaining sessions also included a fourth stage, consisting in the application of the formal law treatment. Agents could reject the arbitration outcome by appealing to a costly fixed law. At the cost of 10% of the final joint capital—a fictional measure of the costs related to formal legal litigation—the law divided the joint venture capital according to initial investment shares. This rule dates back to sixteenth century maritime trade expeditions, if not Hellenistic and Roman merchants (Benrud, 2009), and is known as liquidating dividend policy. It reflects a standard practice for joint-venture dissolution in several national legal systems, including that of Ethiopia.⁷³ It represents an alternative idea of fairness that may appeal higher investors that feel entitled to a higher portion of the joint capital. But it is also in sharp contrast with the concepts of distributive justice and egalitarianism that typically characterize rural communities (see Platteau, 2000). In other words, it allows us to study the effect of introducing a proportional split rule—also known as “liberal egalitarianism” (Cappelen et al., 2003)—in a context of “strictly egalitarian” norms (with biases), often enforced by the customary judges.⁷⁴

4.5 Empirical strategy

We investigate the research questions outlined at the end of section 4.2 as follows. First, we verify the presence of systematic bias against customarily disfavored agents and in favor of privileged ones, without a legal fallback (RQ1). Second, we look at these potential biases in the presence of a costly fixed law, expecting them to attenuate (RQ2). Third, we make use of the egalitarian redistribution in favor of (exogenously determined) lower investors within each joint venture to verify whether the competing formal law draws the arbitration decisions of customary judges closer to the law itself (RQ3). And fourth, we investigate the behavioral changes of agents in terms of ultimatum offers and rejections, with and without legal fallback (RQ4).

⁷³ The provisions of the Ethiopian Commercial Code referring to the dissolution of partnerships and joint ventures are provided under Articles 258(1) to 279(3).

⁷⁴ Strict egalitarianism requires that ‘all inequalities should be equalized’ (Cappelen et al., 2003: 818). In our context, it means that each partner in the joint venture would receive an equal share of the joint capital (in the absence of a legal fallback, this was the case in over 50% of the arbitrations). Instead, liberal egalitarianism accepts inequality as long as it stems from choices under individual control. In our context, it is closely represented by the fixed law, in which each partner receives a portion of joint capital proportional to their initial investment share.

In the absence of the legal fallback, we expect customary judges to discriminate against less “powerful” agents (e.g., women), and favor “embedded” agents (e.g., known plaintiffs). Also, we expect these biases to decrease once a non-discriminatory legal alternative is introduced. We test RQ1 and RQ2 by observing the arbitration outcomes relative to individual characteristics of agents, through the following regressions estimated under the two separate treatments:

$$Payout_i = \alpha + \beta Female_i + \gamma Age_i + \zeta Education_i \quad (1)$$

$$+ \theta HighInvestor_i + \kappa Known_i + \varepsilon_i$$

$$Payout_i = \alpha + \beta Female_i + \gamma Age_i + \zeta Education_i \quad (2)$$

$$+ \theta HighInvestor_i + \kappa Known_i + \lambda (Known_i \times HighInvestor_i) + \varepsilon_i$$

where $Payout_i$ refers to the post-arbitration payout of individual i , as a fraction of the overall mean payout, $Female_i$, Age_i and $Education_i$ indicate respectively the gender, age, and education level of the respondent, $HighInvestor_i$ takes value of 1 if the agent is the higher investor in the joint venture, and $Known_i$ is a dummy taking value of 1 if the agent is known to the customary judge assigned to the arbitration. Standard errors are clustered at the customary judge level. In the absence of a legal fallback, we expect biases against women ($\beta < 0$) and in favor of known plaintiffs ($\kappa > 0$), especially if they are favored by the strictly egalitarian custom (i.e., we expect known lower investors to be favored). Once the formal law is introduced, we expect such biases to attenuate.

As long as customary judges impose some sort of egalitarian redistribution in their arbitrations, as we expect, the experimental design exogenously imposes customary bias against half the agents—the higher investors in the joint venture.⁷⁵ We proceed to assess the robustness of the previous finding by testing whether there is exogenously imposed customary discrimination and whether it is reduced by the introduction of the legal fallback (RQ3). Assuming that the legal fallback is a credible threat in the eyes of arbitrators, the optimal strategy in our experiment would be to reallocate to lower investors the “legal cost” that higher investors would lose by applying the law, therefore redistributing in the direction of the strict egalitarian norm, but not beyond reserve utility of higher investors. We estimate a specification with only the law dummy as a regressor (OLS), and gradually include other individual level variables to reach the following full specification:

⁷⁵ Even though we consistently call the deviation of arbitration outcomes from a proportional split “customary bias”, we do not intend to assign a negative meaning to this. We are aware that this “bias” can be seen as a form of strictly egalitarian redistribution that is “hindered” by the introduction of a liberal egalitarian formal law. We return to this interpretation in the conclusions.

Formal Law and Customary Change

$$\begin{aligned} CustomaryBias_{is} = & \alpha + \beta Law_s + \gamma HighInvestor_{is} + \lambda(Law_s \times HighInvestor_{is}) \\ & + \zeta JVinequality_{is} + \theta UltimatumBias_{is} + \iota X'_{is} + \varepsilon_{is} \end{aligned} \quad (3)$$

where $CustomaryBias_{is}$ represents the deviation of each arbitration i from a proportional split in percentage points (i.e. the rule of law) for each arbitration⁷⁶, Law_s is a dummy taking value of 1 if the session included the formal law option, $HighInvestor_{is}$ is a dummy taking value of 1 if the ultimatum game sender is the higher investor in the joint venture, $JVinequality_{is}$ is the deviation of joint venture investment shares from 50-50 in percentage points, $UltimatumBias_{is}$ represents the ultimatum offer deviation from a proportional split in percentage points, and X'_{is} is a vector of other individual and game characteristics. Other notations have the same meaning as in (1). We expect $\beta < 0$.

Next, we look at the effect of introducing the legal fallback on the behavior of agents. Higher investors should benefit from the introduction of a law that imposes splits according to initial investment shares (RQ4). In terms of negotiation language, the introduction of the formal law increases the best alternative to a negotiated outcome (BATNA) of higher investors. This increases their bargaining power, even if that BATNA is not invoked in equilibrium. Whether they are willing to use the increased bargaining power deriving from the legal fallback depends, however, on the expected social cost of such action. Higher investors may forego the benefits of formal legislation in the presence of reputational concerns or expected social sanctioning. We measure such shift through the following fully specified equation:

$$\begin{aligned} UltimatumBias_{is} = & \alpha + \beta Law_s + \gamma HighInvestor_{is} + \lambda(Law_s \times HighInvestor_{is}) \\ & + \zeta JVinequality_{is} + \iota X'_{is} + \varepsilon_{is} \end{aligned} \quad (4)$$

where $UltimatumBias_{is}$ represents again the deviation of the ultimatum offer of sender i (N=266) from a proportional split in percentage points, and all other notations have the same meaning as before. If social sanctioning or the threat of it is not stringent, we expect $\beta < 0$ and particularly $\lambda < 0$. Yet, a sufficiently high social cost of deviation from the norm may reverse these expectations. In our setting, the identity of everyone's joint venture partner is kept anonymous, but the identity of agents is revealed to customary judges if the ultimatum offer is rejected—actions are observed by customary judges. If agents have reputational concerns particularly at heart, they may wish to signal their distaste for the law, and indicate their intention not to make use of it by making more strictly egalitarian ultimatum offers once the law is introduced. These concerns

⁷⁶ We regress arbitrations (129), rather than participant (258), as otherwise the same ruling would appear twice, for each individual in the pair.

would be particularly expected from agents that stand to gain from the law, i.e., higher investors, and not from those that would lose from it. In such case, $\beta = 0$ and $\lambda > 0$.

As additional controls to RQ4 we look at the actual use of the formal law by agents. If agents that stand to gain from the law face particularly strong concerns with respect to social sanctioning, we should expect them to make lesser use of the formal law—even when its outcome would be preferable to that of the customary arbitration. We estimate a linear probability model with gender as a regressor, and gradually include other variables to reach the following full specification:

$$\begin{aligned} LawEnforced_i = & \alpha + \beta Female_i + \gamma Age_i + \zeta Education_i \\ & + \theta HighInvestor_i + \kappa Known_i + \lambda (Known_i \times HighInvestor_i) + \varepsilon_i \end{aligned} \quad (5)$$

where $LawEnforced_i$ takes value of 1 if agent i applies the formal law, and all other notations have the same meaning as before. We expect $\beta = 0$ and $\theta > 0$ in the absence of reputational concerns and social sanctioning expectations, and $\beta < 0$ and $\theta = 0$ if they are present.

Finally, we investigate the probability of ultimatum offer rejection and application of the formal law. In fact, the absence of change in ultimatum offers need not to derive from the implicit costs of deviating from the custom. An alternative (heuristic) explanation would be that agents do not perceive the changes in the BATNA induced by the formal law, and make ultimatum offers according to some simple rule of thumb. Whether or not changes in the bargaining power are perceived by agents can be seen by looking at the ultimatum offers, but also the rejection rates in the ultimatum game. If agents disfavored by the formal law anticipate a shift in arbitration outcomes once the law is introduced, their willingness to accept an ultimatum offer may increase. On the other hand, it is possible that the legal fallback increases the prevalence of disputes, particularly if agents previously disfavored by the custom take advantage of their increased bargaining power (RQ4). We estimate a linear probability model with only the law dummy as a regressor, and gradually include other variables to reach the following full specification:

$$\begin{aligned} Dispute_{is} = & \alpha + \beta Law_s + \gamma HighInvestor_{is} + \lambda (Law_s \times HighInvestor_{is}) \\ & + \theta UltimatumBias_{is} + \zeta JVinequality_{is} + \iota X'_{is} + \varepsilon_{is} \end{aligned} \quad (6)$$

where $Dispute_{is}$ is a dummy taking value of 1 if the joint venture i resulted in an arbitration⁷⁷, and other notations have the same meaning as in (1) and (2). We expect $\gamma < 0$ and $\lambda > 0$ once the

⁷⁷ Again, regressing at the joint venture level (266), instead of the individual level (532), avoids double counting outcomes.

interaction term is introduced, and $\beta \neq 0$, where the sign of the coefficients depends on the dominating effect.

4.6 Results and discussion

In total our experimental setup induced 129 pairwise disputes arising from rejected ultimatum offers, involving 258 out of 532 villagers. 144 agents entered a controversy in the baseline with customary arbitration only, the remaining 114 being from the treatment with the additional possibility of appealing to a fixed law. First, we test the assumption that customary courts may disfavor less powerful agents (RQ1) (Asfaw and Satterfield, 2010; Sandefur and Siddiqi, 2013), and that biases against them may be reduced by the introduction of a formal legal fallback (RQ2) (Aldashev et al., 2012a, 2012b). We regress the 258 payouts resulting from arbitration by a customary judge, on the individual characteristics of agents (Table 2). In the absence of law, we find that post-arbitration payouts are more than 31% lower for female players (Table 2, column 1). Also, arbitration outcomes are significantly higher for known lower investors—favored by the strictly egalitarian customary norm (Table 2, column 2).

***Result 1:** Arbitration outcomes result in lower payout to women and higher payout to known plaintiffs with lower initial investment shares.*

Most importantly, these statistical differences disappear once the legal fallback is introduced. The post-arbitration payouts for agents participating to the treatment with the legal fallback are not lower for women and not higher for known plaintiffs (Table 2, columns 3 and 4). Column 5 in Table 2 presents the pooled regression, and clearly marks how the lower payout to women vanishes once the law is introduced. Moreover, column 6 presents the results of a Heckman selection model. The reason for presenting this is twofold. First, it allows to control for self-selection into litigation, and second, it presents the process of selection into litigation through the first stage. We find that our results are robust to the selection model and that the coefficients remain relatively stable. Following intuition, joint ventures with a higher difference between the lower and higher investor increase the probability of litigation, while ultimatum offers that are closer to a proportional split reduce it. The coefficient of the inverse Mill ratio is, however, statistically insignificant, meaning that there is no detectible self-selection into litigation and that it is safe to estimate the effects using OLS (such as in previous columns).

***Result 2:** Arbitration biases disappear once the legal fallback is introduced.*

Institutional Change and Economic Development

Table 2: The law reduces customary biases

	Payout relative to mean payout					
	Customary only		Customary + Law		Pooled	
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.313**	-0.289**	0.146	0.142	-0.289**	-0.310**
	(0.138)	(0.126)	(0.119)	(0.122)	(0.125)	(0.149)
Age	0.008	0.008	0.003	0.003	0.008	0.007**
	(0.005)	(0.005)	(0.004)	(0.004)	(0.005)	(0.003)
Education	-0.002	-0.001	-0.007	-0.008	-0.001	-0.007
	(0.022)	(0.021)	(0.016)	(0.017)	(0.020)	(0.022)
Higher investor	0.234***	0.701	0.557***	0.516***	0.701	0.690***
	(0.053)	(0.424)	(0.055)	(0.132)	(0.422)	(0.200)
Known	-0.062	0.169**	-0.009	-0.038	0.169**	0.175
	(0.172)	(0.079)	(0.073)	(0.055)	(0.079)	(0.141)
Higher investor × Known		-0.539		0.063	-0.539	-0.525**
		(0.454)		(0.202)	(0.451)	(0.215)
Law					0.121	0.122
					(0.314)	(0.304)
Law × Female					0.431**	-0.202
					(0.174)	(0.183)
Law × Age					-0.005	0.442**
					(0.006)	(0.196)
Law × Education					-0.006	-0.004
					(0.026)	(0.005)
Law × Higher investor					-0.186	-0.004
					(0.441)	(0.029)
Law × Known					-0.207**	-0.179
					(0.096)	(0.244)
Law × Higher investor × Known					0.602	0.596**
					(0.493)	(0.275)
First Stage (litigation = 1)						
JV inequality						0.064***
						(0.016)
Ultimatum bias						-0.019***
						(0.007)
Joint venture capital						0.000
						(0.001)
Joint venture profit						-0.102
						(0.295)
Number of counteroffers						0.041
						(0.031)
Winning risk game 1						0.017
						(0.120)
Winning risk game 2						-0.232
						(0.172)
Female						0.229
						(0.203)
Age						0.004
						(0.005)
Education						0.051
						(0.032)
Known						-0.190
						(0.139)
Law						-0.415***
						(0.117)
Mill's lambda						-0.123
						(0.148)
Observations	144	144	114	114	258	532
R ²	0.098	0.132	0.348	0.349	0.228	

Notes: OLS, standard errors clustered at the judge level (in parentheses). Confidence: *** ↔ 99%, ** ↔ 95%, * ↔ 90%.

Result 1 had been documented by previous literature both within and outside Ethiopia (Asfaw and Satterfield, 2010; Sandefur and Siddiqi, 2013). Result 2, instead, represents a novel finding. It indicates that customary dispute resolution institutions are capable of responding to increases in competitiveness by the formal law. In response to the introduction of a legal fallback, customary judges change their verdicts, reducing discrimination against agents disfavored by the custom. On the other hand, their capacity to enforce strict egalitarian norms is weakened by the liberal egalitarian law—payouts of lower investors are on average almost 56% lower, instead of 23%.

Result 2 is all the more salient as in our experiment agents disfavored by customary outcomes make limited direct use of the formal law. In fact, only 12 plaintiffs ultimately appealed to the rule of law. In other words, over 91% of the agents settled for a share of capital below that enforceable through the rule of law. Aldashev et al. (2012a, 2012b) predict that, at least partially, this is the consequence of customary change in the direction of the law. As they put it, ‘the “magnet” effect of the law is triggered by the preoccupation of village elders to maintain their authority and to retain enough potential claimants within the purview of their informal jurisdiction’ (Aldashev et al., 2012b: 193). We will return to this at the end of this section.

Next, we explicitly test if the legal fallback draws the decisions of customary judges significantly closer to the formal law (RQ3). Figure 1 shows that in the absence of the legal fallback over half of the arbitrations result in an egalitarian redistribution. Once the formal law is introduced very few arbitrations apply the strict egalitarian rule. We take the deviation of arbitration outcomes from a split proportional to investment shares (i.e. the law) as dependent variable. We find that arbitration outcomes are on average 10 percentage points closer to initial investment shares when the formal law is introduced (Table 3, column 1). This “magnet effect” is robust to introducing experimental and individual controls (Table 3, columns 2 to 4). Importantly, arbitration outcomes are significantly different from the proportional split rule imposed by the law ($p=0.000$), even when the law is introduced: on average, lower investors receive 36% of the joint capital (instead of 45% without the law), against an average initial investment of only 33%.

Result 3: Customary judges (partially) adapt arbitration outcomes in response to the formal law.

Institutional Change and Economic Development

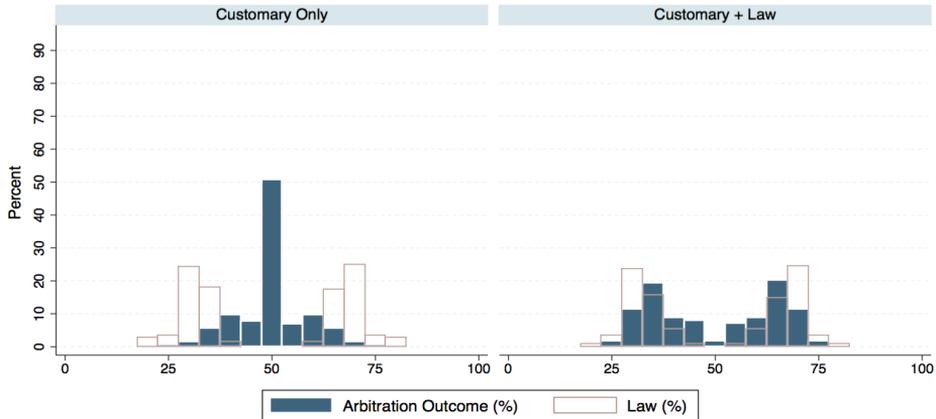


Figure 1: Post-arbitration outcomes vs. proportional splits (law)

Table 3. Customary judges adapt arbitrations in the presence of the legal fallback

	Arbitration: deviation from proportional split (law)			
	(1)	(2)	(3)	(4)
Law treatment	-0.102*** (0.012)	-0.107*** (0.015)	-0.105*** (0.015)	-0.104*** (0.017)
Higher investor sender		0.047** (0.020)	0.047** (0.020)	0.043* (0.022)
Law × Higher investor sender		0.011 (2.103)	0.011 (2.103)	0.011 (2.329)
Ultimatum offer deviation from law		0.316*** (0.097)	0.316*** (0.097)	0.274*** (0.094)
Joint venture inequality			-0.025 (0.149)	-0.081 (0.126)
Additional controls	No	No	No	Yes
Observations	129	129	129	129
R ²	0.475	0.560	0.560	0.591

OLS, standard errors clustered at the judge level (in parentheses). Additional controls in (4): Age, Education, Female, Joint venture capital, Joint venture profit, Number of counteroffers, Win risk game 1, Win risk game 2. *** ↔ 99%, ** ↔ 95%, * ↔ 90%.

Next, we proceed to investigate the behavior of agents across the two treatments (RQ4). Higher investors acquire greater bargaining power once the legal fallback is introduced, as the law grants them a portion of joint venture capital equivalent to the initial investment share. As a result, we could expect higher investors to make less egalitarian ultimatum offers. Yet, if they fear reputational loss and social sanctioning, they may not be willing to use their improved bargaining position, especially if their actions are observable. In our experiment, joint venture partners are unknown to each other, but the arbitration process is non-anonymous, and customary judges play the role of community “observers” of the actions of senders. In the presence of egalitarian customary norms that counter the law, those who stand to benefit from the law may wish to signal their intention not to make use of it, by making more egalitarian ultimatum offers once the law is

introduced. Figure 2 and Figure 3 show two different ways of representing the ultimatum offers made by higher and lower investors across the two treatments. Interestingly, the legal fallback increases the probability that higher investors make an equal split offer (Figure 2). Strictly egalitarian offers made by higher investors increase from 13.3% to 27.8% ($p=0.039$). Table 4 shows that on average the legal fallback does not significantly change the deviation of ultimatum offers from the law (Table 4, column 1). Higher investors make ultimatum offers that are on average 10 percentage points closer to initial investment shares compared to lower investors. Yet, once the law is introduced they increase the relative distance from the law by 3 percentage points (Table 4, columns 2 to 4). In other words, our results go in the opposite direction of what the changes in the BATNA would suggest. This leaves the net social welfare impact of the introduction of the formal law indeterminate. While the improvement in the arbitration outcomes of the customary authorities is a positive effect of the introduction of the formal law, the fact that higher investors make ultimatum offers that are less advantageous to themselves is an unintended effect.

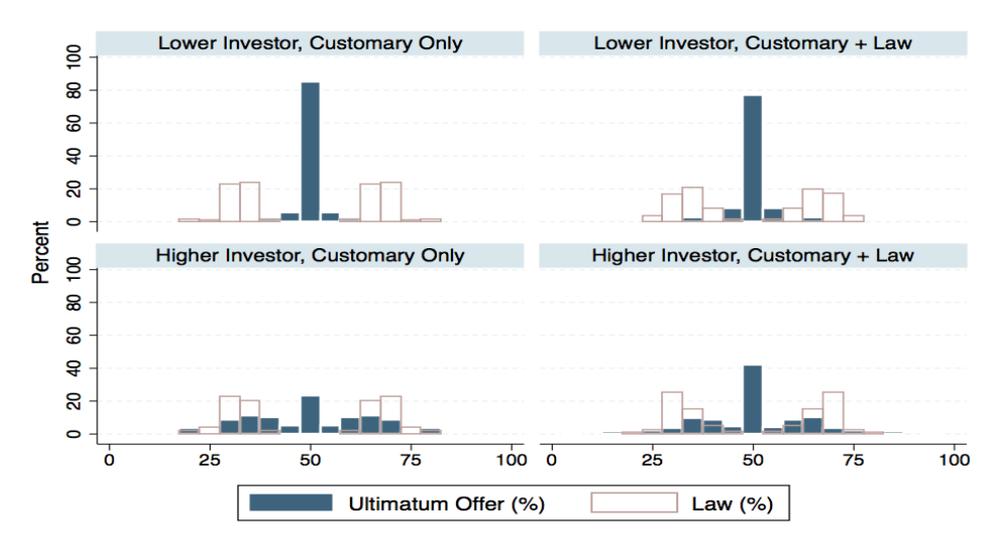


Figure 2: Ultimatum offers by investor type and treatment


Figure 3: Ultimatum offers by investor type and treatment

Table 4. Higher investor senders do not take advantage of the law

	Ultimatum: deviation from proportional split (law)			
	(1)	(2)	(3)	(4)
Law treatment	-0.007 (0.011)	-0.022 (0.014)	-0.011 (0.008)	-0.008 (0.010)
Higher investor sender		-0.097*** (0.012)	-0.103*** (0.012)	-0.099*** (0.012)
Law × Higher investor sender		0.039** (0.019)	0.037** (0.018)	0.032* (0.018)
Joint venture inequality			0.783*** (0.099)	0.793*** (0.095)
Additional controls	No	No	No	Yes
Observations	266	266	266	266
R ²	0.002	0.208	0.333	0.352

Notes: OLS, standard errors clustered at the judge level (in parentheses). Additional controls in (4): Age, Education, Female, Joint venture capital, Joint venture profit, Number of counteroffers, Win risk game 1, Win risk game 2. Confidence: *** ↔ 99%, ** ↔ 95%, * ↔ 90%.

The fact that higher investors make offers that are less advantageous to themselves once the formal law is introduced is coherent with the presence of reputational concerns and social sanctioning expectations. Higher investors may be signaling their intention not to make use of the law and instead comply to the egalitarian custom, even though the law favors them. To further probe this possible mechanism, we look at the likelihood of application of the law. If applying the law does not imply any social cost, it would be reasonable to expect higher investors to be significantly more likely to apply the law than lower investors. In fact, the latter are disfavored by the law, and may only apply it as a way to punish their joint venture partner—as the law removes 10% of the joint endowment. This is, however, costly for them too. Similarly, in the absence of

social costs we would expect women to apply the formal law at least as often as men, even if in equilibrium the customary biases against them are removed by the “threat of law”. Table 5 shows instead that higher investors are not more likely than lower investors to apply the law, and that women are 12% less likely to make use of the formal law than men.

Result 4: The formal law does not alter ultimatum offers. Disfavored senders (higher investors) actually make offers less advantageous to themselves and women are less likely to apply the formal law than men.

Table 5: Agents disfavored by the custom are not more likely to enforce the law than those disadvantaged by it

	Law enforcement (=1)			
	(1)	(2)	(3)	(4)
Female	-0.138*** (0.044)	-0.141*** (0.046)	-0.128*** (0.047)	-0.129** (0.062)
Age	-0.001 (0.002)	-0.000 (0.002)	-0.000 (0.002)	0.001 (0.003)
Education	-0.015* (0.008)	-0.014* (0.008)	-0.014* (0.008)	-0.011 (0.009)
Higher investor		-0.073 (0.057)	-0.038 (0.089)	-0.043 (0.091)
Known			0.079 (0.091)	0.115 (0.091)
Higher investor × Known			-0.061 (0.117)	-0.064 (0.110)
Additional controls	No	No	No	Yes
Observations	114	114	114	114
R ²	0.025	0.039	0.047	0.116

Notes: OLS, robust standard errors (in parentheses). Additional controls in (4): Age, Education, Female, Joint venture capital, Joint venture profit, Number of counteroffers, Win risk game 1, Win risk game 2. Confidence: *** ↔ 99%, ** ↔ 95%, * ↔ 90%.

Finally, we look at the behavior of ultimatum receivers. If ultimatum receivers anticipate the shift in arbitration outcomes, their willingness to accept such offers may change even if senders do not take direct advantage of the legal fallback. Particularly, lower investors may not want to frustrate higher investors by refusing partially redistributive offers, motivating them to apply the costly law. We find that receivers reject significantly less ultimatum offers in the presence of the formal law (Table 6, column 1). This result is driven entirely by reduced rejection rates of lower investors, previously favored by the egalitarian custom, with no significant variation in rejection rates attributable to higher investor receivers (Table 6, columns 2 to 4).

Result 5: The likelihood of litigation is reduced by the legal fallback, but only that of customarily favored agents (lower investors).

Table 6: The legal fallback reduces the ultimatum offer rejections by lower investors

	Ultimatum rejections (=1)			
	(1)	(2)	(3)	(4)
Law treatment	-0.164** (0.064)	-0.280*** (0.086)	-0.258*** (0.080)	-0.263*** (0.081)
Higher investor receiver		-0.023 (0.078)	-0.006 (0.076)	0.003 (0.080)
Law × Higher investor receiver		0.260** (0.119)	0.267** (0.118)	0.261** (0.120)
Joint venture inequality			0.022*** (0.008)	0.021*** (0.008)
Additional Controls	No	No	No	Yes
Observations	266	266	266	266
R ²	0.027	0.056	0.085	0.106

Notes: LPM, standard errors clustered at the judge level (in parentheses). Additional controls in (4): Age, Education, Female, Joint venture capital, Joint venture profit, Number of counteroffers, Win risk game 1, Win risk game 2. Confidence: *** ↔ 99%, ** ↔ 95%, * ↔ 90%.

4.7 Conclusions

This paper builds upon the work of Sandefur and Siddiqi (2013), and explicitly investigates the indirect effects that increased competitiveness of formal law may have on customary arbitration outcomes. Similar to them, we find that customary arbitration outcomes systematically discriminate against women, and favor well-embedded agents. We contribute to the literature on the interaction between customary institutions and formal law, by showing that not only agents but also customary judges respond to incentive structures. We do so by observing the ultimatum game decision of local villagers in rural Ethiopia, as well as the arbitration choices of real local customary judges, ruling over controversies born from the ultimatum game itself. Introducing a costly legal fallback reduces arbitration biases and draws the decisions of customary judges significantly closer to the formal law. In contrast, agents favored by the law do not take direct advantage of the increased BATNA offered by the legal fallback. Instead, they make offers that are less advantageous to themselves. In equilibrium, only a fraction of them make direct use of the formal law and women are less likely to apply the law than men.

If the formal legislation does not depart too radically from the custom (Aldashev et al., 2012a), customary dispute resolution institutions may have a role to play in shifting preexisting norms towards the desired outcome: in areas where formal legal institutions have limited outreach and the custom is dominant, most socioeconomic gains of increased competition between formal law and customary institutions may arise from changes in the latter, and not from plaintiffs seeking justice under the rule of law. Legislators should not overlook the potential contribution of customary legal institutions to change the custom. On the other hand, our experimental setting shows that formal legislation may limit the redistributive functions of customary legal institutions that enforce local egalitarian norms.

Appendix

Table A1: Experimental procedure

Stage 0

0.1 Each participant is assigned a random endowment of either 80 or 120 tokens.

Stage 1

1.1 Each participant is brought to a private location within the premises of the experimental field by an enumerator. Enumerators explain the first risk game to their participant, handing in the endowment of either 80 or 120 tokens.

1.2 The enumerator explains the rule of the risk game until the respondent is able to convey them back correctly. Then a trial round is performed.

1.3 After the trial round, the respondent is asked to explain the outcome of the ‘investment’. If the explanation is correct, the actual game can take place, otherwise the enumerator explains the rules again until full understanding.

1.4 Once the first risk game is completed, the enumerator states the new endowment and proceeds with the second risk game. Steps 1.2 and 1.3 are thus repeated.

1.5 Once the second risk game is completed, the enumerator states the final private endowment of the participant. Each enumerator writes the amount on a ‘private endowment card’, signs it and hands it to the participant. The enumerator explains that the card represents the tokens obtained by the participant during the experiment, and states the value of the endowment in local currency. Once the participant has understood the value of the card, the enumerator seals it into a closed envelope.

1.6 At the end of the first stage, participants are reminded that their choices and earning are anonymous and that their participation to the next stages is tied to the redelivery of their sealed envelope the next day. Participants can go home.

Stage 2

2.1 The next day, each participant is randomly paired with another participant with higher or lower endowment. Participants within each pair are not known to each other and are brought to separate private locations by different enumerators. Once there, enumerators explain the rules of the joint-venture stage, separately to both participants.

2.2 Participants are told that they will merge their private endowment with an anonymous partner, and perform the same risk games as the previous time. This time, however, they will need to compromise on the investment choice. Enumerators state the value of the joint endowment and remind each participant about their private share within the joint venture.

2.3 Pairs bargain over the same risk games of step 1.1 to 1.4. For each risk game, enumerators record and deliver messages between the joint-venture partners, to reach a consensus over the investment choice. If a consensus is not reached by bargaining round 9 of each game, the enumerator may take the average of the two proposals, rounded down to the nearest available option, as investment choice.

2.4 After the two risk games are completed, the enumerator states the final joint-venture endowment of the pair. Each enumerator writes the amount on two ‘joint endowment cards’, signs them and hands them in

Institutional Change and Economic Development

to each participant in the pair. The enumerator explains that that card represents the total tokens obtained by the pair during the experiment.

Stage 3

3.1 Each enumerator randomly assigns one participant to be the sender, and one to be the receiver in an ultimatum game. Enumerators explain the rules of the game, and state that the game will determine which portion of the joint endowment will be theirs to take home. They state that if the receiver does not accept the offer of the sender, the litigation will be sent to a local customary judge.

3.2 Prior to making the ultimatum offer, both participants are informed about the name of the customary judge that would rule over the controversy in case the ultimatum offer is rejected. They are asked to state whether they know that customary judge, and are told that in case an arbitrage is required, the judge will be informed of their name and game history.

3.3 Throughout Stage 2, four customary judges have been instructed on the rules of the game. Each judge is asked to explain the rules of the two risk games, as well as the joint venture stage. If the explanation is not correct, the enumerator explains the rules again until full understanding. Before presiding an ultimatum game arbitrage, customary judges are asked to sign an informed consent and confidentiality notice, requiring them to maintain the anonymity of game participants.

3.4 If the ultimatum offer is accepted by the receiver, the game terminates. Otherwise, the selected customary judge receives a game information sheet containing the names and game history of both participants, and independently reaches a final verdict.

Stage 4 (only for the ‘Customary + Law’ treatment)

4.1 Participants to the ‘+ Law’ treatment are informed of the legal fallback during step 3.1. Similarly, judges are informed during step 3.3. Enumerators explain that the arbitration verdict can be overruled by a costly fixed law: at the expense of 10% of the joint endowment, the law imposes a split according to initial investment shares.

4.2 After the customary judge has emitted the verdict, each participant to the ‘+ Law’ treatment is asked whether they accept or reject the verdict. If at least one pair member rejects the verdict, the costly fixed law is applied.

Table A2: First risk game lottery choices

Selected ball value	Winning probability	Highest possible loss	Standard deviation	Expected value	N
1	8/8	0	0.0	1.0	9
2	7/8	1	1.0	1.6	15
4	6/8	2	2.4	2.6	39
8	5/8	4	5.1	4.1	130
16	4/8	8	10.1	6.1	143
32	3/8	16	19.0	8.1	125
64	2/8	32	33.6	8.1	51
128	1/8	64	52.3	0.1	20

Notes: Winning values are in tokens. Each lottery comprises 8 balls with value $B_n = 2^{n-1}$. Agents select the desired winning ball value B_s . If the ball extracted $B_e \geq B_s$, players win B_s , if $B_e < B_s$, they lose B_e .

Variable definitions

Female: A dummy variable taking value of 1 if the respondent is female, 0 otherwise

Age: Age of respondent i in years, rounded down to the last birthday

Married: A dummy variable taking value of 1 if the respondent is currently married, 0 if otherwise

Orthodox: A dummy variable taking value of 1 if the respondent is Orthodox Christian, 0 if otherwise

Education: A variable indicating the number of completed years of education of respondent i

Household size: The number of people sharing the same roof and sharing the same pot

Non-farm income: The estimated percentage of yearly household income not deriving from farm activities

Informal safety nets: A dummy variable taking value of 1 if the participant belongs to at least one informal institution between Debo, Eqqub, and Iddir⁷⁸, 0 otherwise

First risk game: The value of the selected ball in the first risk game

Second risk game: The fraction of endowment chosen for investment in the second risk game

Private Endowment: The value (tokens) of the endowment possessed by participant i at the end of the “private investment” stage

Joint venture endowment: The value (tokens) of the endowment jointly possessed by each pair of joint venture participants at the end of the “joint venture” stage

Sender: The participant that makes an ultimatum offer to the receiver, with respect to how to split the joint endowment

Receiver: The participant that can accept or reject the ultimatum offer made by the sender

Law treatment: A dummy variable taking value of 1 if the game session included both the customary judge arbitration and the proportional law as outside options to the ultimatum game

Higher investor: A dummy taking value of 1 if respondent i has a higher than 50% share of the joint venture capital

Arbitration offer deviation from law: The deviation of the arbitration decision from a proportional split (the law), in percentage terms with respect to the joint endowment

Ultimatum offer deviation from law: The deviation of the ultimatum offer made by the sender from a proportional split (the law), in percentage terms with respect to the joint endowment

Joint venture inequality: The ratio between the private endowment of the lower investor and the private endowment of the higher investor in each joint venture

⁷⁸ Debo is a local labor exchange arrangement; Eqqub is a rotating savings and credit association, and Iddir is a funeral association functioning as informal insurance arrangement.

Chapter 5

Market Exposure and Risk Aversion

Evidence from Landed Farm Households in Ethiopia

Abstract

Risk preferences are important drivers of many relevant economic decisions of farm households. Particularly, high risk aversion has been identified as a potential trigger of ‘poverty traps’ for farm households in developing countries. Understanding whether risk preferences can be changed is an important prerequisite to design policies aiming to increase the productivity and efficiency of farm households. This paper analyzes the effect of market exposure on risk preferences for a relatively large sample of landed farm households in Ethiopia. We measure risk preferences using experimental and survey data, and relate them to actual market experience of the households. We use an instrumental variable approach to address the issue of endogeneity due to selection into trade. We find that market exposure attenuates risk aversion—farm households with a high degree of market exposure are more risk tolerant. The results are robust to using alternative specifications, controlling for internal mobility and out-migration, and for violations to rational choice.

Publication status: Melesse, B. M. & Cecchi, F. (2015). Market Exposure and Risk Aversion: Evidence from Landed Farm Households in Ethiopia. Working paper.

5.1 Introduction

Risk preferences play an important role in virtually all economic decision of farm households, like the adoption of improved agricultural technologies, production of different crops, and long-term investments. Poor farm households in developing countries are often characterized by a high degree of risk aversion (Binswanger, 1980; Dercon, 1996; Wik et al., 2004; Humphrey and Verschoor, 2004; Harrison et al., 2010; Franken et al., 2014). Their allocation decisions typically reflect not only highest expected value opportunities, but also algorithms that minimize risk exposure. Accordingly, high expected value opportunities often go unrealized. Ample empirical evidence shows that risk-averse farmers are reluctant to adopt innovations, such as improved agricultural technologies (Lamb, 2003; Alem et al., 2010; Fafchamps, 2010). Farmers may thus be locked into persistent risk-induced poverty traps (Dercon and Christiaensen, 2011).

To date, much of the empirical research on risk in developing countries has focused predominantly on the characterization of risk preferences of farm households and their coping strategies (for a review and more discussion, see Dercon, 2008). While the endeavor of measuring and characterizing individual risk preferences is certainly useful, understanding the drivers of changes in risk preferences is perhaps more relevant from a policy perspective. Traditionally, interventions aiming to change the high risk aversion nature of farm households have focused mainly on affecting their decision incentives and constraints. Facilitating farm households' access to insurance markets has typically been the common approach of such interventions. The basic idea is that insurance markets can help farm households to 'sell' and transfer risks, correspondingly relaxing their constraints. Evidently, a small but emerging experimental literature has reported results that indicate access to insurance markets induces farm households to take more risk, allocating more resources to high-return but high-risk agricultural investment and production choices (e.g., Leiva and Skees, 2008; Karlan et al., 2014).

While indisputably valuable, these findings reflect changes in behavior, but do not address the issue of risk preferences *per se*. Economic literature has long suggested that individual preferences are endogenous (e.g., Stigler and Becker, 1977; Becker and Mulligan, 1997; Netzer, 2009), and that markets can directly affect preferences (e.g., Bowles, 1998; Henrich et al., 2010; Falk and Szech, 2013). In this paper, we investigate whether exposure to market induces changes in risk preferences of landed farm households in Ethiopia. In theory, there

are sound arguments for why market exposure may change risk preferences. For example, market exposure may constitute a natural ‘shock’ that provides new experience to induce fundamental changes in risk preferences (Becker and Mulligan, 1997; Bowles, 1998; Netzer, 2009). Empirical evidence also supports this claim and shows that risk preferences evolve. This evidence comes from diverse domains of ‘shocks’, including conflicts or violent trauma (Voors et al., 2012; Callen et al., 2014), and natural disasters (Eckel et al., 2009; Cameron and Shah, 2013; Hanaoka et al., 2014).

The negative correlation between market exposure and risk aversion has been documented in a variety of settings. For example, Fellner and Maciejovsky (2007) find that high risk aversion and low levels of market activity are systematically related, while Franken et al. (2014) show that risk aversion is associated with the choice of more conservative marketing alternatives among American agricultural producers. While suggestive, these studies do not provide conclusive evidence about the causality underpinning such relationships. Arguably, market exposure may be endogenous to risk preferences—more risk-averse agents may self-select out of market exchanges, and vice versa.

We extend the literature beyond documenting simple correlations, and address the issue of causality. Particularly, we make use of the low-mobility of Ethiopian farm households with usufruct land rights to instrument market exposure with distance to nearest daily markets. Our research site, characterized by self-subsistence and extremely limited integration into markets, provides a distinct setting for such study. We employ both context-free experimental and context-specific survey measures of risk preferences. The field experiments provide us an objective and incentive compatible measure of risk preferences with real payoffs at stake. Our risk preference measures from the survey place the analysis in the context of real market decision environment, and facilitate the generalizability of our results. Market data were obtained from surveys of actual marketing behavior of farm households.

Our results reveal that market exposure is a key determinant of risk preferences of farm households. Precisely, market exposure has a robust negative effect on risk aversion of farm households. Next to self-selection into trade, we thoroughly address the issues of internal mobility and selective out-migration, which could make establishing causality difficult in studies of this type. Our finding has important implications for policy interventions aiming to spur rural development in poor countries. If market exposure influences individual risk

preferences, the penetration of markets in remote areas may be an important channel to reduce the productivity lag of rural households, directly helping them exit the poverty trap.

The remainder of the paper proceeds in the following order. Section 2 rationalizes the nexus between market exposure and risk preferences. Section 3 summarizes our data and identification strategy. We discuss and interpret our results in Section 4. Section 5 subjects our result to a series of sensitivity checks to probe their robustness, while Section 5 briefly presents concluding remarks.

5.2 Risk preferences and market exposure

Recent years have witnessed a paradigm shift on the role of markets—from a simple neutral means to allocate resources to active mechanisms that also mediate and influence behavior and preferences of market participants. Different authors have documented empirical evidence suggesting that markets play an important role in shaping and influencing preferences in several ways. For example, Henrich et al. (2010) find that exposure to markets may promote fairness, cooperation, and the propensity to trust others. Similarly, Falk and Szech (2013) show that markets influence deeper behaviors of participants.

Market exposure may also affect risk preferences. Recent evidence suggests that people's risk preferences can change over time.⁷⁹ As long as risk preferences change in response to individual experience, there are good reasons to believe that market exposure may play a particularly important role in shaping and influencing individual preferences towards risk and uncertainty. Markets stimulate signaling and adaptive learning (Bowles, 1998; Spence, 2002; Denrell, 2007), mitigate bounded rationality and asymmetric information (Slembeck and Tyran, 2004), and encourage impersonal transactions (Fafchamps, 2011) through impersonal institutions (Kumar and Matsusaka, 2009). Moreover, experimental studies in different settings report that market experience facilitates the development of rational choice (List and Millimet, 2008; Cecchi and Bulte, 2013).

The effects of market exposure on risk preferences may also explain assertions that are claimed to imply irrational choices in the literature. For example, List (2003, 2011) shows that market experience plays a significant role in eliminating behavioral anomalies like the 'endowment effect'—the fact that individuals demand much more to give up a good than they

⁷⁹ For example, Cameron and Shah (2013) report an increase in risk aversion after exposure to natural disasters in Indonesia. On the other hand, Voors et al. (2012) and Hanaoka et al. (2014) document a decrease in risk aversion (i.e., increased risk tolerance) after exposure to civil conflicts in Burundi and natural disasters in Japan, respectively. Eckel et al. (2009) also report a decrease in risk aversion after exposure to Hurricane Katrina.

would pay to acquire it. Such seemingly anomaly in behavior may look less contradictory when risk aversion is taken into account. Without assuming perfect information, risk-averse agents may discount the prospect that the alternative may not be as valuable as expected. In other words, the uncertainty about the true returns of an unknown alternative may hinder its uptake (Conley and Udry, 2010). The difference between willingness to accept (WTA) and willingness to pay (WTP) may thus wash out when both are weighted on the costs associated to the (perceived) risks. If so, it is possible that differences in WTA/WTP ratios of more and less market exposed agents reflect changes in risk preferences induced by market experience. In fact, List (2004) finds that agents with intense market experience behave largely in accordance with risk-neutral neoclassical predictions, whilst inexperienced agents tend to make risk-averse choices that are better explained through prospect theory (Tversky and Kahneman, 1979).

5.3 Data and empirical strategy

5.3.1 Setting

We conduct a relatively simple experiment involving 179 villages in the fall of 2013. The villages were randomly selected from the districts of Womberma, Bure and Jabitehinan of the Amhara region, Ethiopia. 532 farm households were in turn randomly sampled from these villages to participate in the experiment. Our subjects were the main income earners in the family (household heads). These individuals often assume the primary responsibility for making important economic decisions, and are more likely to confront major risky choices of their respective households on a continual basis—their livelihood depends on how well they deal with risks. All the households in this study possess a usufruct title on the land they farm. In Ethiopia, land use rights are legally bound to permanent physical residence in the community (Deininger and Jin, 2006), which reduces the mobility of landed household systematically. This issue will be discussed in greater detail in Section 5.

5.3.2 Measuring risk preferences

To facilitate understanding, our experiment builds on Gneezy and Potters (1997)—the simplest “risk game” used in the experimental literature (for a review and more discussion, see Charness et al., 2013). Table A1 in the appendix provides the detailed procedure of the game. Subjects were allotted an initial endowment of US\$ 2.8, or around twice the average

daily income in the area.⁸⁰ Subjects were subsequently asked to decide on how much of their endowment to invest in a lottery with 50% probability of doubling and 50% probability of halving the invested amount. The design of the experiment was in such a way that an increase in expected returns could be procured only by increasing risk of losing a large amount. While the literature commonly interprets experimental play as reflecting underlying preferences (e.g., Henrich et al., 2001; Tanaka et al., 2010), we seek to adopt the simplest risk game used in the literature to attenuate the likelihood of capturing other confounding constructs, like expectations, beliefs and shifts in the constraints of our subjects. The experiment was also completely decontextualized and participants did not face any information constraint and uncertainty related to the parameters of the game. Following this, we interpret our risk measure as reflecting risk preferences of participants.

Great effort was made to ensure that participants understood the details of the experiment. Subjects were fully informed about the objective probabilities of winning and losing, and about the corresponding size of gains and losses, using several examples before the actual experiments. Subjects were made to perceive the same probabilities, and no time limit was imposed. Choices were also private and could not be observed by other subjects. In order to encourage our subjects to consider the decision tasks seriously and reveal their true preferences, we made the implied opportunities from our experiments very significant to the participants—the average payoff (58 Ethiopian Birr) was considerably above the opportunity cost of their time.

Risk preferences have also been shown to be context specific, and some scholars argue that risk preference measures should be framed in a relevant context (see MacCrimmon and Wehrung, 1990; Dohmen et al., 2011). Despite the fact that experimental approaches produce objective and incentive compatible measures of risk preferences, their drawback is that they are context free. In an effort to complement for this limitation, we elicit an additional measure of risk aversion using a survey question in the context of a market decision (see next subsection).

5.3.3 Other variables

We collected data on market exposure and a range of relevant socioeconomic variables from our households. The data contain detailed information on various important aspects of the households (see Table 1).

⁸⁰ US\$1 was approximately equal to 18.213 Ethiopian Birr in 2013.

As has already been mentioned, we measure risk preferences in two complementary ways, using a standard risk investment experiment and a survey question. The first two rows in Table 1 present information on these alternative measures of our dependent variable—risk aversion. The first row contains risk preferences as elicited from our experiments. We use the proportion of the endowment which has not been committed to the risky investment relative to the total endowment as an index of relative risk aversion for each household.⁸¹ It is important to note that our main discussion is based on this measure of risk preferences. We present the distribution of the risk aversion index in Figure 1. The figure reveals substantial heterogeneity in risk preferences across our sample farm households. While all our subjects chose to invest some positive amount of their endowment in the risk game (i.e., the maximum of the risk aversion index is less than one), only a small fraction of them chose to put their entire endowment in the risk investment.

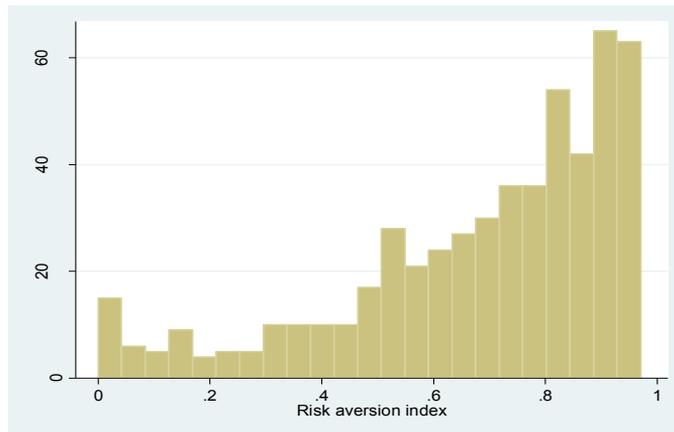


Figure 1: Histogram distribution of the risk aversion index for our sample

The second row presents information on self-reported risk preferences obtained from a simple survey question, which we use as an alternative measure for risk preferences in our robustness checks. We asked respondents to indicate whether they would dare to trade with somebody they didn't know. We then generated a dummy indicator for risk aversion, taking the value 1 if the household did not dare to trade with strangers, 0 otherwise. In both measures, a larger score is associated with more risk aversion. As is evident from the table, the averages of the

⁸¹ The common practice in the literature is to measure risk loving propensity of individuals by using the proportion of the endowment committed to the risky investment relative to the total endowment. We simply prefer to measure the risk aversion coefficient for consistency. The results would basically be the same in magnitude but of opposite sign, as the two indices add up to unity.

two indices indicate that farm households are considerably risk averse. This is consistent with earlier evidence on risk natures of farm households in Asia and Africa alike (see Binswanger, 1980; Wik *et al.*, 2004; Humphrey and Verschoor, 2004; Harrison *et al.*, 2010). We also probed the extent to which the two risk aversion indices capture the same construct, and found a strong but imperfect correlation between them with a coefficient of 0.691 (p-value = 0.000).

The third row in Table 1 gives information on our key explanatory variable—market exposure—proxied by the average number of transactions of the household in a typical month. It only includes transactions which actually occur at the market. The average household conducts about 7 transactions per month. Given our research question, looking at the number of transactions instead of volumes of sales of agricultural products offers several advantages. First, it is market experience that would more likely affect risk preferences, and market experience is more likely to correlate with frequency of transactions rather than with volumes of transactions. Second, households are both buyers and sellers in the market, and number of transactions captures both sides of the market. But commercialization index captures only the selling-side of the market, since it is often constructed as a proportion of the volume of the sales of a household relative to its total agricultural produce. Third, it can help to avoid potential skewedness of the data that might happen due to outliers in the volumes of the transactions. This is related to the fact that the percentage of per capita consumption from own agricultural production becomes significantly small for farmers who engage in more commercialized agriculture.

Table 1 also contains other explanatory variables. Household size was measured by the number of people living under the same roof, sharing production and consumption activities. Livestock was measured using tropical livestock units.⁸² The average household has about 6 family members. It is headed by a person who is about 44 years of age with more than two years of completed schooling. It works on about 8 timads⁸³ of cultivable land with a livestock herd of 7 tropical livestock units. The average household is resided at an altitude of 1979 meters above sea level and at a distance of 17 kilometers from the nearest daily market center. The annual average per capita income (measured in Ethiopian Birr per person) for the sampled households is about 3188. Finally, the average village is a home for about 113 farm households, our proxy for village population.

⁸² Tropical livestock unit is a common unit used to quantify a wide range of various livestock species to a single figure to get the total amount of livestock owned by a household. We employed a tropical livestock unit applicable for SSA.

⁸³ One timad is the land area ploughed by a pair of oxen in a day, and approximately equals 0.25 hectare.

Table 1. Summary statistics

Variables	Description	Mean	Std. Dev.	Min	Max
Risk aversion index	Risk aversion coefficient (game)	0.69	0.246	0	0.971
Risk aversion indicator	Risk aversion index (survey)	0.57	0.495	0	1
Market exposure	Market transactions per month	6.46	4.096	0	18
Distance to market (km)	Distance to the nearest daily market from the household residence	17.40	11.672	1	57
Altitude	Home altitude above sea level	1979.84	225.013	1517	2938
Male	Dummy for male household head	0.92	0.171	0	1
Age	Age in years	43.54	11.448	21	76
Education	Years of education of household head	2.78	3.393	0	13
Household size	People within the household	6.26	2.088	1	17
Land size (timad)	Total cultivable land size	8.21	4.163	2	36.5
Livestock (TLU)	Total Tropical Livestock Units	7.15	4.458	0	35.5
Per capita income	Household per capita income	3187.90	4336.485	520	46150
Village total households	Total number of households in a village	113.31	64.582	45	299

Figure 2 shows the association between the risk aversion index and market exposure. The figure reveals the presence of a clear negative correlation between risk aversion and market exposure, suggesting that market exposure may play a role to account for the systematic heterogeneity in risk preferences. However, such correlation between risk aversion and market exposure should not necessarily be interpreted as causation. Causal inference necessitates accounting for several potential sources of biases (see below). To attenuate these concerns, we employ below an instrumental variables approach in our analysis.

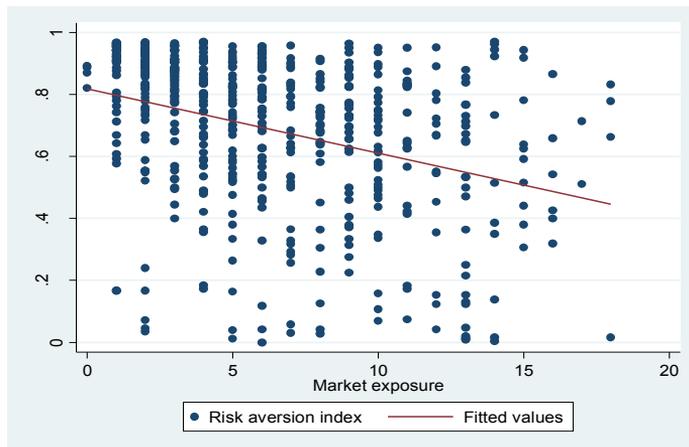


Figure 2: Risk aversion and market exposure

5.3.4 Empirical strategy

Since we are interested to estimate the effect of market exposure on risk preferences of farm households, we set out to estimate the following regression equation:

$$R_i = \beta_0 + \beta_1 M_i + \beta_2 X_i + \varepsilon_i \quad (1)$$

where R_i represents our coefficient of relative risk aversion from our experiments for household i , M_i is our measure of market exposure for household i , X_i is a vector of household covariates, and ε_i is the random error term of the equation. Based on the literature, we identify household covariates that explain risk preferences of individuals, including household size, age, education, a dummy for gender, household land size and livestock ownership.⁸⁴ Because Ra_i as our dependent variable is continuous and censored at 0 from below and at 1 from above, we estimate Eqn. (1) using both ordinary least squares (OLS) and Tobit models to inspect the robustness of our results.

As mentioned before, market exposure may be endogenous in our risk preference model of Eqn. (1). For example, the causal relationship may go in the reverse direction. Perhaps, there may also be omitted variables, such as entrepreneurship, that might trigger some self-selection process by driving both risk preference and market exposure. We adopt an instrumental variables (IV) approach and identify exogenous variation in our market exposure variable to attenuate potential endogeneity concerns. Specifically, we estimate the following two-stage model that simultaneously helps to identify the determinants of market exposure and its causal impact on risk preferences of our households:

$$R_i = \beta_0 + \beta_1 M_i + \beta_2 X_i + \varepsilon_i, \text{ and} \quad (2)$$

$$M_i = \alpha_0 + \alpha_1 Z_i + \alpha_2 X_i + u_i \quad (3)$$

where Z_i is a vector of valid instrumental variables, u_i is the random error term of the first-stage regression, and other variables are as explained in Eqn. (1). As often, the real challenge is to find credible instrumental variables Z_i . In principle, appropriate instruments are variables that explain the market exposure of households, but do not affect risk preferences of these households through any other alternative channel, except via market exposure. This amounts to the condition that the instrumental variables must not be correlated with the error term of our risk preference model.

⁸⁴ At this stage, we exclude income level from our covariates to attenuate the problem of multicollinearity, as many of our explanatory variables are known determinants of rural income. But we include it latter as a robustness analysis.

We got some guidance from the literature in our efforts to identify valid instrumental variables. For example, prior research establishes that geographical factors are key determinants of market exposure decisions of farm households (Barrett, 2008). Distance to potential markets is among the most important geographical factors that determine transaction costs associated with accessing markets (Renkow et al., 2004; Bedemo et al., 2013). While distances to markets are exogenous to a large extent to decisions of farm households, one potential concern is that distance to markets may interact with migration, and thus be endogenous as well. Accordingly, issues of mobility and selective migration might present natural confounds against causal inference from our results. We probe the robustness of our results for this possibility as a part of our robustness analysis in Section 5. On the other hand, there is no obvious argument that systematically links risk preferences and distances to markets. In light of this discussion, distance to daily markets measured from household residence can go well with the admissibility of the exclusion restriction to provide valid instruments for market exposure.

We use a series of tests to inspect whether distance to markets is an appropriate instrument for market exposure. First, we estimate an OLS model explaining variation in market exposure and check whether distance to markets significantly correlates with market exposure in our data. Second, we estimate OLS and Tobit models in an effort to explain variation in our risk aversion index and check that distance to markets does not enter significantly when we control for market exposure. Third, we probe more formally whether the 2SLS model meets the test statistics requirements of the over-identification test. We employ “*altitude*” at the residence of the household as an additional instrument. Altitude can influence access to markets as it captures differences in terrain and hence the quality of the paths/roads to markets. However, it is not a *priori* evident that the exclusion restriction is satisfied with respect to altitude (Pighin et al., 2012). We view this as empirical question and let the data speak. Of course, we also explore our instruments separately to identify the specific instrument that predicts market exposure as a robustness analysis. Finally, we carry out a series of sensitivity checks to establish the robustness of our results.

5.4 Results and discussion

We proceed in several steps to establish the causal effect of market exposure on risk preferences. Tables 2 and 3 provide a series of diagnostic simple regression results. Table 2 summarizes results showing the effects of market exposure on risk preferences based on

estimates of simple OLS and Tobit regressions. In both models, we start by estimating more parsimonious specifications (columns (a)) with market exposure as the only regressor, and gradually include other household covariates to reach the full specification of each model (columns (c)). The most parsimonious specifications of the models suggest a negative and significant correlation between risk aversion and exposure to markets. In the columns designated by (b), we go on to controlling for distance to markets and altitude at household level in both models—two variables that will later be used as instruments for market exposure. While the magnitude of its coefficient decreases from -0.021 to -0.017, market exposure still remains significant at 1% level. Market exposure again remains significant even when we control for other household covariates to reach at the most elaborate specifications of both models (columns (c)), though its coefficient decreases further to -0.015. On average, one standard deviation increase in market exposure results in a 1/3 of a standard deviation decrease in risk aversion.

Our regression results from the most elaborate specifications of the models show that risk preferences of farm households also appear to be associated to a range of other household covariates. The results for these covariates are also consistent with the literature. Male-headed households are less risk averse compared to female-headed households. This may reflect obvious gender differences in economic resources and opportunities. Risk aversion also decreases with education. This is presumably because those who have higher education levels are more likely to have high logical comprehension and thus strong propensity to engage in taking substantial calculated risks. Similarly, risk aversion diminishes with household size—perhaps indicating large social networks and better opportunities to engage in diversified activities to cope with higher risk taking behavior.⁸⁵ The age of the head also matters for risk preferences of a farm household, as risk aversion increases significantly with age. We speculate that this result indicates that older people tend to be conservative as they have low earning power and less time to recover from losses caused by their involvement in potentially risky ventures. The results for all covariates are qualitatively similar in the two models.

Unfortunately, per capita income does not correlate significantly with risk aversion. This merits some discussion. Risk aversion and market exposure are expected to independently correlate with income. Indeed, there is a strong relationship between market exposure and

⁸⁵ Alternatively, the psychology literature has long associated the role of family size in risk behavior with dependency-linked behaviors—large families are expected to have lower dependency-linked behaviors and hence larger tendency to accept more risk than small families (e.g., Jamieson, 1969).

income in our data (see Table 3 below). While it is highly likely that income is also partly endogenous, one would argue that the income effect on risk aversion, and the endogeneity of income, might be more important than the independent relationship between market exposure and risk aversion if the simultaneity issue is addressed in our model. To attenuate this concern, we estimate our risk aversion model excluding the market exposure variable. However, per capita income does not still enter significantly in the model (see Table A2 in the appendix). As such, we argue that the effect of market exposure on risk aversion is independent to the correlation between market exposure and income (see also Section 5).

Table 2. Market exposure reduces risk aversion

Variable	Risk aversion index (OLS)			Risk aversion index (Tobit) ^a		
	(a)	(b)	(c)	(a)	(b)	(c)
Market exposure	-0.021*** (0.003)	-0.017*** (0.004)	-0.015*** (0.005)	-0.021*** (0.002)	-0.017*** (0.005)	-0.015*** (0.005)
Male			-0.147*** (0.049)			-0.147*** (0.049)
Age			0.004*** (0.001)			0.004*** (0.001)
Education			-0.032*** (0.003)			-0.032*** (0.003)
Household size			-0.009* (0.005)			-0.009* (0.005)
Land size			0.002 (0.003)			0.002 (0.003)
Livestock			-0.001 (0.002)			-0.001 (0.002)
Log per capita income			0.012 (0.011)			0.012 (0.011)
Village total households			0.0001 (0.0001)			0.0001 (0.0001)
Distance to market (km)		0.0013 (0.002)	0.0007 (0.001)		0.0014 (0.004)	0.0007 (0.001)
Altitude		0.0000 (0.000)	0.0000 (0.000)		0.0000 (0.000)	0.0000 (0.000)
Constant	0.818*** (0.019)	0.762*** (0.064)	0.771*** (0.116)	0.818*** (0.019)	0.761*** (0.064)	0.771*** (0.115)
Adjusted R ²	0.12	0.12	0.38			
F-test	71.69 (1, 530)	24.14 (3, 528)	30.82 (11, 520)			
P-value	0.000	0.000	0.000			
Pseudo R ²				3.33	3.38	13.25
Log likelihood				23.53	23.99	123.60
Prob. > χ^2				0.000	0.000	0.000
N	532	532	532	532	532	532

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; km stands for kilometer; N stands for the number of observations; ^a 1 left-censored observation at 0; 531 uncensored observations.

Would these results be interpreted as causal effects? Arguably, gender, age, education and family size are exogenous with respect to contemporary individual risk preferences and allow us to give a causal interpretation to our regression results. However, as has already been indicated, market exposure may be endogenous. To attenuate this concern, we use an instrumental variable (IV) approach to identify an exogenous variation in our market exposure variable. We use distance to markets and altitude at household level as our instruments for market exposure. Before running our IV models, we now turn to subjecting our instruments to statistical tests to probe their credibility for being valid instruments. Note that the first criterion for valid instruments is that they should not be significantly correlated with the dependent variable. This criterion is effectively met in our data as our instrumental variables are not correlated with our risk aversion index across different specifications (see Table 2).

The second criterion for valid instruments is that they should be significantly correlated with the endogenous explanatory variable. In order to inspect our choice of the instruments in light of this criterion, we next aim to explain what factors determine market exposure of farm households. To do so, we estimate both OLS and Tobit models.⁸⁶ The results of the regressions are presented in Table 3. The results support our choice of the instrumental variables. As expected, both distance to markets and altitude are negatively and significantly associated with market exposure. While distance to markets mirrors transaction costs, altitude may reflect quality of routes to markets. Education level of the head, land size, livestock and per capita income are other significant determinants of market exposure. Education, land size and per capita income are positively correlated with market exposure of a household as expected. However, livestock enters negatively in the market exposure model, which may be due to attenuation bias caused by the aggregation in its measurement based on the tropical livestock unit that gives equal weight to different species.

⁸⁶ Tobit model estimation has been justified since market exposure is censored at zero.

Table 3. Distance to market and altitude significantly correlate with market exposure

Variable	Market Exposure			
	OLS		Tobit ^b	
	(a)	(b)	(a)	(b)
Male		-0.073 (0.459)		-0.105 (0.453)
Age		-0.007 (0.008)		-0.006 (0.008)
Education		0.004* (0.027)		0.006* (0.026)
Household size		0.071* (0.043)		0.071* (0.042)
Land size		0.363*** (0.024)		0.359*** (0.024)
Livestock		-0.057* (0.019)		-0.055* (0.019)
Log per capita income		0.640*** (0.095)		0.655*** (0.094)
Village total households		0.0005 (0.001)		0.0001 (0.001)
Distance to market (km)	-0.282*** (0.008)	-0.171*** (0.010)	-0.288*** (0.009)	-0.176*** (0.010)
Altitude	-0.006*** (0.0002)	-0.003** (0.0001)	-0.006*** (0.0002)	-0.003** (0.0001)
Constant	12.591*** (0.325)	2.629*** (1.076)	12.705*** (0.324)	2.661*** (1.064)
Adjusted R ²	0.68	0.80		
F-test	572.17 (2, 529)	222.33 (10, 521)		
P-value	0.000	0.000		
Pseudo R ²			0.21	0.30
Log likelihood			-1188.41	-1053.76
Prob. > χ^2			0.000	0.000
N	532	532	532	532

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; km stands for kilometer; N stands for the number of observations; ^b4 left-censored observations at 0; 528 uncensored observations.

Against the background of these qualifications, we proceed to estimate 2SLS and IV Tobit models. Table 4 gives the results of these regressions. Most importantly, our instrumented market exposure still enters negatively and significantly in both models of risk aversion. Now, the direct effect of market exposure on risk preferences dominates a potential reverse effect in explaining the correlation between the two variables, as our instrumental variable estimates isolate the effect of market exposure on risk preferences. The coefficient of the instrumented market exposure is slightly larger than its corresponding value in the fully elaborated OLS models, though it is smaller than the coefficient of market exposure in the parsimonious OLS specifications. A potential explanation for such difference is measurement error in market

exposure, as expected in typical retrospective data measurements from rural household surveys. While measurement errors can bias OLS coefficients towards zero (Theil, 1971: p. 608), instrumental variable approaches often mitigate such problems (Gujarati, 2003: p. 527). Supporting this explanation, the standard deviation of the raw market exposure (4.096) is larger than that of its predicted value used in the second stage regression (3.649).⁸⁷

Table 4. Instrumenting market exposure by *distance to market* and *altitude*

Variable	Risk aversion index (2SLS)	Risk aversion index (IV Tobit) ^c
Market exposure	-0.019*** (0.006)	-0.019*** (0.006)
Male	-0.147*** (0.049)	-0.147*** (0.049)
Age	0.004*** (0.001)	0.003*** (0.001)
Education	-0.032*** (0.003)	-0.032*** (0.003)
Household size	-0.008* (0.005)	-0.009* (0.005)
Land size	0.004 (0.004)	0.004 (0.004)
Livestock	-0.001 (0.002)	-0.001 (0.002)
Log per capita income	0.015 (0.012)	0.016 (0.012)
Village total households	0.0001 (0.0001)	0.0001 (0.0001)
Constant	0.793*** (0.104)	0.794*** (0.104)
Sargan test: χ^2 (1)		
	0.24	
p-value		
	0.623	
Wald test of exogeneity:		
	χ^2 (1)	0.42
	prob. > χ^2	0.516
Adjusted R ²	0.39	
Log likelihood		-939.04
Prob. > χ^2	0.000	0.000
N	532	532

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; km stands for kilometer; N stands for the number of observations; ^c1 left-censored observation at 0; 531 uncensored observations.

Finally, all standard tests for valid instruments indicate that our instruments are valid and strong. As already shown in Table 3, both excluded instruments are significant in the first-stage regression, and the associated partial F-statistic (Table 3 OLS (a)) is way above the

⁸⁷ We also present a robustness analysis accounting for alternative interpretation of distance to markets in Section 5 in an effort to eliminate ‘non-trade’ effects of markets on risk preference.

minimum 10 threshold value of the “rule of thumb” for strong instruments (Staiger and Stock, 1997). Table 4 also provides further qualifications for strength of our instrumental variables. The Sargan test for the 2SLS model fails to reject the null hypothesis that our over-identifying restrictions are valid (p-value =0.623), indicating that the applied instruments are appropriate, and we have correctly excluded the instruments from the second-stage of the model. Moreover, the Wald test of exogeneity for the IV Tobit model fails to reject the null hypothesis that the residuals from the first-stage are not correlated with those from the second-stage model (p-value =0.516), suggesting clearly the admissibility of the excluded instruments. We thus contend that our instrumentation strategy is credible, and our results reveal that market exposure is a key determinant of risk preferences of farm households.

5.5 Sensitivity analysis

5.5.1 Accounting for internal mobility and migration

As has been brought forth in the methodology section, internal mobility and migration may cause a concern for self-selection that may impair our causal inference. While distances to markets are to a large extent exogenous to decisions of farm households, one concern is related to the sample selection potentially caused by out-migration. Another potential concern is that distance to markets may interact with internal mobility—i.e. households relocating closer to markets within the same area. Both out-migration and internal mobility may serve as mechanisms for farm households to get near to markets, and households involved in mobility and migration are obviously not randomly drawn from the general population. We thoroughly address this concern as follows.

Selective migration out of rural communities is a well-known threat to causal inference in non-randomized as well as randomized settings. As a first step to address this concern, we purposely base our household sampling on land holding—only sampling landed households. The Ethiopian Land Proclamation has made enjoying secure and continuous land use rights contingent on permanent physical residence in the community (Deininger and Jin, 2006). As a second step, we confirm our intuition that landed households are less prone to migration through focus groups. And finally, we test the robustness of our results to the most conservative assumptions about out-migration patterns, as well as within-migration in the area of study.

We rely on focus group interviews to learn about the nature of migration over the last five years. The focus group discussions were held at a Kebele level. Each village in the Kebele was represented by a discussant. The group of discussants also included a focal discussant from the Kebele administration office. Totally, 179 villages were included in the group discussions, with an average of 10 villages per Kebele. The average five-year rate of out-migration for all Kebeles in our study was found to be 2.52%, ranging from 0.23% to 4.96%. Consistent with our choice to only study landed households, all focus groups cite landlessness as the number one important reason when asked about the causes of out-migration in their community. Elsewhere, evidence also shows that landlessness is the main driver of rural-urban migration in Ethiopia (Ezra and Kiros, 2001; Mberu, 2006; Hunnes, 2012). This has also been corroborated by our data from the focus group discussions: 98% of the out-migrated households either did not have land at all or had less than 0.25 hectares of land before their migration.⁸⁸

We present the relationship between the out-migration rates and average distance to markets at Kebele level in Figure 3. From the figure, we read a tendency for a positive association between distances from markets and out-migration. In other words, it seems that more remote municipalities have higher rates of out-migration. However, this correlation is not statistically significant ($p = 0.3751$).

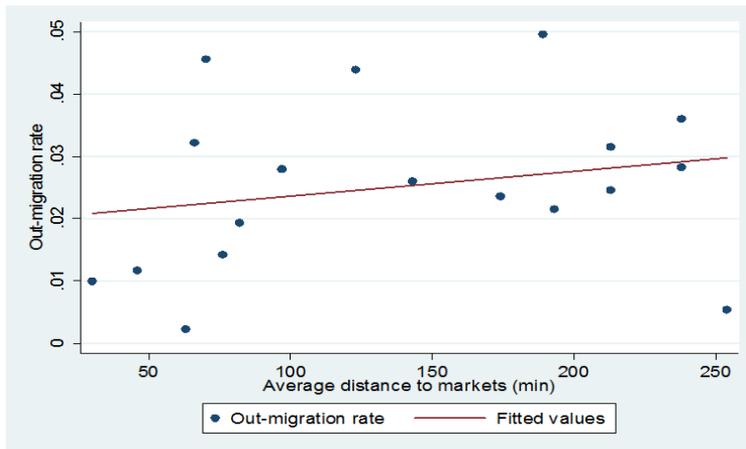


Figure 3: Out-migration rates and average distance to markets

⁸⁸ The threshold, 0.25 hectares of land, is effectively less than a quarter of the average household land holding in our sample area. In fact, this provides support to the negative-selection hypothesis of migration that households with strong incentive to out-migrate are those with no or small landholdings (Chiquiar and Hanson, 2005).

We proceed to test the impact of selective migration on our results more formally. Facing a missing data problem, we use the out-migration rates of each Kebele to place bounds on our estimates of the effects of market exposure on risk aversion. Specifically, we examine if our results would survive the most conservative check for out-migration by hypothetically creating households. The number of households to be artificially created is determined by the out-migration rates of the respective Kebeles, and this process results in 20 such new households. The working assumption is that out-migrated households are more likely to be the high risk tolerant ones (i.e., low risk averters). Accordingly, we proceed by assigning the lowest value of risk aversion index (i.e., zero) to these households. All other variables are valued at their respective averages.⁸⁹ This approach allows us to place bounds on our estimates without making additional untestable distributional assumptions (Gerber and Green, 2012).

We estimate our standard models after making these adjustments to our data. The 2SLS and the IV Tobit results are similar, and we report only the 2SLS results hereafter. These results are presented in Table 5. Column (1a) contains second-stage results. The results remain statistically robust. Our results also survive the most conservative sensitivity check for out-migration over the past ten years. We assumed stable out-migration rates over the past ten years, and doubled the hypothetically generated households to 40. The associated second-stage results are reported in Table 5 (column (2a)), and they remain quite robust. Given the relative low out-migration rates and the narrow range of feasible outcomes of our risk aversion index, these bound estimates can be quite useful to facilitate inference about the sensitivity of our results. The corresponding first-stage results are given in columns (1b) and (2b). The key test statistics also remain quite admissible to substantiate the credibility of our instrumentation approach.

⁸⁹ As indicated, many of the out-migrants are landless. According, evaluating all covariates at their means is a generous valuation.

Market Exposure and Risk Aversion

Table 5. Robustness analysis for out-migration

Variable	2SLS (1): 5 Years		2SLS (2): 10 Years	
	Risk aversion index		Risk aversion index	
	(a)	(b)	(a)	(b)
Market exposure	-0.015*** (0.005)		-0.014*** (0.006)	
Male	-0.139** (0.056)	-0.073 (0.469)	-0.131** (0.065)	-0.070 (0.460)
Age	0.004*** (0.001)	-0.006 (0.008)	0.003*** (0.001)	-0.007 (0.008)
Education	-0.031*** (0.003)	0.019* (0.027)	-0.031*** (0.004)	0.020* (0.026)
Household size	-0.011** (0.005)	-0.017 (0.042)	-0.011* (0.006)	-0.016 (0.041)
Land size	0.003 (0.004)	0.401*** (0.024)	0.003 (0.005)	0.401*** (0.023)
Livestock	-0.0002 (0.002)	-0.032* (0.019)	-0.0002 (0.003)	-0.032* (0.018)
Instruments				
Distance to market (km)		-0.203*** (0.008)		-0.203*** (0.008)
Altitude		-0.003*** (0.0001)		0.003*** (0.0001)
Constant	0.869*** (0.077)	7.995*** (0.723)	0.838*** (0.086)	7.995*** (0.710)
Sargan test:				
$\chi^2(1)$	0.20		0.16	
p-value	0.656		0.692	
First stage F-test		261.19		270.81
Market exposure		(8, 543)		(8, 563)
P-value	0.000	0.000	0.000	0.000
Adjusted R ²	0.31	0.79	0.25	0.78
N	552	552	572	572

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; km stands for kilometer; N stands for the number of observations.

Next, we probe the robustness of our results for internal mobility, i.e., mobility within the community. Some of the group discussants revealed that it is not uncommon for households to move within their villages and/or Kebeles to reside near to markets.⁹⁰ Following this suggestion, a careful investigation of our data has resulted in 92 households who have not resided in the same location in their community since their establishment. In other words, these households have moved within their village and/or Kebele, rather than migrated away from their community. We employ a series of robustness checks to investigate the possibility that these households might have moved to be near to markets, and whether this could affect our results. First, we probe if our results would withstand the most conservative check

⁹⁰ The group discussants in all our Kebeles asserted that households do not migrate to be located away from markets under normal circumstances.

through residing these households to the possible farthest villages to markets in their Kebeles.⁹¹ Second, we replace market distance measured from the residence of households by market distance measured from the center of their village for each household, and estimate our standard models (reported standard errors are clustered at village level). The results for both checks are given in Table 6. The 2SLS results for the first check are shown in column (1a), while column (2a) provides the 2SLS results for the second check. The results are once again robust, and the related statistical tests presented in the respective (b) columns are satisfied well.

Table 6. Robustness analysis for internal mobility

Variable	2SLS (1)		2SLS (2) ^a	
	Risk aversion index		Risk aversion index	
	(a)	(b)	(a)	(b)
Market exposure	-0.013*** (0.005)		-0.018*** (0.005)	
Male	-0.148*** (0.049)	0.315 (0.523)	-0.147*** (0.049)	0.144 (0.414)
Age	0.004*** (0.001)	0.011 (0.009)	0.004*** (0.001)	-0.003 (0.007)
Education	-0.031*** (0.003)	0.042* (0.030)	-0.031*** (0.003)	0.029 (0.038)
Household size	-0.010** (0.004)	-0.057 (0.046)	-0.011** (0.004)	-0.060 (0.057)
Land size	0.001 (0.004)	0.469*** (0.026)	0.005 (0.004)	0.497*** (0.041)
Livestock	-0.0001 (0.002)	-0.020 (0.021)	-0.0003 (0.002)	-0.018 (0.031)
Instruments				
Distance to market (km)		-0.149*** (0.008)		-0.030*** (0.003)
Altitude		-0.004*** (0.0001)		-0.004*** (0.0003)
Constant	0.900*** (0.065)	5.784*** (0.782)	0.903*** (0.065)	6.231*** (0.635)
Sargan test:				
χ^2 (1)	0.42		0.22	
P-value	0.518		0.640	
First stage F-test		199.45		170.44
Market exposure		(8, 523)		(8, 523)
P-value	0.000	0.000	0.000	0.000
Adjusted R ²	0.39	0.75	0.39	0.72
N	532	532	532	532

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; km stands for kilometer; N stands for the number of observations; ^areported standard errors are clustered at village levels.

⁹¹ Some of these households have already resided in the farthest villages to markets in their Kebeles. Their 'mobility' may reflect simply movement from one corner to another of the same village. This is not uncommon in Ethiopia where farmers often seek to reside near to their farm plots.

As alternative robustness check, we drop the 92 households, and estimate our models based on information only from households who have resided at the same place since their establishment. This would help us to assess if our results shake for the exclusion of these households. Column (1a) of Table 7 provides the IV estimation results, and the corresponding first-stage regression results are in column (1b). As can be read, the results remain robust and consistent with our results from Table 4, indicating that our results are not driven by the inclusion of within community migrants. The important test statistics are also quite satisfactory.

Finally, we follow a recent literature facing similar problem of selective mobility (Callen et al., 2014; Hanaoka et al., 2014), and estimate our standard models by including a within-community mobility dummy, indicating that a household has not resided at the same location since its conception, and see if it would enter significantly. We also include an interaction term between the mobility dummy and distance to markets in the market exposure model (first-stage regression) to allow for the possibility that the effect of distance to markets might vary with the status of internal mobility. Column (2a) of Table 7 presents the 2SLS results. However, both the mobility dummy and its interaction with distance to markets do not correlate significantly with market exposure (column (2b)). While it may not be so informative given the first-stage results, we simply include the interaction term in our set of instruments to inspect the sensitivity of the main results.⁹² Our main result remains largely robust (columns (2a)), and key over-identification test statistics also indicate that the over-identification restrictions are valid.

⁹² We run a simple OLS for our risk aversion model including the mobility dummy and its interaction with distance to markets as explanatory variables but neither enters significantly at conventional levels.

Table 7. Robustness analysis for internal mobility

Variable	2SLS (1)		2SLS (2)	
	Risk aversion index		Risk aversion index	
	(a)	(b)	(a)	(b)
Market exposure	-0.016*** (0.005)		-0.016*** (0.004)	
Male	-0.138** (0.055)	0.031 (0.510)	-0.147*** (0.049)	-0.074 (0.479)
Age	0.004*** (0.001)	-0.000 (0.009)	0.004*** (0.001)	-0.006 (0.008)
Education	-0.031*** (0.003)	0.011 (0.028)	-0.031*** (0.003)	0.019 (0.027)
Household size	-0.013*** (0.005)	-0.048 (0.044)	-0.011** (0.004)	-0.017 (0.042)
Land size	0.004 (0.004)	0.392*** (0.025)	0.003 (0.004)	0.401*** (0.025)
Livestock	0.0003 (0.002)	-0.015 (0.020)	-0.0004 (0.002)	-0.032* (0.019)
Migration dummy			-0.040* (0.022)	-0.006 (0.389)
Instruments				
Distance to market (km)		-0.206*** (0.009)		-0.203*** (0.009)
Altitude		-0.003** (0.0001)		-0.003** (0.0001)
Distance to market (km) * Migration				-0.0004 (0.019)
Constant	0.882*** (0.073)	7.744*** (0.780)	0.903*** (0.065)	7.994*** (0.742)
Sargan test:				
χ^2 (1)	0.30			
χ^2 (2)			0.48	
P-value	0.584		0.788	
First stage F-test		228.25		200.49
Market exposure		(8, 431)		(10, 521)
P-value	0.000	0.000	0.000	0.000
Adjusted R ²	0.39	0.80	0.40	0.78
N	440	440	532	532

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; km stands for kilometer; N stands for the number of observations.

5.5.2 Further robustness checks

We estimate several alternative models in an effort to probe further the robustness of our results. In Table 8, we present a series of IV results. We first inspect the robustness of our results to an alternative proxy for risk preference. We use a dummy indicator for risk aversion from a survey question to capture the context of the market decision environment of our subjects. We specifically asked households whether they would dare to trade with whom they

didn't know. Following the binary nature of our dependent variable, we estimate an IV probit model and report marginal effects. As often, we estimate the IV probit model in two steps. In the first step, an OLS model is estimated for market exposure using distance to markets and altitude—our instruments—as exclusion restrictions for identification purpose. We then include both the predicted market exposure and the generated generalized error term as explanatory variables in our probit model. We include the generalized error term to account for any potential endogeneity that might be driven by unobserved heterogeneity. The results are contained in column (1a), while the associated first-stage results of the model are presented in column (1b) (Table 8). There is again a robust correlation between predicted market exposure and our risk aversion indicator. Moreover, the generalized residual term does not enter significantly, implying that there are no common unobserved factors that drive both market exposure and risk aversion.

We provide additional robustness checks in Table 8. In columns (2a) and (3a), we examine if our results are robust to omission of per capita income and inclusion of 'rationality'. As has already been discussed in Section 4, our main result might be contaminated by the so-called multicollinearity statistical problem, as (most of) our covariates are well known and non-trivial predictors of income levels in the literature. We then estimate our standard model excluding per capita income. Columns (2a) and (2b) give, respectively, the second-stage and the corresponding first-stage results, once per capita income is excluded. Altogether, the results indicate that our main results remain largely robust. As per rationality, the literature articulates that rationality of individuals is a strong correlate of their market experience (List and Millimet, 2008; Cecchi and Bulte, 2013). We left rationality out from our initial analysis for this obvious endogeneity problem. We now include "*rationality*" as a covariate in our basic risk aversion models.⁹³ As is evident from the second-stage regressions (column (3a) of Table 8), our results show again a robust correlation between instrumented market exposure and our risk aversion index. Consistent with the literature, market experience and rationality

⁹³ We measure rationality of our subjects through a simple experiment based on the Generalized Axiom of Revealed Preference (GARP), which has been shown to be a necessary and sufficient condition for data to be consistent with rational utility maximization (Varian, 1982). We then employ GARP violations as a nonparametric test to rationality violations. In our design, a GARP violation occurs when a bundle x is chosen when a bundle y is available, where bundle y has at least as much of all goods and strictly more of at least one good than a third bundle z , and z has been directly or indirectly revealed preferred to x . The setup of the game was simple. We just asked our subjects to make their preferred choice from a list of 3 - 5 distinct alternative possibilities of two goods (raw coffee beans and sugar) on seven separate sheets of paper. It requires no ability to do the computation to stay within a budget constraint, and no need to forecast choices and act strategically in any rigorous way.

violations are negatively correlated (first-stage regression). However, rationality does not enter significantly in our risk aversion model.

Table 8. Robustness analysis: Alternative measure of risk preferences and Omitted variable bias

Variable	IV Probit model (1)		2SLS (2)		2SLS (3)	
	Risk aversion indicator		Risk aversion index		Risk aversion index	
	(a)	(b)	(a)	(b)	(a)	(b)
Market exposure	-0.188*** (0.046)		-0.016*** (0.004)		-0.019*** (0.006)	
Male		-0.073 (0.459)	-0.147*** (0.049)	-0.073 (0.478)	-0.146*** (0.049)	-0.122 (0.454)
Age	0.009 (0.006)	-0.007 (0.008)	0.004*** (0.001)	-0.007 (0.008)	0.004*** (0.001)	-0.008 (0.008)
Education	-0.095*** (0.022)	0.004* (0.027)	-0.031*** (0.003)	0.020** (0.027)	-0.032*** (0.003)	0.011 (0.026)
Household size	0.012 (0.036)	0.071* (0.043)	-0.011** (0.004)	-0.017 (0.042)	-0.008* (0.005)	0.059 (0.042)
Land size	-0.012 (0.030)	0.363*** (0.024)	0.003 (0.004)	0.401*** (0.025)	0.003 (0.004)	0.372*** (0.024)
Livestock	-0.005 (0.016)	-0.057* (0.019)	-0.0002 (0.002)	-0.032* (0.019)	-0.001 (0.002)	-0.060** (0.019)
Log per capita income	0.082 (0.093)	0.640*** (0.095)			0.017 (0.012)	0.583*** (0.095)
Village total households	0.001 (0.001)	0.0005 (0.001)	0.0001 (0.0001)	0.0005 (0.001)	0.0001 (0.0001)	0.001 (0.001)
Generalized Residual	0.003 (0.019)					
Rationality					0.001 (0.002)	-0.054*** (0.014)
Instruments						
Distance to market (km)		-0.171*** (0.010)		-0.203*** (0.009)		-0.168*** (0.010)
Altitude		-0.003** (0.0001)		-0.003** (0.0001)		-0.003** (0.0001)
Constant	0.646 (0.703)	2.629*** (1.076)	0.902*** (0.065)	7.995*** (0.737)	0.773*** (0.104)	3.362*** (1.080)
Sargan test:						
χ^2 (1)			0.28		0.26	
p-value			0.595		0.611	
First stage F-test		222.33		251.57		208.59
Market exposure		(10, 521)		(9, 522)		(11, 520)
P-value		0.000	0.000	0.000	0.000	0.000
Adjusted R ²		0.80	0.39	0.79	0.39	0.81
Pseudo R ²	0.26					
Log likelihood	-261.07					
Prob. > χ^2	0.000					
N	516	532	532	532	532	532

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; km stands for kilometer; N stands for the number of observations.

In Table 9, we report a chain of robustness results related to distance to markets—our key instrument. First, we change our measure of distance to markets. We now use distance to markets measured by travel time instead of kilometers. The reason for the change is that distance measured by travel time may better reflect differences in terrain and the quality of roads to markets. Alternatively, it may also help to attenuate potential measurement errors. Column (1a) of Table 9 provides second-stage results. The results are again quite comparable to what we have already obtained in Section 4. The associated first-stage results are given in column (1b). The relevant test statistics also indicate that our instrumentation approach is appropriate.

Another potential concern related to market distance is that permanent markets are often located in small towns. While none of the households in our sample come from such towns, households residing very close to markets, and hence to these towns, may have larger opportunities for social interaction, and may in turn be exposed to broader social interactions. For example, Renkow et al. (2004) contend that market distance is definitely much more than geometry. It governs not only access to market information and infrastructure but also relative opportunities for interaction and communication, suggesting that ‘shortest’ market distances may not capture only market access opportunities. If so, we may not pick only market exposure effects on risk preferences for households residing very close to markets. To attenuate this concern, we throw out 86 households settling at less than 6 km market distance (i.e., within the boundaries of the small towns), and then estimate our model using information only from the remaining households. The results are reported in columns (2a) and (2b), and they are once more robust and consistent with our main results.

Finally, in columns (3a) and (3b), we examine which specific excluded instruments are predicting market exposure. We estimate our basic model by using only “*distance to markets*” as the key excluded instrument. Of course, this entails that we cannot test for the over-identification restriction, as the number of the excluded instrument equals the number of the endogenous variable. We find that our results remain qualitatively unaffected, and it is “*distance to markets*” that largely predicts market exposure status of our farm households.

Table 9. Robustness analysis: Measuring distance to market in time, accounting for alternative interpretations of distance to market, and instrumenting market exposure only by distance to market

Variable	2SLS (1)		2SLS (2)		2SLS (3)	
	Risk aversion index		Risk aversion index		Risk aversion index	
	(a)	(b)	(a)	(b)	(a)	(b)
Market exposure	-0.018*** (0.006)		-0.017** (0.007)		-0.019*** (0.006)	
Male	-0.147*** (0.049)	-0.154 (0.458)	-0.132** (0.051)	0.122 (0.409)	-0.147*** (0.049)	-0.089 (0.462)
Age	0.003*** (0.001)	-0.001 (0.008)	0.003*** (0.001)	-0.010 (0.007)	0.004*** (0.001)	-0.007 (0.008)
Education	-0.032*** (0.003)	0.005* (0.026)	-0.029*** (0.003)	0.007* (0.025)	-0.032*** (0.003)	0.003* (0.027)
Household size	-0.009* (0.005)	0.095* (0.043)	-0.006 (0.005)	0.093** (0.038)	-0.009* (0.005)	0.071* (0.043)
Land size	0.003 (0.004)	0.325*** (0.025)	0.0001 (0.004)	0.253*** (0.029)	0.004 (0.004)	0.371*** (0.024)
Livestock	-0.001 (0.002)	-0.070* (0.019)	-0.0002 (0.002)	-0.062** (0.017)	-0.001 (0.002)	-0.059** (0.019)
Log per capita income	0.015 (0.012)	0.913*** (0.088)	0.007 (0.012)	0.520 (0.084)	0.015 (0.012)	0.645*** (0.096)
Village total households	0.0001 (0.0001)	-0.001 (0.001)	0.0001 (0.0001)	-0.001 (0.001)	0.0001 (0.0001)	0.001 (0.001)
Instruments						
Distance to market (hrs)		-0.087*** (0.005)				
Distance to market (km)				-0.132*** (0.009)		-0.172*** (0.010)
Altitude		-0.003** (0.0001)		-0.003*** (0.0004)		
Constant	0.798*** (0.104)	1.347 (1.032)	0.848*** (0.106)	9.329*** (1.397)	0.797*** (0.104)	1.896*** (1.046)
Sargan test:						
χ^2 (1)	0.26		3.44			
p-value	0.611		0.638			
First stage F-test		224.18		158.26		243.38
Market exposure		(10, 521)		(10, 435)		(9, 522)
P-value	0.000	0.000	0.000	0.000	0.000	0.000
Adjusted R ²	0.39	0.81	0.35	0.78	0.39	0.80
N	532	532	446	446	532	532

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; hrs stands for hours; km stands for kilometer; N stands for the number of observations.

5.6 Conclusions and discussion

“Much of the academic research on risk has its balance wrong, focusing too much on admittedly fascinating issues related to risk management and coping mechanisms, but not enough on its implications and the scope for interventions” (Dercon, 2008: p. ii98). In fact, while risk aversion has long been identified as having a detrimental effect on economic

development, the scope for policy intervention in this domain has remained limited for a long time. The conventional thinking in the existing literature on risk is to envisage farm households as inherently risk averse. This logic has shifted the focus of the analysis in the literature to examining risk management and coping strategies used by the households (Dercon, 2002; World Bank, 2014). Whether policy can have a role in shaping risk preferences of farm households, and thereby their investment decisions, largely depends on our understanding of what determines risk preferences and thus their change.

The obvious endogeneity between risk preferences and the extent to which a person is exposed to markets has limited the capacity of previous studies to draw clear conclusion with respect to the directionality of their relationship. One way of addressing this would be to exogenously impose market exposure—*à la* John List. A seminal example for this strategy has been provided by a recently published paper by Bryan et al. (2014). They randomly assign an incentive to poor households in rural Bangladesh to temporarily out-migrate, and show that out-migration increases consumption at the origin as well as the rates of out-migration even after the incentive is removed. They conclude that out-migration is very profitable and underutilized, and that risk aversion may play a fundamental role in such underutilization. In this paper, we provide complementary evidence based on a non-randomized experimental approach—using an instrument to identify causality. We make use of the low-mobility of farm households with land titles in rural Ethiopia to instrument frequency of market transactions with distance from the nearest daily market.

We find evidence that market exposure is a key determinant of risk preferences of farm households. Specifically, we show that market exposure attenuates risk aversion, and farm households with higher degree of market exposure tend to be less risk averse. Our results are robust to controls for various observable attributes of our farm households, and different sensitivity checks and specifications. Particularly, we address the concern of selective migration thoroughly, and it does not appear to drive our results. The external validity of these results is clearly limited by our use of only landed households, and cannot be verified in our setting. However, the households in our sample account for about 25% of the total households in our study villages, and we have little reason to believe that different dynamics would be at play for landless households.

Overall, our results indicate that risk preferences evolve in response to market exposure, and suggest that the claim that farm households are inherently risk averse may be excessive.

Rather, high risk aversion is a trait that can be modified through increasing exposure to markets, and policy can play some role in influencing risk preferences of farm households. The penetration of markets in remote areas may be more than a means to better allocate resources. It may be an important channel to reduce risk aversion, thereby reducing the productivity lag of rural households, and helping them to exit the poverty trap. This paper thus sheds new light on an alternative mechanism that may potentially link market exposure to economic growth—via attenuating risk aversion.

Appendix

Table A1: Experimental procedure

Steps in the risk game

1. Each participant is brought to a private location within the premises of the experimental field by an enumerator and assigned an initial endowment of 100 tokens. Each token is worth 0.5 Birr.
 2. Enumerators explain the rules of the risk game until the respondent is able to convey them back correctly as follows.
 3. The decision task consists of a choice in which you will decide on a certain amount of your endowment to involve in a risky investment. The investment you make can affect your endowment both negatively and positively. The final pay-out you will earn depends on the choice you make and its outcome during the decision task activities.
 4. Enumerators present the lottery bag and take out all the balls. The investment is very simple; do you notice the red and white balls? There are 8 equal balls. The four balls have white color; the remaining four balls have red color.
 5. This is how the risk game works. From your endowment, you can decide on any amount between 0 and 100 to be used in the risk game, but not more than 100. If you extract a ball with a red color from the lottery bag, the sum you decided to use for this investment will be doubled. But if you extract a ball with a white color, the sum you decided to use for this investment will be halved. In essence, you have $\frac{1}{2}$ chance of doubling the money you decide to use and $\frac{1}{2}$ chance of losing half of it.
 6. Remember, there is no right or wrong in this game, it's about what you want to do. Then a trial round is performed.
 7. After the trial round, the respondent is asked to explain the outcome of the 'investment'. If the explanation is correct, the actual game can take place, otherwise the enumerator explains the rules again until the participant understands the game.
 8. Payments are made after all subjects from the same village complete the game.
-

Institutional Change and Economic Development

Table A2. Per capita income does not significantly correlate with risk aversion even when market exposure is excluded

Variable	Risk aversion index (OLS)
Male	-0.146*** (0.050)
Age	0.004*** (0.001)
Education	-0.031*** (0.003)
Household size	-0.010** (0.005)
Land size	-0.003 (0.003)
Livestock	0.0003 (0.002)
Log per capita income	0.003 (0.010)
Village total households	0.0001 (0.0001)
Distance to market (km)	0.003 (0.001)
Altitude	0.0000 (0.000)
Constant	0.733*** (0.116)
Adjusted R ²	0.37
F-test	32.35 (10, 521)
P-value	0.000
N	532

Standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$; km stands for kilometer; N stands for the number of observations.

Market Exposure and Risk Aversion

Chapter 6

Synthesis

6.1 Introduction

Today, few economists would dispute the primacy of institutions in the development process. We know a good deal about the rationales to believe that institutions are the essential underpinning of economic development, not to mention the vast empirical literature that supports this assertion. During more than two decades of renewed interest in institutions in development economics, scholars have sought answers to general questions, like: How do institutions affect development outcomes? Yet, our understanding on how to bring about institutional change and improvement is still quite imperfect (Evans, 2004; Dixit, 2009). Again, putting in place good institutions that have undergirded the growth of the developed world has not always produced desired results in developing countries (Rodrik, 2006). That is probably so because these institutions have elements that are highly specific to a country's local conditions, which make understanding interaction and micro-foundations of institutions very important—to which each chapter of this thesis has endeavored to contribute.

This thesis has attempted to provide a comprehensive analysis and discussion of the nexus between institutional change and economic development, underlining the dynamics and interdependence between formal and informal institutions. To do so, it has exploited a rich dataset collected through a combination of field experiments and survey approaches. This thesis generates several important insights that contribute to understanding of the complex dynamics and interactions that undergird the relationship between institutional change and economic development, highlighting what may work and may not. Taken to its logical conclusion, the main message that comes across from the different chapters is that formal and informal institutions interact out of entrenched corners with both constructive and deleterious repercussions for economic development. The following sections present the main findings from each chapter and related key debates in the literature on the topic, discussing their implications for both policy and future research.

6.2 Formalizing land rights in Africa: ‘The wrong prescription for the wrong malady’⁹⁴?

Formalization of customary land rights is often carried out based on the presumption that full-fledged land rights provide incentives and opportunities for economic development. Given plausible theoretical justifications, one would expect the empirical evidence on positive impacts of such programs to be robust and pervasive. In many parts of the world, land titling programs have led to enhanced tenure security and increased economic gains. Unfortunately, land titling initiatives have produced relatively disappointing outcomes in Africa. A careful reading of the empirical literature from many African countries reveals that land reforms have not lived up to expectations (e.g., Migot-Adholla et al., 1991; Place and Hazell, 1993; Brasselle *et al.*, 2002; Jacoby and Minten, 2007).

Why do outcomes of formalizing land rights from Africa defy conventional wisdom? While theoretical explanations for formalizing land rights cannot be dismissed outright for potential identification problems in the empirical literature, there is ample cause for concern on the ground. Attentively, we contend that the reasons are down to the history and roles of customary land institutions of the continent. In many African countries, customary tenure systems appear to provide basic land rights that suffice to induce farmers to make necessary land-improving investments (Bromley, 2008; Toulmin, 2009; Fenske, 2011). Given that customary tenure systems provide sufficient security, it takes a small step to conclude that there is not ‘enough’ latent demand for formal land titling at least at this stage of development. This in turn leads to the conclusion that land titling programs were ‘wrong prescriptions for the wrong malady’ in the first place. Evidently, all studies that failed to find significant effects of land titling on economic outcomes have reiterated the role of customary institutions for the weak results. For example, Migot-Adholla et al. (1991, p. 170) conclude that land registration is “unlikely to be economically worthwhile for much of Sub-Saharan Africa at this stage of its economic development”. This same fundamental point is also echoed by Anne-Sophie Brasselle et al.: “By reviewing the justifications for scepticism in the recent literature, we are able to conclude that the traditional village order, where it exists, provides the basic land rights required to stimulate small-scale investment” (Brasselle et al., 2002, p. 1). Furthermore, many analysts argue that these indigenous tenure arrangements are dynamic

⁹⁴ Bromley (2008).

in nature and evolve to accommodate new demand for land and (further) security (Migot-Adholla et al., 1991; Sjaastad and Bromley, 1997; Place, 2009).

Notwithstanding the prevalence of functional indigenous tenure systems, many other considerations may also partially explain the disappointing results. In Africa, land is more than a simple production factor—it is a symbol of wealth and social status, and a basis for complex relations of alliance and reciprocity (Cotula et al., 2004). Land rights have a complex and deep social dimension, and land reforms cannot be viewed as a simple technical matter (Lemel, 1988; Galiani and Schargrotsky, 2010). As such, calling for formalization of customary land rights in vain may not only fail to produce desired outcomes but also create parallel systems that increase insecurity through additional sources of opportunities and uncertainties (Atwood, 1990). Moreover, formalization of land rights per se cannot bring about the necessary level of tenure security and economic outcomes. Specifically, formalization will do little good if the legal foundations of an economy are tenuous and the authority structure that promises effective enforcement of the rights is weak to do so (Bromley, 2008). Where customary institutions function effectively, they do so because they are subject to clear rules and socially embedded enforcement mechanisms. Problems of design and implementation of the titling programs themselves may also blur the lines between what works and what does not. Most of the African land reforms were politically motivated and aimed at defusing social unrest and meeting aid conditions (Place, 2009). Practically, they have inclined to mean an organized redistribution of landholdings or rights. Land reforms have also been portrayed as an example of naïve top-down approaches which, rather than responding to real grassroots needs, prescribe “solutions” without a thorough diagnosis (Easterly, 2008, Toulmin, 2009). Finally, program implementation costs in some countries are found to be very high to the extent that they bore no reasonable relationship to land values (see Jacoby and Minten, 2007; Deininger and Feder, 2009).

In Chapter 2, we look at the impact of the Ethiopian land registration and certification program on agricultural productivity, and show that land titling programs produce significant positive effects, provided that there is demand for more tenure security and implementation procedures of titling schemes are based on the recognition of local contexts to overcome practical challenges. In Ethiopia, customary land arrangements are not prevalent due to the discontinuous evolution of its tenure system. Moreover, the threat of land redistribution by the state has been the main source of tenure insecurity for Ethiopian farmers, explaining the presence of high demand for land registration. Interestingly enough, other studies of the

Ethiopian land registration and certification program find results that are consistent with theory and cast new light on the existing evidence on the importance of such institutions for agricultural development in Africa (see Chapter 2).

From a policy perspective, one important insight that comes out from our results and discussions is that formalization of land rights is a meaningful institutional development where indigenous tenure systems ceased to have legitimacy and the threat of eviction is tangible. Simplistic and blanket calls for formalization of land rights that dismiss the role of competing interests and indigenous tenure arrangements will not yield expected outcomes. It is also imperative to distinguish between ‘formalizing land rights’ and ‘tenure security’. “Formal title is not always necessary or sufficient for high levels of tenure security” (Deininger, 2003, p. 39). After all, individual land titling is not the only option for enhancing tenure security, as indigenous tenure arrangements, where they exist, provide basic levels of security to induce farmers to make the necessary agricultural investment. Neglecting such evidence through one-size-fits-all prescriptions may reduce rather than enhance *de facto* tenure security. Finally, the fact that the economic benefits of land titles depend on the pre-existing land tenure systems and details of the titling program points to the relevance of accounting for heterogeneous effects. Therefore, decentralized titling programs are more likely to understand and correctly assess local conditions, as well as to concentrate their efforts on constraints that bind most the tenure system.

In light of our discussion, future research in this field should look at how local/micro variations in the custom affect demand for formal land rights (formal institutions in general). The ability to identify or induce exogenous variation in the quality of the customary system will be the key for success to any future research endeavor in this direction. Alternatively, we believe that the evolution and dynamics of customary land rights merits careful investigation before setting out for formalizing customary land rights. Equally, understanding the extent to which economic development is affecting the level of tenure security entailed by customary land systems is important to identify the scope for improvement and find alternative arrangements for combining customary and formal tenure systems out of entrenched corners. We also know very little about the longer-term impacts of land registration programs. Moreover, even when positive impacts of titling programs are reported in the literature, their economic relevance is not often closely scrutinized (see Jacoby and Minten, 2007). We believe that all these are promising areas for future research.

6.3 Land rights and women's empowerment

Land titling programs are usually implemented with the view to spur agricultural development through increasing investment and productivity incentives as well as enhancing access to credit. The indirect spillover effects of land reforms often receive little attention and are not subject to thorough investigation. In Chapter 3, we extend our analysis of the impacts of formalizing land rights beyond their classic direct benefits to the nexus between land rights and the distribution of power in the household organization—one important spillover effect of land certification programs. We learned that access to land rights via joint land certification (i.e. titles containing the names of both spouses) increases the bargaining position of women in the household. These effects are in line with collective intrahousehold models, where control of household resources matter for the distribution of bargaining power within the household and its decision outcomes (see Manser and Brown, 1980; Bourguignon and Chiappori, 1992). These results are also important since current development approaches stress the centrality of empowering women as both a development in itself—for women are among the adversely disadvantaged, and a policy instrument to enhance household welfare—because resources controlled by women have multiplier effects on the wellbeing of other household members (Sen, 1999). Taken together, the results of Chapters 2 and 3 reveal that the Ethiopian land certification program has generated significant efficiency and equity gains.

The findings of Chapter 3 have intriguing policy implications. They imply that even oppressive or inequitable social norms can be changed through introducing measures that alter their status quo conditions and incentive structures. For instance, empowering women, where strong custom of male dominance in virtually all important aspects of household decision-making is prevalent, involves changing unequal institutional relationships that sustain the culture of inequality. Unfortunately, these traditional norms are sensitive in nature and least amenable for direct state reform measures as they are typically supported by deep-rooted customs and values. Yet, policy interventions, such as improving women's access to land rights, can be useful instruments to change these enduring social norms that create relations of power and disadvantage between men and women within the household and beyond. Moreover, acknowledging spillover effects of titling programs would allow full exploitation of their potential benefits. It is also imperative to analyze not only whether expected direct benefits can rationalize land formalizing programs but also whether their intervention can be implemented equitably and its institutions can be sustained over time.

An increase in the bargaining position of women is expected to have multiplier effects on household welfare via its effect on the way households spend their resources. Particularly, resources controlled by women are more likely to be allocated on basic consumption goods, children's nutrition and education. In addition, more equal intrahousehold relations would importantly affect agriculture productivity in rural areas. This represents an important opportunity for future research. Future work should also look at beyond conventional benefits of land titling programs and consider interaction effects. Perhaps, the impacts of land titling programs on inequality, performance of local formal and informal land administration and conflict-settling institutions should merit future investigation.

6.4 Interaction between formal and informal institutions

The interaction between formal and informal institutions is often modeled in the literature as either substitutes or complements. In Chapter 4, we departed from this standard approach and studied the impact of the expansion of the realm of the statutory law on mediation outcomes of customary legal authorities. In so doing, we examined the spillover effects of the increased penetration of the formal law, but not its direct impacts. Those who were disadvantaged by the custom have the option to challenge mediation outcomes of the customary judges via appealing to the formal law. In this context, customary judges face a trade-off between making their intrinsic and preferred arbitration outcome that largely reflects the custom to which they are often a guardian and deviation from this preferred outcome so as to retain disputants within the realm of their jurisdiction, which brings them social prestige. Confronted with this trade-off, customary judges have been shown to respond to the threat of 'competition' from the formal law by moving their dispute-resolution decisions in the direction of the formal law itself. This implies that the extent to which customary judges act dictatorially crucially depends on the extent to which they enjoy a monopoly power. These results are consistent with the theoretical predictions of the dynamic model developed by Aldashev et al. (2012a, 2012b), featuring customary judges who respond strategically to the trade-off between jurisdictional erosion by the formal law and loss of social prestige due to plaintiffs appealing against their rulings. It is clear that such penetration of the formal law has a positive welfare effect. However, this is not the end of the story from a social welfare perspective, as the formal law also affects the behavior of adherents of the customary law, and there appear effects operating in opposite directions (see below).

These results may have important implications for policy. First, the formal law may be instrumental to transform customary legal institutions towards a direction wished by the legislator. Second, and contrary to a widespread view, customary institutions themselves may have a role to play in shifting preexisting customs towards a desired outcome. While customary authorities may intrinsically prefer custom-compliant rulings, they do not also necessarily face the social sanctioning associated with deviating from their preferred decisions. As a result, customary authorities may objectively deviate from their preferred custom-compliant rulings so as to avoid loss of reputation from their reversal in case some plaintiffs decide to appeal to the formal law. In areas where formal legal institutions have limited outreach and the custom remains strong, important benefits of increased penetration of the formal law may arise from changes in customary legal institutions, rather than from plaintiffs seeking justice under the rule of law. Thus, the returns from increasing the integration between state law and other normative orders would be high in these areas. In fact, attempts to delegitimize customary institutions, expecting that formal laws would replace them, may also be costly if people have internalized the customary institutions under the existing social order to such a degree that they dislike strong departures from them.

A clear limitation of Chapter 4 is that it does not adopt a holistic approach to analyze the full implications of the penetration of the formal law. Future research on the penetration of the formal law into places where the customary law remains strong should account for the dynamics of the interactions between the two. Clearly, more systematic empirical work is needed to test the extent to which moderately formal laws may have a greater “magnet effect” on the custom than radical ones (Aldashev et al., 2012a, 2012b). This has been shown to crucially depend on the social prestige derived by customary authorities from dispute adjudication, the cost of deviating from their preferred judgment, the embeddedness of the custom in social interactions and the credibility of the formal law as a default option. Again, future empirical work should investigate the practical relevance of these conditions and associated thresholds. In addition, complementing the evidence from the field experiment with naturally occurring data, which can account for the uncertainty of the outcomes of the actual formal law, could provide another opportunity for future research. After more evidence is documented, we should work towards a generalized theory about the interaction between formal and informal institutions.

6.5 Customary institutions are 'malleable'

While customary institutions are a pervasive feature of the institutional framework in all economies, they have special importance in Africa. They enjoy the allegiance of the overwhelming majority of the population, and they permeate not only communities but also the state to the point where they often undermine the credibility of formal institutions (Platteau, 2009; Michalopoulos and Papaioannou, 2014). Their role has been a subject of bitter controversy from a long-term development point of view. There is a strong debate about whether such customary institutions are compatible with modern economic development. Clearly, two views are prevalent within the literature. On the one hand, there are those who believe that customary institutions inevitably evolve to give way to institutions, rules and values that are more congenial to modern economic development as the expansion of economic opportunities unfolds itself (Lewis, 1955; Boserup, 1981). A central tenet of this view is that customary institutions are dynamic and have profoundly evolved in tandem with social, economic, political and cultural changes over time. On the other hand, there are others who see customary institutions as incompatible with modern development and argue that customary institutions, norms and values should be transformed in order to allow rapid development (Rostow, 1963; Harrison, 1987; Platteau, 2000).

The current dominant view among economists is that, while they can facilitate order and prosperity at early stages of development, customary institutions can hardly be a foundation for long-term development, especially when the economy gets matured and its transactions increased in complexity. Evidence also shows that customary institutions attenuate growth incentives through redistributive norms that favor egalitarianism and disfavor competition (Platteau, 2000; Di Falco and Bulte, 2011), albeit there is nothing strictly egalitarian about the outcomes of these norms—as they also tend to discriminate against women and the disadvantaged (see Chapter 4). Yet, there are several approaches to thinking about why such inefficient institutions might persist. Commonly, customary institutions persist because they turn out to be a second-best strategy for individuals who prefer, but cannot achieve, a formal institutional solution (e.g., Platteau, 2009). Acemoglu et al. (2005) contend that bad institutions are kept in place, clearly not for the benefit of society as a whole, but for the benefit of special groups with political power. Alternatively, there is also increasing recognition that some customary institutions are propelled by noneconomic reasons as they take on symbolic value with a coterie of people's customs, values and beliefs (Williamson, 2000).

For these and other reasons, transforming customary institutions so as to effectively meet the central challenges of long-term economic development has always been a vital topic in economics. In the concluding chapter of his seminal book, *Institutions, Social Norms, and Economic Development* (2000), Jean-Philippe Platteau, for example, stressed the importance of transforming the African customary institutions if the continent is to see any meaningful improvement in its long-term economic performance: “In countries of Sub-Saharan Africa (SSA), the legacy to be defeated is one of absence of a strong state, an independent and effective bureaucracy, and a civil society differentiated along class lines rather than on the basis of highly segmented ethnic and other traditional affiliations” (Platteau, 2000, p. 335). Nevertheless, creating and introducing formal institutions to replace existing customary institutions has always presented a big challenge as they are firmly embedded in the practices, customs, values and beliefs of the people.

Notwithstanding the evidence that customary institutions ‘evolve’ spontaneously over time, the question of paramount importance from the point of view of long-term economic development is to what extent they can be purposefully changed or reconfigured. As such, the notion that ‘customary institutions are malleable’ is here clearly intended to be metaphoric. Insofar as the intention of transforming customary institutions is confined to incremental or progressive change, Chapters 3 and 4 illustrate that deliberative choice mechanisms can be implicated to engineer changes in such institutions. The results specifically show that customary institutions respond to exogenous changes in formal institutions that affect their status quo conditions (Chapter 3) or the realm of their jurisdiction (Chapter 4).

What can we learn from these results for policy? One important lesson is that changes in customary institutions can be consciously devised. Looking at the mechanisms of the changes, informal institutional change may be more effectively initiated by providing ‘appropriate’ incentives than by introducing the ‘desired’ kinds of institutions. This approach may offer a better path of institutional development than the often unsuccessful radical approach in many African countries, where customary institutions remain dominant. However, more empirical and experimental evidence is needed, especially with respect to long-run responses of informal institutions to the increasing penetration of formal alternatives. Understanding the mechanism that customary institutions have managed to evolve naturally to accommodate new developments would provide relevant lesson to design effective institutional change and thus future work should focus on unraveling this mechanism. There is also a need to

document empirical evidence on the interplay between the micro-mechanisms through which customary institutions persist and the channels they influence behavior.

6.6 Endogenous risk preferences

In Chapter 5, we investigate the impact of market exposure on risk preferences of poor farm households. We learn that risk preferences evolve in response to market exposure. This finding lends support to the recent small empirical literature that suggests the endogeneity of preferences (see Eckel et al., 2009; Voors et al., 2012; Cameron and Shah, 2013). The result shows that markets are not passive institutions that allocate resources efficiently. They also shape risk preferences of market participants. Penetration of markets in remote areas of developing countries may then induce the possibility of dynamic efficiency gains. It may help to reduce the productivity lag of rural households through attenuating their high risk aversion nature, creating an alternative mechanism that may potentially link market exposure to economic development.

These findings highlight that the high risk aversion nature of poor farmers is a trait that can be affected through policy interventions, such as increasing their exposure to markets. This is a captivating result given the growing interest in policy to link poor farmers in developing countries to market institutions. If the penetration of markets in relatively remote areas of the developing countries is more than a means to better allocate resources, it may generate a chain of externalities that could nurture a self-propelling trajectory of market integration and development at the local level. Policymakers should further consider lifting barriers of market integration in rural areas. They should also pay due attention to the heterogeneity in risk taking across poor farmers when designing policies aiming at stimulating market developments and adoption of different agricultural technologies.

Standard economic theory regards risk preferences as hardwired and fixed characteristics of humans in its textbook model of *homo economicus*. However, a recent but growing literature presents persuasive evidence to the contrary. Notoriously, the evidence suggesting that risk preferences evolve comes from diverse domains of social science. Arguing for a more elaborated endogenous theory of preferences, Samuel Bowles concluded that: [...] The weight of both reason and evidence points strongly to the endogeneity of preferences (Bowles, 1998, p. 103). In light of the growing and compelling evidence for endogenous (risk) preferences, the scope of economic inquiry has been truncated in ways which restrict its explanatory power and policy relevance. If preferences are indeed endogenous, it will be important to

differentiate the effects of institutional changes on incentives and constraints of behavior from those on preferences *per se*. Until recently, economic theory gave little guidance in understanding these effects. Moreover, endogenous (risk) preference may have explanatory power distinct from the institutional changes that set it into motion. It may thus influence the endogenous formation of other preferences in the longer run—via the inertia that it will sustain even after the institutional shock lasts. This is all the more salient in view of the transient nature of institutional shocks, and policymakers should not overlook this implication of endogenous preference formation.

6.7 Unintended consequences of institutional change

As long as economic agents are not automata in responding to structurally-determined incentives, their ideas and interactions have come to matter a great deal for outcomes of institutional changes (Chang, 2007). As a result, outcomes of institutional changes cannot be completely predictable. There often exist unforeseen dynamics, leading to unintended consequences of institutional changes. Unintended consequences may be positive (See Chapter 5) or negative, to which this section is devoted.

The expansion of the realm of the formal law is expected to affect not only the incentive structure of the customary law but also the behavior of its adherents. Specifically, it is supposed to increase the bargaining power of those disadvantaged by the customary system. In the context of Chapter 4, the introduction of the formal legal fallback should benefit higher investors as it enforces the division of the final endowment between the two players according to their initial investment shares. Thus, the introduction of the formal law is said to increase the best alternative to a negotiated outcome of higher investors, even if the best alternative to the negotiated outcome is not invoked in equilibrium. Yet, and quite contrary to our expectations, individuals favored by the law not only failed to take this advantage but also made offers that were less advantageous to themselves. In equilibrium, only a small fraction of them made direct use of the formal law and women were even less likely to apply the law than men. This is clearly a disappointing effect of the introduction of the formal law. While a complete welfare analysis cannot be implicated here, partly because the introduced formal law itself would also have its own direct effect, one clear result is that the introduction of the formal law has had unintended consequences. This result is particularly worrisome as individuals did not face any information constraint and related uncertainty in the outcomes of the formal law—otherwise pervasive features of the outcomes of actual formal legal institutions.

What might explain this unintended effect of the introduction of the formal law? Definitely, behavioral patterns that often characterize poor communities, like altruism and inequality aversion, cannot stand here as plausible explanations for at least why individuals favored by the law made offers that were less advantageous to themselves. Rather, we speculate that these individuals forewent the benefits of the introduction of the formal law due to reputational concerns or expected social sanctioning. The central argument goes as follows. If the custom is strong, expected social cost of applying the legal fallback would be high, as this might be considered as violation of the custom. If so, individuals favored by the formal law may not take this advantage, but rather signal their commitment and adherence to the customary law by even making decisions against the formal law.

Practically, we can go on to listing several real-life examples to elaborate that legal reforms intended to amend or outlaw strong customary rules not only have had no impact on such rules but also generated unintended consequences. For instance, laws enacted to forbid dowry payments in South Asia (India, Bangladesh and Pakistan) not only have destined to failure but increased both the incidence and real value of dowry payments (see Billig, 1992; Rao, 1993; Amin and Cain, 1995; Anderson, 2003). Similarly, legislative reforms introduced to ban female genital mutilation/cutting in many sub-Saharan African countries have generated unintended effects rather than leading to the abandonment of the practice. Particularly, these interventions have pushed the practice towards secrecy and underground, lowered the age of cutting, triggered cross-border migration and resulted in public resentment (see UNICEF, 2010; Shell-Duncan et al., 2013; Camilotti, 2015). The main cause for such unintended effects is often individuals' widespread belief that, given that these laws run counter to deeply-seated customary practices, they are unlikely to be followed by others or backed by appropriate sanctions.

The literature has long emphasized the roles of collection action problems and interest groups for such kinds of unintended effects of institutional changes and survival of inefficient institutions. Olson (1971), for example, believes that a change in any social arrangement necessitate a large degree of organized and effective coordination, which occurs quite infrequently. Taking a different perspective, Kuran (1989) stressed the role of "preference falsification" in a collective choice model that distinguishes between individuals' privately held preferences and those they espoused in public to explain the continuation of hated regimes. A privately hated regime may enjoy widespread public support because of people's reluctance to take the lead in publicizing their opposition. However, collective action problems are not the only reason for the survival of inefficient institutions and unintended consequences of institutional changes. Acemoglu and

Robinson (2006) argue that, despite their adverse effects on aggregate performance, extractive institutions may emerge as equilibrium institutions as they strengthen the position of the groups that hold political power. Interestingly, Friedrich Hayek wrote long ago on the importance of interest groups that adhered to strong customs or beliefs over individuals' wisdom and economic consideration to bring about any change in such institutions: [...] The ultimate decision about what is accepted as right and wrong will be made not by individual human wisdom but by the disappearance of the groups that have adhered to the "wrong" beliefs' (Hayek, 1960, p. 36).

Our tentative results caution policymakers that institutional change is not a simple technocratic exercise. In particular, policymakers have to acknowledge that changing informal institutions represents a challenging process that may involve changing power relations and overcoming path dependency, as well as unforeseen consequences. Unintended (negative) effects may not only attenuate potential welfare gains associated with expansion of formal institutions but may also deteriorate overall welfare conditions. For example, Dixit (2004) compares the outcomes of relational and formal contracting, and demonstrates that opportunities created by the formal law can lower social welfare, because first-best cooperative outcomes may no longer emerge as sub-game perfect equilibria. Given that processes of institutional change are complicated enough, and may even have crucial unforeseen impacts, the debate in the literature that formal and informal institutions are either substitutes or complements might be a red herring. Instead, future research on institutional change and development should underlie the importance of dynamics in institutional change, accounting for interactions of ideas, presumed behavioral repertoires, customary institutions, formal institutional forms and interest groups.

6.8 Concluding remarks

It is easy to state the fundamental challenge confronting current institutional development efforts. There is no proven approach to bring about a required institutional change. Nor when it does occur may we be successful at controlling unintended effects of this change. The literature suggests that good institutions can be acquired, but doing so often requires experimentation, willingness to depart from orthodoxy and attention to local conditions (Rodrik, 2003; Chang, 2007). However, this literature does not inform us precisely how to experiment and account for local conditions. This is exactly what this thesis has attempted to do. This thesis was started with the motivation to study the nexus between institutional change and economic development. Throughout its chapters, special emphasis was placed on institutional change and their outcomes—both institutional and economic outcomes. It has

been shown that an effective institutional development requires a good knowledge of the interaction between formal and informal institutions and the complex dynamics that such interaction entails. Moreover, it should be clear that institutional change is a nonlinear, complex and non-ergodic process, where multiple intended and unintended outcomes are possible.

Yet, there is still a lot to be uncovered about how institutions evolve and interact. Particularly, examining the micro-mechanisms underpinning the emergence, persistence and dynamics of institutions at micro-level would contribute to a better understanding of the complex relationship between institutional change and the dynamics of economic development. Returns to future research on this topic would be very rewarding. Field experimental approaches are also likely to continue as a promising tool to facilitate identification of causal relations and underlying mechanisms for this research effort. At the same time, great care has to be taken in employing field experiments in diverse settings. These approaches inevitably face a trade-off between understanding the specifics of a given context and the generalizability of their findings. And yet, the issue of precise attribution oftentimes remains a latent concern. It is important to be precise about what is actually being measured in decontextualized field experiments. Furthermore, limited cognitive capacities of illiterate subjects in rural areas of developing countries coupled with increasing complexity of experimental modules might pose a grave challenge to attribute what such measures exactly capture and hence, the practical relevance of field experimental approaches. This implies a need for a greater marriage between randomized field experiments and survey work not only for unbiased causal inference but also for ecological validity to real-world causal processes. Equally, complementing quantitative survey-experimental analyses with more qualitative work (including focus group discussions) could enhance the quality of research on institutions. Importantly, qualitative approaches would increase the quality of quantitative data through improving the choice of relevant variables in a survey work. They could also facilitate identification of micro-mechanisms for institutional change, which otherwise might not identify in quantitative analyses.

References

- Acemoglu, D. & Johnson, S. (2005). Unbundling Institutions. *Journal of Political Economy* 113(5): 949-995.
- Acemoglu, D., Johnson, S. & Robinson, J. A. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review* 91(5): 1369-1401.
- Acemoglu, D., Johnson, S. & Robinson, J.A. (2002). Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. *Quarterly Journal of Economics* 117(4): 1231–1294.
- Acemoglu, D., Johnson, S. & Robinson, J.A. (2005). Institutions as a Fundamental Cause of Long Run Growth. In Aghion, P. & Durlauf, S. (Eds.), *Handbook of Economic Growth*, Volume 1A, Chapter 6. Elsevier.
- Acemoglu, D. & Robinson, J. (2006). De Facto Political Power and Institutional Persistence. *AEA Papers and Proceedings* 96(2): 325-330.
- Agarwal, B. (1994). Gender and Command over Property: A Critical Gap in Economic Analysis and Policy in South Asia. *World Development* 22(10):1455-1478.
- Agarwal, B. (1997). "Bargaining" and Gender Relations: Within and Beyond the Household. *Feminist economics* 3(1): 1-51.
- Aghion, P. & Howitt, P. (1998). *Endogenous Growth Theory*. Cambridge, MA: MIT Press.
- Aldashev, G., Chaara, I., Platteau, J.-P. & Wahhaj, Z. (2012a). Formal Law as a Magnet to Reform Custom. *Economic Development and Cultural Change* 60(4): 795-828.
- Aldashev, G., Chaara, I., Platteau, J.-P. & Wahhaj, Z. (2012b). Using the Law to Change the Custom. *Journal of Development Economics* 97(2): 182-200.
- Alem, Y., Bezabih, M., Kassie, M. & Zikhali, P. (2010). Does Fertilizer Use Respond to Rainfall Variability? Panel Data Evidence from Ethiopia. *Agricultural Economics* 41(2): 165-175.
- Allendorf, K. (2007). Do Women's Land Rights Promote Empowerment and Child Health in Nepal? *World Development* 35(11): 1975-1988.

- Alston, L. J., Libecap, G. D. & Schneider, R. (1996). The Determinants and Impact of Property Rights: Land Certificates on the Brazilian Frontier. *Journal of Law, Economics, & Organization* 12(1): 25-61.
- Amin, S. & Cain, M. (1995). The Rise of Dowry in Bangladesh. In Jones, G., Caldwell, J., Douglas, R. & D'Souza, R. (Eds.), *The Continuing Demographic Transition*. Oxford: Oxford University Press.
- Anderson, S. (2003). Why Dowry Payments Declined with Modernization in Europe but Are Rising in India. *Journal of Political Economy* 111(2): 269-310.
- Andre, C. & Platteau, J. P. (1998). Land Relations Under Unbearable Stress: Rwanda Caught in the Malthusian Trap. *Journal of Economic Behavior & Organization* 34(1): 1-47.
- Angeles, L. (2011). Institutions, Property Rights, and Economic Development in Historical Perspective. *Kyklos* 64(2): 157 -177.
- Angrist, J. D. (2004). Treatment Effect Heterogeneity in Theory and Practice. *The Economic Journal* 114 (494): C52-C83.
- Aoki, M. (2001). *Toward a Comparative Institutional Analysis*. MIT Press.
- Asfaw, T. & Satterfield, T. (2010). Gender Relations in Local Level Dispute Settlement in Ethiopia's Zeghie Peninsula. *Research in Human Ecology* 17(2): 160-174.
- Ashraf, N. (2009). Spousal Control and Intra-household Decision Making: An Experimental Study in the Philippines. *American Economic Review* 99 (4): 1245-1277.
- Attanasio, O. & Lechene, V. (2002). Tests of Income Pooling in Household Decisions. *Review of Economic Dynamics* 5(4): 720–748.
- Atwood, D.A. (1990). Land Registration in Africa: The Impact on Agricultural Production. *World Development* 18 (5): 659-671.
- Baker, B. (2013). Where Formal and Informal Justice Meet: Ethiopia's Justice Pluralism. *African Journal of International and Comparative Law* 21(2): 202–218.
- Baland, J., Guirkinger, C. & Mali, C. (2011). Pretending to be Poor: Borrowing to Escape Forced Solidarity in Cameroon. *Economic Development and Cultural Change* 60(1): 1-16.
- Banerjee, A. V. & Duflo, E. (2009). The Experimental Approach to Development Economics. *Annual Review of Economics*: 1(1): 151-178.
- Banerjee, A. V, Gertler, P. J. & Ghatak, M. (2002). Empowerment and Efficiency: Tenancy Reform in West Bengal. *Journal of Political Economy* 110(2): 239-280.
- Barrett, C. (2008). Smallholder Market Exposure: Concepts and Evidence from Eastern and Southern Africa. *Food Policy* 33(4): 299–317.

- Barro, R. (1991). Economic Growth in a Cross-Section of Countries. *Quarterly Journal of Economics* 106 (2): 407-433.
- Baser, O. (2006). Too Much Ado about Propensity Score Models? Comparing Methods of Propensity Score Matching. *Value in Health* 9(6): 377-385.
- Basu, K. (2000). *Prelude to Political Economy: A Study of the Social and Political Foundations of Economics*. Oxford University Press.
- Bateman, I. & Munro, A. (2005). An Experiment on Risky Choice Amongst Households. *Economic Journal* 115(502): C176-C189.
- Bates, R. H. (2006). Institutions and Development. *Journal of African Economies* 15: 10-61.
- Becker, G. (1991). *A Treatise of the Family*. Harvard University Press.
- Becker, G. S. & Mulligan, C. B. (1997). The Endogenous Determination of Time Preference. *Quarterly Journal of Economics* 112 (3): 729-758.
- Bedemo, A., Getnet, K. & Kassa, B. (2013). Determinants of Labor Market Exposure Choice of Farm Households in Rural Ethiopia: Multinomial Logit Analysis. *Journal of Economics and Sustainable Development* 4(1):133-141.
- Bennett, T. W. (2006). Comparative Law and African Customary Law. In Reimann, M. & Zimmermann, R. (Eds.), *Oxford Handbook of Comparative Law*. Oxford University Press.
- Benrud, E. (2009). The Historical Evolution of Dividends. In Baker, H. K. (Ed.), *Dividends and Dividend Policy*. Wiley Online Library.
- Berkowitz, D., Pistor, K., Richard, J.F. (2003). The Transplant Effect. *American Journal of Comparative Law* 51: 163-204.
- Bernard, T., Taffesse, A. S. & Gabre-Madhin, E. (2008). Impact of Cooperatives on Smallholders' Commercialization Behavior: Evidence from Ethiopia. *Agricultural Economics* 39(2): 147-161.
- Besley, T. (1995). Property Rights and Investment Incentives: Theory and Evidence from Ghana. *Journal of Political Economy* 103(5): 903-937.
- Bezabih, M., Kohlin, G. & Mannberg, A. (2011). Trust, Tenure Insecurity and Land Certification in Rural Ethiopia. *Journal of Socio-Economics* 40(6): 833-843.
- Bezabih, M., Holden, S. & Mannberg, A. (2012). The Role of Land Certification in Reducing Gender Gaps in Productivity in Rural Ethiopia. Centre for Land Tenure Studies Working Paper 01/12, Norwegian University of Life Sciences, Norway.
- Billig, M.S. (1992). The Marriage Squeeze and the Rise of Groomprice in India's Kerala State. *Journal of Comparative Family Studies* 23(2): 197-216.

- Binswanger, H. (1980). Attitudes toward Risk: Experimental Measurement in Rural India. *American Journal of Agricultural Economics* 62(3): 395–407.
- Bloom, H. S. (1995). Minimum Detectable Effects: A Simple Way to Report the Statistical Power of Experimental Designs. *Evaluation Review* 19(5): 547-556.
- Boettke, P. J., Coyne, C. J. & Leeson, P. T. (2008). Institutional Stickiness and the New Development Economics. *American Journal of Economics and Sociology* 67(2): 331-358.
- Boserup, E. (1981). Population and Technological Change. Chicago: University of Chicago Press.
- Bosker, M. & Garresten, H. (2012). Economic Geography and Economic Development in Sub-Saharan Africa. *World Bank Economic Review* 26 (3): 1–43.
- Bouquet, E. (2009). State-led Land Reform and Local Institutional Change: Land Certificates, Land Markets and Tenure Security in Mexican communities. *World Development* 37(8): 1390-1399.
- Bourguignon, F. & Chiappori, P. (1992). Collective Models of Household Behavior: An Introduction. *European Economic Review* 36(2-3): 355-364.
- Bowles, S. (1998). Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions. *Journal of Economic Literature* 36 (1): 75-111.
- Brasselle, A.-S., Gaspart, F. & Platteau, J.-P. (2002). Land Tenure Security and Investment Incentives: Puzzling Evidence from Burkina Faso. *Journal of Development Economics* 67(2): 373-418.
- Bromley, D. W. (2008). Formalizing Property Relations in the Developing World: The Wrong Prescription for the Wrong Malady. *Land Use Policy* 26 (1): 20–27.
- Brousseau, E., Garrouste, P. & Raynaud, E. (2011). Institutional Changes: Alternative Theories and Consequences for Institutional Design. *Journal of Economic Behavior & Organization* 79: 3-19.
- Bryan, G., Chowdhury, S. & Mobarak, A. H. (2014). Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh. *Econometrica* 82(5): 671-1748.
- Buchanan, J. M. (1964). What should Economists Do? *Southern Economic Journal* 30: 213-222.
- Caliendo, M. & Kopeining, S. (2008). Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys* 22 (1): 31–72.

- Callen, M., Isaqzadeh, M., Long, J. D. & Sprenger, C. (2014). Violence and Risk Preference: Experimental Evidence from Afghanistan. *American Economic Review* 104(1): 123-148.
- Cameron, L. & Shah, M. (2013). Risk-taking Behavior in the Wake of Natural Disasters. NBER Working Paper No. 19534.
- Camilotti, G. (2015). Interventions to Stop Female Genital Cutting and the Evolution of the Custom. Evidence from Age at Cutting in Senegal. Center for Research in the Economics of Development, University of Namur.
- Cappelen, A. W., Hole, A. D., Sorensen, E. O. & Tungodden, B. (2003). The Pluralism of Fairness Ideals: An Experimental Approach. *American Economic Review* 97(3): 818-831.
- Carlsson, F., He, H., Martinsson, P., Qin, P. & Sutter, M. (2012). Household Decision Making in Rural China: Using Experiments to Estimate the Influences of Spouses. *Journal of Economic Behavior & Organization* 84(2) 525– 536.
- Carlsson, F., Martinsson, P., Qin, P. & Sutter, M. (2013). The Influence of Spouses on Household Decision Making under Risk: An Experiment in Rural China. *Experimental Economics* 16(3): 383–401.
- Caselli, F. (1997). Rural Labor and Credit Markets. *Journal of Development Economics* 54: 235-260.
- Cecchi, F. & Bulte, E. (2013). Does Market Experience Promote Rational Choice? Experimental Evidence from Rural Ethiopia. *Economic Development and Cultural Change* 61(2): 407-429.
- Chamberlin, J., Pender, J. & Yu, B. (2006). Development Domains for Ethiopia: Capturing the Geographical Context of Smallholder Development Options. DSGD Discussion Paper 43, International Food Policy Research Institute (IFPRI), Washington, DC.
- Chang, H-J. (2007). Stranger Than Fiction? Understanding Institutional Changes and Economic Development. Policy Brief no. 6, United Nations University.
- Charness, G., Gneezy, U. & Imas, A. (2013). Experimental Methods: Eliciting Risk Preferences. *Journal of Economic Behavior & Organization* 87 (1): 43-51.
- Chetty, R. & Szeidl, A. (2007). Consumption Commitments and Risk Preferences. *Quarterly Journal of Economics* 122 (2): 831-877.
- Cherry, T. L., Frykblom, P. & Shogren, J. F. (2002). Hardnose the Dictator. *American Economic Review* 92(4): 1218-1221.

- Cherry, T. L., Kröll, S. & Shogren, J. F. (2005). The Impact of Endowment Heterogeneity and Origin on Public Good Contributions: Evidence from the Lab. *Journal of Economic Behavior & Organization* 57(3): 357-365.
- Chiquiar, D. & Hanson, G. H. (2005). International Migration, Self-selection, and the Distribution of Wages: Evidence from Mexico and the United States. *Journal of Political Economy* 113 (2): 239–281.
- Chirayath, L., Sage, C. & Woolcock, M. (2006). Customary Law and Policy Reform: Engaging with the Plurality of Justice Systems. World Development Report Background Paper.
- Coase, R. (1960). The problem of Social Cost. *Journal of Law and Economics* 1(3): 1-44.
- Conley, T. G. & Udry, C. R. (2010). Learning about a New Technology: Pineapple in Ghana. *The American Economic Review*, 100(1): 35-69.
- Cotula, L., Toulmin, C. & Hesse, C. (2004). Land Tenure and Administration in Africa: Lessons of Experience and Emerging Issues. London: International Institute for Environment and Development (iied).
- Couyoumdjian, J. P. (2012). Are Institutional Transplants Viable? An Examination in Light of the Proposals by Jeremy Bentham. *Journal of Institutional Economics* 8(4): 489-509.
- Crewett, W. & Korf, B. (2008). Ethiopia: Reforming Land Tenure. *Review of African Political Economy* 35(116): 203-220.
- Crook, R. C. (2004). Access to Justice and Land Disputes in Ghana's State Courts: The Litigants' Perspective. *Journal of Legal Pluralism and Unofficial Law* 50(1): 1-28.
- Danziger, S., Levav, J. & Avnaim-Pesso, L. (2011). Extraneous Factors in Judicial Decisions. *Proceedings of the National Academy of Sciences of the United States of America* 108(17): 6889-6892.
- Deere, C. D. & Doss, C. R. (2006). Gender and the Distribution of Wealth in Developing Countries. Research paper No. 2006/115, UNU-WIDER, United Nations University.
- Deere, C. D. & Leon, M. (2003). The Gender Asset Gap: Land in Latin America. *World Development* 31(6): 925–947.
- Dehejia, R.H. & Wahba, S. (1999). Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs. *Journal of the American Statistical Association* 94(448): 1053-1062.
- Deininger, K. (2003). Land Policies for Growth and Poverty Reduction. Washington, DC: The World Bank.

- Deininger, K. & Chamorro, J. S. (2004). Investment and Equity Effects of Land Regularization: The Case of Nicaragua. *Agricultural Economics* 30(2): 101-116.
- Deininger, K. & Castagnini, R. (2006). Incidence and Impact of Land Conflict in Uganda. *Journal of Economic Behavior & Organization* 60(3): 321-345.
- Deininger, K. & Jin, S. (2006). Tenure Security and Land-related Investment: Evidence from Ethiopia. *European Economic Review* 50(5): 1245–1277.
- Deininger, K. & Feder, G. (2009). Land Registration, Governance and Development: Evidence and Implications for Policy. *World Bank Research Observer* 24(2): 233-266.
- Deininger, K., Ali, D. A., Holden, S. T. & Zevenbergen, J. (2008). Rural Land Certification in Ethiopia: Process, Initial Impact and Implications for other African Countries. *World Development* 36(10): 1786–1812.
- Deininger, K., Ali, D. A. & Alemu, T. (2011). Impacts of Land Certification on Tenure Security, Investment, and Land Market Participation: Evidence from Ethiopia. *Land Economics* 87(2): 312-334.
- Deininger, K., Ali, D. A. & Alemu, T. (2013). Productivity Effects of Land Rental Market Operation in Ethiopia: Evidence from a Matched Tenant–landlord Sample. *Applied Economics* 45(25): 3531-3551.
- Deininger, K., Ali, D. A. & Yamano, T. (2008). Legal Knowledge and Economic Development: The Case of Land Rights in Uganda. *Land Economics* 84(4): 593–619.
- Demsetz, H. (1967). Toward a Theory of Property Rights. *American Economic Review* 57: 347-359.
- Denrell, J. (2007). Adaptive Learning and Risk Taking. *Psychological Review* 114(1): 177–187.
- Dercon, S. (1996). Risk, Crop choice, and Savings: Evidence from Tanzania. *Economic Development and Cultural Change* 44(3): 485-513.
- Dercon, S. (2002). Income Risk, Coping Strategies and Safety Nets. *World Bank Research Observer* 17 (2): 141-166.
- Dercon, S. (2008). Fate and Fear: Risk and its Consequences in Africa. *Journal of African Economies* 17(suppl 2): ii97-ii127.
- Dercon, S. & Christiaensen, L. (2011). Consumption Risk, Technology Adoption and Poverty Traps: Evidence from Ethiopia. *Journal of Development Economics* 96(2): 159–173.
- Dhillon, A. & Rigolini, J. (2011). Development and the Interaction of Enforcement Institutions. *Journal of Public Economics* 95: 79–87.

- Diamond, J.M. (1997). *Guns, Germs and Steel: The Fate of Human Societies*. New York: W.W. Norton & Co.
- Di Falco, S. & Bulte, E. H. (2011). A Dark Side of Social Capital? Kinship, Consumption, and Savings. *Journal of Development Studies* 47(8): 1128–1151.
- DiPrete, T. A. & Gangl, M. (2004). Assessing Bias in the Estimation of Causal Effects: Rosenbaum Bounds on Matching Estimators and Instrumental Variables Estimation with Imperfect Instruments. *Sociological Methodology* 34(1): 271-310.
- Dixit, A. (2004). *Lawlessness and Economics: Alternative Modes of Governance*. Princeton: Princeton University Press.
- Dixit, A. (2009). Governance Institutions and Economic Activity. *American Economic Review* 99 (1): 3-24.
- Dohmen, T., Huffman, D., Schupp, J., Falk, A., Sunde, U. & Wagner, G. G. (2011). Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences. *Journal of the European Economic Association* 9(3):522–550.
- Donovan, D. A. & Assefa, G. 2003. Homicide in Ethiopia: Human Rights, Federalism, and Legal Pluralism. *American Journal of Comparative Law* 51(1): 505-552.
- Duflo, E. (2003). Grandmothers and Granddaughters: Old Age Pension and Intra-household Allocation in South Africa. *World Bank Economic Review* 17(1): 1-25.
- Duflo, E., Glennerster, R. & Kremer, M. (2008). Using Randomization in Development Economics Research: A Toolkit. In Schultz, T. P. & Strauss, J. A. (Eds.), *Handbook of Development Economics*, Volume 4, Chapter 61. Elsevier.
- Easterly, W. (2006). *The white Man's Burden: Why the West's Efforts to Aid the Rest have Done So Much Ill and So Little Good*. New York: Penguin Press.
- Easterly, W. (2008). Institutions: Top Down or Bottom Up? *American Economic Review: Papers & Proceedings* 98(2): 95–99.
- Easterly, W. & Levine, R. (1997). Africa's Growth Tragedy: Policies and Ethnic Divisions. *Quarterly Journal of Economics* 112(4): 1203-1250.
- Easterly, W. & Levine, R. (2003). Tropics, Germs and Crops: How Endowments Influence Economic Development. *Journal of Monetary Economics* 50: 3–39.
- Eckel, C. C., El-Gamal, M. A. & Wilson, R. K. (2009). Risk Loving after the Storm: A Bayesian-network Study of Hurricane Katrina Evacuees. *Journal of Economic Behavior & Organization* 69(2): 110–124.
- Eckel, C. C. & Grossman, P. J. (2002). Sex Differences and Statistical Stereotyping in Attitudes Toward Financial Risk. *Evolution and Human Behavior* 23(4): 281-295.

- Ege, S. (1997). The Promised Land: The Amhara Land Redistribution of 1997. Working Papers in Ethiopian Development No. 12, Norwegian University of Science and Technology, Dragvoll.
- Evans, P. (2004). Development as Institutional Change: The Pitfalls of Monocropping and the Potentials of Deliberation. *Studies in Comparative International Development* 38(4): 30–52.
- Ezra, M. & Kiros, G. E. (2001). Rural out-migration in the Drought-prone Areas of Ethiopia: A Multi-level Analysis. *International Migration Research* 35(3): 749-771.
- Fafchamps, M. (2010). Vulnerability, Risk Management and Agricultural Development. *African Journal of Agricultural and Resource Economics* 5(1): 243-260.
- Fafchamps, M. (2011). Development, Social Norms, and Assignment to Task. *Proceedings of the National Academy of Sciences of the United States of America* 108(suppl. 4): 21308–21315.
- Fafchamps, M. & Quisumbing, A. R. (2002). Control and Ownership of Assets within Rural Ethiopian Households. *Journal of Development Studies* 38(6): 47-82.
- Fafchamps, M. & Quisumbing, A. R., 2005a. Marriage, Bequest, and Assortative Matching in Rural Ethiopia. *Economic Development and Cultural Change* 53(2):347-380.
- Fafchamps, M. & Quisumbing, A. R. (2005b). Assets at Marriage in Rural Ethiopia. *Journal of Development Economics* 77 (1): 1-25.
- Falk, A. & Szech, N. (2013). Morals and Markets. *Science* 340: 707-711.
- FAO (2002). Gender and Land Rights Database. World Census of Agriculture 2001-2002, Food and Agriculture Organization of the United Nations, Rome.
- Fehr, E. & Gächter, S. (2000). Cooperation and Punishment in Public Goods Experiments. *American Economic Review* 90(4): 980-994.
- Fellner, G. & Maciejovsky, B. (2007). Risk Attitude and Market Behavior: Evidence from Experimental Asset Markets. *Journal of Economic Psychology* 28 (3): 338–350.
- Fenske, J. (2011). Land Tenure and Investment Incentives: Evidence from West Africa. *Journal of Development Economics* 95(2): 137-156.
- Field, E. (2005). Property Rights and Investment in Urban Slums. *Journal of the European Economic Association* 3(2-3): 279-290.
- Franken, J. R. V., Pennings, J. M. E. & Garcia, P. (2014). Measuring the Effect of Risk Attitude on Marketing Behavior. *Agricultural Economics* 45(5): 1–11.
- Frohlich, N., Oppenheimer, J. & Kurki, A. (2004). Modeling Other-Regarding Preferences and an Experimental Test. *Public Choice* 119(1-2): 91-117.

- Future Agricultures (2006). Land, Land Policy and Smallholder Agriculture in Ethiopia. Policy Brief 001.
- Galiani, S. & Schargrodsky, E. (2010). Property Rights for the Poor: Effects of Land Certification. *Journal of Public Economics* 94(9): 700-729.
- Gallup, J., Sachs, J. & Mellinger, A. (1999). Geography and Economic Development. Working paper no. 1. Center for International Development at Harvard University.
- Gangl, M. (2004). Rbounds: Stata module to Perform Rosenbaum Sensitivity Analysis for Average Treatment Effects on the Treated: Department of Economics. Boston College.
- Gantner, A., Güth, W. & Königstein, M. (2001). Equitable Choices in Bargaining Games with Joint Production. *Journal of Economic Behavior & Organization* 46: 209-225.
- Gedzi, V. S. (2012). Women's Property Relations After Intestate Succession PNDC LAW 111 In Ghana. *Research on Humanities and Social Sciences* 2(9): 211-220.
- Gneezy, U. & Potters, J. (1997). An Experiment on Risk Taking and Evaluation Periods. *Quarterly Journal of Economics* 112(2): 631-645.
- Goldstein, M. & Udry, C. 2008. The Profits of Power: Land Rights and Agricultural Investment in Ghana. *Journal of Political Economy* 116(6): 981-1022.
- Gopal, G. & Salim, M. 1998. Gender and Law Eastern Africa Speaks. Washington DC: World Bank.
- Gowok, S. M. (2008). Alternative Dispute Resolution in Ethiopia - A Legal Framework. *African Research Review* 2(2): 265-285.
- Greene, W. H. & Hensher, D. A. (2010). Modeling Ordered Choices: A Primer. New York: Cambridge University Press.
- Greif, A. (2006). Institutions and the Path to the Modern Economy: Lessons from Medieval Trade. Cambridge University Press.
- Greif, A. & Laitin, D. D. (2004). A Theory of Endogenous Institutional Change. *American Political Science Review* 98(4): 633-652.
- Grzymala-Busse, A. (2010). The Best Laid Plans: The Impact of Informal Rules on Formal Institutions in Transitional Regimes. *Studies in Comparative International Development* 45:311-333.
- Gujarati, D. (2003). Basic Econometrics, 4th Edition. New York, NY: McGraw Hill Inc.
- Güth, W., Schmittberger, R. & Schwarze, B. (1982). An Experimental Analysis of Ultimatum Bargaining. *Journal of Economic Behavior & Organization* 3(4): 367-388.

- Guttman, M., Sage, C. M. & Mathiey, C. (2004). Ethiopia Legal and Judicial Sector Assessment. World Bank.
- Hall, R.E. & Jones, C.I. (1999). Why Do Some Countries Produce So Much More Output Per Worker Than Others?. *Quarterly Journal of Economics* 114: 83–116.
- Hallward-Driemeier, M. & Gajigo, O. (2013). Strengthening Economic Rights and Women's Occupational Choice: The Impact of Reforming Ethiopia's Family Law. Policy Research Working Paper 6695, Development Research Group, World Bank.
- Hanaoka, C., Shigeoka, H. & Watanabe, Y. (2014). Do Risk Preferences Change? Evidence from Panel Data before and after the Great East Japan earthquake. Working Paper.
- Harrison, G., Humphrey, S. & Verschoor, A. (2010). Choice under Uncertainty: Evidence from Ethiopia, India and Uganda. *Economic Journal* 120(543): 80–104.
- Harrison, G.W. & List, J.A. (2004). Field Experiments. *Journal of Economic Literature* 42: 1009-1055.
- Harrison, P. (1987). *The Greening of Africa*. London: Paladin Grafton Books.
- Hayek, F. A. (1948). *Individualism and Economic Order*. Chicago, IL: University of Chicago Press.
- Hayek, F. A. (1960). *The Constitution of Liberty*. Chicago, IL: The University of Chicago Press.
- Hayes, J., Roth, M. & Zepeda, L. (1997). Tenure Security, Investment and Productivity in Gambian Agriculture: A Generalized Probit Analysis. *American Journal of Agricultural Economics* 79(2): 369-382.
- He, H., Martinsson, P. & Sutter, M. (2012). Group Decision Making Under Risk: An Experiment with Student Couples. *Economics Letters* 117(3): 691–693.
- Heckman, J., Ichimura, H. & Todd, P. (1997). Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program. *Review of Economic Studies* 64(4): 605-654.
- Heckman, J., Ichimura, H. & Todd, P. (1998). Matching as an Econometric Evaluation Estimator. *Review of Economic Studies* 65(2): 261-294.
- Helmke, G. & Levitsky, S. 2004. Informal Institutions and Comparative Politics: A Research Agenda. *Perspectives on Politics* 2(4): 725-740.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H. & McElreath, R. (2001). In Search of Homo Economicus: Behavioral Experiments in 15 Small-scale Societies. *American Economic Review* 91(2): 73–78.

- Henrich, J., Ensminger, J., McElreath, R., Barr, A., *et al.* (2010). Markets, Religion, Community Size, and the Evolution of Fairness and Punishment. *Science* 327(5972): 1480-1484.
- Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D. & Ziker, J. (2006). Costly Punishment Across Human Societies. *Science* 312(5781) 1767-1770.
- Hicks, J.R. (1969). *A Theory of Economic History*. New York: Oxford University Press.
- Hirano, K., Imbens, G. & Ridder, G. (2003). Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score. *Econometrica* 71 (4): 1161-1189.
- Hoff, K. & Stiglitz, J. (2001). Modern Economic Theory and Development. In Meier, G. & Stiglitz, J., *Frontiers of Development Economics*, pp. 389-460. New York: Oxford University Press.
- Holden, S. T., Deininger, K. & Ghebru, H. (2009). Impacts of Low-cost Land Certification on Investment and Productivity. *American Journal of Agricultural Economics* 91(2): 359-373.
- Holden, S. T., Deininger, K. & Ghebru, H. (2011). Tenure Insecurity, Gender, Low-cost Land Certification and Land Rental Market Participation in Ethiopia. *Journal of Development Studies* 47(1): 31-47.
- Holden, S. T. & Tefera, T. (2008). From Being Property of Men to Becoming Equal Owners? Early Impacts of Land Registration and Certification on Women in Southern Ethiopia. Research report, Land Tenure and Property Administration Section, UNHABITAT.
- Holden, S. T. & Yohannes, H. (2002). Land Redistribution, Tenure Insecurity, and Intensity of Production: A Study of Farm Households in Southern Ethiopia. *Land Economics* 78(4): 573-590.
- Holt, C. A. & Laury, S. K. (2002). Risk Aversion and Incentive Effects. *American Economic Review* 92(5): 1644-1655.
- Humphrey, S. & Verschoor, A. (2004). Decision-making under Risk among Small Farmers in East Uganda. *Journal of African Economies* 13(1): 44-101.
- Hunnes, D. E. (2012). Understanding Rural-to-urban Migration in Ethiopia: Driving Factors, Analytical Frameworks, and Recommendations. *Journal of Global Health Perspectives* 1: 1-10.
- Jacoby, H. G. & Minten, B. (2007). Is Land Certification in Sub-Saharan Africa Cost-Effective? Evidence from Madagascar. *World Bank Economic Review* 21(3): 461-485.

- Jacoby, H. G., Li, G. & Rozelle, S. (2002). Hazards of Expropriation: Tenure Insecurity and Investment in Rural China. *American Economic Review* 92(5): 1420-1447.
- Jamieson, B. D. (1969). The Influences of Birth Order, Family Size and Sex Differences on Risk-taking Behavior. *British Journal of Social and Clinical Psychology* 8(1): 1-8.
- Joireman, S.F. (2001). Property Rights and the Role of the State: Evidence from the Horn of Africa. *Journal of Development Studies* 38 (1): 1-28.
- Joireman, S. F. (2008). The Mystery of Capital Formation in Sub-Saharan Africa: Women, Property Rights and Customary Law. *World Development* 36(7): 1233-1246.
- Jones, C. (1997). On the Evolution of the World Income Distribution. *Journal of Economic Perspectives* 11(3): 19-36.
- Kabeer, N. (1999). Resources, Agency, Achievements: Reflections on the Measurement of Women's Empowerment. *Development and Change* 30(3): 435-464.
- Kahneman, D., Knetsch, J. L. & Thaler, R. H. (1990). Experimental Tests of the Endowment Effect and the Coase Theorem. *Journal of Political Economy* 98(6): 1325-1348.
- Kahneman, D. & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47(2): 263-291.
- Karlan, D., Osei, R., Osei-Akoto, I. & Udry, C. (2014). Agricultural Decisions after Relaxing Credit and Risk Constraints. *Quarterly Journal of Economics* 129 (2): 597-652.
- Khandker, S. R., Koolwal, G. B. & Samad, H. A. (2010). Handbook on Impact Evaluation: Quantitative Methods and Practices. Washington, DC: World Bank.
- Kingston, C., Caballero, G. (2009). Comparing Theories of Institutional Change. *Journal of Institutional Economics* 5(2): 151-180.
- Kumar, K. B. & Matsusaka, J. G. (2009). From Families to Formal Contracts: An Approach to Development. *Journal of Development Economics* 90 (1): 106-119.
- Kumar, N. & Quisumbing, A. R. (2012). Beyond "Death Do Us Part": The long-term Implications of Divorce Perceptions on Women's Well-being and Child Schooling in Rural Ethiopia. *World Development* 40 (12): 2478-2489,
- Kumar, N. & Quisumbing, A. R. (2015). Policy Reform towards Gender Equality in Ethiopia: Little by little, the Egg Begins to Walk. *World Development* 67: 406-423.
- Kuran, T. (1989). Sparks and Prairie Fires: A Theory of Unanticipated Political Revolution. *Public Choice* 61: 41-74.
- Kuyu, C. M. 2005. À la Recherche du Droit Africain du XXIe Siècle. Connaissances et Savoirs.

- Lamb, R. (2003). Fertilizer use, Risk and Off-farm Labor Markets in the Semi-arid Tropics of India. *American Journal of Agricultural Economics* 85(2): 359–371.
- Landeo, C. M., Nikitin, M. & Babcock, L. (2007). Split-Awards and Disputes: An Experimental Study of a Strategic Model of Litigation. *Journal of Economic Behavior & Organization* 63(3): 553-572.
- Leiva, A. J. & Skees, J. R. (2008). Using Irrigation Insurance to Improve Water Usage of the Rio Mayo Irrigation System in Northwestern Mexico. *World Development* 36(12): 2663-2678.
- Lemel, H. (1988). Land Certification: Conceptual, Empirical and Policy. *Land Use Policy* 5(3): 273-290.
- Leuven, E. & Sianesi, B. (2003). PSMATCH2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing and Covariate Imbalance Testing.
- Lewis, W. A. (1955). *The Theory of Economic Growth*. London: George Allen & Unwin Ltd.
- List, J. A. (2003). Does Market Experience Eliminate Market Anomalies? *Quarterly Journal of Economics* 118(1): 41-71.
- List, J. A. (2004). Neoclassical Theory versus Prospect Theory: Evidence from the Marketplace. *Econometrica* 72(2): 615-625.
- List, J. A. (2006). Field Experiments: A Bridge Between Lab and Naturally-Occurring Data. *B.E. Journal of Economic Analysis & Policy* 6(2).
- List, J. A. (2011). Does Market Experience Eliminate Market Anomalies? The case of Exogenous Market Experience. *American Economic Review: Papers & Proceedings* 101(3): 313–317.
- List, J. A. (2014). Using Field Experiments to Change the Template of How We Teach Economics. *Journal of Economic Education* 45(2): 81-89.
- List, John A. & Millimet, D. L. (2008). The Market: Catalyst for Rationality and Filter of Irrationality. *BE Journal of Economic Analysis and Policy* 8(1): Article 47.
- List, J. A. & Rasul, I. (2011). Field Experiments in Labor Economics. In Ashenfelter, O. & Card, D. (Eds.), *Handbook of Labor Economics Volume 4a*, pp. 104-228. Elsevier.
- Lucas, R. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics* 22(1): 3-42.
- Luhan, W., Kocher, M. & Sutter, M. (2009). Group Polarization in the Team Dictator Game Reconsidered. *Experimental Economics* 12(1): 26-41.

- Lundberg, S. & Pollak, R. A. (1993). Separate Spheres Bargaining and the Marriage Market. *Journal of Political Economy* 101 (6): 988-1010.
- MacCrimmon, K.R. & Wehrung, D.A. (1990). Characteristics of Risk-taking Executives. *Management Science* 36(4): 422-435.
- Machovec, F. (1995). Perfect Competition and the Transformation of Economics. New York, NY: Routledge.
- Mahmud, S., Shah, N. M. & Becker, S. (2012). Measurement of Women's Empowerment in Rural Bangladesh. *World Development*, 40(3), 610-619.
- Manser, M. & Brown, M. (1980). Marriage and Household Decision-making: A Bargaining Analysis. *International Economic Review* 21(1): 31-44.
- Martínez, C. A. (2013). Intrahousehold Allocation and Bargaining Power: Evidence from Chile. *Economic Development and Cultural Change* 61(3): 577-605.
- Mason, K. O. (1986). The Status of Women: Conceptual and Methodological Issues in Demographic Studies. *Sociological Forum* 1(2): 284-300.
- Mberu, B. U. (2006). Internal Migration and Household Living Conditions in Ethiopia. *Demographic Research* 14(21): 509-540.
- McElroy, M. B. (1990). The Empirical Content of Nash-bargained Household Behavior. *Journal of Human Resources* 25(4): 559-583.
- McElroy, M. B. & Horney, M. J. (1981). Nash-bargained Household Decisions: Toward a Generalization of the Theory of Demand. *International Economic Review* 22(2): 333-349.
- Melesse, B. M. & Bulte, E. (2015). Does Land Registration and Certification Boost Farm Productivity? Evidence from Ethiopia. *Agricultural Economics*. In Press.
- Merry, S. E. (1988). Legal Pluralism. *Law and Society Review* 22(5): 869-896.
- Michalopoulos, S. & Papaioannou, E. (2014). National Institutions and Subnational Development in Africa. *Quarterly Journal of Economics* 129(1): 151-213.
- Migot-Adholla, S., Hazell, P., Blarel, B. & Place, F. (1991). Indigenous Land Rights Systems in Sub-Saharan Africa: A Constraint on Productivity? *World Bank Economic Review* 5(1): 155-175.
- Morozumi, R. (2012). A Test of a Unitary Model on Labour Supply Using the Information of Household Decision-making Systems. *Applied Economics* 44(33): 4291-4300.
- Narayan, D. (2005). Measuring Empowerment: Cross-disciplinary Perspectives. The World Bank, Washington DC.

- Ndulu, B.J. & O’Connell, S.A. (2008). Policy Plus: African Growth Performance 1960–2000. In Ndulu, B.J., O’Connell, S.A., Bates, R.H., Collier, P., Soludo, C.C. (Eds.), *The Political Economy of Economic Growth in Africa 1960–2000*, pp. 3–75. Cambridge: Cambridge University Press.
- Netzer, N. (2009). Evolution of Time Preferences and Attitudes toward Risk. *American Economic Review* 99(3): 937–955.
- Njuki, J., Poole, J., Johnson, N., Baltenweck, I., Pali, P., Lokman, Z. & Mburu, S. (2011). Gender, Livestock and Livelihood Indicators. International Livestock Research Institute (ILRI), Nairobi.
- North, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge, UK: Cambridge University Press.
- North, D. C. (1991). Institutions. *Journal of Economic Perspectives* 5(1): 97-112.
- North, D. C. (1994). Economic Performance Through Time. *American Economic Review* 84: 359-368.
- North, D. C. (2005). *Understanding the Process of Economic Change*. Princeton: Princeton University Press.
- North, D. & Thomas, R. P. (1973). *The Rise of the Western World*. Cambridge: Cambridge University Press.
- Obeng-Odoom, F. (2012). Land Reforms in Africa: Theory, Practice, and Outcome. *Habitat International* 36(1): 161-170.
- Ordioni, N. (2005). Pauvreté et Inégalités de Droits en Afrique: une Perspective “Genrée.” *Mondes En Développement* 129(1): 93-106.
- Olson, M. (1971). *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, Mas : Harvard University Press.
- Orloff, A. S. (1993). Gender and the Social Rights of Citizenship: The Comparative Analysis of Gender Relations and Welfare States. *American Sociological Review* 58(3): 303-328.
- Oxoby, R. J. & Spraggon, J. (2008). Mine and Yours: Property Rights in Dictator Games. *Journal of Economic Behavior & Organization* 65(3-4): 703-713.
- Park, C. (2007). Marriage Market, Parents’ Bargaining Powers, and Children’s Nutrition and Education. *Oxford Bulletin of Economics and Statistics* 69(6): 773-793.
- Pejovich, S. (1999). The Effects of the Interaction of Formal and Informal Institutions on Social Stability and Economic Development. *Journal of Markets & Morality* 2(2): 164-181.

- Pighin, S., Bonini, N., Savadori, L., Hadjichristidis, C., Antonetti, T. & Schena, F. (2012). Decision Making under Hypoxia: Oxygen Depletion Increases Risk Seeking for Losses but not for Gains. *Judgment and Decision Making* 7(4): 472-477.
- Pimentel, D. (2010). Rule of Law Reform Without Cultural Imperialism? Reinforcing Customary Justice Through Collateral Review in Southern Sudan. *Hague Journal on the Rule of Law* 2(1): 1-28.
- Place, F. (2009). Land Tenure and Agricultural Productivity in Africa: A Comparative Analysis of the Economics Literature and Recent Policy Strategies and Reforms. *World Development* 37(8): 1326-1336.
- Place, F. & Hazell, P. (1993). Productivity Effects of Indigenous Land Tenure Systems in Sub-Saharan Africa. *American Journal of Agricultural Economics* 75(10): 10-19.
- Place, F. & Migot-Adholla, S. E. (1998). The Economic Effects of Land Registration on Smallholder Farms in Kenya: Evidence from Nyeri and Kakamega Districts. *Land Economics* 74(3): 360-373.
- Platteau, J.-P. (1994). Behind the Market Stage Where Real Societies Exist: Part II—The Role of Moral Norms. *Journal of Development Studies* 30(3):753–815.
- Platteau, J.-P. (2000). Institutions, Social Norms and Economic Development. Amsterdam: Harwood Academic Publishers.
- Platteau, J.-P. (2009). Institutional Obstacles to African Economic Development: State, Ethnicity, and Custom. *Journal of Economic Behavior & Organization* 71(3): 669-689.
- Polinsky, M. A. & Shavell, S. (2007). The Theory of Public Enforcement of Law. In Polinsky, M. A. & Shavell, S. (Eds.), *Handbook of Law and Economics Vol. 1*, Pp. 403–454. Elsevier.
- Quisumbing, A. R. & Maluccio, J. A. (2003). Resources at Marriage and Intrahousehold Allocation: Evidence from Bangladesh, Ethiopia, Indonesia, and South Africa. *Oxford Bulletin of Economics and Statistics* 65(3): 283–327.
- Rahmato, D. (2004). Searching for Tenure Security? The Land System and New Policy Initiatives in Ethiopia. Discussion Paper No. 12, Forum for Social Studies (FSS), Addis Ababa.
- Rahmato, D. (2009a). Peasants and Agrarian Reforms: The Unfinished Quest for Secure Land Rights in Ethiopia. In Ubink, J. M., Hoekema, A. J. & Assies, W. J. (Eds.), *Legalizing Land Rights: Local Practices, State Responses and Tenure Security in Africa, Asia and Latin America*, Pp. 33-57. Leiden University Press.

- Rahmato, D. (2009b). Land Rights and Tenure Security: Rural Land Registration in Ethiopia. In Ubink, J. M., Hoekema, A. J. & Assies, W. J. (Eds.), *Legalizing Land Rights: Local Practices, State Responses and Tenure Security in Africa, Asia and Latin America*, Pp. 59-95. Leiden University Press.
- Rao, V. (1993). The Rising Price of Husbands: A Hedonic Analysis of Dowry Increases in Rural India. *Journal of Political Economy* 101(4): 666-677.
- Renkow, M., Hallstrom, D. G. & Karanja, D. D. (2004). Rural Infrastructure, Transactions Costs and Market Exposure in Kenya. *Journal of Development Economics* 73(1): 349–367.
- Richman, B. D. (2012). Norms and Law: Putting the Horse Before The Cart. *Duke Law Journal* 62(1): 739-766.
- Rodrik, D. (1990). How Should Structural Adjustment Programs be Designed? *World Development* 18(7): 933-947.
- Rodrik, D. (1999). Where Did All the Growth Go? External Shocks, Social Conflict and Growth Collapses. *Journal of Economic Growth* 4: 385–412.
- Rodrik, D. (2000). Institutions for High-Quality Growth: What They Are and How to Acquire Them. *Studies in Comparative International Development* 35(3): 3–31.
- Rodrik, D. (2003). What Do We Learn from Country Narratives? In Rodrik, D. (ed.), *In Search of Prosperity: Analytic Narratives on Economic Growth*, PP. 1-19. Princeton, NJ: Princeton University Press.
- Rodrik, D. (2006). Goodbye Washington Consensus, Hello Washington Confusion. *Journal of Economic Literature* 44: 973–87.
- Rodrik, D., Subramanian, A. & Trebbi, F. (2004). Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development. *Journal of Economic Growth* 9: 131–165.
- Roland, G. (2004). Understanding Institutional Change: Fast-Moving and Slow-Moving Institutions. *Studies in Comparative International Development* 38(4): 109-131.
- Romer, P. M. (1986). Increasing Returns and Long Run Growth. *Journal of Political Economy* 94(5): 1002–1037.
- Romer, P. M. (1990). Endogenous Technical Change. *Journal of Political Economy* 98: S71-S102.
- Rosenbaum, P.R. (2002). *Observational Studies*. New York: Springer.
- Rosenbaum, P. R. & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70(1): 41-50.

- Rosenbaum, P.R. & Rubin, D.B. (1985). Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score. *The American Statistician*, 39(1): 33-38.
- Rostow, W. W. (1963). *The Economics of Take-off into Sustained Growth*. New York: St. Martin's Press.
- Rutström, E. E. & Williams, M. B. (2000). Entitlements and Fairness: An Experimental Study of Distributive Preferences. *Journal of Economic Behavior & Organization* 43(1): 75-89.
- Sacco, R. (2008). *Anthropologie Juridique, Apport à une Macro-Histoire du Droit*. Dalloz.
- Sachs, J. D. (2003). *Institutions Don't Rule: Direct Effects of Geography on Per Capita Income*. NBER Working Paper 9490, National Bureau of Economic Research, Cambridge.
- Sachs, J. & Malaney, P. (2002). The Economic and Social Burden of Malaria. *Nature* 415: 680-685.
- Sachs, J. D. & Warner, A. M. (1997). Sources of Slow Growth in African Economies. *Journal of African Economies* 6(3): 335-376.
- Sandefur, J. & Siddiqi, B. (2013). *Delivering Justice to the Poor: Theory and Experimental Evidence from Liberia*. Working Paper.
- Schmitt, P. M. (2004). On Perceptions of Fairness: The Role of Valuations, Outside Options, and Information in Ultimatum Bargaining. *Experimental Economics* 7(1): 49-73.
- Sen, A. (1990). *Gender and Cooperative Conflicts*. In Tinker, I. (Ed.), *Persistent Inequalities: Women and World Development*. Oxford University Press.
- Sen, A. (1999). *Development as Freedom*. New York: A. Knopf Publisher.
- Shell-Duncan, B., Wander, K., Hernlund, Y. & Moreau, A. (2013). Legislating Change? Responses to Criminalizing Female Genital Cutting in Senegal. *Law & society review* 47(4): 803-835.
- Sjaastad, E. & Bromley, D. W. (1997). Indigenous Land Rights in Sub-Saharan Africa: Appropriation, Security and Investment Demand. *World Development* 25(4): 549-562.
- Slembeck, T. & Tyran, J.-R. (2004). Do Institutions Promote Rationality? An Experimental Study of the Three-door Anomaly. *Journal of Economic Behavior & Organization* 54(3): 337-350.
- Smith, A. (1776). *The Wealth of Nations*. New York, NY: The Modern Library.
- Smith, J. & Todd, P. (2005). Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators? *Journal of Econometrics* 125(1-2): 305-353.

- Smith, V. L. (1982). Microeconomic Systems as an Experimental Science. *American Economic Review* 72(5): 923-955.
- Solow, R. (1956). A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics* 70(1): 65–94.
- Spence, M. (2002). Signaling in Retrospect and the Informational Structure of Markets. *American Economic Review* 92(3): 434-459.
- Staiger, D. & Stock, J. (1997). Instrumental Variables Regression with Weak Instruments. *Econometrica* 65(3): 557-586.
- Stigler, G. J. & Becker, G. S. (1977). De Gustibus Non Est Disputandum. *American Economic Review* 67(2): 76-90.
- Tamanaha, B. (2008). Understanding Legal Pluralism: Past to Present, Local to Global. *Sydney Law Review* 30(1): 375-411.
- Tanaka, T, Camerer, C. & Nguyen, Q. (2010). Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam. *American Economic Review* 100(1): 557-571.
- Theil, H. (1971). Principles of Econometrics. New York, NY: John Wiley & Sons, Inc.
- Thomas, D. (1990). Intrahousehold Resource Allocation: An Inferential Approach. *Journal of Human Resources* 25 (4): 635–664.
- Toulmin, C. (2009). Securing Land and Property Rights in Sub-Saharan Africa: The Role of Local Institutions. *Land Use Policy* 26(1): 10-19.
- UNICEF (2010). Legislative Reform to Support the Abandonment of Female Genital Mutilation/Cutting. New York: UNICEF.
- Varian, H. R. (1982). The Nonparametric Approach to Demand Analysis. *Econometrica* 50(4): 945–73.
- Voors, M. & Bulte, E.H. (2014). Conflict and the Evolution of Institutions: Unbundling Institutions at the Local Level in Burundi. *Journal of Peace Research* 51: 455-469.
- Voors, M., Nillesen, E., Verwimp, P., Bulte, E., Lensink, R. & Van Soest, D. (2012). Violent Conflict and Behavior: A Field Experiment in Burundi. *American Economic Review* 102 (2): 941–64.
- Wantchekon, L. & Guardado, J. R. (2011). Methodology Update: Randomised Controlled Trials, Structural Models and the Study of Politics. *Journal of African Economies* 20(4): 653-672.
- Weldon, S. L. (2002). Beyond Bodies: Institutional Sources of Representation for Women in Democratic Policymaking. *Journal of Politics* 64(4): 1153-1174.

- Wiig, H. (2013). Joint Titling in Rural Peru: Impact on Women's Participation in Household Decision-making. *World Development* 52: 104-119.
- Wiig, H., Braten, R. & Fuentes, D. O. (2011). The impacts of Land on Women's Empowerment in Peruvian Communities. *A Background Paper for the World Development Report 2012*, World Bank, Washington DC.
- Wik, M., Kebede, T. A., Bergland, O. & Holden, S. T. (2004). On the Measurement of Risk Aversion from Experimental Data. *Applied Economics* 36(21): 2443-2451.
- Williamson, C. R. (2009). Informal Institutions Rule: Institutional Arrangements and Economic Performance. *Public Choice* 139: 371-387.
- Williamson, O. (2000). The New Institutional Economics: Taking Stock, Looking Ahead. *Journal of Economic Literature* XXXVIII: 596-613.
- Wojkowska, E. (2006). Doing Justice: How Informal Justice Systems Can Contribute. UNDP Oslo Governance Centre.
- World Bank (2002). Empowerment and Poverty Reduction: A Sourcebook. Washington, DC.
- World Bank (2006). Land Policy and Administration. Washington DC: World Bank.
- World Bank (2014). Risk and Opportunity: Managing Risk for Development. World Development Report 2014, World Bank.
- Yngstrom, I. (2002). Women, Wives and Land Rights in Africa: Situating Gender Beyond the Household in the Debate over Land Policy and Changing Tenure Systems. *Oxford Development Studies* 30(1): 21-40.
- Zenger, T. R., Lazzarini, S. G. & Poppo, L. (2002). Informal and Formal Organization in New Institutional Economics. *Advances in Strategic Management* 19: 277-306.

Summary

This thesis studies the intricate relationship between institutional change and economic development. Institutions have shown to provide the essential underpinnings of differences in economic growth rates and development paths across countries. Evidently, a large empirical research has established a persuasive link between institutions and economic development. With the claim that institutions matter for economic development has settled, the issue of building necessary institutions for development has become a first-order relevant theme in development economics research. In this thesis, we have sought to shed some light on this theme at a micro-level of analysis. To do so, we rely on a relatively large dataset that we collected employing experimental and survey approaches in Ethiopia. Chapter 1 sets the stage for the thesis. It introduces key concepts and theoretical foundations of the research topic, outlining the main research questions and the methodology of the thesis. It also elaborates the emphasis given to institutions in the economic growth process over time.

Chapter 2 looks at the impact of the Ethiopian land registration and certification program on agricultural productivity, using propensity score matching to account for selection bias. We have been able to provide confirmation for the hypothesis that formal land rights are important for economic development. Land registration and certification has produced robust positive effects on farm productivity. A simple cost-benefit analysis shows that benefits significantly exceed the cost of implementation of the program. Based on reasoning that eliminates alternative mechanisms, we contend that the assurance effect of secured rights is the most likely channel for our results. After a thorough review of the land titling literature in Africa, we conclude that formalization of land rights is a meaningful institutional development when there is demand for more tenure security and its implementation is based on the recognition of existing informal land rights and local conditions.

Chapter 3 discusses and investigates yet another impact of land titling programs on development outcomes. We connect land rights with the distribution of power in the household organization by studying the impact of joint land certification (i.e., titles containing the names of both spouses) on the bargaining position of women in the household and beyond. We use both non-parametric and parametric approaches in our analysis and present a

detailed counterfactual analysis to control for selection bias. The evidence presented attests that joint land certification increases the bargaining position of women, but not equally across all dimensions. The effects are found to be stronger for indices of empowerment that indicate women's participation and roles outside the home. Considering the results of Chapters 2 and 3 together, we conclude that the Ethiopian land certification program has produced significant efficiency and equity gains.

In Chapter 4, we study how an increased threat of competition from the formal law affects the performance of customary legal institutions. We investigate the impacts of introducing a formal law on the arbitration decisions of real customary judges. We find that introducing a legal fallback reduces arbitration biases and draws the decisions of customary judges significantly closer to the formal law itself. This implies that the extent to which informal institutions behave dictatorially reflects not only the custom—their source of power—but also the contemporary institutional context in which they operate. We interpret the results as evidence that customary judges respond to incentives, and that they can be manipulated purposely through social engineering efforts. As such, local customary institutions, which are often seen as governance problems, can be part of the solution—they may have a role to play in shifting preexisting customs towards a desired outcome. In contrast, individuals disfavored by the custom not only failed to take advantage of their increased bargaining power but also made decisions that are less beneficial to themselves. This unintended effect of the penetration of the law is consistent with the presence of strong social sanctioning against norm-deviating behavior. We then speculate that, in societies where formal legal institutions have limited outreach and the custom remains strong, most benefits of increased competition between formal law and customary legal institutions may come from changes in the latter, rather than from plaintiffs seeking justice under the rule of law.

Chapter 5 focuses on analyzing the effect of market exposure on risk preferences of poor rural households in Ethiopia. We find evidence that indicates market exposure influences and shapes risk preferences of farm households. Precisely, we show that market exposure attenuates risk aversion, and farm households with higher degree of market exposure tend to be more risk tolerant. As such, penetration of markets in remote areas may be more than a means to better allocate resources. Markets may affect the way individuals make decisions through affecting their risk preferences. We then conjecture that a rapid market penetration may offer an alternative mechanism that may potentially link market exposure to economic growth—via attenuating risk aversion. It may specifically induce the possibility of dynamic

efficiency gains, through a chain of externalities that could sustain a self-propelling trajectory of market integration and development at the local level. Finally, Chapter 6 provides a synthesis of the core chapters. It discusses key findings and scientific insights as well as directions for future research.

Overall, the results of this thesis reveal that formal and informal institutions interact in a variety of ways with both constructive and deleterious repercussions for economic development. The debate in the literature that formal and informal institutions are either substitutes or complements may be ill-conceived. Instead, a practically relevant institutional development effort requires thinking beyond institutional designs and forms. There is a special need to recognize the importance of the complexity and dynamics of the interaction between formal and informal institutions. In particular, a better understanding of such interaction effects can help to attenuate the risk of missing many of the important incentives and constraints that underlie economic behavior, especially in many African countries where informal institutions often undermine the credibility of formal institutions. Moreover, the outcomes of institutional changes may be unpredictable—some expected outcomes may not be happening while other unintended outcomes occurring. Equally, institutional changes may be resisted for many reasons, including conflicts with the custom, the interests of current beneficiaries, or the sheer force of habits and routines—‘institutional stickiness’. In this light, failure in institutional changes may have more to do with whether or not individuals have had the incentive to mutually form expectations to accept institutions than with whether or not specific institutional changes are ‘correct’.

Samenvatting

Dit proefschrift behandelt het verband tussen institutionele verandering en economische ontwikkeling. Instituties spelen een belangrijke rol bij het verklaren van verschillen in economische groei en ontwikkelingstrajecten tussen landen. Empirisch onderzoek heeft een duidelijk verband aangetoond tussen instituties en economische ontwikkeling. Nu het duidelijk is dat instituties een rol spelen in ontwikkeling, is het bouwen van de instituties die noodzakelijk zijn voor ontwikkeling een belangrijk onderzoeksthema in de ontwikkelingseconomie. Middels dit proefschrift dragen we bij aan dit thema met een analyse op microniveau. Daarvoor gebruiken we een relatief grote dataset die met gebruik van experimentele en enquête methoden verzameld is in Ethiopië. Hoofdstuk 1 legt de basis voor het proefschrift. Hierin worden sleutelbegrippen en de theoretische onderbouwing van het onderzoeksonderwerp geïntroduceerd met een uiteenzetting van de onderzoeksvragen en methodologie van het proefschrift. Verder wordt de nadruk die gelegd wordt op de rol van instituties binnen het economische groeiproces verder uitgewerkt.

In Hoofdstuk 2 wordt gekeken naar de impact van het land registratie en certificatie programma op de productiviteit in de landbouw. Hierbij wordt Propensity Score Matching toegepast om mogelijke selectie effecten te voorkomen. Mijn bevindingen bevestigen de hypothese dat formele landrechten belangrijk zijn voor economische ontwikkeling. Registratie en certificatie van land leidt tot robuuste positieve effecten voor de landbouwproductiviteit. Een eenvoudige kosten-batenanalyse toont aan dat de voordelen significant groter zijn dan de implementatiekosten van het programma. We beredeneren dat alternatieve mechanismen uitgesloten zijn, en concluderen hieruit dat het verzekeringseffect van landrechten het meest waarschijnlijke mechanisme is achter onze resultaten. Na een grondige bespreking van de literatuur over landrechten in Afrika concluderen we dat formalisatie van landrechten een belangrijke institutionele ontwikkeling is indien er vraag is voor meer zekere landrechten en de implementatie gebaseerd is op erkenning van bestaande informele rechten en lokale omstandigheden.

In Hoofdstuk 3 wordt een ander effect van landregistratieprogramma's op ontwikkelingsuitkomsten besproken en onderzocht. We brengen landrechten in verband met

de verdeling van macht binnen het huishouden, door het effect te bestuderen van gezamenlijke landcertificatie (ofwel eigendomsaktes met de namen van beide partners) op de onderhandelingspositie van vrouwen binnen en buiten het huishouden. We gebruiken zowel niet-parametrische als parametrische methoden in onze analyse en gebruiken een middels Propensity Score Matching geconstrueerde controlegroep om te controleren voor selectie effecten. Het gepresenteerde bewijs bevestigt dat gezamenlijke certificering de onderhandelingspositie van vrouwen versterkt, maar niet over alle dimensies. De effecten zijn sterker voor indicatoren van empowerment op het gebied van participatie en rollen buiten het huis. De resultaten van hoofdstuk 2 en 3 tezamen genomen stellen ons in staat te concluderen dat het Ethiopische land certificatie programma significante positieve effecten heeft gehad, zowel op het gebied van efficiëntie als op het gebied van eerlijkheid.

In Hoofdstuk 4 bestuderen we hoe toegenomen concurrentie van het formele recht de prestaties van het gewoonrecht beïnvloedt. We onderzoeken de impact van het introduceren van formeel recht op de arbitragebeslissingen van rechters binnen het gewoonrecht. We vinden dat de introductie van een wettelijke terugvalmogelijkheid partijdigheid in arbitrage vermindert en de beslissingen binnen het gewoonrecht dichter bij het formele recht brengt. Dit impliceert dat de mate waarin informele instituties zich dictatoriaal gedragen niet de gebruiken waar ze hun macht aan ontleneen weerspiegelt, maar ook de institutionele context waarbinnen zij opereren. We interpreteren dit als bewijs dat rechters binnen het gewoonrecht reageren op prikkels, en dat de maatschappij in dit opzicht dus maakbaar is. Zodoende kunnen de instituties binnen het gewoonrecht, die doorgaans gezien worden als obstakels voor goed bestuur, deel zijn van de oplossing, aangezien ze een rol kunnen spelen in het verschuiven van bestaande gebruiken naar een gewenste uitkomst. Echter, individuen die door het gewoonrecht in het nadeel zijn gesteld, maakten geen gebruik van hun verbeterde onderhandelingspositie en maakten bovendien beslissingen die niet in hun eigen voordeel waren. Dit onbedoelde effect van de uitbreiding van het bereik van de wet is consistent met de aanwezigheid van sterke sanctionering van normafwijkend gedrag. Vervolgens speculeren we dat in samenlevingen waar de formele wet een beperkt aanwezig is en het gewoonrecht een sterke positie heeft, de voordelen van toegenomen concurrentie tussen formeel recht en gewoonrecht eerder komen van veranderingen in het laatste, dan van het zoeken van rechtshulp binnen het formele recht.

Hoofdstuk 5 richt zich op het effect van markttoegang op risicovoorkeuren van arme rurale huishoudens in Ethiopië. We vinden bewijs dat suggereert dat het hebben van markttoegang

invloed heeft op risicovoorkeuren van huishoudens. We tonen aan dat markttoegang risico aversie vermindert, en dat huishoudens met betere markttoegang geneigd zijn meer risico tolerant zijn. Zodoende kan het creëren van markten in afgelegen gebieden meer zijn dan alleen een middel om hulpbronnen beter toe te wijzen. Markten kunnen de manier waarop mensen beslissingen nemen, beïnvloeden doordat zij de risicovoorkeuren van mensen aanpassen. We vermoeden dat het aanpassen van risicovoorkeuren een alternatief mechanisme is dat de verspreiding van de markt verbindt met economische groei. Het biedt de mogelijkheid voor winsten in dynamische efficiëntie door een keten van externe effecten dat een zelf-versterkend traject van markt integratie en ontwikkeling op het lokale niveau op gang kan starten. Tot slot biedt Hoofdstuk 5 een synthese van de romp hoofdstukken. Besproken worden de hoofdresultaten en wetenschappelijke inzichten naast mogelijke richtingen voor toekomstig onderzoek.

De resultaten in dit proefschrift tonen aan dat formele en informele instituties op verschillende wijzen op elkaar inwerken, met zowel positieve als negatieve effecten voor economische ontwikkeling tot gevolg. Het is goed mogelijk dat de vraag in de literatuur of formele en informele instituties complementen dan wel substituten zijn ondoordacht is. Voor institutionele ontwikkeling met praktische relevantie is het juist nodig om verder te denken dan institutionele ontwerpen en vormen. Er bestaat een behoefte om het belang van de complexiteit en dynamiek van de wisselwerking tussen formele en informele instituties te erkennen. In het bijzonder kan een beter begrip van deze wisselwerking bijdragen aan het verminderen van het risico dat belangrijke prikkels en beperkingen die ten grondslag liggen aan economisch gedrag over het hoofd worden gezien, vooral in Afrikaanse landen, waar informele instituties de geloofwaardigheid van formele instituties vaak ondermijnen. Bovendien kunnen de uitkomsten van institutionele veranderingen onvoorspelbaar zijn – het is mogelijk dat verwachte uitkomsten niet gerealiseerd worden terwijl andere ongewenste uitkomsten zich wel manifesteren. Tegelijkertijd kan er om vele redenen verzet optreden tegen institutionele verandering, zoals conflicten met informele rechten, belangen van huidige begunstigen, of de brute kracht van gewoontes en routines – ‘*institutional stickiness*’. Zo bekeken heeft het falen van institutionele verandering misschien meer te maken met de vraag of individuen de prikkel hebben gehad om verwachtingen op te bouwen aangaande de gezamenlijke aanvaarding van instituties, dan met de vraag of de specifieke institutionele veranderingen de ‘juiste’ waren.

Acknowledgements

I would like to express my appreciation to a longlist of people from whom I have benefited much in one way or another while conducting this study. Prof. Erwin Bulte, my supervisor and promotor, deserves to be in the first row of this queue. Erwin, thank you for allowing me a perfect academic freedom to explore the dense ‘forest’ by myself but making sure that I am on the right track with your exceptional guidance. I have had the great pleasure of working with you. I could never have imagined having a better mentor for my PhD study.

I am profoundly grateful to those who contributed to the chapters of this thesis, either by collaborating on co-authoring papers or giving their insightful comments. I cheerfully acknowledge Erwin Bulte, Francesco Cecchi, Adane Dabissa, Eskender Andualem, Hans-Peter Weikard, Janneke Pieters, Jeroen Klomp, Koen Leuveld (particularly for translating the summary to Dutch), Niccolo Meriggi, Rein Haagsma, Maarten Voors, Michel Handgraaf, Robert Lensink and my thesis committee members, Justus Wesseler, Stein Holden, Jean-Marie Baland and Patricio Dalton, as well as my officemates over the last four years, John, Prossy, Aisha and Fausta. I would also like to cordially thank all my fellow PhD students, colleagues and the admin staff of the Development Economics Group.

My sincere thanks go to the more than one thousand respondents, who participated so patiently in the lengthy survey questionnaires and experimental games. My local research assistants were an awesome crew. I would like to take this opportunity to offer my sincere gratefulness to them. Special mention goes to Awoke Biset, Gebregiorgis and Agegnehu for their unreserved efforts in coordinating and leading the field teams.

I seek to make a separate paragraph here for my friends. A heartfelt word of thanks goes to my lifetime friends: Dessie, Worku, Gebiaw, Endalew, Akalie, Chombie, Bayuh and Gizachew. I am ever grateful to you all for everything that we have shared since the early years of our schooling, and will always treasure our great memories. Enough cannot be said about Hibret, who has been such a fabulous companion over the years. Who else can deserve more than Hibret to have such a decent shock absorbing listener! The conspiracy with Yohannes has been an important source of refreshment after a demanding work. I also thank

the 3:00 pm coffee buddies from the ‘other side’ of the building: Tsion, Beshir, Wudu and Tariku. Often, I couldn’t wait for our coffee break, though we sometimes had a two months’ time for an appointment. Yibeltal, Melkam, Biresaw, Enyachew, Desalegn, Akalu, Baye, Nebyou, Yihune, Melkamu Sewunet (my best high school teacher) and Asmare Yalew have been a great source of motivation over the years. I thank you all for your great friendships.

My time at Wageningen has also been enjoyable due to the many friends and groups that have come an important part of my non-academic life. I am very grateful to my Brothers and Sisters in Christ at Mahibere Kidusan Europe Center in general and Holland Sub-center in particular. I believe we make a very unique family in Christ that we shall eternally cherish. This has been made possible thanks to Mahibere Kidusan, to which I am personally indebted so much. I would not have been the same person without Mahibere Kidusan. I must also thank my spiritual fathers, Aba Kelemework, Kesis Dr. Ashenafi and Aba Gebre Tsadik, for their prayers and spiritual support. I would like to thank all my brothers and sisters whom I met at our Sunday School in Wageningen University over the last four years; you all deserve to be mentioned in this thesis, but listing your names would make several paragraphs.

I preserve my special thanks to my family. Words cannot express how grateful I am to my parents, Biset Melesse and Workalem Mekonnen. You have been central to keep me motivated as much as humanly possible with your much needed encouragement and belief. I would also like to thank my brothers Awoke, Alemnew, Amlaku and Tadlo, and my sisters Jerusalem, Abebech and Kalekidan, for their care and support. I would not have managed to remain motivated without your cheerfulness. A recent addition to the family, Amha, has fast become a source of joy. Asmare, Yekoye, Enat, Maru and Temesgen, thank you for all that you have done for me all the way. I am profoundly grateful to my parents-in-law, Gashe and Etiye, for your blessings and keeping me in your prayers. I also sincerely thank my brothers-in-law, Ghion, Haile and Merkogn, for your kind encouragements. I have put my very special thanks off for awhile now, but it comes time to try to thank my beloved wife, Engdawork who also designed the cover page of the thesis, for her love, understanding and encouragement. Enguye, you have been always there for me.

Finally, I would like to thank my late grandfathers and grandmothers, who ignited the light for the family through sending my parents to school, while they themselves were ‘illiterate’ farmers. I would have sought to see your proud smile more than anything. Particularly, Tilaye, I wish I could see you more. This one goes for you.

Mequanint Biset Melesse
Completed Training and Supervision Plan
Wageningen School of Social Sciences (WASS)



Wageningen School
of Social Sciences

Name of the learning activity	Department/Institute	Year	ECTS*
A) Project related competences			
Advanced Econometrics (AEP 60306)	Wageningen University	2012	6
Panel Data Econometrics – Microeconomics	NAKE	2012	3
Behavioral and Experimental Economics (ECH51306)	Wageningen University	2012	6
Game Theory with Applications	WASS	2012	1.5
Introduction to R for Statistical Analysis	PE&RC	2013	0.6
Advanced Microeconomics (ECH 32306)	Wageningen University	2014	6
B) General research related competences			
Information Literacy PhD including EndNote Introduction	WGS	2011	0.6
Research Proposal	WASS	2012	4
MSc Thesis Supervising DEC-80433	Wageningen University	2013	1
Scientific Writing	WGS	2013	1.8
Techniques for Writing and Presenting a Scientific Paper	WGS	2013	1.2
Career Perspectives	WGS	2015	1.6
Project & Time Management	WGS	2015	1.5
C) Career related competences/personal development			
Regular PhD Meetings	DEC	2011 - 2015	2
‘Market Exposure and Risk Aversion: Evidence from Landed Farm Households in Ethiopia’	SEEDEC (Bergen, NHH)	2014	1
‘Joint Land Titling and Women’s Empowerment: Evidence from Ethiopia’	CSAE (Oxford)	2015	1
‘Joint Land Certification Programs and Women’s Empowerment: Evidence from Ethiopia’	World Bank (Washington DC)	2015	1
‘Formal Law and Customary Change: A Lab-in-Field Experiment in Ethiopia’	ICAE (Millan, IAAE)	2015	1
Total			40.8

*One credit according to ECTS is on average equivalent to 28 hours of study load.

The research described in this thesis was financially supported by the Dutch Organization for Scientific Research (NWO), grant # 453-10-001.