Food Security, Land Rights, Agriculture and Conflict:

An empirical analysis of household and civil conflict in sub-Saharan Africa

Naureen Fatema

Department of Economics

McGill University, Montreal

October 2021

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

ABSTRACT

This dissertation aims to advance existing knowledge and understanding of localized conflict in sub-Saharan Africa by examining underlying mechanisms through which household and community behavior and culture determine the path of local conflict and by evaluating policy models for predicting local conflict. It comprises four individual essays, of which the former two use household survey data from three territories in the North Kivu province of eastern Democratic Republic of Congo (DRC) while the latter two use subnational level, disaggregated, conflict data, georeferenced at half decimal degrees (approximately 55 by 55 km at the equator) across 48 countries in sub-Saharan Africa.

The study begins by empirically analyzing the role of household behavior on the root causes of interhousehold conflict through two key aspects of household decision making – food security and property rights. First, it explores the impact of household food sufficiency and food sharing behavior on interhousehold conflict and finds that food sufficiency reduces local conflict only in the presence of food sharing behavior. Thus, it adds to the broader food security-conflict literature by showing that food sharing may be a potential behavioral channel through which food security reduces conflict. Next, it examines the impact of land title on the risk and cost of interhousehold conflict. The study finds that while land title reduces the initial risk of interhousehold conflict, it does not impact a household's costs in the event of a conflict. Owing to the physical and geo-political challenges to data collection from conflict prone societies, the use of household survey data is rare in the Discipline. By exploiting a survey specifically designed to document interhousehold conflict, these two essays provide insights on the behavior of households from an active conflict zone and on the nature, outcomes and costs of interhousehold conflict.

Next, the study continues to community behavior by exploring pathways through which cultural evolution emanating from agricultural practices affect localized conflict within historical and modern-day rice growing communities of sub-Saharan Africa. Findings show that historically rice growing regions, that continue to cultivate rice at present, experience lower conflict than all other agricultural regions of sub-Saharan Africa. This contributes to the emerging literature on the role of Africa's historical legacies in shaping its path of contemporary conflict.

Finally, the study compares the predictive performance of a traditional logistic model with four machine learning models in order to evaluate the optimal policy option for predicting local conflict in sub-Saharan Africa. It finds that there exists a trade-off between preselected metrics of performance evaluation across the models. Hence, the optimal policy choice depends on the goals and priorities of the policy maker.

Overall, this dissertation contributes to the conflict literature by uncovering channels through which informal institutions in sub-Saharan Africa such as cultural traits, behavioral norms, and precolonial history affect the path of local conflict. Further, it contributes to conflict policy by evaluating and proposing the optimal model for conflict prediction depending on policy goals. A deeper understanding of the root causes of local conflict, supplemented by superior models for predicting local conflict, can ultimately help avert local conflicts before they escalate to national crises.

RÉSUMÉ

Cette thèse vise à faire progresser les connaissances et la compréhension existantes des conflits localisés en Afrique subsaharienne en examinant les mécanismes sous-jacents par lesquels le comportement et la culture des ménages et des communautés déterminent la voie des conflits locaux et en évaluant les modèles de politiques afin de prédire les conflits locaux. Cette thèse comprend quatre chapitres individuels, dont les deux premiers utilisent des données d'enquête auprès des ménages de trois territoires de la province du Nord-Kivu de l'est de la République démocratique du Congo (RDC), tandis que les deux derniers utilisent des données de conflit au niveau infranational, désagrégées, géoréférencées à des demi-décimales (environ 55 par 55 km à l'équateur) dans 48 pays d'Afrique subsaharienne.

Cette thèse commence avec une analyse empirique du rôle du comportement des ménages sur les causes profondes des conflits entre ménages à travers deux aspects clés de la prise de décision des ménages – la sécurité alimentaire et les droits de propriété. Tout d'abord, on explore l'impact de la suffisance alimentaire des ménages et du comportement de partage de nourriture sur les conflits entre ménages et on constate que la suffisance alimentaire ne réduit les conflits locaux qu'en présence d'un comportement de partage de nourriture. Ainsi, ce chapitre ajoute à la littérature plus large sur la sécurité alimentaire et les conflits en montrant que le partage de la nourriture peut être un canal comportemental potentiel par lequel la sécurité alimentaire réduit les conflits. Ensuite, la chapitre suivant examine l'impact du titre foncier sur le risque et le coût des conflits entre ménages. L'étude révèle que, bien que le titre foncier réduise le risque initial de conflit entre ménages, il n'a pas d'incidence sur les coûts d'un ménage en cas de conflit. En raison des défis physiques et géopolitiques liés à la collecte de données auprès de sociétés sujettes aux

conflits, l'utilisation des données d'enquête auprès des ménages est rare dans la discipline. En exploitant une enquête spécialement conçue pour documenter les conflits entre ménages, ces deux chapitres fournissent des informations sur le comportement des ménages d'une zone de conflit active et sur la nature, les résultats et les coûts des conflits entre les ménages.

Après cela, la thèse aborde le comportement communautaire en explorant les voies par lesquelles l'évolution culturelle émanant des pratiques agricoles affecte les conflits localisés au sein des communautés rizicoles historiques et modernes d'Afrique subsaharienne. Les résultats montrent qu'historiquement, les régions rizicoles, qui continuent à cultiver du riz à l'heure actuelle, connaissent moins de conflits que toutes les autres régions agricoles d'Afrique subsaharienne. Cela contribue à la littérature émergente sur le rôle des héritages historiques de l'Afrique qui influencent le cheminement des conflits contemporains.

Enfin, la thèse compare la performance prédictive d'un modèle logistique traditionnel avec quatre modèles d'apprentissage automatique afin d'évaluer les options de politiques optimales pour prédire les conflits locaux en Afrique subsaharienne. Le chapitre constate qu'il existe un compromis entre les mesures présélectionnées de l'évaluation des performances dans les modèles. Par conséquent, le choix optimal de la politique dépend alors des objectifs et des priorités du décideur.

Dans l'ensemble, cette thèse contribue à la littérature sur les conflits en découvrant les canaux par lesquels les institutions informelles en Afrique subsaharienne telles que les traits culturels, les normes comportementales et l'histoire précoloniale affectent la voie des conflits locaux. En outre, il contribue à la politique de conflit en évaluant et en proposant le modèle optimal pour la prédiction des conflits en fonction des objectifs des politiques. Une compréhension plus approfondie des causes profondes des conflits locaux, complétée par des modèles supérieurs pour prédire les conflits locaux, peut finalement aider à prévenir les conflits locaux avant qu’ils ne dégénèrent en crises nationales.

ACKNOWLEDGEMENTS

I would like to thank McGill University for the following financial support: Graduate Mobility Award, Graduate and Postdoctoral Studies; Graduate Students Excellence Award, Department of Economics; Principal's Graduate Fellowship Award, Department of Economics; Ragan Graduate Award; and WYNG Trust Fellowship Award. I am grateful to my supervisors, Franque Grimard and Robert D. Cairns for their endless support, discussions and guidance throughout the program. I am also thankful to all those who have taught me, and especially to Sonia Laszlo, Jennifer Hunt, Matthieu Chemin, and Jim-Engel Warnick for invoking my interest in microempirical analysis.

I am grateful to the Center on Conflict and Development at Texas A&M University for granting me access to their survey data. I thank members of the Households in Conflict Network (HiCN) for their invaluable feedback in the early stages of the second chapter. I would also like to thank all participants at the Sustainability and Development Conference (2018) hosted at the University of Michigan, participants at the Agricultural and Applied Economics Association (AAEA) conferences, 2015, 2016, and 2018, and anonymous reviewers at *World Development* and *Sustainability* for their feedback.

I thank all my friends whose feedback has helped me grow professionally and personally over the years. Above all, I shall remain eternally grateful to my family members without whose relentless support this dissertation would not have been possible.

PREFACE AND AUTHOR CONTRIBUTIONS

This dissertation consists of four essays all of which are original scholarship and distinct contributions to the conflict literature. The work in the thesis is primarily my own.

I am the first author of the first essay, “Givers of great dinners know few enemies: The impact of household food sufficiency and food sharing on low-intensity interhousehold and community conflict in eastern Democratic Republic of Congo”. This essay is co-authored with Dr. Shahriar Kibriya, research director for the Center on Conflict and Development at Texas A&M University. The primary data used in paper was collected as part of the Best Practices in Coffee and Cacao (BPCC) project of the Center. Dr. Kibriya was in charge of acquiring funds for the project from the United States Agency for International Development (USAID) and the Howard Buffet Foundation in addition to survey design, supervision of data collection, and overall management of the project. I provided inputs in the survey questionnaire, helped with data cleaning, conducted all empirical analyses, literature review, and drafted the initial paper. The manuscript was finalized through joint discussions and is currently under a Revise and Resubmit status at a Development Journal.

I am the sole author of the second and third essays. The second essay titled, “Can land title reduce low-intensity interhousehold conflict incidences and associated damages in eastern DRC?”, has been published in *World Development* (2019), DOI: 10.1016/j.worlddev.2019.104612. The manuscript for the third essay, “Rice, culture, and conflict in sub-Saharan Africa”, is currently being prepared for submission.

The fourth essay, “Prevention is better than cure: Machine learning approach to conflict prediction in sub-Saharan Africa”, is co-authored with Dr. Mark Musumba and Dr. Shahriar

Kibriya. Conceptualization, methodology design, writing and preparation of the manuscript for submission was jointly conducted by all three authors. In addition, I was fully responsible for obtaining, cleaning, and overseeing the data, and for all statistical analysis using Stata. Dr. Musumba was responsible for all statistical analyses requiring Python. Dr. Kibriya was solely responsible for project funding. The essay has been published in the journal *Sustainability* (2021), DOI: 10.3390/su13137366.

LIST OF ABBREVIATIONS

ACLED	Armed Conflict Location and Event Data
ADF	Allied Democratic Forces
APS	Average Precision Score
ATT	Average Treatment Effect on the Treated
BPCC	Best Practices in Coffee and Cacao Production
CDF	Congolese Franc
DRC	Democratic Republic of Congo
DRE	Doubly Robust Estimator
FAO	Food and Agriculture Organization
FARDC	Armed Forces of the Democratic Republic of the Congo
FDLR	Democratic Forces for the Liberation of Rwanda
GAEZ	Global Agro-Ecological Zones
GDP	Gross Domestic Product
GIS	Geographic Information System
IFAD	International Fund for Agricultural Development
IFPRI	International Food Policy Research Institute
IIASA	International Institute for Applied Systems Analysis
IMF	International Monetary Fund
IPWRA	Inverse Probability Weighting Regression-Adjustment
ML	Machine Learning
MLP	Multi-Layer Perceptron
NGO	Non-Governmental Organization
NNM	Nearest Neighbor Matching
NRC	National Research Council
OLS	Ordinary Least Square
PRIO	International Peace Research Institute of Oslo
PSM	Propensity Score Matching
RCT	Randomized Control Trials
ROC	Receiver Operating Characteristic
SAGE	Center for Sustainability and the Global Environment
SMOTE	Synthetic Minority Over-sampling Technique
SPAM	Spatial Production and Allocation Model
SPEI	Standardised Precipitation-Evapotranspiration Index
SSA	Sub-Saharan Africa
SVM	Support Vector Machines
UCDP	Uppsala Conflict Data Program
UNFAO	United Nations Food and Agriculture Organization
UNICEF	United Nations Children's Fund
UNOCHA	United Nation’s Office for the Coordination of Humanitarian Affairs
WFP	World Food Programme

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

For more than half a century the continent of Africa has been marred by conflict, affecting two thirds of the sub-Saharan Africa (SSA) region since 1980 (Bellows & Miguel, 2009). Conflicts cost the economy and society immeasurable human suffering, loss of lives and livelihoods, loss in investment and state capacity, forced displacement and mass migration, and an eventual breakdown of social cohesion, institutions, and norms. An estimated 750,000 to 1.1 million people were killed in conflict battles in Africa between 1989 and 2010 (McGuirk & Burke, 2017). A recent study of 45 sub-Saharan African countries between 1989 and 2019, conducted by the International Monetary Fund (IMF), reports that the region remains conflict prone despite downward trends in conflict since the late nineties (Fang, Kothari, McLoughlin, & Yenice, 2020). The study finds that approximately 30 percent of the countries of SSA were affected by some conflict in 2019 though the nature of conflicts has shifted from traditional civil wars to non-state-based conflicts in recent years.

Economists and political scientists offer four main hypotheses to explain the conflict prone nature of Africa: natural resource dependence, weak institutions, ethnic fractionalization and polarization, and poverty (Blattman & Miguel, 2010; Collier & Hoeffler, 1998, 2002; Couttenier & Soubeyran, 2013; Esteban & Ray, 2011; Esteban, Mayoral & Ray, 2012; Fearon & Laitin, 2003; Hegre & Sambanis, 2006; Laitin, 2007; Michalopoulos & Papaioannou, 2011; Miguel, Satyanath & Sergenti, 2004; Montalvo and Reynal-Querol, 2005; Nunn, 2009; Sambanis & Elbadawi, 2000, etc.). Yet, the conflict literature has been underrepresented on a few accounts. First, most analyses and numbers presented are derived from aggregated statistics that make headlines in news media.

Local violence, while underreported and often missing the news radar, is more common and possibly more costly than civil wars (Fearon & Hoeffler, 2014). Scholars largely agree that civil conflict and large-scale violence often originate from and are shaped by local conflict rooted in the behavior of individuals, households, or communities, and their interactions within social norms (Autesserre, 2010, 2014; Blair, Blattman, & Hartman, 2017; Kalyvas, 2006). Weakness of social, political and legal institutions, a common characteristic of developing nations, manifests itself through barriers to access to resources, local power dynamics, and differences in cultural norms and values among groups often leading to social conflicts. These conflicts are usually concentrated locally, within or between villages, and do not always involve mass violence and casualties. Yet they pose an underlying threat to human security and wellbeing, prove economically and socially costly to monitor and can potentially escalate to wide-scale unrest (Barron, Kaiser & Pradhan, 2009; Blair, Blattman, & Hartman, 2017). Second, the physical and political barriers to collecting information from conflict and post-conflict societies result in a lack of availability of high quality, disaggregated data at lower levels of geographical specifications and partially contribute to this neglect of local events in news headlines. Consequently, micro-empirical conflict analysis has suffered a general lack of attention to the causes and consequences of such localized conflict events. Third, despite an increasing number of studies on the root causes of conflict in Africa, most theories have largely ignored advances in behavioral economics, particularly the relevance of culture on economic outcomes (Alesina & Giuliano, 2015; Blattman & Miguel, 2010; Guiso, Sapienza & Zingales, 2006; Nunn, 2009). As a result, with a few exceptions, the existing conflict literature is mostly limited to cross country studies of civil wars and war related casualties where the state or state institutions are the main actors (Justino, 2011) and little is known about local conflict dynamics rooted in microeconomic behavior.

This dissertation aims to advance the existing knowledge and understanding of localized conflict through an examination of underlying mechanisms through which household and community behavior, culture and local institutions determine local conflict and an evaluation of conflict prediction models based on policy goals. Within the context of this dissertation, *local conflict* is used to refer to all individual or social tensions, acts of violence and/or non-violence, as well as isolated and mass casualties between microeconomic agents that occur within a specified locality (such as within and between communities). It should be emphasized that in some cases these conflicts may begin and end in innocuous disputes while others have the potential to escalate to social unrest with grave consequences. At times, the term *low-intensity* conflict is used to emphasize the nature of conflict under examination and to distinguish it from the traditional usage of the word conflict in reference to large scale national violence or civil conflict. Microeconomic agents are defined as civilians, households, communities, state and non-state actors, rebel forces and militia. Examples of conflict events comprise disagreements and minor altercations, major disputes, riots and protests, conflict with government and/or rebel forces, violence against civilians, battles between armed groups and fatal exchanges. A broad definition of (local) institutions is used to refer to the humanly devised constraints that structure human interaction, including formal constraints such as (local) rule of law as well as informal constraints such as (local) norms of behavior and (local) cultures (North, 1994). Though the notion of culture is implicitly included within informal constraints, this dissertation follows empirical studies on culture and institutions (*see* Alesina & Giuliano, 2015) and prefers the usage of the term culture to refer to decision-making heuristics that include values, beliefs and social norms (Nunn, 2012) to draw a clear distinction from formal institutions that are governed by the polity. When used explicitly, the word *culture* is defined as “those customary beliefs and values that ethnic, religious, and social groups transmit fairly unchanged from generation to generation" (Guiso, Sapienza & Zingales, 2006).

The examination of local conflict begins with analyses of household behavior and decision-making in Chapters 2 and 3, then proceeds to an analysis of community behavior and cultural norms in Chapter 4 and ends with a system for predicting local conflict in Chapter 5. The next two chapters use primary survey data from 1763 households from three territories (Beni, Lubero, and Rutshuru) of North Kivu province in eastern Democratic Republic of Congo (DRC). Both chapters study the impact of household behavior and decisions, within the context of local norms and institutions, on low-intensity, interhousehold conflict. These include all conflict exchanges between households within a community, ranging from aggressive, yet non-violent expressions of behavior (such as arguments, altercations, and verbal disagreements) to major disputes and acts of violence that may result in a substantial loss of physical and human capital, including fatalities. Chapter 2 examines the impact of household food sufficiency and food sharing behavior on low-intensity, interhousehold conflict. Food sufficiency, defined as having sufficient food available within the household, is differentiated from food security, which is a more complex, multidimensional phenomenon. This chapter explores whether food sufficient households are less likely to experience low-intensity, local conflict and if this likelihood is affected by a culture of sharing food with other households. After a series of tests, the chapter concludes that food sufficiency reduces low-intensity, interhousehold conflict only in the presence of food sharing behavior. Findings suggest that food sufficiency alone cannot reduce low-intensity, interhousehold conflict unless accompanied by benevolent cultural norms.

To supplement the understanding of household behavior and decision-making, Chapter 3 examines whether the possession of land titles can reduce the likelihood of low-intensity, interhousehold conflict and protect households from adverse consequences in the event of a conflict. Results indicate that though land title can reduce the initial risk of low-intensity, interhousehold conflicts, possessing land title does not ensure fewer adverse consequences should such conflicts arise. This is relevant since the current land tenure system in eastern DRC is governed by the coexistence of both formal and informal local institutions. Customary laws based on traditional norms prevail in rural and indigenous areas while the formal state law is more common near city centers. Thus, the household decision to obtain legal title in DRC is, besides usual informal market constraints in developing countries, also a function of the preference for informal over formal local institutions. The evidence from this paper may have implications for land reform policy in conflict and post-conflict societies. Findings suggest that formalization of intricate land titling system that provides citizens with land title is not a panacea for all conflict related adversities. Land reform programs in conflict prone societies can benefit from being supplemented with prediction policies to provide accurate early warning for local conflict such that they can be avoided, or with mitigation policies to help vulnerable households recover from conflict related adversities. In addition, local institutions should be able to adjudicate and adjust for the dual laws.

Following the impact of household behavior on low-intensity, localized conflict, Chapter 4 proceeds to explore channels through which community behavior, cultural norms, and local institutions influence local conflict. The debate on geographical versus institutional pathways to economic development has recently infiltrated the civil conflict literature (Acemoglu, Johnson, & Robinson, 2001, 2002; Easterly & Levine, 2003; Engerman and Sokoloff, 2000; Gallup & Sachs, 2000; Nunn, 2009; Tabellini, 2007). Proponents of the geography hypothesis stress the importance of factors, such as climate, natural resources, access to coastline and sea, navigable rivers, landlock, conditions favorable for agriculture, terrain ruggedness and disease environment in determining current development and conflict. Other schools of thought argue that historical events such as colonial rule (Acemoglu et al., 2001, 2002; Engerman & Sokoloff, 1997, 2002; Nunn, 2009) and slave trade (Nunn, 2008; Nunn & Wantchekon, 2011; Zhang, Xu & Kibriya, 2021) impact long term development outcomes through their effect on local institutions, culture and norms of behavior (see Nunn, 2009 for a comprehensive review). This chapter builds on the philosophy that domestic institutional structure and cultural norms leave a lasting impact on local conflict. Thus, it contributes to the broader debate by analyzing a different historical event - the introduction of rice to Africa – and documenting the process by which a history of rice farming, through its effect on local institutions and cultural norms, may have determined the path of localized conflict in present day sub-Saharan Africa. Detailed accounts, as may be found in Paul Richard's (1986, 1996) books, describe culture and community values in African rice farmers and plantations. Recent social scientists have begun exploring how these coordinated efforts may have evolved into a culture of trust, cooperation, and interdependence over time that may result in a tendency to avoid conflict. The underlying argument in this chapter is that, historically, rice farming cultures involved labor exchanges and coordination over multiple time-sensitive steps of rice plantation. Conflict proved costly, returns to cooperation were high and over time this evolved into cultural norms and local institutions that supported exchange and peaceful coexistence among the community. Assuming these norms and institutions persisted over time, as repeatedly suggested in the economic development literature, these norms of behavior should be reflected in present day conflict outcomes.

To empirically test for correlations between rice growing traditions and lower prevalence of conflict in SSA, this chapter uses cross sectional conflict data from the *Armed Conflict Location* *and Event Data Project* (ACLED) version 7, developed by the International Peace Research Institute of Oslo, Norway, and the University of Uppsala, Sweden (referred to as PRIO/Uppsala) and obtained from Raleigh et al. (2010). This dataset offers several advantages over similar available data for analyzing local conflict. First, it catalogs information on all known conflict events and records a range of violent and non-violent events by date, location, and actors involved (such as governments, rebels, militias, ethnic groups, active political organizations, and civilians). This allows for the possibility of analyzing conflict events of all intensities, including ones that miss news headlines. Second, the data are both high frequency (collected daily) and high resolution (available at half degree by half degree). This level of disaggregated data in specific geographical locations provides valuable information on localized conflict events instead of aggregating information over larger geographical boundaries. Third, local conflict outcomes may be intrinsically affected by sociopolitical correlates within administrative boundaries (such as countries, states or villages), thus potentially confounding estimates. The availability of data within arbitrary geographical boundaries can circumnavigate this issue. Results from this chapter show that contemporary local conflict in sub-Saharan Africa is negatively correlated with a culture of rice farming, especially in regions with a prior history of rice farming. This provides suggestive evidence in support of the theory that historical events can affect long term conflict through strong, local institutions. It further implies that geographical factors, such as agricultural suitability of the soil to rice farming, affect long term conflict through its influence on a history of rice farming and the evolution of a cooperative culture and peaceful institutions within rice growing communities. This finding is consistent with the "history matters versus geography matters" debate which agrees that the most significant impact that geography has on economic development is through its influence on history (Nunn, 2009).

Chapter 4 attempted to uncover underlying relationships between cultural correlates and local conflict using conventional econometric approaches. Using the same data but with modern machine learning approaches, Chapter 5 aims to predict local conflict in SSA. While causal mechanisms provide scholarly insight on the root causes and dynamics of local conflict, tools to predict the location of the next conflict may be more useful from a policy perspective. All conflict, local or national, is socially and economically costly to monitor and mitigate. Accurate prediction offers policy makers a more pragmatic approach to identify conflict threats by looking beyond p-values and allowing for the efficient allocation of scarce, societal resources such as security or policing forces. It should be recognized that machine learning does not only provide a different tool from causal inference; it solves a different problem: the objective of causal inference is to provide unbiased, consistent and efficient parameter estimates whereas the objective of prediction is to identify models that make the best out-of-sample prediction by fitting complex and flexible functional forms to the data without simply overfitting (Mullainathan & Speiss, 2017; Athey, 2015). As such, the parameters from the machine learning model are neither designed to be consistent nor to be equivalent to the causal inference parameters. In addition, traditional causal analysis of conflict puts emphasis on in-sample prediction power with extreme caution towards not accepting a false hypothesis. However, successful conflict prediction requires out-of-sample predictions along with caution towards rejecting the true hypotheses.

A few recent studies have used machine learning to predict conflict in Africa (*for example*, Cederman and Weidmann, 2017; Chadefaux 2017; Perry, 2013; van Weezel, 2017), but these are either limited by the scope (such as inclusion of a handful of countries or use of aggregated data), the scale of conflict studied, or the number of predictors included. The current chapter offers an advantage over these studies by including continent-wide, disaggregated data and a rich set of over

100 initial predictors guided by the causal inference conflict literature. Though machine learning algorithms are typically designed to analyze a large number of observations, they can also be applied to a large number of independent variables (thin data) as executed in this chapter. The ACLED data spans over 40 countries in sub-Saharan Africa and allows the examination of localized conflict at a fine grain resolution within each country and the prediction of conflict incidences at a sub-national level. Conflict prevention or mitigation goes beyond civil war deterrence and the promotion of peace. Often, less intense and local events are warning signs which, if detected, can prevent potential civil wars. The analysis compares and contrasts between a conventional logit model and four machine learning algorithms using three metrics of measurement to evaluate model performance: accuracy, precision and recall. *Accuracy* refers to the overall prediction accuracy and is a measure of the proportion of correctly predicted outcomes. *Precision* (positive predictive values) is the percentage of results that are relevant, in this case conflict. *Recall* (sensitivity) is the ratio of the true positive instances, conflict, that are correctly detected by the model. Recall is a measure of completeness, while precision is the measure of exactness. This chapter serves as a pilot study to assess whether newer machine learning models outperform a conventional logistic model for predicting local conflict. Results show that not all models perform better than the logit but the ones that do so outperform the logit model. The overall conclusion is that the answer depends on whether the ultimate policy goal is to identify as many potential conflict zones as possible, at the cost of a few false positives, or to ensure that most conflict zones are correctly predicted, with some areas being overlooked. The chapter discusses this trade-off that policymakers need to assess based on the existing infrastructure, available resources, and a cost-benefit analysis of devising preventative measures. Policymakers are skeptical about implementing prevention measures as they are viewed as riskier or even futile in

the presence of erroneous prediction models. The results from this study suggest that it might be worthwhile to invest in updating the models using real-time data to design a conflict early warning system. ACLED already collects information on localized conflict events on a daily basis. Incorporating this real-time conflict data could enhance the application and delivery of timely prediction models that can prove crucial to averting local conflict, reduce or eliminate costs due to conflict, and most importantly, save human lives.

Increasing availability of unconventional, disaggregated conflict data, and a growing emphasis on incorporating studies from other Social Sciences, such as Psychology, History and Cultural Anthropology have led recent economists towards understanding the root causes of conflict in Africa (for example, Besley & Reynal-Querol, 2014; Michalopoulos & Papaioannou, 2014, 2016; Moscona, Nunn & Robinson, 2020; Nunn, 2012). This dissertation adds to the budding literature by examining various aspects of behavior and culture on localized conflict in Sub-Saharan Africa. Overall, there appears to be ample evidence that household and community behavior, values, decision-making heuristics, and cultural norms play a key role in shaping low-intensity, localized, conflict dynamics. The findings are consistent with African historiography which deemphasizes the role of formal, domestic institutions on economic and political development as one moves away from the cities. Instead, the emphasis is on understanding ethno-cultural traits, customary laws and pre-colonial history (Michalopoulos & Papaioannou, 2013; Herbst, 2000). Understanding root causes of local conflict is not only of paramount importance for advancing the frontiers of knowledge, but also to inform development policy for managing conflict in developing countries. The findings from this dissertation imply two key policy directions. First, promoting awareness on the virtues of benevolent behavior and engaging households and communities in local team-building activities that encourage coordination, cooperation, and benevolence can potentially avert and resolve localized conflict. Second, investing in a conflict early warning system with real-time data to predict local conflict accurately before it escalates to a national scale can help reduce the economic and social costs of conflict. Deeper understanding of the root causes of conflict can inform the nature of data required for better predictions. The behavioral drivers of local conflict uncovered in this study suggest that local conflict prediction models would benefit from the inclusion of local cultural values, traditions and norms.

Chapter 2

GIVERS OF GREAT DINNERS KNOW FEW ENEMIES: THE IMPACT OF HOUSEHOLD FOOD SUFFICIENCY AND FOOD SHARING ON LOW-INTENSITY INTERHOUSEHOLD AND COMMUNITY CONFLICT IN EASTERN DEMOCRATIC REPUBLIC OF CONGO

2.1 Introduction

Historical accounts of food shortages causing conflict can be traced back to the Russian, English and French Revolutions of the 17th and 18th century. In modern times, prevalence of hunger has been documented to drive violent behaviour and conflict between and within communities through environmental, social, economic, and political channels (see for example, Bora et al., 2010). Due to the complexity of establishing a direct relationship between hunger and conflict, the more popular academic approach of investigation has been through the aforementioned channels and almost entirely confined to macro or district level analyses of violent armed combat. Examples include the causal linkage between climate change and conflict with food shortage as an underlying cause (Barnett & Adger, 2007; Burke, Miguel et al., 2009; Miguel, Satyanath, & Sergenti, 2004; Salehyan, 2008); poverty and grievance driven by hunger and malnutrition, and causing civil conflict (Collier, 2004; Pinstrup-Andersen & Shimokawa, 2008); and volatility in food prices and acute food shortages triggering conflict (Arezki & Brückner, 2011; Bellemare 2015; Berazneva & Lee, 2013; Bessler, Kibriya et al., 2016; Bush & Martiniello, 2017). While these studies establish hunger as one of the drivers of violent combat at a national or subnational level, recent scholars acknowledge a need to capture behavioural differences emanating from food insecurity and how they shape conflict responses for households (for example, Justino, Bruck & Vervimp, 2013; Martin-Shields & Stoetz, 2018; van Weezel, 2017). Other studies that attempt to address this literary gap include violence exposure and household food deprivation (Mercier, Ngenzebuke, & Verwimp, 2017); and household resilience and food security (Brück, D'Errico, & Pietrelli, 2018). We strengthen this collection of scholarship by exploring the links between having sufficient food and the act of sharing food with others (referred to as *benevolence* throughout this article) on interhousehold and community level low-intensity conflict[1] with survey data collected from 1763 households of North Kivu in eastern Democratic Republic of Congo (DRC). Additionally, we contribute to the economic anthropology of human behaviour by explaining the findings through rational choice theories.

Our conjecture is that food sufficient and benevolent households may be able to avert low-intensity conflict incidents with other households and community members. Accordingly, we investigate two specific questions, i) Are food sufficient households less likely to engage in low-intensity interhousehold and community level conflict; and ii) Can food sufficient households avert interhousehold and community level conflict by sharing their food with other households? Contextually, North Kivu, DRC provides an ideal opportunity to answer these queries due to their pervasive food insecurity, localized micro-level conflict and fragile socio-political conditions.

Given that food sufficient and food insufficient households may be systematically different, we employ quasi-experimental estimation technique of propensity score matching (PSM) to estimate the effects of food sufficiency on household and community level conflict. It is possible

[1] We define low-intensity household conflict as aggressive, though not always violent, behaviour at the interhousehold or community level. These may include any form of arguments, disputes, altercations or violence

that inherent differences between households in our sample that are food sufficient and those that are not leads to any observed correlation between food sufficiency and conflict. Matching solves this problem by discarding these drastically different observations and comparing "like to like" (similar to a randomized experiment). We test the robustness of our findings with different matching techniques and tests of covariate balance as well as estimating our results using a doubly robust estimator. Our quasi-experimental setup offers several benefits. First, we avoid the requirement of baseline data on households who have become food insufficient (Imbens & Woolridge, 2009). Second, we ensure that the comparison of the outcome variable, conflict, is undertaken between households with similar characteristics (Dehejia & Wahba, 2002). Third, when comparing sub-populations of households with similar characteristics, covariates are independent of households that are not food sufficient (Imbens & Woolridge, 2009).

Our initial set of results show that a household's food sufficiency status reduces its probability of conflict with other households and groups within the community by roughly 10 percentage points. However, upon conditioning on benevolence, we find that food sufficient households that share food with others have a 13.8 percentage point lower probability of conflict on average while the effects disappear for food sufficient households that do not share their food. Thus, we conclude that food sufficiency reduces low-intensity interhousehold and community conflict only in the presence of benevolent behaviour. Although we took measures to control for various sources of bias, we show extreme caution to claim causality. However, at a minimum, our results establish a micro-foundational linkage between food sufficiency, food sharing behaviour and household level conflict which is scarce in the literature. We explain the more conflict prone nature of the food insufficient households through Thompson's (1961) the "Moral Economy", where North Kivu peasants may have chosen an ethical path driven by a normative notion of justice. However, benevolent food sufficient households avoid confrontation through gift a reciprocity culture described by both Mauss, (1950) and Malinowski (1922).

Our attempt stands to make a few contributions to the rational choice and conflict literature as well as the food policy debate. First, our initiative documents micro-level information of categorized disputes between neighbors, extended family members, pastoralists, and government and rebel forces which remain largely unreported. Understanding the behaviour of households and how they can drive interhousehold conflict can assist policy makers to avoid such low-intensity conflict which has the potential to aggravate into large scale wars. Second, to the best of our knowledge, this is the first attempt to empirically examine the effects of having sufficient food and of sharing food, on interhousehold and community level conflict. Based on our results, it appears that the act of sharing food within communities can result in lower incidents of low-intensity conflict. This act of benevolence can reduce friction within communities and promote peace within fragile societies. Third, we provide empirical evidence of household level rational choice conflict as well as reciprocal culture of gift and favours in the African peasant society previously absent from the literature.

The remainder of the paper is organized as follows: section 2 describes the context and study justification; section 3 explains the sampling strategy and data and describes the variables; section 4 develops an empirical model and identification strategy. Section 5 presents the results and analysis while section 6 offers summary of results, discussion and concluding remarks.

2.2 Study context and justification

2.2.1 Study context

Despite being one of the most resource rich countries, the Democratic Republic of Congo is plagued by food insecurity, inequality and poverty, unstable governments, weak property rights, rebel groups and competition over resources. About 70 percent of the population is engaged in agriculture, mostly for subsistence (IFAD, extracted April 2016). DRC was ranked 7^{th} out of 178 countries on the 2017 Fragile States Index, placing it in the highest category of risk ("very high alert") and reflecting widespread conflict and insecurity (WFP 2018). Of D.R.C.'s population of 74.88 million, 63.6 percent live below the poverty line and lack access to adequate food (WFP 2016).

After serving as a Belgian colony for almost a century (1870 - 1960), DR Congo gained independence in 1960. However, the period following independence has been marked by extreme corruption, exploitation and political instability. Between 1990 and 1994 civil war broke out in the neighboring country of Rwanda which left a lasting impact on DRC. Following the Rwandan genocides of 1994, a lot of the marginalized population fled to eastern DRC (then known as Zaire) to refugee camps established along the border. Rwandan militia forces followed them into DRC and this entry ignited the Congolese wars. Between 1996 and 1997 Rwandan and Ugandan armed forces formed a coalition to overthrow the government of Zaire (under Mobutu's rule) in an attempt to control mineral resources, thus leading to the first Congolese war. They succeeded in overthrowing the government but the new leader, Laurent-Désiré Kabila urged the armed forces to leave the country. Although the armed forces left DRC, newly formed rebel groups from Rwanda and Uganda instigated the second Congolese war in 1998 in an attempt to overthrow

Kabila. While the second civil war officially ended in 2003, unrest continues between the military of DRC and Rwanda, and the rebel forces of the Democratic Forces for the Liberation of Rwanda (FDLR) remaining in DRC.

Within DRC, North Kivu residents had the least access to assets and were one of the most affected by the economic collapse (Marivoet & De Herdt, 2019). Consequently, at present, North Kivu poses the greatest threats to political stability in DRC (see Stearns, 2012; Vlasseroot & Huggins, 2005; and Vlassenroot and Raeymaekers, 2008 for a detailed account of the conflict in North Kivu). Citizens have a lack of food access, social governance and cohesion that are sowing the seeds of micro level interhousehold and community conflicts. Our field studies show that semi-violent and non-violent altercations are common among fellow villagers, government and supporters of rebel groups, pastoralists and famer groups, extended family members and community members at large. Thus, given prevalent hunger, ongoing history of conflict and current social tensions, North Kivu provides an ideal yet unfortunate setting for this study.

2.2.2 Study justification

Quite a few studies in interdisciplinary development journals report food insecurity driving conflict. However, studies related to our specific effort is lacking since we approach conflict from a largely non-armed and interhousehold level. Food insecurity has been shown to initiate feelings of horizontal inequality, grievances and discontent (Humphreys & Weinstein, 2008, Qstby, 2008; Stewart, 2011); while even illusions of food security (or such programs) have been noted to provide a comforting sense (White et. al, 2016). Nutrition and health studies also show that lack of food and hunger is related to poor mental health, depression, anger and aggression (Chilton & Sue, 2007; Carter et al., 2011; Bushman et. al., 2014; Heflin et al. 2005). Recent exploration in the development literature by Rojas & Guardiola (2017) show that hunger depresses people's subjective wellbeing. On the other hand, evidence from Nepal and South Sudan suggest that food security can enhance a feeling of equality and harmony at a community level (McCandless, 2012). Conversely, food insecurity can provide individuals and households with both material and non-material incentives to engage in any form of anti-social behaviour (Justino, 2011; Martin-Shields & Stoetz, 2018).

Though we study micro level low-intensity, mostly non-violent conflict, because of the relative lack of knowledge in this area, we refer to the broader literature on violent conflict driving possible food insecurity through different channels. Food secure households in an impoverished society are likely to have better access to education and employment which increases the opportunity cost of joining a movement (Taeb, 2004). Food insecurity can also cause undue competition for resources such as water and land which may lead to personal (Messer, 1998; Cohen & Pinstrup-Anderson, 1999) and community level conflict (Homer-Dixon, 1999; Kahl, 2006). Lack of access to land and water resources often create conflict between farmers and pastoralists (Hendrix & Salehyan, 2010; Schomerus & Allen, 2010; Turner et al. 2011). While such conflict between pastoralists and farmers due to land encroachment and water resources are more common against a backdrop of hunger (Raleigh, 2010), food security ensures less cattle raiding and altercation over resources (Schomerus & Allen, 2010). Conflict between agricultural communities and rebel groups over food and resource at both community and individual level is quite common in African societies (Macrae & Zwi 1992; Richards, 1998; Winne, 2010).

While the aforementioned literature on civil conflict provides valuable insights between the links of food security and different types of violence, it is largely silent on social altercations at a lower level that may be caused by basic food insufficiency. We propose that households that are food sufficient will be less prone to low-intensity interhousehold and community conflict. Our conjecture is furthered by introducing food sharing as a connection in this linkage. We define food sufficiency as never having difficulty in providing food to all family members in the six months prior to the survey. Low-intensity interhousehold and community conflict are defined as experiences of interpersonal or community level conflicts, disputes, disagreements, and social altercations, often non-violent in nature, reported by surveyed households.

We choose to study food sufficiency[2] over food security for the following reasons. Household food security is a multidimensional phenomenon that is difficult to capture without a detailed survey dedicated specifically to that purpose. In addition, food security can affect household conflict through multiple channels, thereby making causal exploration challenging and prone to multiple sources of bias. Instead, we use a binary response to measure one aspect of household food security – whether the household had sufficient food for the entire family over a six-month period. We draw motivation from FAO's Coping Strategy Index (CSI) (Maxwell et al., 2003) which states, "Clearly, food security is about much more than just how much people have to eat…Yet, having "enough" food to eat is clearly the most important outcome of being food secure, and while physiological requirements differ, people largely know whether they have "enough" or not".

Based on the food security and conflict literature discussed above, we argue that food sufficient households are less prone to grievances, greed, psychosocial frustration, anger and emotional stress than their food insufficient counterparts. By feeling content, such households would have lower motivation and aggravation of engaging in conflict. In addition, we propose that if food sufficient households show benevolence towards others, they may also be able to avoid

[2] Our underlying assumption is all households need a standard amount of bare minimum food to sustain themselves; thereby we use the term food sufficiency.

interpersonal conflict. These households may express their content through acts of kindness by helping others with food thereby further reducing their chances of getting involved in such interpersonal altercations.

To be circumspect about potential measurement and endogeneity bias, we employ a cautious research design. Our survey instrument was designed to specifically inquire about conflict experiences such as inheritance disputes, disagreement with pastoralists, disputes with other households, conflict over community resources such as the Virunga Park[3], etc.[4]. Given the way we define food sufficiency and the nature of conflicts explored, it is unlikely that such incidences would affect households' likelihood of having sufficient food over a sustained period. Citing some examples, conflict occurring over Virunga National Park resources[5] has a very limited probability to cause household food insufficiency. While violent conflict occurring from inheritance with immediate family may cause food shocks, we specifically inquire about disputes (alluding to a lower level conflict) over inheritance that is unlikely to cause food insufficiency within a six-month period. Similarly, for every other low-level conflict we explore, food insufficiency during a six-month period is improbable. Hence, our cautious approach and the categories of interhousehold and community conflict considered abate reverse causality suspicions to a large extent[6].

[3] Africa's first national park overseeing the North Kivu region which is a considered a bio-diversity hot spot.

[4] A more detailed description of the incidences considered is depicted in the variable section.

[5] A common cause for community level conflict due to its natural and wildlife resources and conservation

[6] As alternate measures of food sufficiency, we also asked households whether any member had ever gone hungry in the past five days; as well as whether any children in the household had ever gone hungry in the past five days. We used these two variables to test whether conflict affected household food sufficiency. Much to our assurance, this is not the case. If conflict was the driver of household food insufficiency, then it should have affected food insufficiency in the past five days, but this is not what the results show.

2.3 Data description

2.3.1 Survey design and data collection

During July 2014, The Howard G. Buffett Foundation funded and initiated the data collection for this research through Texas A&M University, as part of its Best Practices in Coffee and Cacao Production (BPCC) Project. The authors of this paper contributed to the survey design and information collection procedure that ensured pertinent sample population and specific survey questions related to this study. Data for this study was collected from the province of North Kivu, Eastern DRC.

The present administrative unit of the region is divided into six territories or zones. Our survey was conducted in three of these territories – Beni, Lubero and Rutshuru. Since precise population densities are not known and could not be incorporated in the sampling procedure, we used a grid-based randomization technique to make the study sample as representative of the population as possible by ensuring each grid in the selected region had equal likelihood of being studied. High-resolution maps from the United Nation's Office for the Coordination of Humanitarian Affairs (UNOCHA) were used to divide each region into 5 by 5 km squares. If a square had at least one village, it was assigned a unique number (Figure 2.3-2.5 in Appendix). Thus 626 unique numbers were assigned corresponding to populated squares with 190 in Beni, 272 in Lubero, and 164 in Rushuru territory. The statistical software "R" was used to generate random numbers to select squares for village sampling. Squares that could not be surveyed for any reason were replaced with the next number. While omitting squares with high levels of conflict from our sample could raise concerns for biased estimates, the actual number of squares that had to be abandoned for such reasons was trivial, and hence not an issue in this study. Village selection used proportional weighting within each square. If a square had three or less villages, all villages were surveyed. If a square had between four and nine villages, three were selected at random; while for squares that had over ten villages, four were chosen at random. The random selection procedure was executed by assigning numbers to each village and using a random number generator to select the village to be studied.

Local extension agents were employed as enumerators for data collection. A household, the unit of analysis for the study, was defined as a group of people sleeping under the same roof and eating together. Enumerators were instructed to interview all households from selected villages. A strict starting location was not enforced since the sample design included the entire village. If the decision maker was absent at the time of visit, the enumerators were asked to move on to the next house and return later. Households for which vital information was missing were dropped from the analysis. Through this process, we obtained a full sample of data from 1763 farming households.

Structured questionnaires were used to gather information on household socio-economic and demographic structure, food sufficiency measures, conflict experiences, land access patterns, access to markets and knowledge, access to basic services, cooperative membership and social cohesion and empowerment. The questionnaire was translated to French, a language commonly used in North Kivu, and pilot tested before actual surveys took place. The responses were translated back to English before being coded. The interviews took place in a one-on-one setting to maintain confidentiality of the participants. Due to the low education levels and high rate of illiteracy in the region, interviewers sought oral consent by guaranteeing the respondents confidentiality and ensuring their names were not recorded. Each participant was distinguished by unique identification numbers. Respondents did not receive any compensation for participating in the study.

2.3.2 Variables

The outcome variable of interest is low-intensity interhousehold and community level conflict[7] experienced by households. To measure conflict, households were asked if they experienced any of the following types of conflict in the past six months: a) conflict with neighbors and fellow villagers; b) disagreement involving Virunga National park; c) landholder reclaimed occupied land; d) border conflict with landholder; e) dispute among non-dwelling family members f) occupied land granted to a new tenant; g) disagreement with pastoralists; h) conflict over community resources and agricultural inputs; i) resource conflict with rebel forces; j) land conflict with rebel forces; k) land conflict with government; l) resource conflict with government forces; m) other kinds of conflict with government forces; and n) any other kind of conflict that they were asked to specify. Focus group discussions with community members prior to the household interviews helped identify the above-mentioned types of low-intensity interpersonal conflict as the most prevalent in our study areas.

We examine three variations of the conflict outcome in all analyses. First is the overall *probability of conflict* which is the probability that a household experiences any of the conflict mentioned above. The second, *probability of conflict with individual households*, is the probability that a household experiences conflict with neighboring households, fellow villagers, landholders, non-dwelling relatives and/or pastoralists. The third is the *probability of conflict with groups*, which is the probability that a household experiences any conflict with the community, government forces or rebel forces.

[7] We refer to low-intensity interhousehold and community conflict as “conflict” for the sake of brevity and fluency.

The main explanatory variable is household level food sufficiency. We asked households, "how often have you had difficulty feeding your entire family in the last six months?" Respondents could choose between three options, namely, "often", "sometimes" or "never". For our analysis, we categorize a household as *food sufficient* if it responded "never"; and *food insufficient* if it responded "often" or "sometimes". Provided that a household would not gain anything by claiming to be food sufficient or insufficient, we rule out the possibility of misreported responses. In addition, our summary statistics show that around 56 percent of the households claim to be food insufficient, which is consistent with reported household surveys conducted by WFP (2014) and UNICEF (2010) in DRC and North Kivu. To measure "benevolence", we asked households if they had helped others with food in the past six months. Households that answered positively were classified as *benevolent* and households that responded negatively were categorized as *non-benevolent*.

While it is impossible to rule out the presence of omitted variables from survey data, we include a large set of control variables from relevant literature to match households. We also include community fixed effects to capture any differences in communities and macro level shocks that could affect households. Control variables included community specifications and basic household demographics such as religion, household size, number of adult males in the household, education, income, access to markets and information, access to water and cooking fuel, social empowerment and voice in the community, land ownership status, and membership in cooperatives. Household size is included since larger households may have a greater likelihood of being involved in situations of conflict or depending upon adult members will have varying degree of food sufficiency. Education, which may reduce both food insufficiency and conflict, is accounted for through the years of education of the most highly educated member of the household.

Assuming diminishing marginal return to education, the variable is included in both linear and quadratic forms. The link between poverty and conflict has long been established in the conflict literature. Hence, we control for household income; household influence; access to basic services such as drinking water and cooking firewood; and access to information and technologies such as radio/television/cell phone/internet, as well as access to bicycle or motorized vehicles as these may provide a means to acquire information about markets or current situations of conflict.

2.3.3 Descriptive statistics

Table 2.1 presents a cross tabulation of the types of conflict incurred by households by food sufficiency status. Panel A shows around 50 percent of the sample households experienced some form of conflict; 43 percent of these are food sufficient while 56 percent are food insufficient. Panel B shows detailed accounts of the types of conflict experienced by households. Approximately 41 percent were involved in conflicts with other households, while 9 percent incurred conflict with the community.

Table 2.1: Detailed account of conflict reported by households

	HH claims to be food sufficient	HH claims to be food insufficient	Total number of HH
Panel A: Conflict experience of households			
Number of HH that did not experience any conflict	438	482	920
Number of HH that experienced at least one conflict	328	515	843
Total number of HH	766	997	1763
Panel B: Type of conflict reported by households			
Conflict with neighbors and fellow villagers	129	249	378
Conflict with landholder	100	243	343
Inheritance dispute among non-dwelling family members	73	96	169
Disagreement with pastoralists	127	193	320
Total number of reports of conflict with individual HH	*429*	*781*	*1210*
Land and resource conflict with rebel forces	61	135	196
Land and resource conflict with government forces	11	37	48
Conflict over community resources including Virunga Park	9	14	23
Others	15	36	51
Total number of reports of conflict with groups/community	*96*	*222*	*318*

Source: Authors' calculations based on the survey data.
Note: A single household may report multiple incidents of conflict. 'Other' forms of conflict reported include, theft, robbery, sorcery, etc. HH refers to household in the table.

Table 2.2 presents a comparison and t-test of means for conflict outcomes and socioeconomic characteristics of food sufficient and insufficient households. The mean values of the three conflict variables can also be interpreted as the proportion of households that experienced conflict. The monthly per capita income for a typical household is 17,600 Congolese Francs (CDF)[8]. This translates to less than US $1/day, which is below the World Bank's 2013 estimate of international poverty line of US $1.90/day (World Bank, 2016).

[8] 1 USD=925 CDF at the time of the study.

Table 2.2: Summary statistics of main variables

Variable	Food sufficient households (N=762)	Food insufficient households (N=1001)	All households (N=1763)
Dependent variables			
Probability of conflict	0.46***	0.54	0.50
Probability of conflict with individual households	0.36***	0.45	0.41
Probability of conflict with groups	0.09	0.09	0.09
Independent variables			
Household size	5.45*	5.23	5.33
Number of adult males	2.24**	2.10	2.16
Household education (number of years)	9.48***	8.83	9.11
Household education squared	111.67***	99.42	104.67
Household income (`000 CDF/capita)	19.3*	16.4	17.6
Household has written land claim (yes=1)	0.37***	0.43	0.40
Access to technology and markets (yes=1)	0.84***	0.69	0.75
Lack of extension services (yes=1)	0.60	0.62	0.61
Cooperative membership (yes=1)	0.23	0.21	0.21
Access to safe drinking water (yes=1)	0.63	0.63	0.63
Inadequate access to cooking fuel (yes=1)	0.56	0.64	0.61
Leadership position (yes=1)	0.69***	0.55	0.61
Household is benevolent with food (yes=1)	0.78***	0.63	0.69

Source: Authors' calculations based on the survey data.

Notes: We used t-tests to test for equal means between food sufficient and insecure households. *, **, and *** indicate significance at 10%, 5% and 1% levels respectively. Community and religion specific dummies have been omitted from the table to save space. CDF=Congolese Franc. It should be noted that a household may experience multiple instances of conflict.

The table clearly shows that food sufficient households are different from food insufficient households in terms of socioeconomic and demographic characteristics. For example, the average food sufficient household is significantly larger, comprises of more adult males, has attained a higher level of education and earns more household income than food insufficient households. Furthermore, food sufficient households have significantly greater access to communication, technology and transportation (e.g., mobile phones, radio, television, internet, bicycles or motorcycles). They are also more likely to hold influential positions in the community and exhibit benevolence towards others. Access to agricultural extension services, access to cooking fuel and membership in cooperatives were higher but statistically insignificant for the average food sufficient household. Around two thirds of all households help others with food.

2.4 Empirical framework

In an ideal world where food sufficiency was randomly assigned to households, estimating average treatment effects would give us the causal impact of being food sufficient on conflict. However, such an experiment that entails artificially ensuring food sufficiency for randomly assigned households is neither possible nor ethical. Since we cannot randomize an intervention to avoid selection bias, we are left with quasi-experimental techniques (see Cook, Shadish, & Wong, 2008) to improve (if not isolate) the estimates of the effect of food sufficiency on conflict. Two prominent approaches, instrumental variables and regression discontinuity, would be useful methods but are difficult to employ. Valid instruments are difficult to identify (Imbens & Woolridge, 2009). Some possibilities exist, e.g. natural disasters, but require assumptions such as exogeneity of the instrument, that are particularly difficult to justify in this context. Regression discontinuity is another option but requires consistent decision-making around some arbitrary cutoff. In our case, food insufficiency is unlikely to be allocated in such manner. In addition, as Table 2.2 shows, the inherent differences between food sufficient and insufficient households in our sample points to potential selection bias which may lead to biased estimates if compared directly. Therefore, we employ a third quasi-experimental approach - propensity score matching (PSM) – whereby observable differences between food sufficient and insufficient households that may confound the estimates are statistically balanced to neutralize any selection bias, thus allowing us to isolate the effects of food sufficiency on conflict.

To summarize, propensity score in this study is the conditional probability that a household will be food sufficient, given its vector of observed covariates. A logit model is used to estimate the propensity score. PSM pairs each food sufficient household with food insufficient households with similar observable characteristics before estimating the average treatment effect on the treated (ATT) as the difference in mean outcomes between the two groups. This can be expressed as follows:

$$ATT = \left[\left(E(Y_1 |\, T = 1) - E(Y_0 |\, T = 0)\right) - (E(Y_0 |\, T = 1) - E(Y_0 |\, T = 0))\right] \qquad (1)$$

where T is a binary 'treatment variable' equal to 1 if the household is food sufficient and 0 otherwise; Y_1 is the conflict outcome for a food sufficient household and Y_0 is the outcome for the same household had it not been food sufficient. Though Y_0 is the counterfactual, which is not observed in reality, given proper matching food insufficient households can serve as an appropriate proxy. Since PSM methods are sensitive to the exact specification and matching method (Imbens, 2004; Caliendo and Kopeinig, 2008), we employ three commonly used algorithms to ensure the robustness of PSM estimates: i) nearest neighbor matching (NNM) using three neighbors with replacement; ii) Kernel based matching using the Epanechnikov Kernel function with a bandwidth of 0.06 (Heckman, Ichimura & Todd ,1997); and iii) radius matching with a caliper of 0.001. The choice of variables included in the estimation is guided by economic theory, previous research, and the literature on matching (see Dehejia & Wahba, 2002; Heckman, Ichimura & Todd, 1997, 1998; Abadie & Imbens, 2006; and Caliendo & Kopeinig, 2008). A more detailed explanation of the PSM methodology and assumptions used is documented in Appendix B.

We also employ heterogeneous treatment effect by observable characteristics (Crump et al., 2008; Imbens & Woolridge, 2009) to estimate whether household reduction in conflict can be attributed to food sharing behaviour. To do this, we divide the full sample into two subsamples, based on whether or not the household shares food with others, before estimating two separate ATTs for each subsample. The difference of the subsample ATTs provides the heterogeneous treatment effects (see Kibriya, Zhang & Xu, 2017; Xie, Brand, & Jann, 2012; Verhofstadt & Maertens, 2014) and is expressed as follows:

$$ATT_{diff} = E[(Y_1 - Y_0)|\, T = 1, B)] - E[(Y_1 - Y_0)|\, T = 0, B)] \qquad (2)$$

where B=1 if the household displays food sharing behaviour and 0 otherwise.

2.5 Results and analysis

2.5.1 Determinants of household food sufficiency

Table 2.3 presents the results from a logit model to determine the likelihood of being food sufficient, given observable characteristics of the household.

Table 2.3: Logit estimates of the determinants of household food sufficiency

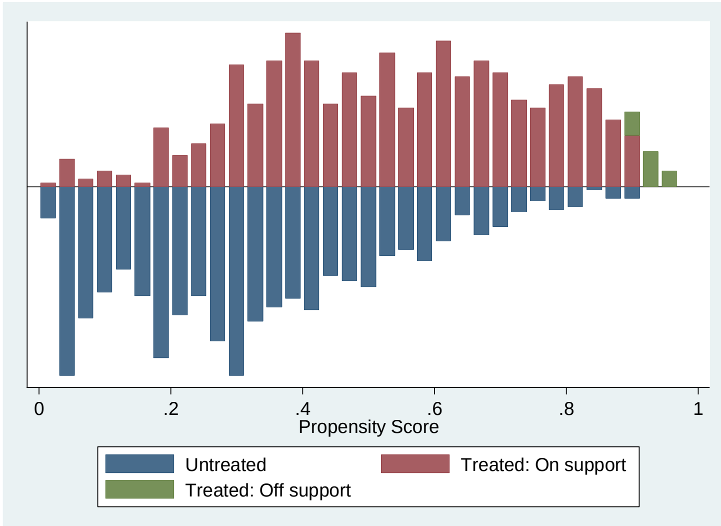
Variable	Coefficient	Standard error	Marginal effect
Dependent variable: =1 if household is food sufficient =0 otherwise			
Household size	-0.024	(0.031)	-0.007
Number of adult males	0.013	(0.061)	0.003
Household education	-0.052	(0.0404)	-0.012
Household education squared	0.006**	(0.002)	0.001
Household income (`000 CDF/capita)	0.004**	(0.001)	0.000
Household has written land claim (yes=1)	-0.057	(0.131)	-0.014
Access to technology and markets (yes=1)	0.644***	(0.151)	0.154
Lack of extension services (yes=1)	-0.433***	(0.139)	-0.103
Cooperative membership (yes=1)	-0.024	(0.148)	-0.006
Access to safe drinking water (yes=1)	0.243*	(0.134)	0.058
Inadequate access to cooking fuel (yes=1)	-0.456***	(0.127)	-0.109
Leadership position (yes=1)	0.826***	(0.257)	0.197
Constant	-2.261***	(0.491)	
Community fixed effects	Yes		
Religion controls	Yes		
Summary Statistics			
Pseudo R^2	0.18		
LR chi-square (36)	395.090***		
Log-likelihood ratio	-894.610		
Percentage correctly predicted	70.53%		
Number of observations	1,605		

Source: Authors' calculations based on the survey data.
Note: *, **, and *** indicate significance at 10%, 5% and 1% levels respectively. Community and religion controls are not shown for brevity.

It appears that certain traits are more likely to make a household food sufficient. For example, education enables a household to make informed decisions about agricultural practices (such as crop diversification or technology adoption); higher income allows households to not only purchase more food but to invest in agriculture; increased access to information and communication technologies may reduce information asymmetry as well as transaction cost for farmers; extension services from government or non-government organizations may make farming households more aware of new technologies and ways to use them. All these factors may further enhance agricultural income and productivity and thus ensure food sufficiency. Given the large fraction of rural households that use fuelwood for cooking it is not surprising that access to cooking fuel increases the probability of being food sufficient. Finally, holding important positions in the community can help households gain access to credit and services and increase social capital which can improve food sufficiency.

2.5.2 Impact of food sufficiency on conflict

Figure 2.1 shows the distribution of propensity scores between food sufficient and food insufficient households. Visual inspection of the density distributions of propensity scores for the two groups shows that there is much overlap between the estimated scores. Thus, the common support assumption is satisfied. Furthermore, there is sufficient difference in the distribution of propensity scores between the two groups to justify using matching. Figures 6-8 in the Appendix offer further visual proof of the quality of matching.



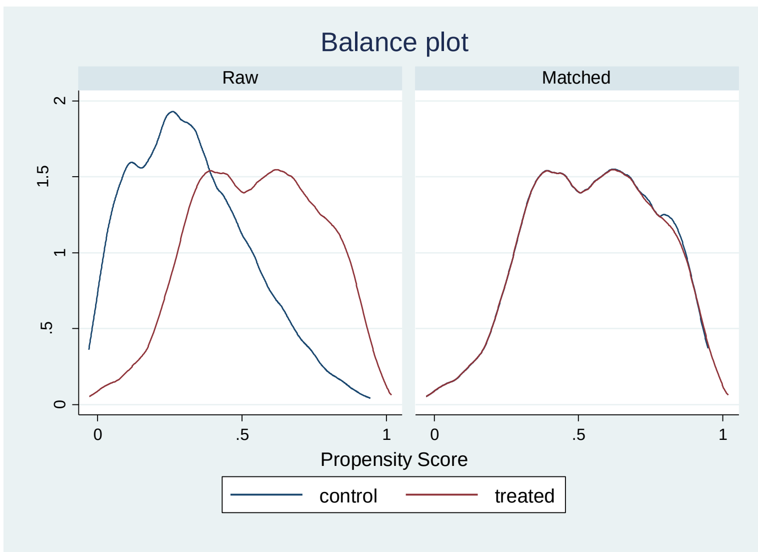


Figure 2.1: Distribution of propensity scores and the region of common support

Note: The propensity scores range from 0.016 to 0.967 with a mean value of 0.420 and a standard deviation of 0.233. Food sufficient households have propensity scores ranging between 0.024 and 0.967 with a mean score of 0.550 and standard deviation of 0.211 while food insufficient households have propensity scores ranging between 0.016 and 0.899 with a mean of 0.326 and standard deviation of 0.200. Thus, the region of common support as dictated by the minima and maxima criteria lies between 0.024 and 0.899. About 8.7% of households whose propensity scores fell outside this range were dropped from the analysis.

Table 2.4 summarizes the ATT estimates for the three conflict outcomes using each of the three matching algorithms. Consistent across the table, food sufficient households are less likely to engage in conflict than they would have been had they not been food sufficient.

Table 2.4: Average treatment effect on the treated (ATT) estimates of food sufficiency on conflict

Outcome Variable	Treatment variable: food sufficiency		
	Nearest-neighbor matching (3)	Kernel matching	Radius matching
Probability of conflict	-0.095*** (0.027)	-0.101*** (0.030)	-0.076 *** (0.042)
Probability of conflict with individual households	-0.089*** (0.0271)	-0.095 *** (0.033)	-0.100*** (0.046)
Probability of conflict with groups	-0.040* (0.023)	-0.033* (0.018)	-0.026* (0.030)

Source: Authors' calculations based on the survey data.

Note: *, **, and *** indicate significance at 10%, 5% and 1% levels respectively. All estimates shown are average treatment effect on the treated. Abadie and Imbens (2006) robust standard errors reported for nearest neighbor matching while bootstrapped standard errors with 100 replications of the sample are reported for kernel and radius matching. Kernel matching uses a bandwidth of 0.06 while radius matching uses a caliper of 0.001. Number of observations=1605 for all matching algorithms.

Depending on the algorithm used, food sufficient households are between 7.6 and 10.1 percentage points less likely to experience conflict than their food insufficient counterparts. Recalling that the average household has a 50 percent probability of experiencing conflict, a 10 percentage point reduction leads to an effect size of about 20 percent reduction in the probability of conflict. Disaggregating by conflict type, we find that food sufficiency reduces the probability that a household will engage in conflict with other households by about 9 to 10 percentage points. Given that households have a 36 percent probability of engaging in conflict with other households, a 9 to 10 percentage point reduction amounts to between 22 and 24 percent reduction in probability of conflict with individual households. Finally, food sufficient households have a 3 to 4 percentage points lower likelihood of engaging in community conflict. Though at first glance it might appear that the effect size is smaller for conflict with groups, accounting for the fact that households have a 9 percent probability of engaging in conflict with groups to begin with, this effect size actually amounts to between 33 and 44 percent reduction in the probability of conflict with groups. The lower level of significance for the coefficients on conflict with groups may have been driven by the relatively fewer number of observations in this category. These results support our expectation that controlling for socioeconomic differences, food sufficient households experience lower levels of conflict with other households and with groups within the community. This is most likely because food sufficiency reduces cause for grievance and general frustrations which can translate to aggressive behavior.

2.5.3 The heterogeneous effect of food sufficiency conditional on benevolence

In this section, we investigate the heterogeneous effects of being food sufficient, conditional on food sharing behavior. In particular, we test whether sharing food with others affects the probability of conflict for food sufficient and food insufficient households differently. Before delving into regressions, we display the summary statistics for our main conflict variables by household food sufficiency as well as food sharing behavior or benevolence status in Table 2.5.

Table 2.5: Summary of conflict by household food sufficiency and benevolence

	Household has sufficient food (1)		Household does not have sufficient food (2)		Difference in means between food sufficient and food insufficient households (3)	
Conflict measure	Benevolent (a)	Non-benevolent (b)	Benevolent (a)	Non-benevolent (b)	Benevolent (a)	Non-benevolent (b)
Probability of conflict	0.39***	0.58	0.50	0.54	***	-
Probability of conflict with individual households	0.37***	0.52	0.45	0.49	***	-
Probability of conflict with groups	0.09***	0.18	0.17	0.17	***	-

Source: Authors' calculations based on survey data.

Notes: We use t-tests to test for equal means for both benevolent and non-benevolent households, for a given food sufficiency level; and between food sufficient and food insufficient households, for a given benevolence level. *, **, and *** indicate significance at 10%, 5% and 1% levels respectively. The asterisks in column (1a) show that food sufficient households that are benevolent experience significantly lower levels of conflict than food sufficient households that are not benevolent. The absence of asterisks in column (2a) shows that the mean levels of conflict for benevolent and non-benevolent households that are food insufficient are similar. Similarly, the asterisks in column (3a) show that food sufficient households that are benevolent experience significantly lower levels of conflict that food insufficient households; while column (3b) shows that if the household is not benevolent, there are no significant differences in the mean level of conflict experienced between food sufficient and food insufficient households.

A preliminary comparison shows that food sufficient households that are benevolent experience significantly lower conflict than food sufficient households that are not benevolent. In contrast, if the household does not have sufficient food, there is no significant difference between benevolent and non-benevolent households. The last two columns show that among benevolent households, food sufficient ones have a lower probability of conflict than food insufficient ones. However, in the absence of benevolence, the difference does not appear to be significant. Since these differences in means could occur if food sufficient and insufficient households were systematically different, we proceed with a propensity score matching analysis.

To conduct this estimation, we subsample the data into households that show benevolence towards others, and households that do not. Figures 2.9 and 2.10 in the Appendix show the propensity score distributions for the matched and unmatched samples in the benevolent and non-benevolent groups respectively. For each subsample, we estimate a separate ATT and compare the results. This allows us to compare the conflict outcome for food sufficient households with the same households had they not been food sufficient, conditional on benevolence. Table 2.6 shows the results of the estimation.

Table 2.6: Effect of food sufficiency conditional upon benevolence of household.

Outcome Variable	Matching Algorithm		
	NNM (3)	KM	RM
Panel A: Effect of food sufficiency given household is benevolent			
Probability of conflict	-0.106**	-0.138***	-0.081*
	(0.045)	(0.042)	(0.046)
Probability of conflict with individual households	-0.110**	-0.124***	-0.083*
	(0.045)	(0.042)	(0.045)
Probability of conflict with groups	-0.036*	-0.053*	-0.026*
	(0.032)	(0.030)	(0.031)
Number of Treated	521	521	298
Number of Controls	585	585	585
Panel B: Effect of food sufficiency given household is not benevolent			
Probability of conflict	-0.019	-0.025	0.139
	(0.067)	(0.061)	(0.088)
Probability of conflict with individual households	-0.019	-0.019	0.136
	(0.068)	(0.061)	(0.088)
Probability of conflict with groups	-0.060	-0.052	-0.058
	(-0.060)	(0.046)	(0.060)
Number of Treated	144	143	63
Number of Controls	315	315	315

Source: Authors' own calculations based on survey data.

Note: All coefficients reported show average treatment effect on the treated. Robust standard errors in parenthesis. *, **, and *** denote significance at or below 1%, 5%, and 10% levels. Number of treated refer to the number of treated that fall in the region of common support. NNM=nearest neighbor matching using three nearest neighbors with replacement. EKM=Epanechnikov kernel matching with a bandwidth of 0.06. RM=radius matching using a caliper of 0.001. IPW-RA= inverse probability weighted regression analysis.

Panel A shows that conditional on benevolence, food sufficiency statistically significantly reduces conflict for the average household. This result holds across all matching techniques. Depending on the algorithm used, a food sufficient household that shows benevolence can expect a reduction in the probability of conflict by 8.1 and 13.8 percentage points for all kinds of low-intensity conflict; between 8.3 and 12.4 percentage points in case of conflict with individual households; and between 2.6 and 5.3 percentage points in case of conflict with groups. However, the results in panel B show that none of the coefficients are statistically significant. This implies that if a food sufficient household does not show benevolence, the effect of food sufficiency on conflict disappears and in such a case food sufficient and food insufficient households have an equal probability of experiencing conflict.

2.5.4 Sensitivity analysis and selection on unobservables

We run a series of check and balance tests to ensure that the assumptions of propensity score matching are met and that the quality of matching is reliable. These tests, presented in the appendix, include a 'Placebo' regression (Table 2.8 in Appendix); covariate balance test for the matching process (Table 2.9 in Appendix); performance comparison between nearest neighbor, radius, and Epanechnikov Kernel matching algorithms (Table 2.10 in Appendix); covariate balance test for the matching process given benevolence (Table 2.11 in Appendix) and non-benevolence (Table 2.12 in Appendix); and comparison of the matching quality indicators in benevolent and non-benevolent sub-samples (Table 2.13). From the Placebo regression we fail to reject the null hypothesis of unconfoundedness. This suggests that there are most likely no omitted variables correlated with being food sufficient and validates our assumption on selection of observables. The covariate balance tests show us that matching reduces the difference in means between treated and control groups for all covariates. Furthermore, the percentage bias in means between treated and control groups post matching is below the predetermined threshold of 25 percent (Rubin, 2001). Comparison of matching quality indicators tells us that the indicators perform better after matching. Between a low pseudo R^2, high *p*-values and a reduction in bias post-matching the tests assure us that matching has successfully balanced the distribution of covariates in treated and control groups. Overall, the results satisfy us that the assumptions and conditions of propensity score matching have been met and that proper matching quality has been ensured.

Next, we test the sensitivity of our results using a doubly robust estimator (DRE). DRE requires specifying two separate models – one for treatment (food sufficiency) and one for the outcome (conflict). The advantage of using a doubly robust estimator is that it allows for misspecification in either the treatment model or outcome model. As long as either one of the specifications is correct, DRE will provide unbiased estimates. Following Wooldridge (2010), we use the inverse probability weighting regression-adjustment (IPWRA) combination as the DRE. IPWRA estimators use weighted regression coefficients to compute averages of treatment-level predicted outcomes, where the weights are the estimated inverse probabilities of treatment. The contrasts of these averages estimate the treatment effects. Table 2.7 presents the results from the doubly robust estimation procedure using the IPWRA. The doubly robust estimates of the average treatment effects of being food sufficient are very similar to the results from the matching algorithms in Table 2.4. On average, food sufficiency reduces the likelihood that a household experiences conflict by about 10 percentage points for overall conflict; 9.5 percentage points for conflict with other households and 3.6 percentage points for conflict with groups within the community. The similarity in results from the doubly robust estimation and propensity score matching assures us of reliable estimates.

The doubly robust estimation from the impact of food sufficiency given benevolence is shown in the fourth column. The estimates are same as the propensity score estimates shown in Table 2.6. This result further substantiates our previous finding that conditional on benevolence, food sufficiency reduces conflict for households.

Table 2.7: Doubly Robust Estimation and Rosenbaum critical level of hidden bias results.

Outcome Variable	Treatment: food sufficiency		Treatment: food sufficiency given benevolence	
	IPWRA	Critical level of hidden bias (Γ)	IPWRA	Critical level of hidden bias (Γ)
Probability of conflict	-0.101*** (0.031)	5.50	-0.138*** (0.033)	2.05
Probability of conflict with individual households	-0.095*** (0.031)	1.65	-0. 124*** (0 .033)	1.65
Probability of conflict with groups	-0.0360* (0.020)	3.25	-0.053* (0 .025)	3.65
Number of observations	1605		1106	

Source: Authors' calculations based on the survey data.

Note: *, **, and *** indicate significance at 10%, 5% and 1% levels respectively. IPWRA refers to inverse probability weighted regression analysis. AI robust standard errors are reported. Critical level of hidden bias (Γ) refers to the Rosenbaum bounds for hidden bias using Hodges-Lehmann point estimates. Critical level results refer to propensity score matching using kernel estimation. Results from other matching methods are similar.

Finally, we test for potential selection on unobservable covariates using the Rosenbaum bounds for hidden bias (Rosenbaum, 2002). For example, if household members show aggressive behavior both in pursuing measures to achieve food sufficiency as well as in their preference for violence, our estimates may be biased by the unobservable presence of aggression. The Rosenbaum bound (Γ) measures how big this difference in unobservables need to be to make ATT estimates insignificant. We use the Hodges-Lehmann point estimates. Γ is the odds ratio of being assigned to treatment, in this case food sufficiency. When Γ =1, the assumption is no potential hidden bias implying an equal probability of being food sufficient or food insufficient for given observable characteristics. Under this assumption, the results are similar to our estimates. Table 2.7 shows the critical levels of hidden bias for our estimates. In general, the values of Γ range between 1.65 and 5.5. For example, when the outcome of interest is probability of conflict, as long as the odds of being assigned to food sufficiency is 5.5 or lower, the ATT will not be sensitive to hidden bias. Put in another way, it implies that the unobserved covariates would have to increase the odds of being food sufficient by 450 percent to overturn the significance of our ATT estimates. On conditioning upon benevolence, the results are more sensitive to hidden bias with Γ being just over 2. The critical levels of hidden bias for probability of conflict with individuals and probability of conflict with groups is similar across the two treatments (food sufficiency and food sufficiency given benevolence). For both treatments matched households with similar observable characteristics would have to differ by a factor of 1.65 (65 percent) in unobserved covariates to overturn the significance of the ATT estimates on probability of conflict with individual households. In case of probability of conflict with groups the odds of being food sufficient would need to be more than 3.25 (225 percent) while the odds of being food sufficient, given benevolence, would need to be 3.65 (265 percent) to affect the significance of the estimates. In comparison it appears that the probability of conflict with groups is less sensitive to potential hidden bias than the probability of conflict with individual households. It should be noted however that a lower value of Γ does not imply that the one would not observe an effect on the probability of conflict with individual households in the presence of hidden bias. The Rosenbaum bounds presents a worst-case scenario that assumes treatment assignment is influenced by the presence of unobservable covariates (Li, 2013). The results show that the probability of conflict with individual households is more sensitive to potential hidden bias that could affect a household’s odds of being treated. Even so, contemporary studies using the same methodology report similar values of Γ (e.g., Ogutu, Okello & Otiendo, 2014; Chiputwa, Spielman & Oaim, 2015; Wendimu, Henningsen & Gibbon, 2016). Overall, these results somewhat assure us that any unobservable characteristics would have to cause matched households to differ substantially for the ATT estimates to be affected by potential hidden bias.

2.6 Summary, Discussion and Concluding Remarks

By exploiting survey data of 1763 households collected from three territories in the North Kivu province of eastern DRC, we study the impact of food sufficiency and foods sharing on low-intensity interhousehold and community conflict. Since food sufficient households may be systematically different from food insufficient households, we use the quasi-experimental method of propensity score matching to control for any pre-existing differences. This allows us to compare conflict experiences of a food sufficient household with essentially the same household had it not been food sufficient, thus allowing us to plausibly isolate the effect of food sufficiency on household conflict. By exploiting heterogeneous treatment effects, we find empirical evidence to support that food sufficiency can reduce the probability of conflict for households only in the presence of benevolence. Food sufficient households that show benevolence towards others reduce their overall probability of conflict by an average of 13.8 percentage points; a reduction of up to 12.4 percentage points in the probability of conflict against individual households and a reduction of up to 5.3 percentage points in the probability of conflict against groups within the community.

Potential biases were accounted for through various econometric approaches. The assumption of selection on observables is addressed through a placebo regression, while the overlap assumption is assessed through normalized differences in means and graphical representation of propensity score distributions. The inverse probability weighted regression analysis is used as a doubly robust estimator to check the robustness of our estimates. Finally, the Rosenbaum bounds for hidden bias is used to test for any potential bias arising from unobservable confounders. Although we take extreme caution to claim causality, our checks and balance tests do not indicate concern for violations of the assumptions used, suggesting that a causal claim of our finding is plausible, at the least.

Our explorations isolate food insufficient households and show that they are more prone to conflict with government and park officials, NGOs, neighbors, rebel groups and fellow villagers. For the impoverished and deprived, these quantitative results validate the notions of rational choice theory presented in the "Moral Economy" (Thompson, 1963). For food insufficient households lacking institutional support, engaging in conflict is a rational option due to lack of attainable resources. Grievances towards the social hierarchy and governance systems further drive these actions (Thompson, 1963). These situations are similar to the European food riots of 17th and 18th century when absence of welfare, regulation, safety nets and strict social hierarchy evoked aggressive reactions from the food insecure. While our investigation does not explore communal responses or collective action, the premises remain similar with individual households. These incidences of house level conflict are not random events driven by sudden impulse but may be linked to "subsistence ethic" and "notion of right".[9]

Extending from the theory of moral economic choices, the second part of our results contributes to the philosophical and social norms of gift exchange. The act of gifting food makes food sufficient households immune to social aggression. Thompson's idea of moral economy is

[9] Scott (1977) and O'Brien and Li (2006) also provided similar situation of collective peasant protests for Vietnam, Burma and China albeit with a different approach.

based on social justice, which is, ideally, built on fairness and mutual benefits. They can be practiced in closed communities and difficult to envision in modern market economies regulated by institutions. North Kivu's failed market conditions with pervasive scarcity initiates an informal economy which depends on barter, social justice and benevolence (or fairness). Gift in such contexts is tantamount to Mauss's (1970) Polynesian cultural interactions or Malinowski's (1922) Papua New Guinean social exchanges. While Mauss and Malinowski differ in the motives of the giver, they both agree that social giving can gel kinship bonds which serves socio-economic and political functions. At a household level, gifting can promote social fairness and justice and ultimately serve mutual interests of both the food sufficient and insufficient. Food sharing thus promotes the greatness of the giver and establishes a higher social hierarchy for the household. On the contrary, a food sufficient household which does not share its food stands to lose its honour and social hierarchy. The givers of food can present food either from their kindness or a selfish motive to avoid conflict. The receivers of food are then bound to reciprocity (Malinowski 1922) or obligation (Mauss 1970) to remain peaceful with the givers while seeking justice from others. The social norm of reciprocity and obligation thus avoids conflict for food sufficient households who share their food. The natural extension of this research will be to investigate such phenomena within villages or tribes that would shed light on communal and tribal conflict.

While the existing literature mostly uses cross country or district level data for analyses of civil wars and conflicts, we shed light on the facets of interhousehold and community conflicts that most frequently do not make headlines and are subsequently ignored. Our findings also advance the understanding of the intricate relationship between food sufficiency, kinship and conflict at the micro level. Food aid programs have been documented to have mixed effects on conflict (Barrett, 2001; Nunn & Qian, 2014). Our approach of analysing the connection between household level food sufficiency and food sharing tendencies with low-intensity local conflict can offer new insights to program implementers and evaluators of international food policy. Our findings show that food sufficiency alone cannot reduce low-intensity interhousehold and community level conflict unless accompanied the idea of kinship. As such, our results illuminate the need for food aid and other relevant food policies to incorporate communal affiliation. Thus, policymakers may find that encouraging food sharing practices within society is a useful and effective tool that can complement food security, poverty alleviation and conflict reduction initiatives.

References

Abadie, A., & Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica*, *74*(1), 235-267.

Arezki, M. R., & Bruckner, M. (2011). *Food prices and political instability* (No. 11-62). International Monetary Fund.

Barnett, J., & Adger, W. N. (2007). Climate change, human security and violent conflict. *Political geography*, *26*(6), 639-655.

Barrett, C. B. (2001). Does food aid stabilize food availability? *Economic development and cultural change*, *49*(2), 335-349.

Bellemare, M. F. (2015). Rising food prices, food price volatility, and social unrest. *American Journal of Agricultural Economics*, *97*(1), 1-21.

Berazneva, J., & Lee, D. R. (2013). Explaining the African food riots of 2007–2008: An empirical analysis. *Food Policy*, *39*, 28-39.

Bessler, D. A., Kibriya, S., Chen, J., & Price, E. (2016). On Forecasting Conflict in the Sudan: 2009–2012. *Journal of Forecasting*, *35*(2), 179-188.

Bora, S., Ceccacci, I., Delgado, C., & Townsend, R. (2010). World Development Report 2011: Food sufficiency and Conflict. *Washington, DC: Agriculture and Rural Development Department, World Bank*.

Brinkman, H. J., & Hendrix, C. S. (2011). *Food insufficiency and violent conflict: Causes, consequences, and addressing the challenges*. World Food Programme.

Brück, T., d'Errico, M., & Pietrelli, R. (2018). The effects of violent conflict on household resilience and food security: Evidence from the 2014 Gaza conflict. *World Development*

Burke, M. B., Miguel, E., Satyanath, S., Dykema, J. A., & Lobell, D. B. (2009). Warming increases the risk of civil war in Africa. *Proceedings of the national Academy of sciences*, *106*(49), 20670-20674.

Bush, R., & Martiniello, G. (2017). Food Riots and Protest: Agrarian Modernizations and Structural Crises. *World Development*, *91*, 193-207.

Bushman, B. J., DeWall, C. N., Pond, R. S., & Hanus, M. D. (2014). Low glucose relates to greater aggression in married couples. *Proceedings of the National Academy of Sciences*, *111*(17), 6254-6257.

Caliendo, M., and S. Kopeinig. 2008. Some Practical Guidance for the Implementation of Propensity Score Matching. Journal of Economics Surveys 22 (1): 31–72.

Carter, K. N., Kruse, K., Blakely, T., & Collings, S. (2011). The association of food security with psychological distress in New Zealand and any gender differences. *Social science & medicine*, *72*(9), 1463-1471.

Chilton, M., & Booth, S. (2007). Hunger of the body and hunger of the mind: African American women's perceptions of food insecurity, health and violence. *Journal of nutrition education and behavior*, *39*(3), 116-125.

Chiputwa, B., Spielman, D. J., & Qaim, M. (2015). Food standards, certification, and poverty among coffee farmers in Uganda. *World Development*, *66*, 400-412.

Cohen, M. J. (1993). Hunger 1994: transforming the politics of hunger. *Bread for the World Institute, Silver Spring, MD*.

Cohen, M. J., & Pinstrup-Andersen, P. (1999). Food sufficiency and conflict. *Social Research*, 375-416. Food sufficiency, Justice and Peace (2002). FAO.

Collier, P., & Hoeffler, A. (2004). Greed and grievance in civil war. *Oxford economic papers*, *56*(4), 563-595.

Cook, T. D., Shadish, W. R., & Wong, V. C. (2008). Three conditions under which experiments and observational studies produce comparable causal estimates: New findings from within-study comparisons. *Journal of policy analysis and management*, *27*(4), 724-750.

Crump, R. K., Hotz, V. J., Imbens, G. W., & Mitnik, O. A. (2008). Nonparametric tests for treatment effect heterogeneity. *The Review of Economics and Statistics*, *90*(3), 389-405.

Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and statistics*, *84*(1), 151-161.

Fearon, J. D., Humphreys, M., & Weinstein, J. M. (2009). Can development aid contribute to social cohesion after civil war? Evidence from a field experiment in post-conflict Liberia. *The American Economic Review*, *99*(2), 287-291.

Global Humanitarian Report-An Overview 2015. UN Office for the Coordination of Humanitarian Affairs

Heckman, J. J., Ichimura, H., & Todd, P. (1998). Matching as an econometric evaluation estimator. *The Review of Economic Studies*, *65*(2), 261-294.

Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The review of economic studies*, *64*(4), 605-654.

Heflin, C. M., Siefert, K., & Williams, D. R. (2005). Food insufficiency and women's mental health: findings from a 3-year panel of welfare recipients. *Social science & medicine*, *61*(9), 1971-1982.

Hendrix, C. S., & Salehyan, I. (2010). After the Rain: Rainfall Variability, Hydro-Meteorological Disasters, and Social Conflict in Africa.

Hendrix, C., & Brinkman, H. J. (2013). Food insufficiency and conflict dynamics: Causal linkages and complex feedbacks. *Stability: International Journal of Security and Development*, *2*(2).

Homer-Dixon, T. F. (1994). Environmental scarcities and violent conflict: evidence from cases. *International security*, *19*(1), 5-40.

Homer-Dixon, T. F. (1999). Environment, scarcity, and conflict. *Princeton University*.

Humphreys, M., & Weinstein, J. M. (2008). Who fights? The determinants of participation in civil war. *American Journal of Political Science*, *52*(2), 436-455.

Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and statistics*, *86*(1), 4-29.

Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of economic literature*, *47*(1), 5-86.

Justino, P. (2011). The impact of armed civil conflict on household welfare and policy responses. *Securing Peace: State-Building and Economic Development in Post-Conflict Countries*, *19*.

Justino, P., Brück, T., & Verwimp, P. (Eds.). (2013). *A micro-level perspective on the dynamics of conflict, violence, and development*. Oxford University Press.

Kahl, C. H. (2006). *States, scarcity, and civil strife in the developing world*. Princeton University Press.

Kibriya, S., Xu, Z. P., & Zhang, Y. (2017). The negative consequences of school bullying on academic performance and mitigation through female teacher participation: evidence from Ghana. *Applied Economics*, *49*(25), 2480-2490.

Leuveld, K. et al. (Forthcoming). The impact of agricultural extension and input subsidies on knowledge, input use and food security in Eastern DRC

Li, M. (2013). Using the propensity score method to estimate causal effects: A review and practical guide. *Organizational Research Methods*, *16*(2), 188-226.

Macrae, J., & Zwi, A. B. (1992). Food as an instrument of war in contemporary African famines: a review of the evidence. *Disasters*, *16*(4), 299-321.

Malinowski, B. (1984) [1922]. *Argonauts of the Western Pacific: An account of native enterprise and adventure in the archipelagoes of Melanesian New Guinea*. Prospect Heights, Ill.: Waveland Press.

Marivoet, W., & De Herdt, T. (2019). What Happens with Household Assets during Economic Collapse? The Case of the Democratic Republic of the Congo (1975–2010). *The Journal of Development Studies*, *55*(4), 680-701.

Martin-Shields, C. P., & Stojetz, W. (2018). Food Security and Conflict: Empirical challenges and future opportunities for research and policy making on food security and conflict. *World Development.*

Mauss, M. (1970). *The Gift: Forms and Functions of Exchange in Archaic Societies*. London: Cohen & West.

Mauss, M. 1990 [1950] The Gift: The form and reason for exchange in archaic societies. W.D. Halls, trans. Norton & Company, Inc. New York.

Maxwell, D., Watkins, B., Wheeler, R., & Collins, G. (2003). The coping strategies index: A tool for rapidly measuring food security and the impact of food aid programs in emergencies. *Nairobi:*

CARE Eastern and Central Africa Regional Management Unit and the World Food Programme Vulnerability Assessment and Mapping Unit.

McCandless, E (2012). *Peace Dividends and Beyond: Contributions of Administrative and Social Services to Peacebuilding*. New York: UN/PBSO.

Mercier, M., Lionel Ngenzebuke, R. L. Verwimp, P. (2017) Violence exposure and deprivation: Evidence from the Burundi civil war.

Messer, E., Cohen, M. J., & d'Costa, J. (1998). *Food from peace: Breaking the links between conflict and hunger* (Vol. 24). Intl Food Policy Res Institute.

Miguel, E., Satyanath, S., & Sergenti, E. (2004). Economic shocks and civil conflict: An instrumental variables approach. *Journal of political Economy*, *112*(4), 725-753.

Nunn, N., & Qian, N. (2014). US food aid and civil conflict. *The American Economic Review*, *104*(6), 1630-1666.

O'brien, K. J., & Li, L. (2006). *Rightful resistance in rural China*. Cambridge University Press.

Ogutu, S. O., Okello, J. J., & Otieno, D. J. (2014). Impact of information and communication technology-based market information services on smallholder farm input use and productivity: The case of Kenya. *World Development*, *64*, 311-321.

Østby, G. (2008). Polarization, horizontal inequalities and violent civil conflict. *Journal of Peace Research*, *45*(2), 143-162.

Pinstrup-Andersen, P., & Shimokawa, S. (2008). Do poverty and poor health and nutrition increase the risk of armed conflict onset? *Food Policy*, *33*(6), 513-520.

Poverty and household living conditions in the North-Kivu Province. 2009. UNDP Report.

Raeymaekers, T. (2008). Conflict and food sufficiency in Beni-Lubero: back to the future? *Alinovi, Hemrich y Russo (2008), págs*, 169-195.

Raleigh, C (2010). 'Political Marginalization, Climate Change, and Conflict in Sahelian States'. *International Studies Review* 12(1): 69–86.

Richards, P. (1998). *Fighting for the rain forest: war, youth & resources in Sierra Leone* (No. Reprinted Ed.). James Currey Ltd.

Rojas, M., & Guardiola, J. (2017). Hunger and the experience of being well: Absolute and relative concerns. *World Development*, *96*, 78-86.

Rosenbaum, P. R. (2002). Covariance adjustment in randomized experiments and observational studies. *Statistical Science*, *17*(3), 286-327.

Rosenbaum, P. R. (2002). Observational studies. In *Observational Studies* (pp. 1-17). Springer New York.

Rosenbaum, P., & Rubin, D. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika, 70*(1), 41-55.

Rubin, D. B. (2001). Using propensity scores to help design observational studies: application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, *2*(3), 169-188.

Salehyan, I. (2008). From climate change to conflict? No consensus yet. *Journal of Peace Research*, *45*(3), 315-326.

Schomerus, M., & Allen, T. (2010). Southern Sudan at odds with itself: Dynamics of conflict and predicaments of peace.

Scott, J. C. (1977). *The moral economy of the peasant: Rebellion and subsistence in Southeast Asia*. Vol. 315. Yale University Press.

Stearns, J. (2012). North Kivu: The background to conflict in North Kivu province of eastern Congo. Usalama Project, Rift Valley Institute, London, 2012.

Stewart, F. (1998). Food aid during conflict: can one reconcile its humanitarian, economic, and political economy effects. *American Journal of Agricultural Economics*, *80*(3), 560-565.

Stewart, F. (2011). Inequality in Political Power: A fundamental (and overlooked) dimension of inequality. *The European Journal of Development Research*, *23*(4), 541.

Taeb, M. (2004). Agriculture for Peace: Promoting agricultural development in support of peace. The International Fund for Agricultural Development (IFAD), United Nations. IFAD in the Democratic Republic of the Congo.

Thompson, E. P. (1991). The Making of the English Working Class [1963].

Turner, M. D., Ayantunde, A. A., Patterson, K. P., & Patterson III, E. D. (2011). Livelihood transitions and the changing nature of farmer–herder conflict in Sahelian West Africa. *The journal of development studies*, *47*(2), 183-206.

UNICEF, M. (2000). Multiple indicator cluster survey (MICS).

United Nations (2015). North Kivu Factsheet: MONUSCO. United Nations Organisation Stabilization Mission in the Democratic Republic of Congo.

Van Weezel, S. (2017). Estimating the effect of conflict on food supply at the national level.

Verhofstadt, E., & Maertens, M. (2014). Can agricultural cooperatives reduce poverty? Heterogeneous impact of cooperative membership on farmers' welfare in Rwanda. *Applied Economic Perspectives and Policy*, ppu021.

Vlassenroot, K., & Huggins, C. (2005). Land, migration and conflict in eastern DRC. *From the ground up: land rights, conflict and peace in sub-Saharan Africa*, 115-195.

Vlassenroot, K., & Raeymaekers, T. (2008). Crisis and food sufficiency profile: The Democratic Republic of the Congo. *Beyond relief: food sufficiency in protracted crises*, 157-168.

Wendimu, M. A., Henningsen, A., & Gibbon, P. (2016). Sugarcane outgrowers in Ethiopia:"Forced" to remain poor?. *World Development*, *83*, 84-97.

White, S. C., Fernandez, A., & Jha, S. (2016). Beyond the grumpy rich man and the happy peasant: mixed methods and the impact of food security on subjective dimensions of wellbeing in India. *Oxford Development Studies*, *44*(3), 332-348.

Winne, M. (2010). *Food Rebels, Guerrilla Gardeners, and Smart-Cookin' Mamas: Fighting Back in an Age of Industrial Agriculture*. Beacon Press.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

World Bank (2011). *World Development Report 2011: Conflict, Security, and Development*. Washington, DC: World Bank.

World Bank (2016). "Poverty and Shared Prosperity 2016: Taking on Inequality."

World Food Programme (WFP). 2014. Democratic Republic of Congo: Comprehensive Food Security and Vulnerability Analysis (CFSVA). Rome: WFP.

World Food Programme (WFP). 2016. Democratic Republic of Congo: Comprehensive Food sufficiency and Vulnerability Analysis (CFSVA).

World Food Programme (WFP). 2018. Democratic Republic of the Congo. Available online at http://www1.wfp.org/countries/democratic-republic-congo

Xie, Y., Brand, J. E., & Jann, B. (2012). Estimating heterogeneous treatment effects with observational data. *Sociological methodology*, *42*(1), 314-347.

APPENDIX II A

Tables and Figures

Table 2.8: Estimation results from a Placebo regression

Dependent variable is farmer ID	Coefficient	Standard error
Household is food sufficient	-13.02	9.53
Household size	-0.63	2.32
Number of adult males	-11.18	4.34
Household education	-2.53	2.89
Household education squared	0.13	0.17
Household income (`000 CDF/capita)	1.03×10^{-4}	1.00×10^{-4}
Household has written land claim (yes=1)	14.27	9.34
Access to technology and markets (yes=1)	-27.78***	10.72
Lack of extension services (yes=1)	-13.84	10.05
Cooperative membership (yes=1)	12.31	10.73
Access to safe drinking water (yes=1)	51.17***	10.67
Inadequate access to cooking fuel (yes=1)	-12.45	9.46
Leadership position (yes=1)	24.88	18.28
Community fixed effects	Yes	
Religion controls	Yes	

Source: Authors' own calculations.

Note: *, **, and *** indicate significance at 10%, 5% and 1% levels respectively. A 'Placebo' regression is a regression of the treatment variable and all controls on an exogenous dependent variable which is unlikely to be related to the treatment. The dependent variable we chose is the farmer identification number (ID) coded during the interview process. Results show that the coefficient on the treatment variable is not significantly different from zero. While not a proof of the unconfoundedness assumption, failure to reject the null hypothesis of unconfoundedness suggests that there are most likely no omitted variables correlated with being food sufficient. This validates our assumption on selection of observables.

Table 2.9: Balancing properties of covariates before and after matching

Covariate	Sample	Mean				
		Treated	Control	% Bias	% Reduction in bias	Diff: *p*-value
Household size	U	5.45	5.23	9.1		**0.058**
	M	5.33	5.30	1.3	85.5	0.834
Number of adult males	U	2.24	2.10	12.1		**0.011**
	M	2.15	2.23	-6.8	44.3	0.293
Household education	U	9.50	8.83	14.6		**0.004**
	M	9.15	9.54	-8.3	43.3	0.214
Household education squared	U	111.90	99.42	15.5		**0.002**
	M	106.10	115.42	-11.6	25.3	0.095
Household income (`000 CDF/capita)	U	19583	16370	8.1		0.101
	M	20235	20009	0.6	93	0.943
Household has written land claim (yes=1)	U	0.37	0.43	-11		**0.023**
	M	0.4	0.42	-4.4	60.3	0.495
Access to technology and markets (yes=1)	U	0.83	0.69	34.7		**0**
	M	0.80	0.82	-6.3	81.9	0.294
Lack of extension services (yes=1)	U	0.59	0.62	-6.6		0.169
	M	0.60	0.59	1.9	71.9	0.772
Cooperative membership (yes=1)	U	0.22	0.21	4.5		0.35
	M	0.23	0.24	-0.1	97.6	0.987
Access to safe drinking water (yes=1)	U	0.63	0.63	1.1		0.817
	M	0.65	0.66	-3.1	-172.2	0.626
Inadequate access to cooking fuel (yes=1)	U	0.56	0.64	-16.8		**0.001**
	M	0.58	0.56	4	76	0.536
Leadership position (yes=1)	U	0.96	0.91	18.6		**0**
	M	0.94	0.96	-5.4	71.1	0.335

Source: *Authors' calculations based on the survey data.*

Notes: As seen from the table, the means of the treated and control groups are significantly different for most covariates prior to matching. The matching process reduces the difference in means between treated and control groups for all covariates such that there are no significant differences between the means of the two groups after matching. In addition, we test the percentage bias in means between the treated and control groups post matching. Following Rubin (2001), we consider a covariate to be balanced across treated and control groups if the absolute percent standardized difference in mean bias in the matched sample is 25% or less. The absolute percent standardized difference in mean bias between treated and control groups is indeed less than 25% for all covariates in the matched sample. Since 25%

is a rule of thumb, it is assuring to find that the absolute percentage bias in all our covariates is in fact less than 12%. These figures ensure us that the balancing property is satisfied for all covariates of interest.

Table 2.10: Comparing matching quality indicators among the three matching algorithms

Matching algorithm	Pseudo R^2		LR χ^2		$p > \chi^2$		Mean standardized bias		Total % bias reduction
	Before	After	Before	After	Before	After	Before	After	
NNM	0.180	0.010	392.97	19.26	0.000	0.990	14.0	3.2	77.9
EKM	0.180	0.007	392.97	12.17	0.000	1.000	14.0	2.7	82.3
RM	0.180	0.012	392.97	16.96	0.000	0.997	14.0	3.9	75.3

Source: Authors' own calculations using the survey data.

Note: NNM=nearest neighbor matching using three nearest neighbors, with replacement. EKM= Epanechnikov kernel matching with a bandwidth of 0.06. RM=radius matching using a caliper of 0.001. Before and after columns show results before matching and after matching. For all three matching algorithms the standardized mean bias for covariates reduced from 14.0 before matching to a range of 2.7 and 3.9 after matching. The total percentage bias reduced by around 78 to 82 percent. The p-values of the likelihood ratio tests show the joint significance of all covariates in the logit regression after matching. The low values of the pseudo R^2 after matching indicate that there is no systematic difference in the distribution of the treated and control groups. Overall, the low pseudo R^2, the high p-values and the reduction in bias post matching assure us that the propensity score matching has successfully balanced the distribution of covariates in treated and control groups. Although the values are similar for all three methods used, the performance was slightly better for kernel-based matching.

Table 2.11: Covariate balance in treated and control groups for benevolent households

Covariate	Sample	Treated	Control	% Reduction in bias	Diff: *p*-value
Household size	U	5.50	5.18		**0.027**
	M	5.53	5.50	89.7	0.836
Number of adult males	U	2.24	2.14		0.148
	M	2.25	2.31	38.6	0.422
Household education	U	9.73	9.38		0.194
	M	9.73	9.55	46	0.501
Household education squared	U	114.84	107.01		0.099
	M	115.06	110.51	41.8	0.38
Household income (`000 CDF/capita)	U	19553	15483		0.05
	M	19716	27362	-87.9	0.127
Household has written land claim (yes=1)	U	0.39	0.48		**0.004**
	M	0.40	0.41	78.2	0.556
Access to technology and markets (yes=1)	U	0.86	0.76		**0**
	M	0.86	0.86	98	0.929
Lack of extension services (yes=1)	U	0.57	0.60		0.318
	M	0.55	0.58	22.8	0.472
Cooperative membership (yes=1)	U	0.26	0.23		**0.2**
	M	0.26	0.25	60.6	0.642
Access to safe drinking water (yes=1)	U	0.63	0.64		0.695
	M	0.63	0.62	74.9	0.927
Inadequate access to cooking fuel (yes=1)	U	0.57	0.65		**0.003**
	M	0.58	0.53	35	0.077
Leadership position (yes=1)	U	0.97	0.90		**0**
	M	0.97	0.97	96.2	0.802

Source: Authors' calculations based on the survey data.

Note: U=unmatched sample and M=matched sample. For each covariate, the standardized mean percent reduction in bias is calculated using one minus the difference in means between treated and control groups after matching divided by the difference in means between treated and control groups before matching. Bold p-values indicate the difference in means are significant at a level of 10% or lower. Due to space constraints, the means for community and religion dummies have been excluded from the table. The number of observations is 675 for treated and 930 for control groups. The balancing tests presented here are for the onset of conflict outcome using radius-caliper matching. The results are similar for other outcomes and for the other matching algorithms used. Therefore, to save space those are not reported. N=1054. The means of the treated and control groups are significantly different for most covariates prior to matching.

The matching process reduces the difference in means between treated and control groups for all covariates such that there are no significant differences between the means of the two groups after matching.

Table 2.12: Covariate balance in treated and control groups for non-benevolent households

Covariate	Sample	Treated	Control	% Reduction in bias	Diff: *p*-value
Household size	U	5.27	5.32		0.772
	M	5.14	5.25	-87.5	0.673
Number of adult males	U	2.27	2.03		**0.015**
	M	2.27	2.33	73.2	0.640
Household education	U	8.83	7.92		**0.062**
	M	8.79	8.79	99.8	0.997
Household education squared	U	103.33	86.99		**0.043**
	M	103.35	101.89	91	0.887
Household income (`000 CDF/capita)	U	19886	17761		0.656
	M	20208	23647	-61.9	0.662
Household has written land claim (yes=1)	U	0.30	0.35		0.327
	M	0.31	0.34	45.4	0.669
Access to technology and markets (yes=1)	U	0.78	0.57		**0.000**
	M	0.76	0.76	97.6	0.920
Lack of extension services (yes=1)	U	0.65	0.66		0.803
	M	0.65	0.66	53.3	0.926
Cooperative membership (yes=1)	U	0.10	0.17		**0.044**
	M	0.10	0.10	99.8	0.996
Access to safe drinking water (yes=1)	U	0.63	0.61		0.651
	M	0.67	0.68	53.4	0.862
Inadequate access to cooking fuel (yes=1)	U	0.56	0.65		0.068
	M	0.57	0.53	48.5	0.464
Leadership position (yes=1)	U	0.91	0.93		0.474
	M	0.93	0.92	62.9	0.832

Source: Authors' calculations based on the survey data.

Note: U=unmatched sample and M=matched sample. For each covariate, the standardized mean percent reduction in bias is calculated using one minus the difference in means between treated and control groups after matching divided by the difference in means between treated and control groups before matching. Bold p-values indicate the difference in means are significant at a level of 10% or lower. Due to space constraints, the means for community and religion dummies have been excluded from the table. The number of observations is 675 for treated and 930 for control groups. The balancing tests presented here are for the onset of conflict outcome using radius-caliper matching. The results are similar for other outcomes and for the other matching algorithms used. Therefore, to save space those are not reported. N=459. The means of the treated and control groups are significantly different for most covariates prior to matching.

The matching process reduces the difference in means between treated and control groups for all covariates such that there are no significant differences between the means of the two groups after matching.

Table 2.13: Matching quality indicators for benevolent and non-benevolent households

Sample	Pseudo R^2	LR χ^2	$p > \chi^2$	Mean standardized bias	%Bias	Total % bias reduction
Panel A: Household is benevolent						
Unmatched	0.197	301.68	0	14.5	112.5*	
Matched	0.013	18.62	0.993	3.7	26.5*	76.4
Panel B: Household is not benevolent						
Unmatched	0.174	99.62	0	15.1	107.6*	
Matched	0.009	3.5	1	2.8	22.1	79.4

Source: Authors' own calculations using the survey data.

Note: Results shown for Epanechnikov kernel matching with a bandwidth of 0.06. * indicates that %bias is over 25.

Overall, the indicators perform better after matching, thereby ensuring the quality of the matching process in both subsamples.

Figures

Figure 2.2: Map of DRC showing North Kivu

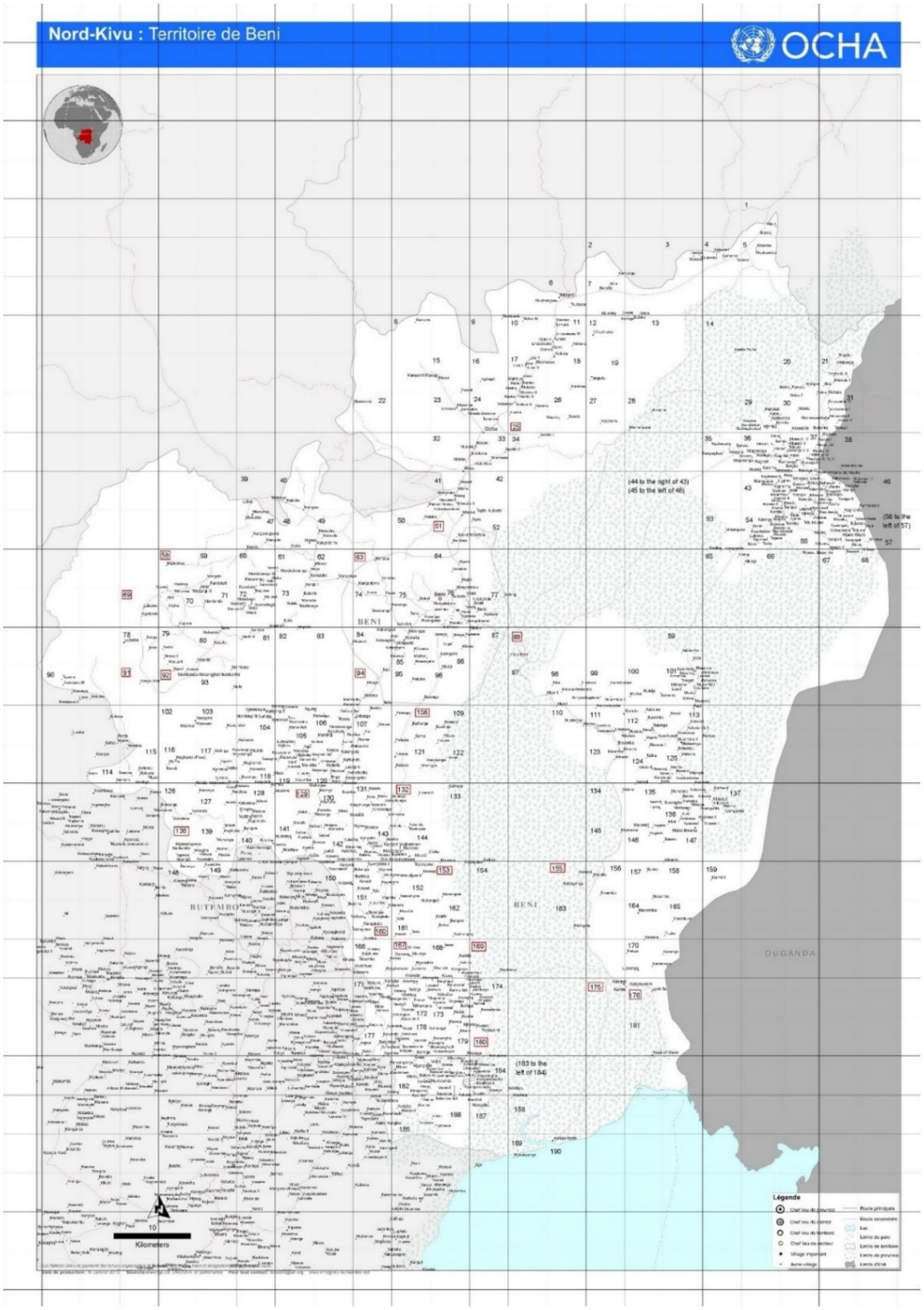


Figure 2.3: Grid map of Beni territory

Source: The United Nations Office for the Coordination of Humanitarian Affairs (OCHA), available at

www.rgc.cd

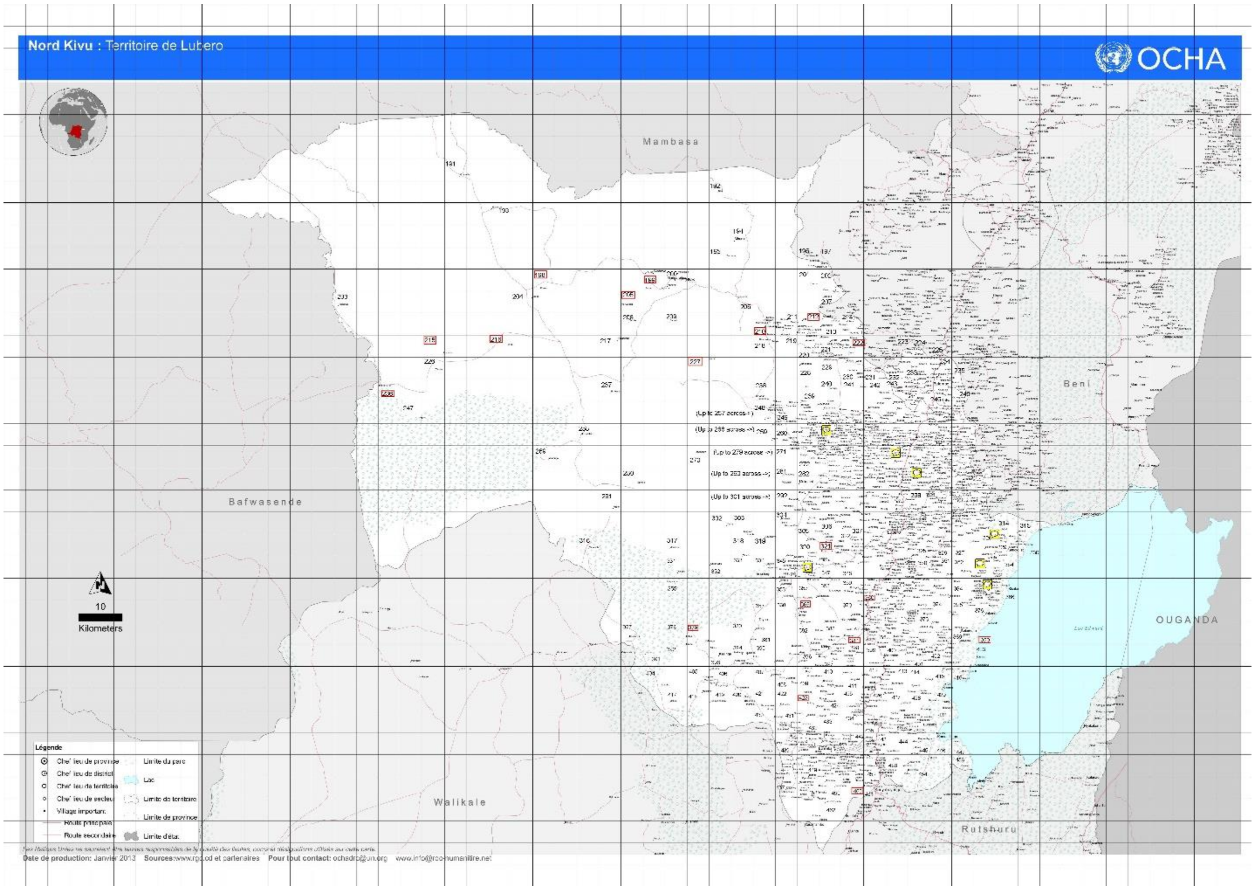


Figure 2.4: Grid map of Lubero territory.

Source: The *United Nations Office for the Coordination of Humanitarian Affairs* (OCHA), available at www.rgc.cd

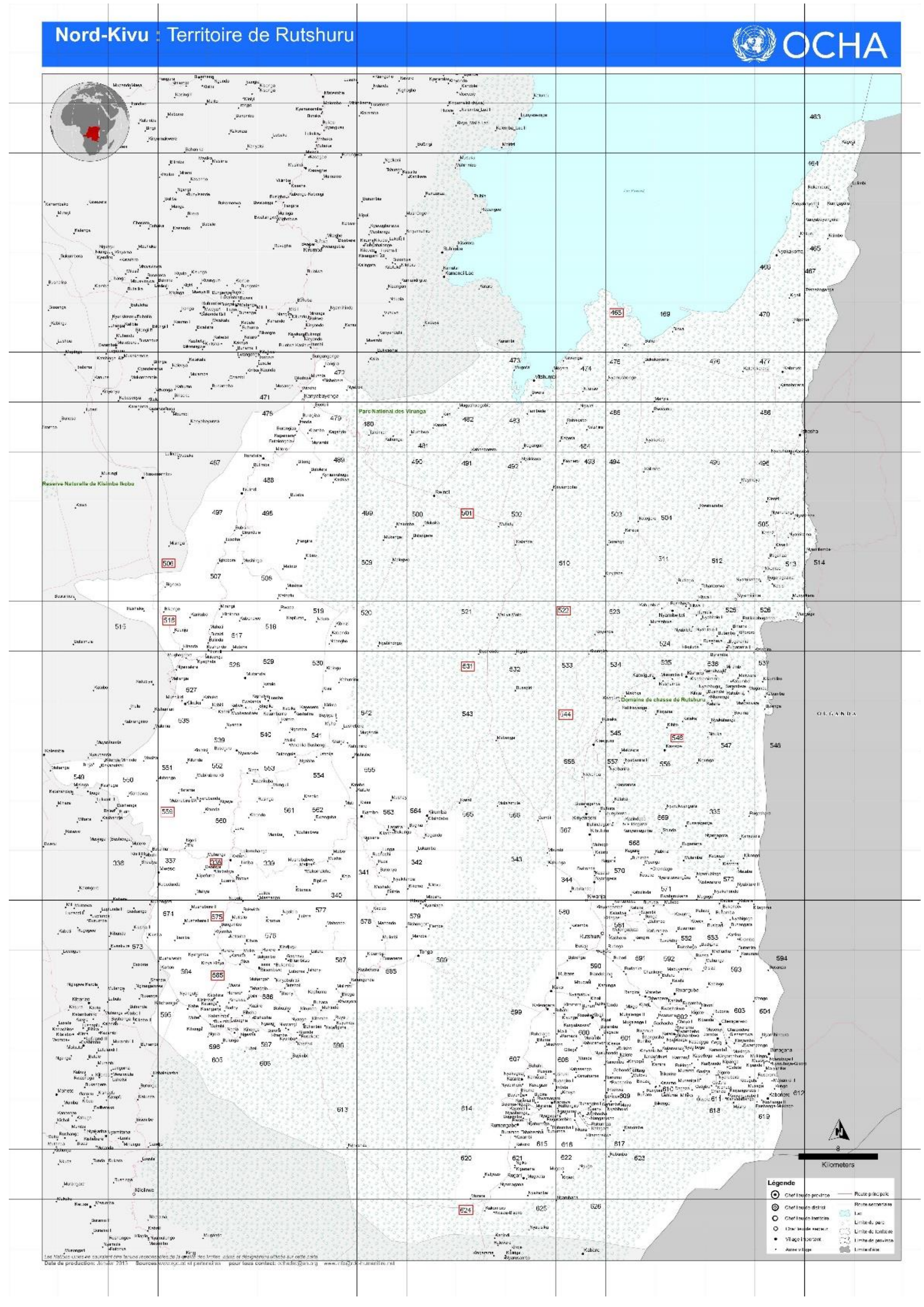


Figure 2.5: Grid map of Rutshuru territory

Source: The *United Nations Office for the Coordination of Humanitarian Affairs* (OCHA), available at www.rgc.cd

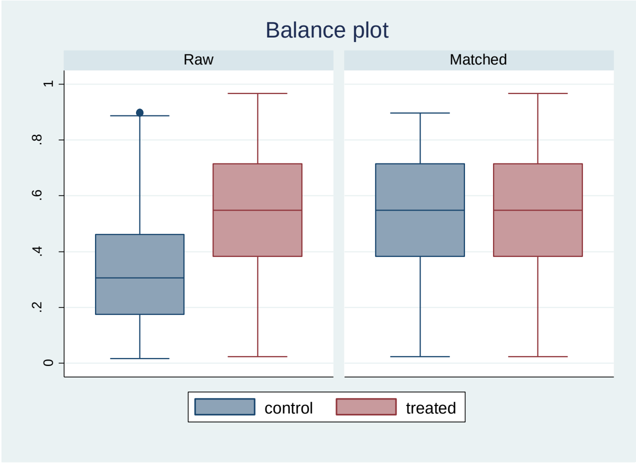


Figure 2.6: Box plot to show distribution of propensity score between treated and control groups before and after matching.

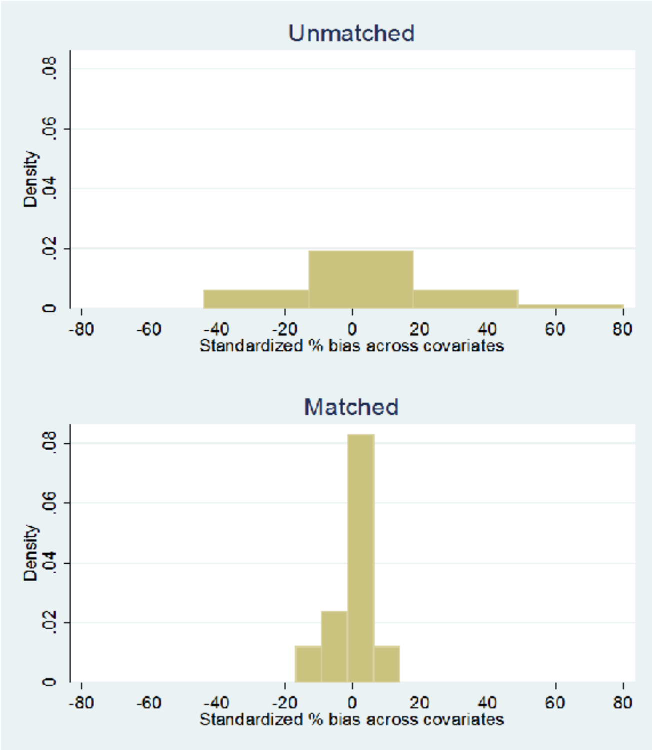


Figure 2.7: Histogram of standardized differences before and after matching

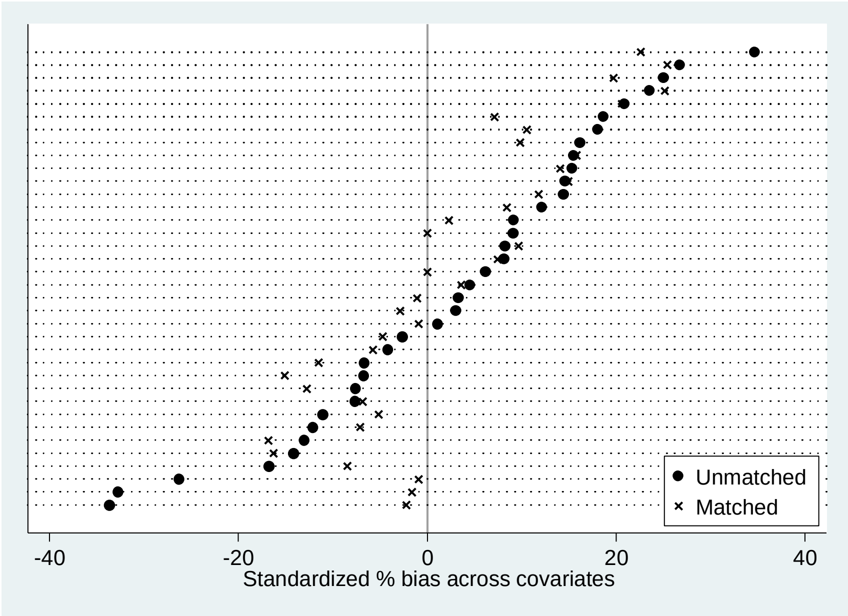


Figure 2.8: Graph of standardized differences before and after matching

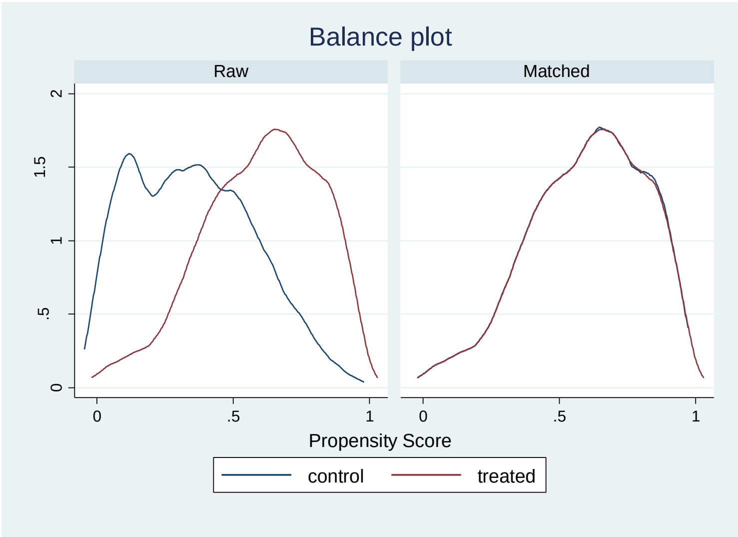


Figure 2.9: Distribution of propensity scores in unmatched and matched samples for benevolent households.

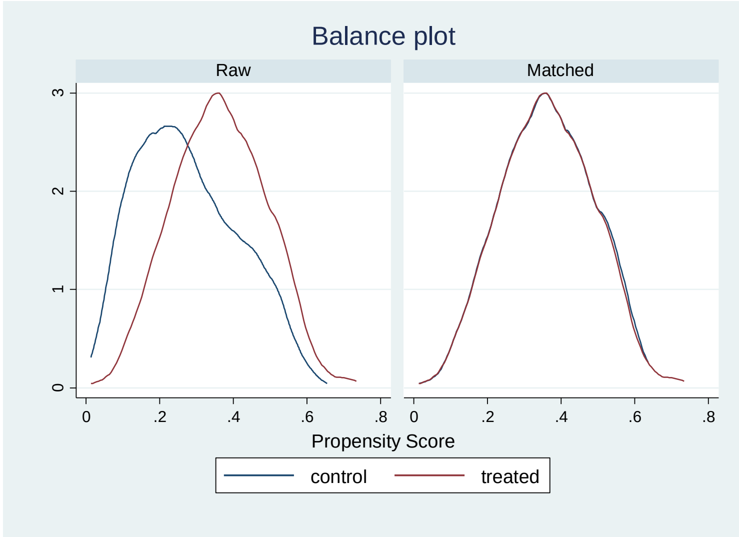


Figure 2.10: Distribution of propensity scores in unmatched and matched samples for non-benevolent households.

APPENDIX II B

Propensity Score Matching and ATT Estimation Details

Let T denote a binary treatment variable (T=1 if the household is food sufficient and T=0 otherwise). Let Y_1 denote the outcome (conflict status) of a household that is food sufficient and Y_0 the outcome for the same household had it not been food sufficient; let X be a vector of observable covariates (background characteristics). If T could be randomly assigned to households, estimating the average treatment effects (ATE) would give us the causal impact of being food sufficient on conflict. However, such an experiment that entails providing food sufficiency to randomly assigned households is neither possible nor ethical. Since we cannot randomize an intervention to avoid selection bias, we are left with quasi-experimental techniques (see Cook, Shadish, & Wong, 2008) to improve (if not isolate) the estimates of the causal effect of food sufficiency on conflict. Two prominent approaches, instrumental variables and regression discontinuity, would be useful methods but are difficult to employ. Valid instruments are difficult to identify (Imbens & Woolridge, 2009). Some possibilities exist, e.g. natural disasters, but require assumptions such as exogeneity of the instrument, that are particularly difficult to justify in this context. Regression discontinuity is another option but requires consistent decision-making around some arbitrary cutoff. In our case, food insufficiency is unlikely to be allocated in such a way. Therefore, we employ a third quasi-experimental approach - propensity score matching - in which all observable confounding factors are statistically balanced to neutralize any potential selection bias, thus allowing us to isolate the causal effects of food sufficiency on conflict.

Intuitively speaking, an unbiased average effect of treatment on the treated (ATT) could be calculated as the difference in mean outcome for the treated given that they received treatment and the mean outcome for the treated had they not received treatment. However, this outcome of the treated had they not received treatment is the counterfactual which cannot be observed in reality. Matching aims to solve this problem by constructing the correct sample counterpart for the missing information on the outcomes of the treated group had they not been treated. In other words, it addresses the 'counterfactual' by pairing each participant in the treated group with similar participants in the control group and then estimating the ATT as the difference in mean outcomes between the two groups. This can be expressed as follows:

$$ATT = E(Y_1 - Y_0 | X, T = 1)$$

$$ATT = E[E(Y_1 | X, T = 1) - E(Y_0 | X, T = 1)]$$

$$ATT = \left[\left(E(Y_1 | T = 1) - E(Y_0 | T = 0)\right) - \left(E(Y_0 | T = 1) - E(Y_0 | T = 0)\right)\right] \qquad (1)$$

Equation 1 shows how the ATT can provide correct estimates by adjusting for selection bias represented by the second term on the right-hand side.

One way to implement matching could be to match treated and control households on every covariate. However, as more variables are added to the analysis, it becomes harder to find exact matches for observations. The propensity score matching technique, proposed by Rosenbaum and Rubin (1983), solves this 'curse of dimensions' by combing all confounders into a single score, the propensity score, and matching observations based on the propensity score alone. In this study, the propensity score is the conditional probability that a household will be food sufficient, given its vector of observed covariates. PSM technique simulates the conditions of a randomized experiment by relying on two assumptions. The first is the assumption of conditional independence (or unconfoundedness) which requires potential outcomes to be independent of treatment, conditional on background variables. Under the conditional independence assumption, the propensity score is defined as the conditional probability of receiving treatment, given pre-treatment characteristics:

$$p(X) = Pr\ (T = 1|X) \qquad (2)$$

For our purposes, the conditional assumption implies that by adjusting for all observable covariates between food sufficient and food insufficient households, we can regard the treatment assignment, food sufficiency, as random and uncorrelated with the conflict outcome. The second assumption of PSM is the common support assumption which states that for each value of X, there is a positive probability of being both treated and untreated, i.e.

$$0 < Pr\ (\mathrm{T} = 1|\mathrm{X}) < 1 \qquad (3)$$

In other words, it assumes that the support of the conditional distribution of the covariates for food sufficient households sufficiently overlaps with the conditional distribution of the covariates for food insufficient households. If these two assumptions hold, then the PSM estimator for ATT is the mean difference in conflict status between food sufficient households matched with food insufficient households based on their propensity scores. This can be expressed as:

$$ATT = E(Y_1|\ T = 1,\ p(X)) - E(Y_0|T = 1,\ p(X)) \qquad (4)$$

Next, we test for the evidence of heterogeneity in the treatment effect by observable characteristics (Crump et al., 2008; Imbens & Woolridge, 2009). Specifically, by employing heterogeneous treatment effect estimation, we test whether food sufficient households that are benevolent towards others experience a further reduction in conflict. This is achieved by dividing the full sample into two subsamples based on whether the household is benevolent and estimating two separate ATTs for each subsample. The difference of the subsample ATTs provides the heterogeneous treatment effects (see Kibriya, Zhang & Xu, 2017; Xie, Brand, & Jann, 2012; Verhofstadt & Maertens, 2014) and is expressed as follows:

$$ATT_{diff} = E[(Y_1 - Y_0)| T = 1, B)] - E[(Y_1 - Y_0)| T = 0, B)] \qquad (5)$$

where B=1 if the household shows benevolence towards others and 0 otherwise.

Propensity scores can be calculated using a logit or probit estimation; we use a logit estimation. Once the propensity scores are generated, households must be matched based on their scores. Since PSM methods are sensitive to the exact specification and matching method (Imbens, 2004; Caliendo and Kopeinig, 2008), we employ three commonly used algorithms to ensure the robustness of PSM estimates. These include nearest neighbor matching (NNM), Kernel based matching and radius matching. NNM matches a food sufficient household to non-food sufficient households that are closest to its propensity score. For nearest neighbor matching, we use three nearest neighbors with replacement since replacement increases the quality of matching, especially when there are fewer close matches. Kernel matching uses a weighted average of all non-food sufficient households to match it with food sufficient households, placing higher weights on households with similar propensity scores. Following Heckman, Ichimura and Todd (1997), we use the Epanechnikov Kernel function with a bandwidth of 0.06. Radius matching algorithm matches each food sufficient household with all non-food sufficient households whose propensity scores fall within the predefined neighborhood of the propensity score of the food sufficient households (known as the caliper). We choose a caliper of 0.001 which is commonly used in the literature.

The choice of variables included in the estimation is guided both by economic theory and previous research as well as the literature on matching (see Dehejia & Wahba, 2002; Heckman, Ichimura & Todd, 1997, 1998; Abadie & Imbens, 2006; and Caliendo & Kopeinig, 2008). In summary, variable selection for matching methods is an iterative process involving a tradeoff between efficiency and bias. Therefore, it is recommended to start with a rich set of explanatory variables that simultaneously affect treatment and outcome and through a process of iteration selecting the set of covariates that gives the best balance in terms of distribution of propensity scores as well as distribution in covariates across the treated and control groups.

One limitation of propensity score matching is that it can only correct for selection on observables but not for potential unobservable confounders. While unobservable variables cannot be controlled for, we use the Rosenbaum test for hidden bias to check how much our estimates may have been affected by unobservable confounders.

Connecting Text I

Chapter 2 began an exploration into the root causes of local conflict through analyses of the impact of household food sufficiency and food sharing behavior on interhousehold conflict using primary survey data from farming households in the North Kivu province of eastern DRC. Using the same survey data, Chapter 3 continues the analysis of household behavior and decision-making by exploring the impact of land title on the prevalence and cost of interhousehold conflict. The findings from the next chapter, in conjunction with the previous chapter, can shed light on potential mechanisms through which household decisions and behavior affect low-intensity, interhousehold conflict in DRC.

Chapter 3

CAN LAND TITLE REDUCE LOW-INTENSITY INTERHOUSEHOLD CONFLICT INCIDENCES AND ASSOCIATED DAMAGES IN EASTERN DRC?

3.1 Introduction

Conflict between households and communities, government, and rebel forces have been widely documented to result in adverse consequences in fragile societies. Such adverse consequences include household members being injured, killed, made refugees, houses being burnt down, theft, ambush, forced displacement, destruction of infrastructure, assets and livelihoods; and leading to fear, distrust, and eventual break down of social cohesion, institutions and norms (Autesserre, 2012; Bellows & Miguel, 2009; Justino, 2009, 2011; Voors et. al, 2012). The existing literature on conflict is mostly limited to country-level analyses (*see* Blattman & Miguel, 2010 for a comprehensive review) where the state or state institutions are the main actors (Justino, 2011) and consequences include large scale violence and massacre. Though scholars agree that civil war violence is often shaped by local conflict, rooted in the behavior of individuals or households and their interactions within social norms (Autesserre, 2010, 2014; Blair, Blattman, & Hartman, 2017; Kalyvas, 2006), the difficulty of collecting data from conflict and post-conflict societies poses a great challenge to micro-empirical research. As a result, little is known of social conflicts at the individual or household level. Using a household survey gives me a unique opportunity to investigate the role of land title on the incidence and adverse consequences of social conflict in a fragile society, with a focus on micro-level conflict experiences that are usually neglected from mainstream reports and analyses.

The survey I use documents both violent and non-violent conflict experienced by 1,582 farming households from the postwar society of North Kivu, in eastern Democratic Republic of Congo (DRC). All forms of disagreements, disputes, protests and violence between households and other members of society (e.g. neighbors, landowners, government and rebel forces) that may lead to adverse consequences, such as loss of livelihood and assets, disruption of services, physical injury, death, etc., are recorded in the survey. I refer to these as *low-intensity local (or interhousehold) conflict* to distinguish it from national or civil conflict (defined as twenty-five or more battle deaths per annum) that have traditionally been referred to as local conflict in most literature (Autesserre, 2012). I explore, empirically, whether possessing land title can i) lower the incidence of low-intensity local conflict between households; and ii) protect households that experience conflict from the adverse consequences or damages associated with conflict.

Despite challenges to data collection, in recent years a few micro-empirical studies have emerged that study the impact of land tenure on social conflict. These studies range from examining the ambiguous effect of land reform programs on reducing insurgencies in Colombia (Albertus & Kaplan, 2013); landholding inequality causing rural unrest (Albertus, Brambor, & Ceneviva, 2018); land conflict reducing rental tendencies (Alston & Mueller, 2010); and weak tenure security, as measured by perceived risk of conflict and expropriation, leading to reduced agricultural productivity (Linkow, 2016). In an examination similar to my first question, Falco, Laurent-Lucchetti & Veronesi (2015) use household data to show that tenure security reduces a farming household's probability of experiencing land disputes by 40% in Ethiopia. Related to my second question is a study by Blattman, Hartman, & Blair (2014), who evaluate the short-term impact of mass education campaigns on resolving disputes and reducing violence in Liberia. The authors explore both violent and non-violent disputes (mostly land related) between friends, neighbors, family members, and strangers; as well as the consequences of disputes, such as property destruction, physical violence, or threats of violence. Two other studies to explore household conflict, Bellows & Miguel (2009), and Voors et. al. (2012), evaluate the consequences of conflict on household behavior. In contrast, I study whether land title affects the consequences of household conflict. I hope to contribute to the relatively new conflict economics literature in several ways. First, the difficulty of data collection from post-conflict societies has resulted in a limited number of endeavors to document and explain conflict from the perspective of households. The survey data used in this study sheds light on different types of local conflict, both violent and non-violent, experienced by households, with detailed accounts of the consequences that resulted from conflict. Most of these conflict events do not appear on police or media reports, thus making the data analysis invaluable. Second, and more important, to the best of my knowledge this is one of the first instances in development literature to investigate how conflict *outcomes* are affected at a micro level. Most micro-empirical studies have focused on the incidence of conflict (e.g., Albertus & Kaplan, 2013; Alston & Mueller, 2010; Blattman, Hartman, & Blair, 2014). Very few studies extend the analysis to what happens after conflict, or the outcomes resulting from conflict. Conflict is a natural occurrence in society which cannot be completely eradicated. However, it does not have to lead to unintended social consequences. In order to shape effective policies to sustainably manage conflict, it is important to understand the factors that can potentially reduce the adverse outcomes of conflict. Since policy makers propose better institutions to resolve conflict in a just manner, the outcomes from conflict is a pertinent yet mostly unexplored question. Results from this study can inform development policy makers to the extent land reform programs may reduce the incidence and adverse consequences associated with such low-intensity household conflict.

To address concerns of selection bias, I opt for a quasi-experimental estimation technique. I use propensity score matching (PSM) to compare conflict outcomes for households with similar observable background characteristics, rendering unbiased estimates (Imbens, 2004, and Imbens & Woolridge, 2009). A rigorous set of tests and sensitivity analyses ensure quality of matching and reliability of estimates. I find that land title reduces a household's probability of experiencing low-intensity conflict by about 10 to 18 percentage points. Previous empirical studies have established that property rights can reduce national or sub-national level conflict (Blattman & Miguel, 2010). I add to this body of work by providing empirical evidence that property rights can also reduce low-intensity, localized, and even non-violent interhousehold conflict. However, I find no evidence that households with land title suffer fewer damages in the event of a conflict.

The remainder of the paper is organized as follows: section 2 describes the study context and theoretical framework; section 3 explains the sampling strategy, data and variables; section 4 develops an empirical model and identification strategy. Section 5 presents the results, followed by a discussion of the main findings; and section 6 concludes the paper.

3.2 Study Context and Theoretical Framework

3.2.1 Land tenure system of DRC

Although vast and rich in natural resources, the Democratic Republic of Congo (DRC) has been subject to civil war and conflict since its inception in 1960. In less than sixty years of independence from Belgium, the country has seen two civil wars. The aftermath of wars, including death, destruction, disease, and displacement, combined with political turmoil has left the country in a fragile state. Located in the heart of Africa, and being the third largest country in the continent, DRC shares its geographical boundary with nine countries most of which have experienced to civil war. This led to a large influx of migrants flowing in to DRC to seek employment, exploit natural resources or to flee crises in their own countries. These foreign nationals were granted citizenship at some stage which was later revoked for political gains (Van Acker, 2005). The presence of multiple ethnicities, politically driven agendas and an ambiguous land tenure system all had repercussions for land access in DRC.

The current land tenure system in DRC is governed by the coexistence of a dual set of laws – a customary/traditional law which is prevalent in rural and indigenous areas and a modern law that is more common in urban areas (*see* Musafiri, 2008, Van Acker, 2005, Vlassenroot & Huggins, 2004, 2005 for a comprehensive understanding of the evolution of land tenure system in DRC). Traditional law, which had prevailed since the times of the earliest settlers in the region, dictate that land is held in common by groups defined by ethnicity, clans, or lineage. Customary chiefs (*mwami*) hold the power to grant non-alienable use rights to any peasant in exchange for an initial payment of *tribute* which is usually renewed on an annual basis (Van Acker, 2005). On the other hand, the modern law known as the General Property Law, established in 1973, makes the state the sole owner of all land but allows individuals to buy private rights without becoming landowners. The General Property Law is summarized in Musafiri (2008), as follows,

"Under formal law, the state owns all the land in the DRC; people and entities desiring use-rights to land can apply for concessions in perpetuity (*concessions perpétuelles*) or standard concessions (*concessions ordinaires*), or own immovable property (*immeuble*). Concessions in perpetuity are available only to Congolese nationals and are transferable and inheritable by Congolese nationals only. Standard concessions can be granted to any natural person or legal entity, whether of Congolese or foreign nationality, for specific time periods, usually up to 25 years with the possibility of renewal. Renewal is usually guaranteed so long as the land is developed and used in accordance with the terms of the concession."

The state also holds the right to terminate concessions in perpetuity through expropriation, usually for public use. Concessions are granted in the form of written documents issued by government officials. Under this law all land must be acquired from the state through an administrative process including official surveys, registration and cadastration (Vlassenroot & Huggins, 2004).

Although the new legislation had intended to abolish customary land rights and provide equal access to all citizens based on individual property rights, through improper implementation, inefficient management, and exploitation of power by elites the traditional system prevailed. To convolute the system further, traditional authorities such as customary chiefs are often appointed as administrative staff to implement state laws, especially in areas with inadequate government officials. This leads to ambiguities in the system and land documents are often issued based on customary laws rather than the state legislation. In many cases, authorities granting the concessions can be bribed to ignore standard titling procedure such as surveying or consulting with locals and to issue some kind of a non-standard document which is not even recorded as per the law (Vlassenroot & Huggins 2005). Furthermore, the formal law only recognizes perpetual or standard concessions and thus rights obtained through transfers of private land remained legally ambiguous (Vlassenroot & Huggins 2005). In addition, groups migrating to a new area are often unable to integrate into new communities or face rejection from existing ethnic groups and thus fail to secure rights to land. (Musafiri, 2009; Van Acker, 2005). Finally, customary chiefs are often politically motivated, which can make them depart from traditional laws and allocate land to the powerful (Vlassenroot & Huggins 2005). Thus, the provision for dual laws created social fragmentation, land alienation and marginalization of the rural population and eventually became the source of conflict between the state and indigenous communities over land ownership and assignment (Musafiri, 2008; Vlassenroot & Huggins, 2005).

3.2.2 Land and conflict in North Kivu

The eastern province of North Kivu has been at the heart of conflict in DRC and still remains a volatile province after a decade since the end of the second Congolese war. One of the predominant causes for this continued violence and local conflict is the competition over land combined with challenges over land access (Vlassenroot & Huggins, 2005). North Kivu has hosted its own share of migrants from the bordering countries of Uganda and Rwanda, both of which have been victims of civil war. The presence of multiple ethnic groups living in North Kivu convoluted the implementation of customary laws. Customary allocations from one community were often rejected by another and many households were excluded from customary structures and left without secure access to land (Van Acker, 2005). Furthermore, powerful rural elites including Rwandan immigrants with ethnic ties in Kivu, gained the support of chiefs and government officials in evicting local communities and purchasing their land. The corrupt political system made this inequality worse by rewarding political loyalty with land titles. The result was that in Kivu most of the land is owned by a handful of elites while most farmers are forced to succumb to insecure land titles and land alienation (Vlassenroot & Huggins 2005). Concentrating land distribution in the hands of the powerful elites meant a collapse of the social structure on which customary laws were based. This social fragmentation, growing inequality and competition over land among households formed the basis of local conflicts in North Kivu. Acts of violence, repression, crop destruction, arrests and formal complaints at the local land registry office all became part of the conflict.

Over the years, numerous armed groups have originated from the province of North Kivu, with over two dozen emerging only in the past two decades. Figure 1 in the Appendix shows a map of DRC and Kivu displaying boundaries shared with other countries as well as the distribution of rebel forces and militias. At present, the armed groups pose the greatest threats to political stability in the DRC (*see* Stearns, 2012; Vlassenroot & Huggins, 2004, 2005; or Vlassenroot & Raeymaekers, 2004, 2008, for detailed accounts of the conflict in North Kivu). The many armed groups that are still active in the province, such as the Democratic Forces for the Liberation of Rwanda (FDLR), the Allied Democratic Forces (ADF) and various Mayi-Mayi militias, continue to rebel against the state authorities and sporadically attack vulnerable civilians. The government's armed forces (FARDC) is also reported to clash with civilians. Thousands of civilians continue to be affected and displaced as a result of these multiple sources of conflict. News and humanitarian agency reports are full of accounts of civilians reporting clashes, looting, extortion, houses being burnt down, having to flee, etc.

The previous section discusses how the existing land tenure system in DRC is governed by dual laws that can lead to ambiguous land rights and access. While land documents may range from state-issued concessions to documents issued by community chiefs, my inquiry for this research is limited to the possible effect of the possession of any land document on a household's experience of low-intensity conflict. Henceforth, I refer to all forms of written documentation of land as *land title* irrespective of the identity of the issuing authority.

3.2.3 Theoretical framework

It is important to establish a general theoretical framework and justification before the empirical exploration of land rights, and possible repercussions on conflict and corresponding damages. As such, I begin with contemporary theories of political science and economics to explain participation in conflict. While political theorists view conflict in terms of motive, economists model conflict with an emphasis on opportunity costs and payoffs (Collier & Hoeffler, 2004). My framework uses both philosophies to explain the propensity to engage in conflict, within the context of land title. I also distinguish between motives and opportunities for perpetrators and victims of conflict. Although these socio-political theories focus on participation in conflict and do not delve into consequences of conflict, in many cases the arguments can be extended to the consequences of conflict as well.

Political theorists of social revolution argue that *motives* to participate in conflict arise from discontent and grievances over one's social and economic status. Humphreys & Weinstein (2008) provide a brief overview of these theories, classifying three pathways through which grievance can lead to conflict. First, conflict may be driven by social class (Paige, 1975; Wickham-Crowley, 1992), whereby citizens who belong to lower strata of society may have a higher propensity to revolt. For example, Wickham-Crowley (1992) shows the Latin American peasants who did not have access to land (such as squatters, sharecroppers or migrant laborers) were more prone to revolting. Second, ethnic and political grievances that separate one group from another may motivate an individual's decision to participate in rebellion (Horowitz, 1985). Third, individual frustrations arising from the inability to express oneself in non-violent ways may lead to rebellion (e.g. Kaplan, 1994). In their own study, Humphreys & Weinstein (2008) find empirical evidence to suggest all three sources of grievance can drive the decision to participate in rebellion. In the context of this paper, land title can provide a basis for grievance along all three categories. First, land title may divide a society into different classes, creating a have and have-not scenario, and thus provide a motive for conflict. Second, if the process of acquiring land title is driven by ethnic or political grounds, it may imply that citizens who are unable to secure title belong to the minority group and/or suffer from political representation. Such scenarios can cause alienated households without title to rebel against households from the privileged groups, with or without land title. Finally, individual frustrations, as well as lack of governance and civil rights, may cause households to use brutal force to obtain land or access to a territory, such as pastoral lands, when non-violent methods fail. Thus, it is reasonable to assume that a household lacking land title will have grievances and is therefore more likely to engage in conflict. On the contrary, a household holding land title will lack the grievance motives and is therefore less likely to perpetrate conflict, though it may fall victim to perpetration by others. General arguments may be formulated that in a perfect scenario where citizens have better income distribution, more respect for the rule of law and stronger entitlement, conflict/disagreement will be addressed quickly and with limited damages. However, in fragile states with weak governance and institutions, grievances can drive households not only to engage in conflict, but to impart and suffer greater damages from the emanating conflict. Even if a household has land title, consequences of a disagreement or aggression may not result in just or amicable solution due to weak socio-political conditions.

To explain participation in conflict from the *economic opportunity* point of view, I expand on two main theories (see Garfinkel & Skaperdas, 2007, for a comprehensive review). First, some of the oldest economic theories on conflict model participation, in terms of the choice between production and appropriation (i.e. grabbing the production of others). These models predict that participation in conflict falls with increasing opportunity cost of appropriation (Garfinkel, 1994; Haavelmo, 1954; Hirshleifer, 1991, 1995; Skaperdas, 1992). For example, individuals may be tempted to join a rebellion or engage in conflict when the forgone income is relatively low. Similarly, any activity that enhances income generating opportunities will make engaging in conflict less lucrative for individuals and households. Property rights can affect participation in conflict by changing the opportunity cost of appropriation. For example, insecure property rights can reduce economic outcome or efficiency by increasing the risk of expropriation; raising the cost of protecting property, reducing gains from trade and failing to act as collateral for credit (Besley & Ghatak, 2009). Empirical studies such as Besley (1995), Deininger et. al (2011), Galiani & Schargrodsky (2010), and Goldstein & Udry (2008) demonstrate that property rights improve economic outcomes, such as investment and income, in developing countries. By increasing expected income, land title increases the opportunity cost of conflict, thus reducing the probability that farming households will choose conflict over production. Increased opportunity cost also implies lesser time and effort exerted on appropriation, disputes or any action that may prolong the conflict. Hence, due to higher opportunity costs (i.e. the possibility of forgoing a peaceful yet lucrative alternative), households that have land title will be much less likely to be perpetrators of conflict. However, in the event that such households get engaged in conflict, the consequent damages suffered may be higher because they have higher wealth levels (more to lose) or lower if the households choose to reach quick settlement due to the high opportunity cost of time. These effects may negate each other so that households with and without title end up with the same level of damages.

Second, following Hirschleifer (1991), the most common theory to explain conflict is a contest model which predicts that the odds of winning increases with the relative effectiveness of the contestant's fighting technology. The contest model assumes that parties expend resources on arming themselves to increase their probability of winning in case of conflict. Since prolonged conflicts can typically result in violence or destruction of assets and output, a negotiated settlement is often the more desirable outcome, for both parties (Blattman, Hartman & Blair, 2014; Fearon, 1998; Garfinkel & Skaperdas, 2006). In the context of this study, it can be assumed that land title serves as the bargaining tool that households use as a device for conveying private information to opponents during the bargaining process. Following the predictions of Hirschleifer's base model, land title may increase a household's odds of winning low-level disputes with other households, while lowering the incentive for other households to enter the contest. However, bargained solutions are more likely to break down because of asymmetric information and commitment problems, arising from the inability of parties to trust each other or stick to an agreement in the absence of enforceable contracts. In such cases, becoming a victim for these households may also imply lack of information. For example, North Kivu's unique "dual rights" system makes such information asymmetry or commitment problems most imminent.

To summarize, the effect of land title on conflict appear to remain consistent through all socio-political and economic lenses. Thereby I formulate my first hypothesis as: *land title holders in North Kivu will have a lower probability of low-intensity, interhousehold conflict*. However, the extent of damages expected by land title holders in the event of a conflict appears to differ across theoretical extensions. With the assumptions of greater income distribution, respect and implementation of rule of law, institutional governance, symmetry of information and enforceable contracts, damages emanating should be limited in the presence of land title. But before arriving at any such hypothesis, it is important to acknowledge that North Kivu is plagued by pervasive inequality, extreme poverty, as well as lack of governance, civil rights, and rule of law. Not only do none of the underlying socio-economic preconditions hold for North Kivu, but it also presents unique information asymmetry through dual rights systems. Hence, I propose my second hypothesis as: *given fragile conditions and under current land allocation systems, land title will not reduce the damages inflicted by the violence in North Kivu*.

3.3 Data Description

3.3.1 Survey design and data collection

The data for this study was collected through household surveys of smallholder farmers in the North Kivu province of eastern DRC. The survey was a component of the Best Practices in Coffee and Cacao Production (BPCC) project, funded by the Howard G. Buffet Foundation and administered through Texas A&M University in the month of July in the year 2014.

The survey was conducted in three territories of North Kivu province - Beni, Lubero, and Rutshuru. A grid-based randomization strategy was used to select villages from each of these territories. Specifically, high resolution area maps from the United Nations Office for the Coordination of Humanitarian Affairs (UNOCHA) were used to divide each territory into 5 km by 5 km squares or "grids" that were then sequentially numbered. Grids were selected using computer generated random numbers. If a chosen grid contained three or fewer villages, all villages were selected for the study. If a grid contained between four and nine villages, three were selected at random. For grids that contained ten or more villages, four random selections were made. Since precise population densities are unknown, this randomization technique was used to generate a representative study sample. Local extension agents were trained to serve as enumerators and instructed to visit every household in the selected villages. If the decision maker was absent at the time of the visit, enumerators were instructed to move on to the next house. In the end, full sample data was collected on 1,582 households.

Enumerators gathered detailed information on any kind of conflict experienced by households, the parties involved in conflict, and the consequences resulting from the conflict. In addition, they gathered information on household land access type and possession of land documents, perceptions of tenure security, farm level characteristics, household demographics and food security status, access to markets and knowledge, and empowerment and voice within the community. Structured questionnaires translated to French (the commonly spoken language in North Kivu) were used to collect this data. Further details on data collection may be found in Fatema & Kibriya (2018).

3.3.2 Variables and descriptive statistics

My unit of analysis is the household. The explanatory variable of interest is *land title*. To measure this variable, households were asked whether they possess written documentation of their land claims. Of the 1,582 households, 912 did not possess any written documentation of their land claim while 670 households did so. Households were also asked to rate, on a scale of one to four, how concerned they were that their land may be expropriated from them. This information was used as a robustness check.

The main dependent variables are the incidence and outcomes of low-intensity, local conflict experienced by households. To measure the incidence of conflict, households were asked to specify whether they experienced any of the following low-intensity conflicts in the past six months: a) conflict with neighbors and fellow villagers, b) disagreement involving Virunga National Park[10], c) conflict with landholder over occupied land, border, or granting land to new tenant d) inheritance dispute between family members, e) disagreement with pastoralists, f) land and resource conflict with rebel forces, g) land and resources conflict with government forces, h) other kinds of conflict with government forces, i) conflict over community resources and agricultural inputs, and j) any other kind of conflict that they were asked to specify.

[10] A 3000 square mile national park declared as a world heritage site by UNESCO. The park has been plagued by poaching and land invasions.

Table 3.1 summarizes the conflict experience of households in the study. Panel A shows the relationship between household land title status and conflict experience. Approximately half the households in the study were involved in some form of conflict in the previous six months. Of these, 324 households had written documentation of their land and 462 did not. The remaining half encountered approximately two different incidents of conflict on average in the past six months.

Panel B displays the categories of conflict encountered. A household could have encountered multiple types of conflict. The most common conflicts reported by households occurred with neighbors and fellow villagers, with landholders, or with pastoralists, with around 300 households reporting each type. These originated from disputes over land boundaries, landholders claiming back land occupied by tenants or granting occupied land to new tenants, and disputes over pastoral land. Other frequently occurring sources of tension included conflict with rebel forces over land and resources, and inheritance disputes among family members residing in separate households (e.g. siblings). A minor fraction reported conflict over community resources and disagreements over Virunga Park. An even smaller fraction was engaged in conflict over land and resources with government forces. Finally, less than two percent households encountered other forms of conflict, such as confrontations with petty thieves or robbers, and allegations of sorcery by fellow villagers.

Table 3.1: Types of conflict reported households

	Households with title	Households without title	All Households
Panel A: Conflict experience of household			
Number of HH that did not experience any of conflict	346	450	796
Number of HH that experienced some type of conflict	324	462	786
Total number of HH	670	912	1582
Panel B: Type of conflict			
Land related conflict with neighbors and fellow villagers	152	204	356
Conflict with landholder	74	237	311
Inheritance dispute among non-dwelling family members	48	109	157

Disagreement with pastoralists	142	168	310
Land and resource conflict with rebel forces	70	108	178
Land and resource conflict with government forces	6	13	19
Conflict over community resources including Virunga Park	38	47	77
Others	12	14	26

Source: Authors' calculations based on survey data

Note: A household may report multiple incidents of conflict.

Low-intensity interhousehold conflict is measured using two variables: *conflict incidence*, an indicator variable which equals one if the household has experienced any land related conflict with other households in the six months prior to the survey, and zero otherwise; and c*onflict intensity*, the number of categories of land related conflict experienced by the household in the past six months. For example, if a household reported conflict with neighbors, pastoralists, and rebel forces, the conflict intensity variable for the household takes a value of three.

Households that experienced at least one conflict incidence were further asked to specify which, if any, of the following outcomes resulted from the conflict: a) loss of a portion of its land; b) loss of land rights over part of land; c) gain of a section of land; d) gain of land rights over part of land; e) requirement to pay compensation, monetary or otherwise; f) receipt of compensation, monetary or otherwise; g) loss of community acceptance and punishment by the village council; h) resolution of conflict without exchange of land or money; i) destruction of agricultural equipment; j) theft or loot of crops; k) theft or destruction of personal property or equipment; l) displacement from home or property; m) loss of off-farm income source; n) gain of off-farm income source; o) increase in market access; p) physical violence to anyone in household; q) death of anyone in household; r) loss of access to certain roads; s) loss of access to usual market; t) destruction of livestock; u) damage or destruction of crops; v) continuing conflict; w) destruction of house; x) extortion of any household member; and y) any other outcome (to be specified).

This information was used to define a variable, *damages*, as the number of damages or losses incurred by the household as a result of the conflict. Responses were also regrouped into the following five categories: i) loss/damage to physical capital; ii) loss/damage to human capital; iii) loss of financial capital; iv) loss of rights or dignity; and v) loss of access. Table 3.6 in the Appendix summarizes the variable definitions based on the survey responses. These variables were used as additional dependent variables to test whether land title was related to any specific type of losses or damages. Loss/damage to physical capital included damage, destruction, and theft of crops, livestock, productive tools, or personal property as well as loss of land. Loss/damage to human capital included physical injury or death of family members resulting from conflict. Financial losses included loss of off-farm income or monetary losses paid in compensation. Loss of rights included loss of community acceptance and punishment by village council for engaging in conflict. Loss of access included damage or destruction of access roads or markets as a result of conflict[11].

Table 3.2 summarizes the outcomes of conflict by land title status of the household. Panel A shows the number of households with gains, losses, or no outcomes from conflict. Approximately half the surveyed households reported no consequences from the conflict, either because they were not involved in conflict to begin with, or because the conflict did not affect them in any way. About 43.9 percent households that suffered damages had land title and roughly 45.6 percent did not. At

[11] An alternative way to capture the extent of damages is to create an *index*, as in Bellows & Miguel (2009), to calculate the average of the responses to conflict damages. This would yield the same results and only the interpretation would be slightly different. For example, on a scale of zero to one, the average household has a conflict victimization index of 0.2.

first glance, the difference between these numbers appear trivial enough and imply that land title does not seem to affect the extent of damages suffered in the event of a conflict.

Panel B summarizes the number of incidents in which the household suffered damages or losses. A small minority reported gains from the conflict, such as financial gains or access to markets or services. About half of the households suffered no loss, consistent with the fact that about half the households were not involved in any conflict. Of the households that suffered losses, the most common number reported was one incident of damage followed by three or more incidents of damages. The total number of incidents reported ranged between zero and fifteen. Less than eight percent of the sample reported suffering more than three different damages/losses from the conflict. Hence households with three or more losses were grouped together for the analysis.

Finally, Panel C further breaks down the damages/losses by category. The figures represent the total number of households that reported a specific type. A household may suffer multiple losses from conflict. The most frequent type reported is the loss of physical capital in the form of loss of land, damage or destruction of assets, etc. with 796 households reporting some loss of physical capital. Of these, 337 had titles over their land and 459 did not. The next loss reported is financial capital with 142 households having lost off-farm income or having been forced to pay compensation. Loss of human capital, access to roads and markets, and loss of rights or dignity are reported by fewer than 100 households each.

Table 3.2: Outcomes of conflict by land title status

	HH with title	HH without title	All HH
Panel A: Outcome of conflict			
Households that had gains from conflict	30	42	72
Households that suffered losses/damages from conflict	294	416	710
Households with no outcome from conflict	346	454	800
Total number of Households	670	912	1582
Panel B: Loss/damage by count			
Households suffered no loss/damage	376	494	872
Households suffered one type of loss/damage	141	179	320
Households suffered two types of loss/damage	72	103	175
Households suffered three or more types of loss/damage	90	140	230
Panel C: Loss/damage by category			
Loss/damage to physical capital	337	459	796
Loss/damage to human capital	34	64	98
Loss of financial capital	50	92	142
Loss of rights or dignity	16	58	74
Loss of access to markets	47	48	95

Source: Authors' calculations based on survey data

Note: One household may report more than one damage. The total number of incidents of damage~~s~~ suffered ranged from 0 to 15. Households with three or more incidents of damage~~s~~ comprised less than 8% of the sample and were grouped together for brevity.

Following theory on the drivers of social conflict, I include all control variables that may be related with social inequality and lead to grievances. These include differences in household income, education, status of food security and land ownership, membership in cooperatives, and access to markets (including credit and insurance), services, and technology, etc. I also include measures of empowerment such as holding leadership positions within the community. Some of these variables, such as income, education and access, may also affect participation in conflict by changing the opportunity cost of participation. For example, Collier & Hoeffler (2004) use per capita income and male secondary school enrollment as proxies for forgone income. The authors argue that focusing on young males is relevant since this is the group from which rebels are recruited. Therefore, I further include number of adult males in my analysis. Finally, I include variables to control for differences in household demographics as well community fixed-effects. This absorbs any variation in conflict outcomes due to community-specific characteristics, such as the quality of local governance and institutions, and essentially compares households within the same community.

Table 3.3 shows the mean values of the main variables, along with a t-test of means between households with and without title. Panel A shows that the average household has around 50 percent probability of experiencing at least one type of conflict. On average, households suffer less than one incident of damage from conflict[12]. On average, households that possess land titles report significantly fewer incidents of land related conflict. There appear to be no other significant differences in the mean values of dependent variables for households with and without land titles.

[12] Since half the study sample do not experience any conflict, the averages of the dependent variables appear to be low. If I restrict my sample to-households that experienced conflict, the average number of incidents reported is 1.7.

Table 3.3: Descriptive Statistics

Variable	Household has land title (N=670)	Household does not have land title (N=912)	All households (N=1597)
Panel A: Dependent variables			
Conflict incidence	0.46	0.49	0.48
Conflict intensity	0.72**	0.91	0.83
Panel B: Control variables			
Household size	5.15***	5.47	5.33
Number of adult males	0.48***	0.44	0.46
Education	9.76***	8.56	9.07
Education squared	115.47***	95.38	103.90
Household income per capita (1000 CDF)	22.8***	14.9	18.3
Access to technology (yes=1)	0.82***	0.69	0.74
Access to credit (yes=1)	0.09	0.08	0.09
Lack of extension services (yes=1)	0.47***	0.66	0.58
Cooperative membership (yes=1)	0.26***	0.17	0.21
Leadership position (yes=1)	0.67***	0.57	0.61
Teacher or doctor	0.36***	0.28	0.31
Food insecure	0.14	0.13	0.13
Rents land	0.17***	0.44	0.33
Purchased land	0.69***	0.16	0.38
Inherited land	0.25***	0.36	0.31
Received farmer training	0.22***	0.17	0.19
Has insurance	0.03***	0.01	0.02

Source: Authors' calculations based on the survey data.

Notes: I used t-tests to test for equal means between households with and without land title. The null hypothesis is that the means are not significantly different from zero. *, **, and *** indicate significance at 10%, 5% and 1% levels respectively. Location and religion specific dummies have been omitted from the table to save space. CDF=Congolese Franc. 1 USD=925 CDF at the time of the study.

Panel B shows that the representative household in my survey has around five members, of which about 46% are adult males. The most educated member of a typical household has an average education of around nine years. The monthly per capita income of an average household is 18,300 Congolese Franc (CDF) or 219,600 CDF per year. This equates to an annual per capita income of about 237 USD, which leaves household members with less than 1 USD per day. Around a third of the households had access to some form of technology such as radio, television, internet or mobile phones, as well as access to motorized vehicles, affecting easy access to markets and information. Only nine percent had access to any form of credit and about two percent have any type of insurance. Almost half the sample had not received any form of extension services (such as communication about production, sales or markets or farm visits) from crop buyers or extension workers. About a fifth each belong to some farmer cooperative and have received some form of farmer training. The majority households have some member who has held an important position in the village council. Around half the households reported to being food insecure at some time, while 13 percent claimed to being food insecure often. Roughly, a third of households purchased, a third inherited, and a third rented their land.

A simple t-test of means reveals that for almost every control variable the mean values are significantly different between households with and without land title. For example, households that possess land title are smaller in size but are composed of more adult members than the average households without land title. The presence of more adult males in the household might make it easier to acquire land documents, assuming that they would be more mobile and able. Similarly, compared to a household without title, the typical household with land title has more than one extra year of schooling and about 7,900 CDF more per capita per year. In addition, these households have a 13 percent greater likelihood of having access to telecommunication and transportation technology. It appears reasonable that wealthier and more educated households, with greater access to technology, would be more informed about the land titling process. Access to technology can also make it easier to acquire the necessary land related documentation. This difference in income, education and technology could also explain why households with land title have greater access to extension services, cooperative membership as well as positions of leadership within the community. Finally, it is hardly surprising that land ownership is associated with land title. Overall, these differences in means points to a potential self-selection bias, since households that choose to acquire titles on their land appear to be systematically different from households that do not.

3.4 Empirical Estimation

3.4.1 Identification Strategy

I estimate the following equation:

$$Y_{ij} = \alpha + \beta_1 T_{ij} + \beta_2 X_{ij} + \mu_j + \varepsilon_{ij}$$

where Y is an indicator of conflict (or consequences of conflict) experienced by household i in community j, T is a dummy variable equal to one if the household has written documentation of its land claims and zero otherwise, X represents the vector of background characteristics, μ_j represents community fixed effects and ε_{ij} is an idiosyncratic error term.

Table 3.3 shows that households with and without land title appear to be systematically different. As a result, an Ordinary Least Square (OLS) estimation technique would lead to biased coefficients of the treatment variable. Though instrumental variables could solve identification challenges, it is difficult to find convincing instruments that can affect a household's probability of possessing land title but not the probability of conflict. A popular approach with cross-sectional data is a randomized experimental design, but that would entail granting land title to random households which could exacerbate social unrest in an already volatile area. Hence, I exploit a quasi-experimental setting and use propensity score matching (PSM) to estimate my coefficients. I use several tests and checks to ensure the quality of matching and robustness of results.

3.4.2 Propensity score matching

Propensity scores are used to overcome selection bias by comparing groups based on *observable* covariates. Following Rosenbaum and Rubin (1983), the propensity score for this study is defined as the household's probability of being assigned to a treatment (i.e., having land title), conditional upon background variables. Formally, let T be a binary treatment variable such that T=1 if the household has land title; and 0 otherwise, Y be the conflict outcome for the household, and X be a vector of observable covariates (background characteristics or 'pretreatment' variables). The propensity score is defined as:

$$p(x) = \Pr\,(T = 1|X = x)$$

In order that the propensity-score matching may lead to what can hold to be unbiased estimates, one must make a few assumptions. The assumption of *conditional independence* or *unconfoundedness* requires potential outcomes to be independent of treatment, conditional on background variables X. Identifying a rich set of covariates to predict assignment to treatment then allows us to view treatment as random. Therefore, adjusting for differences in the covariates should be sufficient to attaining valid causal estimates (Imbens & Woolridge, 2009). Imbens (2004) explains that "the implication of these assumptions is that systematic (for example, average or distributional) differences in outcomes between treated and control units with the same values for these covariates are attributable to the treatment."

The next assumption, the assumption of *common support*, implies that for each value of X there is a positive probability of being both treated and untreated, i.e.

$0 \quad < \Pr(T=1|\ X) < 1$

This assumption is also known as *overlap* since it assumes that the support of the conditional distribution of the covariates for the control group overlaps with the conditional distribution of the covariates for the treatment group. In addition to these two assumptions, the *balancing condition* of PSM states that assignment to treatment is independent of the X characteristics, given the same propensity score. This condition is testable.

If these assumptions hold, then the average treatment effect on the treated (ATT) can be estimated as follows:

$$ATT = E(Y_1|\ T = 1,\ p(X)) - E(Y_0|T = 1,\ p(X))$$

where Y_1 is the conflict outcome of a household that has land title, and Y_0 is the outcome for the same household if it did not have land title. The second term, which is the hypothetical outcome for the treated had they not received treatment, is the counterfactual which is unobserved. However, the matching procedure allows us to assume random assignment to treatment and hence, for the purposes of estimation proxy, households without land titles as the perfect counterfactual for households with titles had they not received title. However, it should be noted that while PSM can help improve precision of estimates of treatment effects, it cannot account for any unmeasured confounders that may exist. Therefore, I use Rosenbaum bounds in my sensitivity analysis to test for any hidden bias resulting from potential unmeasured confounders.

3.4.3 Selection of matching algorithms and variables

Following literature, I use three matching algorithms to ensure the robustness of PSM estimates: nearest neighbor matching (NNM), kernel-based matching and radius matching. For nearest neighbor matching, each household with land title is matched with five households without land title and that have the closest propensity scores to the observations they are matched with.

The matching is conducted with replacement (i.e. an untreated household may be matched to more than one treated household) to increase the potential for finding better matches. For kernel-matching, I use the Epanechnikov Kernel function with a bandwidth of 0.06 (Heckman, Ichimura & Todd, 1997). The radius matching algorithm matches each treated household with all control households in the neighborhood using a predefined radius of propensity score known as the ‘caliper’. Following common practice, I use a caliper of 0.001.

The control variables included in the PSM determine the quality of matching, and hence the bias and efficiency of estimates. At the same time, this variable set must also satisfy the assumptions and conditions of PSM. Finding the variable set to be used is an iterative process. As mentioned in section 3 (b), the control variables for this study are based on the existing theory reviewed in section 2 (a), as well as extensive literature on variable selection using PSM (see Dehejia & Wahba, 2002; Heckman, Ichimura & Todd, 1997, 1998; Abadie & Imbens, 2006; and Caliendo & Kopeinig, 2008).

3.5 Results

3.5.1 Variables associated with land title

Table 3.4 summarizes the results from a logit model of the likelihood of a household to have land title given other observable background characteristics. This helps us determine which variables are most likely to affect selection into receiving land title. The model has a McFadden R^2 of 0.39 and a χ^2 significant at the 1% level indicating a good fitness of the overall model. The model is successful in correctly predicting the land title status of 80% of households with title and about 84% of households without land title.

Table 3.4: Results from logit estimation

VARIABLES	Logit	Marginal effects
Dependent variable: =1 if household has land title =0 otherwise		
Household size	0.0594	0.008
	(0.0458)	(0.006)
Number of adult males	-0.00712	-0.001
	(0.407)	(0.053)
Education	0.0218	0.003
	(0.0565)	(0.007)
Education squared	-0.000940	0.000
	(0.00328)	(0.000)
Household income (1000 Congolese Franc)	0.004*	0.005*
	(0.0021)	(0.0028)
Access to technology (yes=1)	0.0265	0.003
	(0.203)	(0.027)
Access to credit (yes=1)	-0.0272	-0.004
	(0.330)	(0.043)
Lack of extension services (yes=1)	-0.337	-0.044
	(0.215)	(0.028)
Cooperative membership (yes=1)	0.464**	0.061**
	(0.206)	(0.027)
Leadership position (yes=1)	0.0252	0.003
	(0.171)	(0.023)
Teacher or doctor	0.0327	0.004
	(0.191)	(0.025)
Food insecure	-0.535**	-0.070**
	(0.250)	(0.033)
Rents land	-0.748***	-0.098***
	(0.249)	(0.032)
Purchased land	2.491***	0.327***
	(0.219)	(0.024)
Inherited land	-0.262	-0.034
	(0.232)	(0.030)
Received farmer training	-0.356	-0.047
	(0.253)	(0.033)
Has insurance	1.130	0.148
	(0.760)	(0.100)
Constant	18.44	
	(1,890)	
Area and religion dummies	Yes	
Summary Statistics		
McFadden R^2	0.39	

LR χ^2 (90)	770.36***
Log-likelihood ratio	-593.39
Land title holders correctly predicted	80.10
Non-title holders correctly predicted	84.61
Overall percentage correctly predicted	82.69

Note: Standard errors in parenthesis for logit estimate coefficients. Delta-method standard error in parenthesis for marginal effects estimation. *, **, and *** indicate significance at 10%, 5% and 1% levels respectively.

3.5.2 Propensity score estimation

As an initial check of the balance, I visually inspect a graph of propensity scores between treated and control groups (refer to Figure 3.2 in the Appendix). For every interval of propensity score, there exist observations from both treated and untreated groups. In other words, there is a region of common support between the two groups. Table 3.7 (in the Appendix) shows the distribution of the propensity scores for households with and without land titles. Similarly, Figure 3.3 in the Appendix graphs the distribution of propensity scores in the treated and control groups in the matched and unmatched samples. A visual inspection shows that the overlap in the distribution of propensity scores dramatically improves after matching. Together, Figures 3.3, 3.4, and 3.5 in the Appendix visually assures us that the *overlap assumption* is not violated.

3.5.3 Main results and discussion

Table 3.5 summarizes the average treatment effects on the treated using the three matching methods. Panel A shows the effect of land title on the probability of household conflict. The signs on all coefficients in panel A are negative and all, but one, are statistically significant at 90 percent or higher confidence level. The results show that holding everything else constant, land title reduces the probability of land related conflicts between approximately 10 and 18 percentage points. In addition, households that possess land title report between 0.17 and 0.41 fewer cases of land related conflict incidents compared to households without land title. Comparing across the three matching methods, I find that the coefficients are similar in magnitude for nearest-neighbor and kernel matching and largest for radius matching.

Panel B shows the effect of land title on household damages. Land title does not have a statistically significant effect on the total number of damages incurred by a household, regardless of whether the sample in question includes all households or the fraction that have incurred any form of conflict. This result holds across all matching algorithms. When disaggregating conflict by type, I find no evidence that loss/damage to physical, human, or financial capital, and access to markets are affected by land title. Only loss of rights is statistically significantly determined by land title, but this result may not be of economic significance given that less than five percent households reported loss of rights due to conflict. Hence, results for damages from conflict, disaggregated by type, have been omitted from the table.

Table 3.5: ATT estimates using alternate matching methods

Outcome	Treatment: land title		
	Nearest-neighbor matching	Kernel matching	Radius matching
Panel A: Incidence and intensity of conflict			
Conflict incidence	-0.101**	-0.098**	-0.183**
	(0.0563)	(0.0534)	(0.0524)
Conflict intensity	-0.171*	-0.167	-0.406***
	(0.1296)	(0.1318)	(0.1212)
Panel B: Consequences of conflict			
Number of damages suffered (full sample)	-0.152	-0.121	-0.097
	(0.127)	(0.120)	(0.119)
Number of damages suffered conditional on conflict	0.100	0.12	0.015
	(0.132)	(0.131)	(0.160)

Source: Authors' calculations based on the survey data. Bootstrapped standard errors with 100 repetitions reported.

Note: The symbols *, **, and *** indicate significance at 10%, 5% and 1% levels respectively. All estimates shown are average treatment effect on the treated. Bootstrapped standard errors with 100 replications of the sample are reported in parenthesis. Nearest neighbor matching uses the five nearest neighbors, with replacement. Kernel matching uses a bandwidth of 0.06 while radius matching uses a caliper of 0.005. 701 observations were included for the number of damages conditional on conflict. The results of regression analysis with outcomes including each damage category and gains from conflict are statistically insignificant and are omitted from the table.

To summarize the results, I find empirical evidence that households with land title land are less likely to experience low-intensity conflict with other households within the community. However, in the event of a conflict, land title does not affect the extent to which households suffer damages due to the conflict. The first result is consistent with the theoretical expectations and formulated hypothesis. Households of North Kivu that have land title are less likely to perpetrate conflict driven by grievance. Further, conflict may simply not be worthwhile due to the higher valued alternative of production forgone.

The second result may appear less intuitive under the general theoretical framework of secure property rights. However, as discussed in previous sections and formulated in my second hypothesis in section 3.2.3, the underlying socio-political preconditions and assumptions are much different in North Kivu. Lack of governance, along with unstable conditions has plagued North Kivu with conflict for decades. This translated to pervasive poverty, inequality, and weakening of legal institutions needed to enforce contracts. Therefore, the ineffectiveness of land title in reducing damages from conflict can be attributed to weak institutional structure, stratified social structures and asymmetry of information. Legislation and institutional structure delineating property rights are weak in DRC. Because of the presence of a dual system of property rights, land titles are delegated and administered in an ad-hoc manner in multiple layers. Information asymmetry can also prevent citizens from exercising their rights (Deininger, Ayalew, & Yamano, 2008) and thereby fail to limit damages from conflict. Contradictory land use plans and allocation policies, though not uncommon, can create weak implementation, thereby marginalizing a certain sect of the population (Broegaard et al., 2017). Privileged and well-connected citizens can also use these ad hoc complicated systems to their advantage to acquire land through spurious claims (Feder & Nishio, 1998) which may further inequality and subsequent grievances. Jointly, these conditions produce an unconducive environment to limit damages from conflict and provide just solutions, even for households with land title. For instance, as citizens are further marginalized, their propensity to induce conflict in society increases. Even if title holders do not perpetrate conflict themselves, a higher fraction of aggrieved households suggest that eventually conflict is imminent. Once a conflict breaks out, the mounting frustrations of the aggressors combined with the lack of law enforcement agents can result in a prolonged and/or violent encounter whereby damages can no longer be kept under control.

My two findings can be reconciled with the argument that due to lack of discontent as well as high opportunity cost, title holders refrain from perpetrating conflict themselves and getting provoked by others to participate in conflict. It is the households that lack land title who are likely to hold grievances, as well as have a lower opportunity cost of engaging in conflict. These households are more susceptible to getting drawn into conflict and perpetrating conflict themselves. In the event that they succeed in generating conflict, their frustrations translate to aggression and consequently result in damages. In the absence of law and order, both parties involved in conflict can exert aggression and physical violence, so that the presence of land title becomes immaterial in determining how much damage results from the conflict.

3.5.4 Matching quality tests

Several methods are employed to test the balance in the distribution of covariates between treated and control groups after matching. These include comparing normalized differences in mean for each covariate between the treated and untreated groups (see Table 3.8 in Appendix); visually inspecting the percentage standardized bias between covariates in treated and control groups after matching (see Figure 3.5 in Appendix); comparing matching quality indicators across the three matching algorithms (see Table 3.9 in the Appendix); and running a "placebo" regression whereby the treatment and control variables are regressed on a variable which is logically uncorrelated with treatment (see Table 3.10 in the Appendix). More details on each of these tests and results can be found in the Appendix. The results from all the above tests indicate good quality matching.

3.5.5 Sensitivity analyses

To test the sensitivity of my analyses, I estimate the models using a doubly robust estimator where the calculated propensity scores are included as an additional control variable in a linear regression. Table 3.11 in the Appendix shows the results. The results obtained are similar to the ones found earlier. In addition, I test the sensitivity of my estimates using the Rosenbaum bounds for hidden bias (Rosenbaum, 2002). Since PSM matches households based only on observable covariates, potential bias in estimates may arise from selection on unobservables. The Rosenbaum bound (Γ) measures how large the difference in unobservables need to be to in order to render ATT estimates insignificant. Based on the results, I conclude that the findings are robust to potential hidden bias from unobserved covariates. Finally, Table 3.12 in the APPENDIX shows the results from alternate specifications such as running a Poisson analysis and using fear of expropriation instead of land title as a proxy for tenure security. Similar to land title, fear of expropriation fails to account for damages given that the household experienced conflict.

3.6 Conclusion

I study whether land title can reduce the incidence and adverse consequences of low-intensity interhousehold conflict experiences of farming households in North Kivu, DRC. The field survey used in this study allowed for the analysis of unique, micro-level conflict events and disagreement scenarios that do not appear in local news and reports and hence is primarily absent from relevant literature. Moreover, to the best of my knowledge, this is the first study to examine the social consequences associated with household conflict. The results and discussions of this study are quite possibly generalizable for similar societies with fragile socio-political conditions.

My theoretical expectations are based on socio-political (Humphreys & Weinstein, 2008; Paige, 1975; Wickham-Crowley, 1992; Horowitz, 1985) and economic (Haavelmo, 1954; Hirschleifer, 1991, 1995; Grossman, 1992, 1994; Skaperdas, 1992; Garfinkel, 1994) models of conflict. My empirical analysis highlights two revelations. First, possessing land title reduces the probability that a household experiences low- intensity conflict with other households within the community. Second, households that possess land title do not necessarily experience fewer damages in the event of a conflict. I conclude that in fragile societies land title can protect households partially by reducing the household's probability of experiencing conflict in the first place. However, in the event of a conflict, land title cannot protect households from the adverse consequences of conflict.

My findings hold policy implications for sustainable development by demonstrating the need for enhancing and strengthening property rights and promoting a better environment to sustain property legislation. Given that much is written about fragile institutions in Africa, sustainable development policies should look at the long-term effect of programs or policies. In recent times multilateral and bilateral agencies, as well governments, have supported many programs that provide land title to households in developing countries. However, my presented evidence suggests that formalization of intricate land titling system that only provides citizens with land title is not a panacea for all conflict related adversities and thus cannot serve as a stand-alone tool to reduce adversities associated with conflict. Further research is required to study whether supplementing land reform policies with good governance and strong institutions can reduce the adversities associated with household conflict. I propose three main policy recommendations to be considered for further research. First, land titling system should be inclusive, transparent, efficient and devoid of contradictory layers of bureaucracy (Broegaard et al. 2017). Issuing and enforcing land titles should be preceded by a diagnosis of relevant policies, the political environment and analysis of the sustainability component (Deininger & Feder, 2009). In addition, the titling procedure should be cost efficient and easily achievable for all citizens (Arruñada, 2003). These good governance practices may address grievances and inequality issues. Second, the implementation of rule of law must be strengthened to ensure peaceful solutions to any land related disputes or disagreement. Well organized, local institutions must be formed and financed to implement the enacted titles (Bromley, 2009). Such institutions should be able to adjudicate and adjust for the dual laws (Chimhowu & Woodhouse, 2006). These may be augmented by land rights and registration awareness programs, as well as broader public access to land holding data to reduce information asymmetry (Deininger & Feder, 2009). Third, complimentary programs to facilitate better returns from land usage need to be implemented. These may include building infrastructure, credit programs, and access to better agricultural technology and inputs. While investments in infrastructure have been shown to solidify urban property rights (Gulyani & Talukdar, 2008), similar investments may also enhance rural land tenure. Traditional, rural and agricultural development programs involving microcredit and technology to improve returns from farming may also contribute towards more sustainable land titling strategies.

The 2030 Agenda for Sustainable Development, adopted by all United Nations Member States in 2015 states that the Sustainable Development Goals (SDGs) “aim to significantly reduce all forms of violence, and work with governments and communities to find lasting solutions to conflict and insecurity”. Given that conflict is a complex and costly phenomenon that is difficult to eradicate, we need to focus our discussion on how to effectively and sustainably protect vulnerable households with high exposure to conflict and thus steer towards achieving the sustainable development goal of promoting peace, justice, and strong institutions.

References

Abadie, A., & Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica*, *74*(1), 235-267.

Albertus, M., Brambor, T., & Ceneviva, R. (2018). Land inequality and rural unrest: theory and evidence from Brazil. *Journal of Conflict Resolution*, *62*(3), 557-596.

Albertus, M., & Kaplan, O. (2013). Land Reform as a Counterinsurgency Policy Evidence from Colombia. *Journal of Conflict Resolution, 57*(2), 198-231.

Alston, L. J., & Mueller, B. (2010). *Property rights, land conflict and tenancy in Brazil* (No. w15771). National Bureau of Economic Research.

Arruñada, B. (2003). Property enforcement as organized consent. *Journal of Law, Economics, and Organization*, *19*(2), 401-444.

Autesserre, S. (2010). *The trouble with the Congo: Local violence and the failure of international peacebuilding* (Vol. 115). Cambridge University Press.

Autesserre, S. (2014). Going micro: Emerging and future peacekeeping research. *International Peacekeeping*, *21*(4), 492-500.

Bellows, J., & Miguel, E. (2009). War and local collective action in Sierra Leone. *Journal of Public Economics*, *93*(11), 1144-1157.

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

Besley, T. J., & Ghatak, M. (2009). Improvement and extension of property rights. *Handbook of development economics*, *5*.

Blattman, C., Hartman, A. C., & Blair, R. A. (2014). How to promote order and property rights under weak rule of law? An experiment in changing dispute resolution behavior through community education. *American Political Science Review*, *108*(1), 100-120.

Blair, R. A., Blattman, C., & Hartman, A. (2017). Predicting local violence: Evidence from a panel survey in Liberia. *Journal of Peace Research*, *54*(2), 298-312.

Broegaard, R. B., Vongvisouk, T., & Mertz, O. (2017). Contradictory land use plans and policies in Laos: tenure security and the threat of exclusion. *World Development*, *89*, 170-183.

Bromley, D. W. (2009). Formalising property relations in the developing world: The wrong prescription for the wrong malady. *Land Use Policy*, *26*(1), 20-27.

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys*, *22*(1), 31-72.

Chimhowu, A., & Woodhouse, P. (2006). Customary vs private property rights? Dynamics and trajectories of vernacular land markets in Sub-Saharan Africa. *Journal of agrarian change*, *6*(3), 346-371.

Collier, P., & Hoeffler, A. (2004). Greed and grievance in civil war. *Oxford economic papers*, *56*(4), 563-595.

Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and statistics*, *84*(1), 151-161.

Deininger, K., & Feder, G. (2009). Land registration, governance, and development: Evidence and implications for policy. *The World Bank Research Observer*, *24*(2), 233-266.

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.

Di Falco, S., Laurent-Lucchetti, J., & Veronesi, M. (2015). Property rights and conflicts: theory and evidence from the Highland of Ethiopia Preliminary and incomplete. Retrieved from: https://www.tse-fr.eu/sites/default/files/TSE/documents/conf/energy_climat/Papers/laurent-lucchetti.pdf

Fearon, J. D. (1998). Bargaining, enforcement, and international cooperation. *International organization*, *52*(2), 269-305.

Fatema, N., & Kibriya, S. (2018). *Givers of great dinners know few enemies: The impact of household food sufficiency and food sharing on low intensity interhousehold and community conflict in eastern Democratic Republic of Congo (*HiCN Working Paper 267). Retrieved from the Institute of Development Studies (IDS) at the University of Sussex, Households in Conflict Network website: http://www.hicn.org/wordpress/wp-content/uploads/2018/03/HiCN-WP-267.pdf.

Feder, G., & Nishio, A. (1998). The benefits of land registration and titling: economic and social perspectives. *Land use policy*, *15*(1), 25-43.

Field, E. (2005). Property rights and investment in urban slums. *Journal of the European Economic Association*, *3*(2-3), 279-290.

Galiani, S., & Schargrodsky, E. (2010). Property rights for the poor: Effects of land titling. *Journal of Public Economics*, *94*(9), 700-729.

Garfinkel, M. R. (1994). Domestic Politics and International Conflict. *The American Economic Review*, *84*(5), 1294–1309.

Garfinkel, M. R., & Skaperdas, S. (2007). Economics of conflict: An overview. *Handbook of defense economics*, *2*, 649-709.

Goldstein, M., & Udry, C. (2008). The profits of power: Land rights and agricultural investment in Ghana. *Journal of political Economy*, *116*(6), 981-1022.

Gulyani, S., & Talukdar, D. (2008). Slum real estate: The low-quality high-price puzzle in Nairobi's slum rental market and its implications for theory and practice. *World Development*, *36*(10), 1916-1937.

Haavelmo, T. (1954). *A study in the theory of economic evolution* (No. 330.1/H11s). Amsterdam: North-Holland.

Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The review of economic studies*, *64*(4), 605-654.

Heckman, J. J., Ichimura, H., & Todd, P. (1998). Matching as an econometric evaluation estimator. *The review of economic studies*, *65*(2), 261-294.

Hirshleifer, J. (1991). The technology of conflict as an economic activity. *The American Economic Review*, *81*(2), 130-134.

Hirshleifer, J. (1995). Theorizing about conflict. *Handbook of defense economics*, *1*, 165-189.

Horowitz, D. L. (1985). *Ethnic groups in conflict*. Univ of California Press.

Humphreys, M., & Weinstein, J. M. (2008). Who fights? The determinants of participation in civil war. *American Journal of Political Science*, *52*(2), 436-455.

Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and statistics*, *86*(1), 4-29.

Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of economic literature*, *47*(1), 5-86.

Justino, P. (2011). The impact of armed civil conflict on household welfare and policy responses. *Securing Peace: State-Building and Economic Development in Post-Conflict Countries*, *19*.

Kalyvas, S. N. (2006). *The logic of violence in civil war*. Cambridge University Press.

Kaplan, R. D. (1994). The coming anarchy. *Atlantic monthly*, *273*(2), 44-76.

Linkow, B. (2016). Causes and consequences of perceived land tenure insecurity: Survey evidence from Burkina Faso. Land Economics, 92(2), 308-327.

Musafiri, N. P. (2008). Land Rights and the Forest Peoples of Africa: Historical, Legal and Anthropological Perspectives.

Paige, J.M. 1975. *Agrarian Revolutions.* New York: Free Press.

Raeymaekers, T. (2008). Conflict and food security in Beni-Lubero: back to the future? *Alinovi, Hemrich y Russo (2008), págs*, 169-195.

Skaperdas, S. (1992). Cooperation, conflict, and power in the absence of property rights. *The American Economic Review*, 720-739.

Van Acker, F. (2005). Where did all the land go? Enclosure & social struggle in Kivu (DR Congo). *Review of African Political Economy*, *32*(103), 79-98.

Vlassenroot, K., & Huggins, C. (2004). Land, migration and conflict in Eastern DR Congo. *African Centre for Technology Studies*, *3*(4), 1-4.

Vlassenroot, K., & Huggins, C. (2005). Land, migration and conflict in eastern DRC. *From the ground up: land rights, conflict and peace in sub-Saharan Africa*, 115-195.

Vlassenroot, K., & Raeymaekers, T. (Eds.). (2004). *Conflict and social transformation in Eastern DR Congo*. Academia Press.

Vlassenroot, K., & Raeymaekers, T. (2008). Crisis and food security profile: The Democratic Republic of the Congo. *Beyond relief: food security in protracted crises*, 157-168.

Voors, M. J., Nillesen, E. E., Verwimp, P., Bulte, E. H., Lensink, R., & Van Soest, D. P. (2012). Violent conflict and behavior: a field experiment in Burundi. *The American Economic Review*, *102*(2), 941-964.

Wickham-Crowley, T. P. (1992). *Guerrillas and revolution in Latin America: A comparative study of insurgents and regimes since 1956*. Princeton University Press.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

APPENDIX III

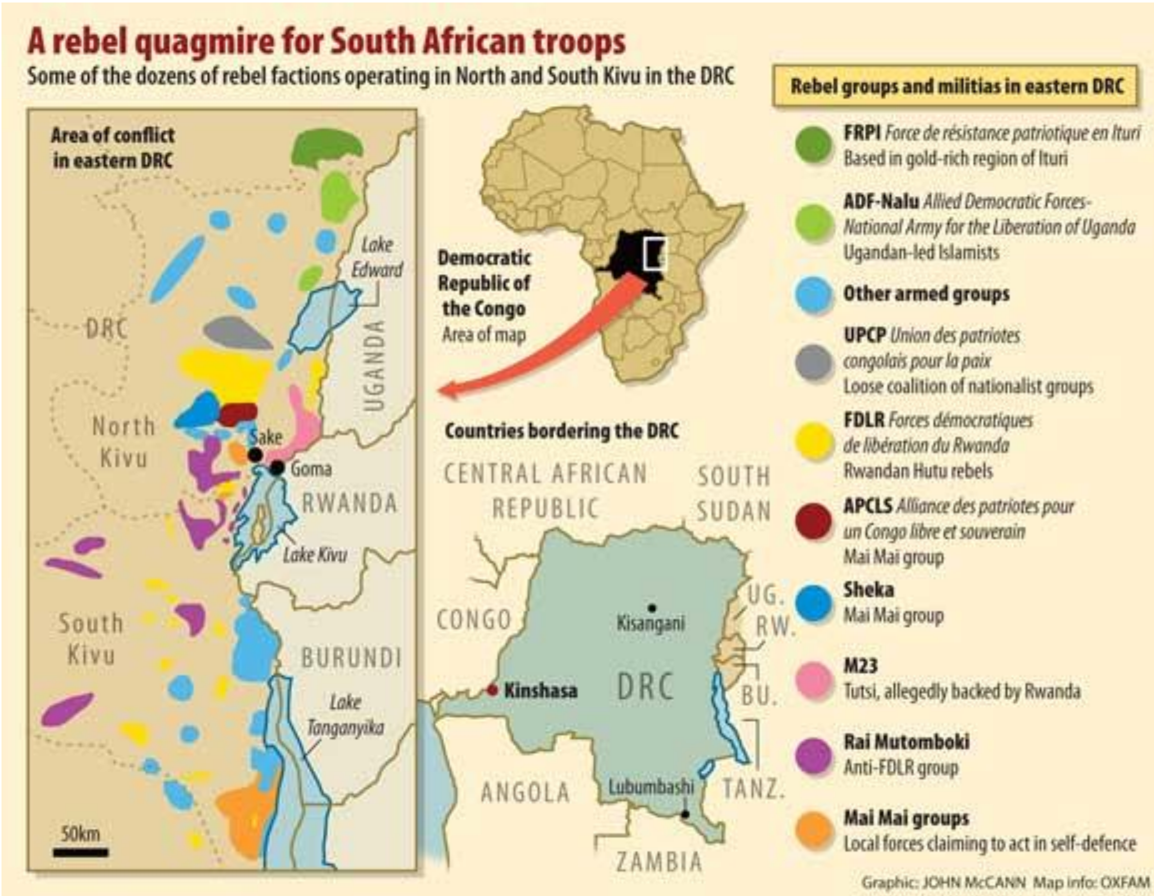


Figure 3.1: Map of DRC and Kivu showing rebel forces

Source: The Mail and Guardian. Available at https://mg.co.za/article/2013-02-15-00-sadc-mission-could-be-suicidal

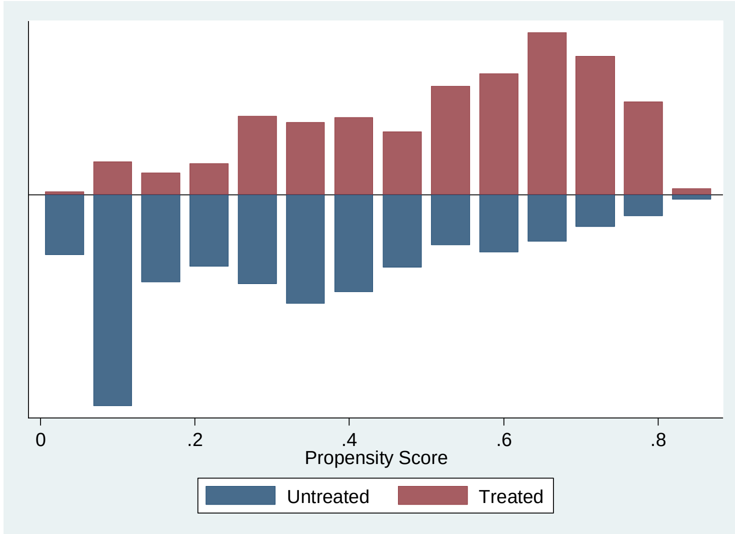


Figure 3.2: Graph of propensity score across control and treated groups

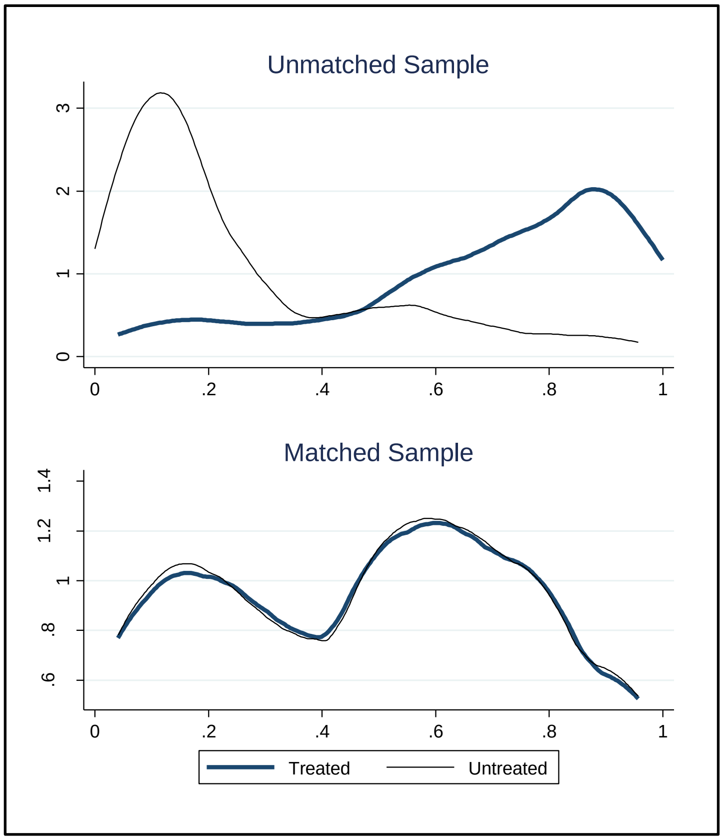


Figure 3.3: Distribution of propensity scores (given conflict) in treated and control groups before and after matching using radius algorithm

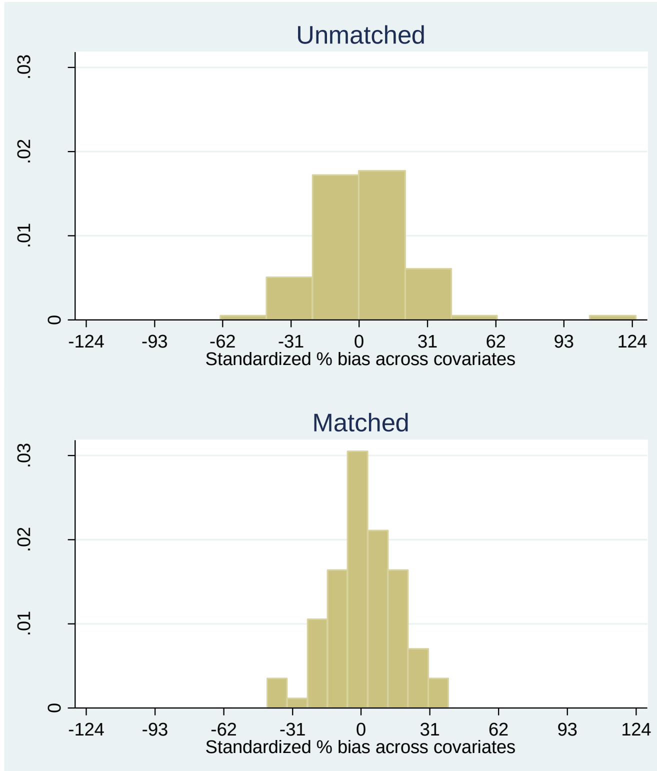


Figure 3.4: Histogram of standardized % bias across covariates before and after matching

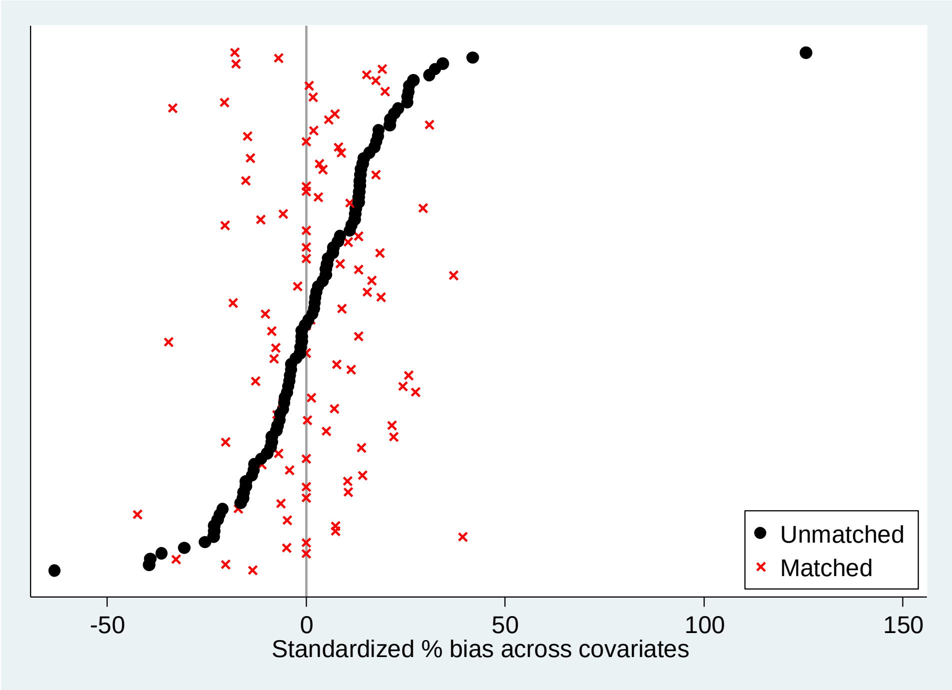


Figure 3.5: Graph of standardized differences

Note: The graph shows the percentage standardized bias between covariates in treated and control groups. The standardized bias approach calculates the difference of sample means in the treated and the matched control groups as a percentage of the square root of the average sample variance in both groups. It can be seen that the standardized bias is closer to 0 in the matched sample. This implies that matching reduced the imbalance in distributions across covariates and provides further support for good quality matching.

Table 3.6: Definition of variables to record damages from conflict

Variable	Survey question: (If in the past six months your household has experienced any form of conflict,) Please state any outcomes that resulted from the conflict (circle any answers that the respondent states):
Loss/damage to physical capital	loss of a portion of land
	destruction of agricultural equipment
	theft or loot of crops
	theft or destruction of personal property or equipment
	displacement from home or property
	destruction of livestock
	damage or destruction of crops
	destruction of house
Loss/damage to human capital	physical violence to anyone in household
	death of anyone in household
Loss of financial capital	requirement to pay compensation, monetary or otherwise
	loss of off-farm income source
	extortion of any household member
Loss of rights or dignity	loss of land rights over part of land
	loss of community acceptance and punishment by the village council
Loss of access	loss of access to roads
	loss of access to usual market

Table 3.7: Propensity score distribution for treated and control households

Propensity score for households	Mean	Std. Dev.	Min	Max	No. of Obs.
With land title	0.695	0.269	0.040	1	614
Without land title	0.228	0.226	0.000	0.969	819

Source: Authors' calculations based on the survey data.

Note: Households with land titles have propensity scores ranging between 0.040 and 1, with a mean value of 0.695 and standard deviation of 0.269. Households without title have propensity scores between 0 and 0.969 with a mean value of 0.228 and standard deviation of 0.226. Therefore, the region of common support lies within a propensity score of 0.040 and 0.969. Only observations whose propensity scores lie within this region are used in my analysis.

Table 3.8: Covariate balance between treated and control groups, before and after matching

Covariate	Sample	Mean Treated	Control	% Bias	% Reduction in bias	Diff: *p*-value
Household size	U	5.15	5.47	-13.60		**0.008**
	M	5.72	5.39	14.1	-4.2	0.320
Number of adult males	U	0.48	0.44	13.60		**0.007**
	M	0.45	0.48	-15.3	-12.7	0.283
Education	U	9.76	8.56	26.00		**0**
	M	9.55	9.65	0.8	97.1	0.957
Education squared	U	115.47	95.38	25.60		**0**
	M	110.96	109.62	1.7	93.4	0.902
HH income (1000 Congolese Franc)	U	22.82	14.90	18.30		**0**
	M	17.83	17.01	1.9	89.7	0.795
Access to technology (yes=1)	U	0.82	0.69	32.50		**0**
	M	0.76	0.68	19.20	41.0	0.217
Access to credit (yes=1)	U	0.09	0.08	5.00		0.328
	M	0.13	0.09	13.10	-163.8	0.425
Lack of extension services (yes=1)	U	0.47	0.66	-39.40		**0**
	M	0.51	0.60	20.20	48.7	0.179
Cooperative membership (yes=1)	U	0.26	0.17	21.20		**0**
	M	0.24	0.21	5.60	73.5	0.709
Authority	U	0.92	0.93	-6.6		0.187
	M	0.90	0.90	0.3	94.9	0.984
Leadership position (yes=1)	U	0.67	0.57	21.00		**0**
	M	0.75	0.60	31.0	-47.2	0.307
Teacher or doctor	U	0.36	0.28	18.10		**0**
	M	0.30	0.37	-14.7	18.6	0.326
Food insecure	U	0.14	0.13	2.20		0.663
	M	0.12	0.18	-18.40	-730.4	0.234
Rents land	U	0.17	0.44	-63.20		**0**
	M	0.24	0.3	-15.50	78.6	0.362
Purchased land	U	0.69	0.16	125.7		**0**
	M	0.43	0.51	-81.0	85.7	0.306
Inherited land	U	0.25	0.36	-23.20		**0**
	M	0.35	0.17	39.3	-69.6	**0.005**
Received farmer training	U	0.22	0.17	13.70		0.006
	M	0.28	0.21	17.5	-27.8	0.3278
Has insurance	U	0.03	0.01	17.30		**0**
	M	0.01	0.00	8.1	53.1	0.319

Source: Authors' calculations based on the survey data.

Note: U=unmatched sample and M=matched sample. This table shows results for number of damages given conflict, using radius matching. Results are very similar with other matching algorithms as well as when the dependent variable is probability of conflict. The table shows results from t-tests of the normalized differences in mean13 for each covariate in the treated and control groups before and after matching. If the difference in mean is not significantly different from zero, it can be argued that the distribution is balanced. The null hypothesis is that the mean values are not significantly different for any covariate in the treated and control groups. Bold p-values indicate the difference in means are significant at a level of 10% or lower. Due to space constraints, the means for community and religion dummies have been excluded from the table. The total number of observations is 701 with 286 for treated and 415 controls. As the table shows, the means are significantly different between the two groups prior to matching, thus implying a potential selection bias. However, the post matching results show that the difference in means is not statistically significant at the 10% level for any variable except inherited land. Furthermore, a bias is calculated for each variable and the change in bias is reported. For each variable, the "bias" is defined as the difference in mean values of the treatment and control group, divided by the square root of the average sample variance in the treatment group and the unmatched sample in the control group. A rule of thumb (Rubin, 2001) is to consider a covariate as balanced if the percent reduction in bias in the matched sample is 25% or less. Results from the table indicate that the percent reduction is less than 25% for each covariate post-matching. These results suggest that matching has indeed sufficiently reduced the pre-existing selection into treatment.

Table 3.9: Matching quality indicators

Sample	Pseudo R^2	LR χ^2	$p > \chi^2$	Mean standardized bias	B	R	%Var
Unmatched	0.392	767.9	0	15.3	155.3*	1.14	60
Nearest-neighbor matched	0.224	172.82	0	9.7	103.7*	2.64*	20
Kernel matched	0.225	175.58	0	9.4	105.3*	3.21*	20
Radius matched	0.303	64.4	0.057	11.9	133.4*	1.22	0

Source: Authors' own calculations using the survey data.

[13] For each covariate, the difference in averages by treatment status, scaled by the square root of the sum of the variances is used as a scale-free measure of the difference in distributions

Note: Epanechnikov kernel matching method used a bandwidth of 0.06. Radius matching used a caliper of 0.001. Before and after refer to results before matching and after matching. Dependent variable is number of damages given conflict.

Table 3.10: Results from placebo regression

Covariates	
Household has land title	0.0309
	(0.0337)
Household size	0.0291***
	(0.00764)
Number of adult males	-0.521***
	(0.0807)
Education	-0.00779
	(0.0107)
Education squared	0.000594
	(0.000637)
Household income	-7.64e-07
(1000 Congolese Franc)	
	(6.72e-07)
Access to technology (yes=1)	0.0616
	(0.0383)
Access to credit (yes=1)	-0.0215
	(0.0521)
Lack of extension services (yes=1)	-0.0312
	(0.0387)
Cooperative membership (yes=1)	0.0462
	(0.0357)
Leadership position (yes=1)	0.111***
	(0.0321)
Teacher or doctor	0.0395
	(0.0344)
Food insecure	-0.0865*
	(0.0459)
Rents land	0.0298
	(0.0456)
Purchased land	0.0439
	(0.0419)
Inherited land	0.0634
	(0.0412)
Received farmer training	-0.0394
	(0.0412)
Has insurance	-0.0589
	(0.106)
Constant	1.943***
	(0.389)
Observations	1424
R-squared	0.19

Source: Authors' calculations based on the survey data. Robust standard errors in parenthesis.

Note: *, **, and *** indicate significance at 10%, 5% and 1% levels respectively. A "placebo" regression is one where the treatment and control variables are regressed on a variable which is logically uncorrelated with treatment, and therefore exogenous to the treatment. The dependent variable in this table is the interviewee's relationship to the first household member introduced in the interview. This is logically uncorrelated with the land title status of the household. Results show that land title has no effect on this exogenous dependent variable. This also suggests that there are most likely no omitted variables correlated with having land title and is tentative validation of selection of observables.

Table 3.11: Doubly robust estimation and Rosenbaum test results

Outcome of interest	(1)	Critical level of hidden bias (Γ)
Conflict incidence	-0.089**	2.15
	(0.032)	
Conflict intensity	-0.203***	1.43
	(0.073)	

Source: Authors' calculations based on the survey data.

Note: *, **, and *** indicate significance at 10%, 5% and 1% levels respectively. Column (1) shows the results for a linear probability model using propensity score as a control variable in the regression equation. Number of observations=1426 for all matching algorithms. For the Rosenbaum bounds, Γ is significant at the 10% level. Dependent variable is conflict. I use the Hodges-Lehmann point estimates. I find that under the assumption of no potential hidden bias, i.e. when Γ =1, the results are similar to my estimates. With land title as the treatment, the values of Γ range between 1.43 and 2.15. This implies that the unobserved covariates would have to increase the odds of having land title by a factor of between 1.43 (43%) and 2.15 (115%) to overturn the significance of my ATT estimates.

Table 3.12: Results from alternate specifications

Measure of property rights	Full sample	HH that experienced conflict
Land title	-0.159**	-0.005
	(0.076)	(0.049)
No fear of expropriation	-0.303***	-0.019
	(0.071)	(0.044)
No fear of transfer rights	-0.136**	-0.088**
	(0.073)	(0.045)
No. of obs.	1433	701

Source: Authors' calculations based on the survey data.

Note: *, **, and *** indicate significance at 10%, 5% and 1% levels respectively. Dependent is the number of damages. Poisson regression results. Robust standard errors in parenthesis.

Connecting Text II

Chapters 2 and 3 used a survey of farming households in eastern DRC to empirically analyze the role of behavior on interhousehold conflict through two key aspects of household decision making – food security and property rights. The use of household survey data from conflict and post-conflict societies is generally lacking due to the physical and political barriers to collecting information from such areas. By exploiting a survey that was specifically designed to capture the conflict experience of households, these two chapters jointly contribute to the micro-empirical conflict literature by providing new insights on the behavior of households from a fragile, active conflict zone. Chapter 4 continues the examination of the root causes of local conflict by extending the analysis from household behavior to community behavior and cultural evolution in agrarian communities across sub-Saharan Africa.

Chapter 4

RICE, CULTURE, AND CONFLICT IN SUB-SAHARAN AFRICA

4.1 Introduction

The debate on geographical and institutional pathways to economic and political development (Acemoglu, Johnson, & Robinson, 2001, 2002; Easterly & Levine, 2003; Engerman and Sokoloff, 2000; Gallup & Sachs, 2000; Michalopoulos & Papaioannou, 2013; Nunn, 2008) has recently infiltrated the civil conflict literature in growing efforts to explain comparative regional conflict across Africa (for example, Besley & Reynal-Querol, 2014; Michalopoulos & Papaioannou, 2011, 2020; Nunn, 2009). One school of thought emphasizes the importance of geographical factors, such as climate, natural resources, access to coastline and sea, navigable rivers, landlock, conditions favorable for agriculture, terrain ruggedness and disease environment in determining conflict in Africa (Collier & Hoeffler, 1998, 2002; Fearon & Laitin, 2003; Harari & La Ferrara, 2018; McGuirk & Burke, 2017; Miguel Satyanath & Sergenti, 2014, etc.). The other posits that historical legacies such as colonial rule, the partitioning of post-independence Africa, and trans-Atlantic and Indian ocean slave trades (Besley & Reynal-Querol, 2014; Michalopoulos & Papaioannou, 2013, 2014, 2016; Nunn & Wantchekon, 2011; Zhang, Xu & Kibriya, 2021; etc.) have shaped the path of conflict in Africa through their effect on institutions (see Nunn, 2009 and Michalopoulos & Papaioannou, 2020 for comprehensive reviews). These studies uncover the relationship between civil-conflict and institutions, economics and the environment in SSA. The institutional and economic drivers along with human induced changes to the environment are largely driven by human norms and behaviors. However, despite growing efforts to understand the root causes of conflict in Africa, most theories have largely ignored advances in behavioral economics and until recently economists have been hesitant to study the relevance of behavior and culture on economic outcomes (Alesina & Giuliano, 2015; Blattman & Miguel, 2010; Guiso, Sapienza & Zingales, 2006; Nunn, 2009; Nunn, 2012). Consequently, the impacts of human behavior and cultural practices on sub-national conflict has hardly been examined with observational data. In recent years, African historiography has challenged the theories accepted universally by economists and political scientists on the role of colonial powers in shaping the path of development in contemporary Africa through the establishment of formal domestic institutions. Instead, there has been an emphasis on the importance of informal institutions such as deep rooted ethnocultural traits that existed centuries prior to colonial rule (Herbst, 2014). As a result, there has been a growing interest to examine the role of culture and heritage through analyses of unique historical events to shed light on contemporary conflict in Africa (Nunn, 2009, 2012; Michalopoulos & Papaioannou, 2016, 2020). Though the notion of culture is implicitly included within informal constraints, this chapter uses the term culture to refer to decision-making heuristics that include values, beliefs and social norms (*see* Alesina & Guiladano, 2015; Nunn, 2012) to draw a clear distinction from formal institutions that are governed by the polity. When used explicitly, the word *culture* is defined as "those customary beliefs and values that ethnic, religious, and social groups transmit fairly unchanged from generation to generation" (Guiso, Sapienza & Zingales, 2006). Through this study, I hope to contribute to this budding avenue of research on the role of culture and heritage in shaping the path of Africa's conflict by introducing into the discussion yet another historical legacy that preexisted long before colonial rule but has thus far been neglected from analyses - the introduction of rice to the continent.

Studies from the Behavioral and Social Sciences such as Sociology, Psychology and Cultural Anthropology have established a relationship between a legacy of rice growing traditions and a culture of cooperation, compliance, interdependence, and collectivism. The underlying mechanism is that rice farming inherently requires cooperation and coordination among farming households thus encouraging the formation of tight reciprocal relationships. For example, recent empirical evidence suggests that individuals and communities from historically rice growing regions tend to display higher propensity towards values such as trust, altruism, and family ties, and a weaker affinity towards intellectual property rights and democracy (Ang & Fredriksson 2017; Jiong, Ang & Frederikkson, 2019; Ang, Madsen, & Wang, 2021; Talhelm et al. 2014; Tsusaka et al., 2015, etc.). Thus, it seems plausible that rice growing communities support exchange and peaceful coexistence that make cooperation valuable, and conflict economically and socially costly within a particular geographical area. Therefore, I hypothesize that, i) enclaves that predominantly cultivate rice develop a culture of cooperation and peace which makes them less prone to conflict; and ii) since cultural evolution persists through time, a heritage of rice farming legacy is correlated with lower prevalence of contemporary conflict. I follow the definition used by Guiso, Sapienza & Zingales (2006) and define *culture* as, "those customary beliefs and values that ethnic, religious, and social groups transmit fairly unchanged from generation to generation." Though the standard definition of *institutions* by Douglass North (1990) includes both formal and informal constraints, this chapter uses the word culture to refer to informal institutions (beliefs, values and norms of behavior) to distinguish it from formal rules of law that are governed by the polity (*see* Alesina & Giuliano, 2015 for further discussion on the distinctions between culture and institutions in the empirical literature).

To empirically test for these relationships, I use widespread data from traditional and contemporary rice and non-rice growing regions of SSA to test whether rice farming legacies are correlated with lower conflict. I combine large datasets geocoded at half decimal degree "grid cells" (approximately 55 by 55 km at the equator). I use the *Spatial Production and Allocation Model* (SPAM) 2005, version 1.0 for cross-sectional agricultural data and the *Armed Conflict Location and Event Data Project* (ACLED) version 7 for conflict data. While ACLED is available as panel data, I mostly restrict the analysis to the year 2005, with a few exceptions for robustness checks, to be able to match it with the available crop data. These high resolution datasets offer a few advantages. First, comparing agriculture and conflict outcomes across countries would be confounded by cross country differences. The data allows for fine grained analysis by comparing similar grid cells within each country. Second, ACLED catalogs the most comprehensive conflict data known, documenting nine types of events ranging from riots and protests to violent battles. This allows for the possibility of exploring specific events that may be more relevant for local agricultural practices and cultural values. Third, rich crop data in the SPAM dataset allows a comparison of rice with other staple crops of the region such as maize, pearl millet, sorghum, cassava, and wheat. The final sample comprises a total of 5930 grids (units) across 48 countries in sub-Saharan Africa.

I first test whether contemporary rice growing regions correlate with lower prevalence of conflict. I find a negative and robust relationship with a sizeable magnitude of over 60 percent reduction in conflict risk in rice producing units. This holds true for three measures of conflict - conflict incidence (risk of any conflict event), intensity (total number of conflict events) and onset (start of a new event). The results hold with the inclusion of an extensive set of political and geographical covariates and country fixed effects. Next, I test whether present day rice farming in historically rice growing regions correlates with lower contemporaneous conflict. This test serves two purposes. One, despite controlling for observable economic, political, and geographical differences, the correlations may suffer from factors unaccounted for including unobservable differences in rice and non-rice growing units. Comparing conflict outcomes in units that predominantly cultivate rice with units that do not, in regions with and without a history of rice farming, offers four outcome scenarios: i) contemporary conflict in present-day rice farming units belonging to regions with centuries-old history of rice cultivation; ii) contemporary conflict in present-day rice farming units belonging to regions without a prior history of rice; iii) contemporary conflict in present-day non-rice farming units belonging to regions that planted rice for centuries but have recently switched to other crops; and iv) contemporary conflict in present-day non-rice farming units from regions that did not traditionally farm rice. This comparison essentially rules out unobserved differences between units. Two, it tests whether a longer history of rice growing traditions (reflecting a deeper culture of cooperation and peace) is correlated with additional reduction in contemporary conflict. I find that historically rice cultivating localities that continue to grow rice at present experience lower conflict than all other farming localities of SSA. My findings provide suggestive evidence that a history of rice growing practices promotes a culture of cooperation, coexistence, and peace. This study contributes to the emerging literature on historical legacies and culture in shaping comparative development by introducing the theoretical framework and preliminary evidence for the channel through which a relatively unexplored African heritage may have influenced its path of conflict. Thus, it contributes to understanding the deep roots of conflict. The study also addresses the broader "geography versus institutions" debate on determining economic and political outcomes. Finally, it adds to the body of empirical evidence from the social sciences on observed relationships between a legacy of rice farming and sociopolitical behavior and outcomes. While these studies have examined the influence of rice growing traditions on individuals' analytical and social abilities, family values, intellectual property rights, democracy, etc. (e.g., Ang & Fredriksson, 2017; Ang, Madsen & Wang 2021; Talhelm et al., 2014) to the best of my knowledge none of these have been undertaken outside Asia and none have examined the relationship between rice farming and continent wide, comprehensive, sub-national conflicts.

The remainder of the paper is organized as follows. Section 2 introduces the relevance and significance of Africa's rice farming heritage by documenting the process by which a millennial old history of rice may have determined the path of sub-national conflict in present day sub-Saharan Africa. Sections 3 and 4 describe the data, variables and methods. Section 5 presents the results from tests of robust correlations between present day rice farming, cultural evolution through a history of rice farming, interactions between the two, and contemporaneous conflict across SSA. Section 6 discusses potential future avenues to disentangle causal relationships between rice farming traditions and cultural values, domestic institutions, and their persistent effects on present day sub-national conflict. Section 7 concludes the paper.

4.2 Theoretical Framework

The theoretical framework for this chapter is based on two chains of arguments: the relationships between rice cultivation and the evolution of a culture of cooperation and peace; and the relationships between cultural values, domestic institutions and conflict. I discuss each below.

4.2.1 Behavioral pathways from a legacy of rice farming to cooperative culture

In recent years, empirical studies across multiple disciplines in the Behavioral and Social Sciences have established a link between a legacy of rice farming and social behavior. This stream of literature originated with the *"rice theory of culture"* hypothesizing that individuals descending from traditionally rice farming communities develop values and habits which continue to shape their thought process and behavior in the modern world (Talhelm et al., 2014). Accordingly, researchers have demonstrated that people from rice growing regions tend to exhibit a culture of compliance, interdependence, and collectivism as observed though their decisions and actions across multiple social phenomena. These arguments are based on aspects of rice cultivation that differentiates itself from other crops via a constant need for individuals and households to collaborate and coordinate among themselves. The reasons for this are as follows. First, rice cultivation requires almost twice as much labor as other crops such as wheat, barley, maize, and millet ("Fei, 1945; Buck, 1935; and Richards, 1987", as cited in Talhelm, 2015). Part of the reason behind extra labor demand is that rice seedlings need to be grown in small plots before being transplanted to farms. In contrast, wheat production requires a relatively short growing season and low labor inputs for most parts (Ang & Fredriksson, 2017). Even scaling up wet-rice agriculture requires increasing inputs of manual labor (Bray, 1983). Second, rice plantation requires peasant households to cooperate with one another and coordinate their actions to cope with the urgent and highly labor-intensive tasks of transplanting and harvesting in due time. While seasonal cultivation is universally time sensitive, the steps required for rice plantation is more sensitive than similar crops such as corn (Fei 1945). The obligation to complete each step within a given window leads to excessive labor demand at critical times during the production cycle, thus creating a labor bottleneck (Richards, 1987). For example, for small-scale rice farmers in Sierra Leone, labor shortage is the most pressing constraint to production (NRC, 1996). One way in which farmers cope with high labor demands, and labor bottlenecks is by forming labor exchange cooperatives. Such labor exchange systems are common in communities that grow rice and even small farmers are required to hire or exchange labor (Bray, 1986). Other forms of coordination between farmers include staggering planting and harvesting dates and planting different varieties thus ensuring that effective rotas for labor exchange and hire can be established and sufficient labor is available for all during harvest season (Bray, 1983; Lewis, 1971; Talhelm et al., 2014). Richards (1987) documents the use of percussion instruments to create a rhythm to help coordinate physical movement between groups while Richards (1996) describes how upland rice farmers in West Africa clear land in blocks such that as many as twenty to thirty households occupy adjacent plots. This coordinated land use among farmers offers certain economies of scale so that more effective land clearing ‘burns’ or better managed protection against rodents and birds, in addition to having extra help available during the labor-intensive steps of rice farming. Third, paddy rice grows best in standing water which requires substantial labor and coordinated efforts to operate elaborate water management systems for both irrigation and drainage in case of excess water (Aoki, 2001; Bray, 1986; Richards, 1987; Talhelm, 2014; Tsusaka et al. 2015; Von Carnap, 2017). Irrigation networks require many man-hours to build, dredge, and drain each year. Repairing and maintaining dams, irrigation channels and drainage systems is both labor intensive and costly for individuals and require collective organization by relatively large communities (Aoki, 2001; Bray, 1986; Mustafa & Qazi, 2007). In addition, since the infrastructure is shared by multiple farms, farmers are forced to coordinate water use. For example, Aoki (2001) describes labor intensive tasks associated with maintaining productive irrigation systems such as cleaning dirt and weeds from water channels or preventing the destruction of the system during floods, as well as coordination in the timing of seedling planting and collective decisions on when to flood and drain the fields. Repair of irrigation channels and embankments is the primary reason for rice farming requiring double the labor compared to millet in West Africa (Richards, 1987). A socially driven structure implies African farmers rely on hand cultivation, and an irrigation pump serving about 40 hectares might be shared by as many as 30-50 households (Moris, 2019) resulting in a greater a need for coordination among those households to share and manage water. Irrigation and drainage systems are usually governed by village communities who often bear the cost of maintenance as well (Aoki, 2001; Bray, 1986; Fei, 1945; Ruan, 2015). While sharing labor and cost over common pool resources such as irrigation systems has the potential hazard of free riding by villagers who refuse to contribute, these free riders are often punished through social exclusion by fellow villagers (Aoki, 2001). Thus, the entire process of rice farming encourages farmers to value cooperation and form tight reciprocal relationships, thus fostering a culture of collectivism and making confrontations and conflict economically and socially costly.

These social theories have been put to empirical test by Talhelm et al (2014) who used a series of laboratory experiments to show that people from rice growing regions of southern China are more interdependent, think holistically or collectively, and tend to avoid conflict in their everyday lives compared to the individualistic, and assertive nature of individuals from wheat growing regions of northern China. Furthermore, individuals from rice growing regions tend to file for less divorces in their marriages and less individual patents in their scientific innovations. The study generated much interest and publicity from scientific communities and the media, thus drawing other researchers to investigate further. For example, using high resolution data from over 2000 counties in China, Jiong, Ang & Frederikkson (2019) find further robust, supporting evidence that counties with a legacy of rice production file for less individual patents. In an earlier study, Ang

& Fredriksson (2017) show that through lower dependence on irrigation, and by requiring less manpower for production, wheat farming creates a lower need for interdependence, cooperation and coordination between members of farming households. Eventually this leads to a culture of weak family ties. Henrich (2014) argues that a predominant culture of wheat farming may explain why the Industrial Revolution originated from Europe and not China. Ang, Madsen & Wang (2021) show that centuries of collaboration among workers in societies with a rice farming legacy leads to the adoption of a collectivist culture that deters the process of democratization.

4.2.2 Pathways from culture and domestic institutions to conflict

There is ample evidence in the interdisciplinary literature to support that historical events play a crucial role in determining the evolution of culture, that cultural evolution is interrelated with domestic institutions (*see* Alesina & Giuliano, 2015 for a review), and that the persistent nature of cultural change, domestic institutions, and their interactions affect long-term conflict outcomes (*see* Nunn, 2012 for a comprehensive review). A traditional Neo-Classical model predicts that a historic event or shock can only result in temporary deviations from the unique steady state equilibrium. Economic theories used to explain the persistence of conflict caused by historical legacies include mechanisms of culture, domestic institutions, their interactions with each other, and the existence of multiple steady state equilibria (Nunn, 2009). Scholars argue that culture is an evolutionary process and by altering the relative costs and benefits of cultural traits, historical events can permanently affect economic and political outcomes. In addition, if these traits are passed down through generations, the effects can persist over time. Furthermore, culture can affect development outcomes permanently through its interaction with other persistent factors such as domestic institutions. For example, Guiso, Sapienza, and Zingales (2008) provide a model for the evolution of trust to demonstrate permanent effects of a cultural shock. In their model with multiple equilibria, children learn cultural values from parents as well as through market transactions. If they learn mistrust from parents, there is no economic activity and no learning but if they inherit the value of trust, there is high trade in the economy and further learning about the true trustworthiness of the population. In such a case, the economy can permanently move to a new steady state equilibrium through cultural evolution. In another model of intergenerational cultural transmission, Bisin & Verdier (2001) study the long run stationary state pattern of preferences in the population, according to various socialization mechanisms and institutions in a setting where parents socialize and transmit their preferences to their offspring. Tabellini (2008) uses a model for the coevolution of culture and institutions to show that a culture that values cooperation prefers institutions that enforce cooperation, thus increasing the returns to cooperation, and reinforcing this cultural preference. Nunn (2012) argues that long term economic growth in Acemoglu et al.'s (2001) colonial origins hypothesis was brought about by both culture and domestic institutions. While the original paper argued that early European settlers built strong domestic institutions that led to long term growth, Nunn adds that the cultural values brought by the settlers to the colonies, such as their beliefs and values on liberty, equality, and the appropriate role of government, had additional impact on long term growth through an effect on the nature of the formal institutions established.

A growing number of interdisciplinary studies in recent years have established a relationship between historical legacies and the development of Africa with the insight that events and features rooted deeply in the history of Africa can explain variations in conflict (*see* Michalopoulos & Papaioannou, 2020 for a comprehensive review). Of particular interest to this chapter are the studies on cultural values, domestic institutions, and contemporaneous conflict in Africa. For example, Besley & Reynal-Querol (2014) examine mechanisms through which events that took place centuries ago influence contemporary economic and political outcomes. By analyzing conflicts between 1400 and 1700, the authors find that historical conflict in Africa is correlated with lower levels of trust and higher levels of contemporary conflict. In another study, Nunn and Wantchekon (2011) show that a legacy of slave trade in Africa is correlated with present day mistrust. Using a number of econometric techniques, the authors identify two causal channels to explain results. First, slave trade permanently altered the cultural values of individuals whose ancestors were exposed to slave trade, making them inherently mistrusting of others. Second, slave trade had a detrimental long-term effect on domestic institutions, which led to further distrust among people at present. Though mechanisms include both cultural values, beliefs and norms that are internal to individuals as well as external domestic institutions, the authors find that the impact from the culture channel is about twice the magnitude of the institutions channel. In another study on the role of formal institutions, culture, and history on the comparative regional development in Africa, Michalopoulos & Papaioannou (2013, 2014) study the role of institutions on comparative regional development within historic homelands of all ethnic groups. The study exploits variations in the arbitrarily drawn colonial borders in post-independence Africa that partitioned ethnic groups in different countries, thus subjecting similar ethnic-specific traits such as culture and history to different legal and political local institutions. The study finds that differences in local institutions do not explain economic performance within ethnicities and emphasizes the need to further examine the interactions between formal institutions and ethno-cultural traits in explaining development in Africa.

The two chains of theoretical arguments outlined above suggest the plausibility that rice growing communities support a culture of exchange and peaceful coexistence that make cooperation valuable, and conflict economically and socially costly within a particular geographical area. Furthermore, since cultural norms and values evolve and persist through time, a longer legacy of rice cultivation should evolve into a deeper culture of peace and greater aversion towards conflict. Before testing for these correlations formally, I present an overview of the history of rice farming in Africa.

4.2.3 *Historical and present-day rice farming in SSA*

While the exact origin and domestication of rice in Africa is debated by historians, it is agreed upon that it originated in West Africa some 2000-3000 years ago (*see* Linares, 2002; NRC, 1996; or Richards, 1996 for history of rice in Africa). Linares (2002) provides historical accounts of vast rice plantations and local rice growing practices in West Africa, as observed and documented by explorers and early settlers throughout the centuries. Most of the evidence suggests that the majority of inhabitants along the coasts and riverbanks of West Africa were involved in vast rice plantations. Accounts mention the Inland Delta of the Upper Niger River (present day Mali), Guinea, Guinea-Bissau, Gambia, Sierra-Leone and the Ivory Coast where local inhabitants were mostly involved in wetland or swamp rice farming using intensive technologies such as building dikes to retain rainwater and transplanting seeds.

Modern day accounts show that rice remains an important crop in most of West Africa, particularly for Guinea, Sierra Leone, Ivory Coast, Guniea-Bissau and Liberia, where rice seeds are stored and maintained as a community activity with no external intervention (Richards, 1996). At present rice is the main staple crop grown by 10-15 million people habilitating along the coast of West Africa, from the Casamance River in Senegal to the Bandama River in the Ivory Coast in addition to being grown as a commercial crop in Ghana and Nigeria" (Linares, 2002). Rice production rates have almost doubled in the past few decades from 1.76 % in 1991–2001 to 3.96 % during 2002–2013 making it the fastest emerging cereal crop in SSA (Zenna, Senthilkumar, & Sie, 2017).

Much of SSA has a semi-arid climate whereby agriculture relies on river flow which is a highly variable source of water (Moris, 2019). Rice cultivation in Africa occurs in one of three ways – dryland or upland cultivation, paddy rice, and "floating" rice. Dryland rice is mostly grown on rainfed fields while paddy or swamp rice grows best in standing water, thus performing well when the water level can be controlled. In contrast floating rice, which often requires canoes for harvesting, can survive in completely inundated basins through their exceptional elongation ability. About 40 percent of the rice produced in Africa relies on rainwater as its only source of water and only about a-sixth of the rice is produced using irrigation with 60 percent of that in just one country—Madagascar (NRC, 1996). Though an estimated 40 million hectares of land is suitable for irrigation, only 7.3 million ha is actually irrigated and only four countries (Madagascar, Nigeria, South Africa, and Sudan) account for the vast majority of irrigated land (Burney& Naylor, 2012).

Thus, given the millennial old legacy of rice farming in Africa, the distinct regional divide between traditionally rice producing regions of West Africa and the rest of Africa, and the prevalence of conflict throughout the region, SSA serves as an appropriate ground for testing whether communities from traditional and modern rice farming cultures are less prone to conflict.

4.3 Data and Variables

I combine large, publicly available datasets on conflict, agriculture, and other variables geocoded at a high resolution. Where data are unavailable at this resolution, I resample it with the help of a GIS expert. The unit of analysis for the study is 0.5 X 0.5 degree grid cells (approximately 55 X 55 km squares at the equator). The final sample comprises 5928 grid cells across 48 countries in sub-Saharan Africa for the year 2005. Using grids as the unit of analysis rules out potential bias from confounders such as political regime, rule of law, or local dynamics that may exist at an administrative level but should not transcend geographical boundaries defined by the grids.

4.3.1 Conflict

To measure conflict events, I use the *Armed Conflict Location and Event Data Project* (ACLED) version 7, developed by the International Peace Research Institute of Oslo, Norway, and the University of Uppsala, Sweden (referred to as PRIO/Uppsala) and obtained from Raleigh et al. (2010). ACLED collects information on political violence from around the world, based on news media from war zones, humanitarian agencies, and research publications. Political violence is defined as the use of force by a group with a political purpose or motivation. The project records a range of violent and non-violent events by date, location, and actors involved (such as governments, rebels, militias, ethnic groups, active political organizations, and civilians). The data is both high frequency (collected daily) and high resolution (available at 0.5X0.5 degrees). This level of disaggregated information in specific geographical locations offers two advantages: i) it allows researchers to investigate local behavior instead of aggregating information over larger geographical boundaries; and ii) it provides information on the dynamic aspects of civil war, such as war onset, via recorded data on changes in location and expansion of events over time (Raleigh et al., 2010). Though ACLED data covers a period from 1997 to 2016, most of the analysis in this paper is restricted to the year 2005 to match the available crop data. Conflict events are recorded as days of political activity. Multiple events may occur on a given day but if similar events are reported between the same parties and in the same location, these are reported as single events. If an event starts on a given day and activity continues over the next three days, these are coded as three separate events so that each day of political activity is on record. On the contrary, if a turmoil lasts a month but has only four days of reported activity, these are recorded as four active days to avoid overcounting (Raleigh & Dowd, 2017). Recorded dates may be exact to the day, month, and year when available, or approximated within the month when the exact day is not known. A time precision variable is used to indicate the accuracy level of the temporal code. Events where sources cannot verify the date within a month of preciseness are not included in the data. Event location data includes the following: i) name of the specific location; ii) geographic coordinates (exact latitude and longitude) of the location; (iii) first, second and third level administrative zones that the specific location is found in, based on available and updated GIS based assignments; iv) the state in which the event occurred; and v) a spatial precision code denoting the level of accuracy of the location information (*see* Raleigh et al., 2010 for detailed information on the data). The availability of localized conflict data, specified by exact location and geographical coordinates, makes it possible to test whether rice cultivating grids are related to local conflict events. To account for the possibility that rice growing cultures may be related to different types of local conflict, a few measures of conflict are used/constructed.

I. **Event:** ACLED records nine types of conflict events: i) battle between two armed violent groups where no change of territory takes place; ii) battle where a non-state actor overtakes territory; iii) battle where government regains control of a territory; iv) establishment of headquarters or base by a non-state group; v) strategic development activities by rebel groups/militia/government where no active fighting takes place; vi) riots/protests; vii) violence against civilians; viii) non-violent transfer of territory; and ix) remote violence. Of these, five are incidences of local conflict including three types of battle events, riots/protests and violence against civilians. These events involve local inhabitants and are thus more likely to be related to the local norms, behaviors and cultures that may have been shaped by current and traditional local agricultural practices. Their definitions, along with their respective codes in the data, are summarized in Table 1. The remaining four include a small number of non-violent events that are not directly related to violence but can be used to capture critical events and strategic moves that may be relevant for wider conflict as well as events of remote violence. Since, the scope of this paper is restricted to local conflict that may plausibly be affected by cultural practices of local inhabitants in rice growing areas, these strategic and remote events are excluded from the analysis.

Table 4.1: Definition and distribution of conflict outcomes in the sample

Variable	Frequency	Std. Err.	[95% Conf. Interval]	
ACLED events used in this study				
E1 (battle between two armed violent groups where no change of territory takes place)	41	12.2777	16.93124	65.06876
E2 (battle where a non-state actor overtakes territory)	730	83.88857	565.5478	894.4522
E3 (battle where government regains control of a territory)	8	3.160837	1.803608	14.19639
E6 (riots/protests)	7	2.644412	1.815989	12.18401
E7 (violence against civilians)	29	8.992873	11.37069	46.62931
Additional conflict variables constructed				
Conflict incidence (probability of any conflict)	243	15.2669	213.0713	272.9287
Conflict intensity (total number of conflict events)	815	91.91505	634.813	995.187
Onset (probability that conflict begins in given grid in 2005)	127	11.149	105.1439	148.8561
Battle incidence (probability of battle)	238	15.11566	208.3678	267.6322
Battle intensity (total number of battle events)	779	87.1201	608.2129	949.7871

Source: Definition of events from ACLED codebook (Raleigh & Dowd, 2017). Figures based on author's calculation from ACLED data for year 2005.

II. **Conflict intensity:** Conflict intensity is defined as the total number of relevant, local conflict events that occurred in a given grid in 2005. This variable is the sum of the five conflict events (discussed above and denoted by Ej below) for each grid for the year 2005.

$$conflict\ intensity = \sum_{j=1,2,3,6,7} \mathbf{Ej}$$

III. **Conflict incidence:** Conflict incidence is an indicator variable that equals one if at least one of the five local conflict event is reported in the given grid in the year 2005, and zero otherwise.

$$conflict\ incidence = \begin{cases} 1, & if\ intensity > 0 \\ 0, & otherwise \end{cases}$$

IV. **Conflict onset:** Conflict onset is an indicator variable that equals one if a grid experienced at least one local conflict event in 2005 conditional upon having experienced no local conflict in the past year; and zero otherwise. Onset is an indicator that a new conflict started in the grid and may be important given that conflicts are often driven by past events (Mach et al., 2019).

$$conflict\ onset = \begin{cases} 1, & if\ incidence_{2005} > 0\ and\ incidence_{2004} = 0 \\ 0, & otherwise \end{cases}$$

V. **Battle intensity:** This is defined as the total number of battle events that occurred in a given grid. Battles are defined in ACLED as "a violent interaction between two politically organized armed groups at a particular time and location." Battle events correspond to the first three types of events (E1 through E3). Thus, battle intensity is defined as,

$$battle\ intensity = \sum_{j=1}^{3} \mathbf{Ej}$$

VI. **Battle incidence:** This is an indicator variable that equals one if at least one battle event is reported in the given grid in the year 2005, and zero otherwise.

$$battle\ incidence = \begin{cases} 1, & if\ battle\ intensity > 0 \\ 0, & otherwise \end{cases}$$

Table 1 above summarizes the definitions and descriptive statistics for all conflict outcomes. It should be noted that conflict is a low-probability event. The probability of observing conflict in a given grid in 2005 is around 9 percent, half of which (4.2 percent) is a new conflict while the remaining comprise a continuation of an existing conflict.

4.3.2 Crops and irrigation

Agricultural data is obtained from the Spatial Production and Allocation Model (SPAM) 2005 version 1.0. This data was developed by the Food and Agriculture Organization (FAO), International Food Policy Research Institute (IFPRI), and the Center for Sustainability and the Global Environment (SAGE) at the University of Wisconsin-Madison. SPAM 2005 uses national and sub-national level crop statistics (averaged and standardized to 2004-2006 FAO country levels) for 42 crops and disaggregates them to produce estimates on physical area, harvested area, production and yield at 5 arc-minute pixels for the year 2005.

SPAM employs a "cross-entropy optimization" approach which essentially combines data from different sources available at various spatial levels and uses an optimization model to disaggregate the data and generate results at a 5 arc-minute grid cell level. National and sub-national crop statistics are averaged and standardized around a three-year average. For example, SPAM 2005 averages crop statistics to 2004-2006 at FAO country levels and disaggregated by production systems (refer to the MapSPAM website for more details or see You & Wood, 2005 and Wood-Sichra, Joglekar & You, 2016). The model uses a variety of inputs such as production statistics, farming systems, satellite image, crop biophysical suitability and potential yields, cropland surface, location of irrigated areas, rural population densities, and crop prices. To distinguish between resource used by each crop (such as seeds, water, fertilizer, pesticides, labor, and machinery) SPAM differentiates four production systems in its estimates. These production systems are differentiated by their water use (irrigated vs rainfed) as well as other input use (high input, low input, and subsistence). This leads to four distinct production systems: irrigated-high input; rainfed-high input; rainfed-low input; and rainfed-subsistence production. The production systems are defined as follows:

i) irrigated-high input production refers to crop area that is equipped with either full or partial control irrigation[14], and that usually uses high level of inputs (such as modern seed varieties, fertilizers) as well as advanced management systems (such as soil/water conservation measures)

ii) rainfed-high input/commercial production refers to cropland that uses rainfed-based agriculture, high-yield varieties and some animal traction and mechanization, applies some fertilizer, chemical pesticides, disease or weed controls and is mostly produced for the market

iii) rainfed, low-input production refers to rainfed crop production, mostly for own consumption, and uses traditional varieties and mainly manual labor without (or with little) application of nutrients or chemicals for pest and disease control

iv) rainfed, low-input/subsistence production refers to crop productions by small-scale farmers under rainfed conditions with low inputs and for own consumption regardless of whether or not the cropland is suitable for cultivation

The final output generated by the model results in predicted estimates on i) physical area; ii) harvested area; iii) production quantity; and iv) yield, categorized by the four production systems of each of the 42 crops (refer to Figure 5 in the Appendix for a schematic diagram of the primary data generation process of SPAM). For example, the data can give us estimates of what proportion of the total rice harvested in a grid was irrigated and how that compares with some

[14] "In areas where farmers' indigenous techniques used recessional (decrue) cultivation following an annual flood, French engineers devised polders and dikes to assist in retaining the river's water. Since such systems do not control the ultimate supply (which may fail in dry years) they have been termed partial control systems in contrast to full control irrigation" (Moris, 2019).

other crop; or how the yield of rice that was rainfed using high inputs compares with rice that was rainfed using low inputs, and so on.

This dataset has a few advantages that make it suitable for this study. First, spatial representation of agricultural production with specific geographical coordinates can reveal geographic patterns that would otherwise be lost in larger mapping units. The granular grid level data on crop distribution and production systems thus allows for a much more fine-grained analysis than national crop statistics could provide. Second, even when national agricultural surveys are available, production data is typically not represented by production system (i.e., irrigated versus rainfed water, high input versus low input levels). This type of representation can allow researchers to determine the specific effects of irrigation or input levels on outcomes of interest. Finally, although the original data is available at a 5 arc-minute resolution (approximately 10X10 km around the equator), the SPAM dataset is designed such that it can be run at higher resolutions (see Wood-Sichra, Joglekar & You, 2016). For the purposes of this study, the SPAM data is resampled at 0.5X0.5 degree resolution (approximately 55X55 km grids around the equator), with the help of a GIS expert, in order to match the crop data to the resolution of the ACLED conflict data.

For each grid, the crop occupying the largest harvested area is identified and used to construct a categorical variable, **Crop**, to denote the main or dominant crop cultivated in the grid. This variable includes the categories rice, wheat, maize, pearl millet, sorghum, cassava, and "all other crops". The top six cereal crops, expressed as the percentage of total grids for it is the dominant crop, include maize (25.5%), cassava (16.5%), sorghum (15.6%), pearl millet (9.7%), rice (6.26%), and wheat (1.75%). Thirty-six other crops make up the dominant crop for the remaining 25% grids. These are classified as "all other crops" in subsequent analyses. Figure 1 below provides a visual representation of the distribution of the main crops of each grid across

SSA. While maize is the most important staple crop for Africa, rice is an important staple in eastern and western Africa (FAO, 2016).

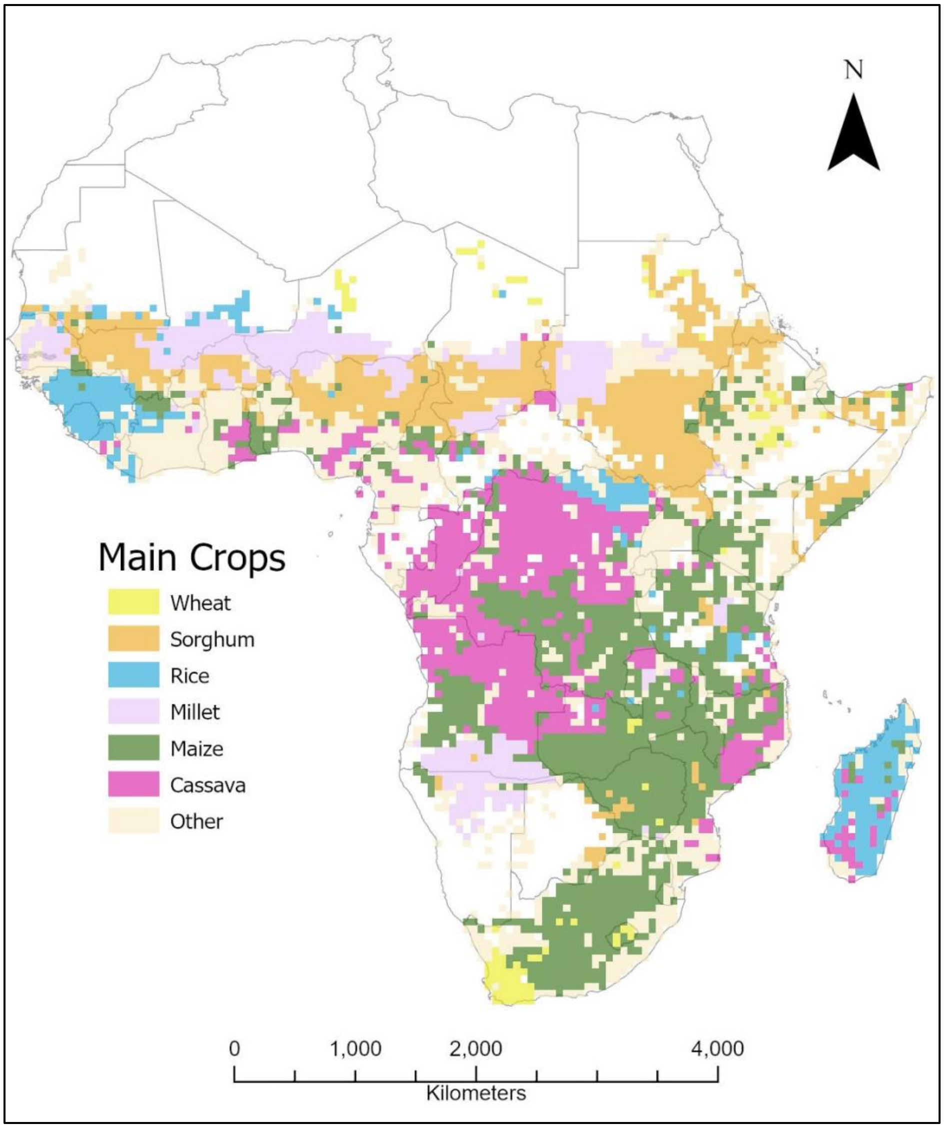


Figure 4.1: Distribution of main crops across sub-Saharan Africa, 2005

Source: Author's own representation based on ACLED and SPAM data

Along with this figure, Table 2 below shows the distribution of the six main crops in the data, both for SSA overall as well as by each of its four sub-regions - east, west, central, and south. Regions are classified as follows. East Africa region includes the countries Burundi, Comoros, Djibouti, Eritrea, Ethiopia, Kenya, Madagascar, Malawi, Mauritius, Rwanda, Seychelles, Somalia, United Republic of Tanzania, and Uganda. West Africa region includes Benin, Burkina Faso, Cape Verde, Ivory Coast, Gambia, Ghana, Guinea, Guinea-Bissau, Liberia, Mali, Mauritania, Niger, Nigeria, Senegal, Sierra Leone, and Togo. Central Africa includes Cameroon, Central African Republic, Chad, Congo, Democratic Republic of the Congo, Equatorial Guinea, and Gabon. Finally, South Africa includes the countries Angola, Botswana, Lesotho, Mozambique, Namibia, South Africa, Swaziland, Zambia, and Zimbabwe. Note that two other countries in the data, Sudan and Western Sahara, which officially belong to North Africa are excluded from this analysis. The Table shows that even at present most of the rice in SSA in grown in the west. Figure 1 above helps identify the exact countries where rice is the dominant crop. Rice appears to be the main crop throughout Madagascar which is the major contributor of rice production in the east. Rice is also common in parts of West Africa in countries such as Guinea, Ivory Coast, Liberia, Mali, Mauritania, Niger, Nigeria, Senegal, and Sierra Leone.

Table 4.2: Distribution of rice and other cereal crops in the sample, by region

Crop	SSA (%)	EA (%)	WA (%)	CA (%)	SA (%)
Rice	6.26	9.21	12.98	3.85	0.18
Wheat	1.75	1.73	0.8	0.81	3.28
Maize	25.2	20.39	6.89	22.49	46.41
Pearl millet	9.68	5.74	21.96	3.77	9.31
Sorghum	15.57	28.53	22.6	9.99	1.76
Cassava	16.53	3.29	4.25	40.38	19.65
Other crops	25	31.1	30.53	18.71	19.4

Next, Figure 4.2 shows the distribution of conflict across grids and overlaps with the main crop for each grid. The first noteworthy feature is that areas predominantly cultivating rice such as Madagascar, parts of West Africa, and a few scattered grids in central-eastern Africa show very few instances of conflict in 2005. In rice growing areas that experience conflict, the number of events is at most ten. For example, the only two grids in the island of Madagascar experiencing any conflict show ten or fewer events in 2005. The West African nations of Guinea and Sierra Leone with five and two grids that experienced conflict respectively, each had ten or less conflict events. Although the Democratic Republic of the Congo in Central Africa experienced nearly sixty conflict events in 2005 (based on the data), the map below shows that only five of these occurred in rice producing grids while all remaining events occurred in grids that predominantly produce maize or cassava. Other major conflict hubs in 2005 that experienced over a hundred conflict events were Sudan in central-northern Africa, where the dominant crop was pearl millet; Zimbabwe in South Africa, where the major crop was maize; and Somalia in East Africa, where the major crop was sorghum.

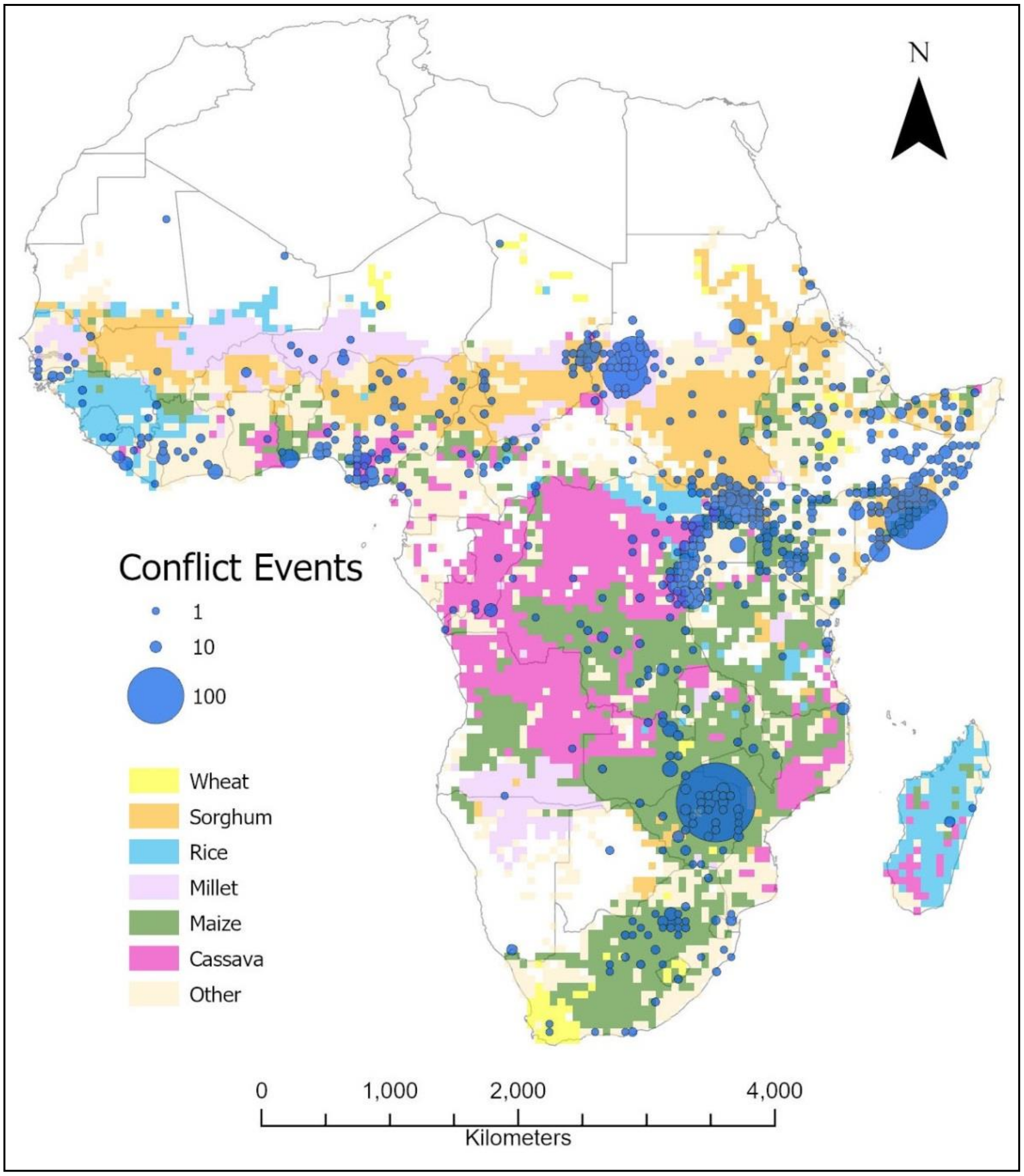


Figure 4.2: Distribution of conflict events and crops across SSA, 2005

Source: Author's own based on ACLED and SPAM data

Next, an examination of the overlap between irrigation technology and conflict across the region is displayed in Figure 4.3 below. The figure shows that areas with the highest fraction of land equipped for irrigation such as Madagascar, South Africa and parts of West Africa overlap with areas of very little conflict. Conversely, areas with the highest concentration of conflict seem to have little to no land equipped for irrigation. While encouraging, it is pertinent to then examine whether rice growing areas are in fact the ones that are more equipped with irrigation technology.

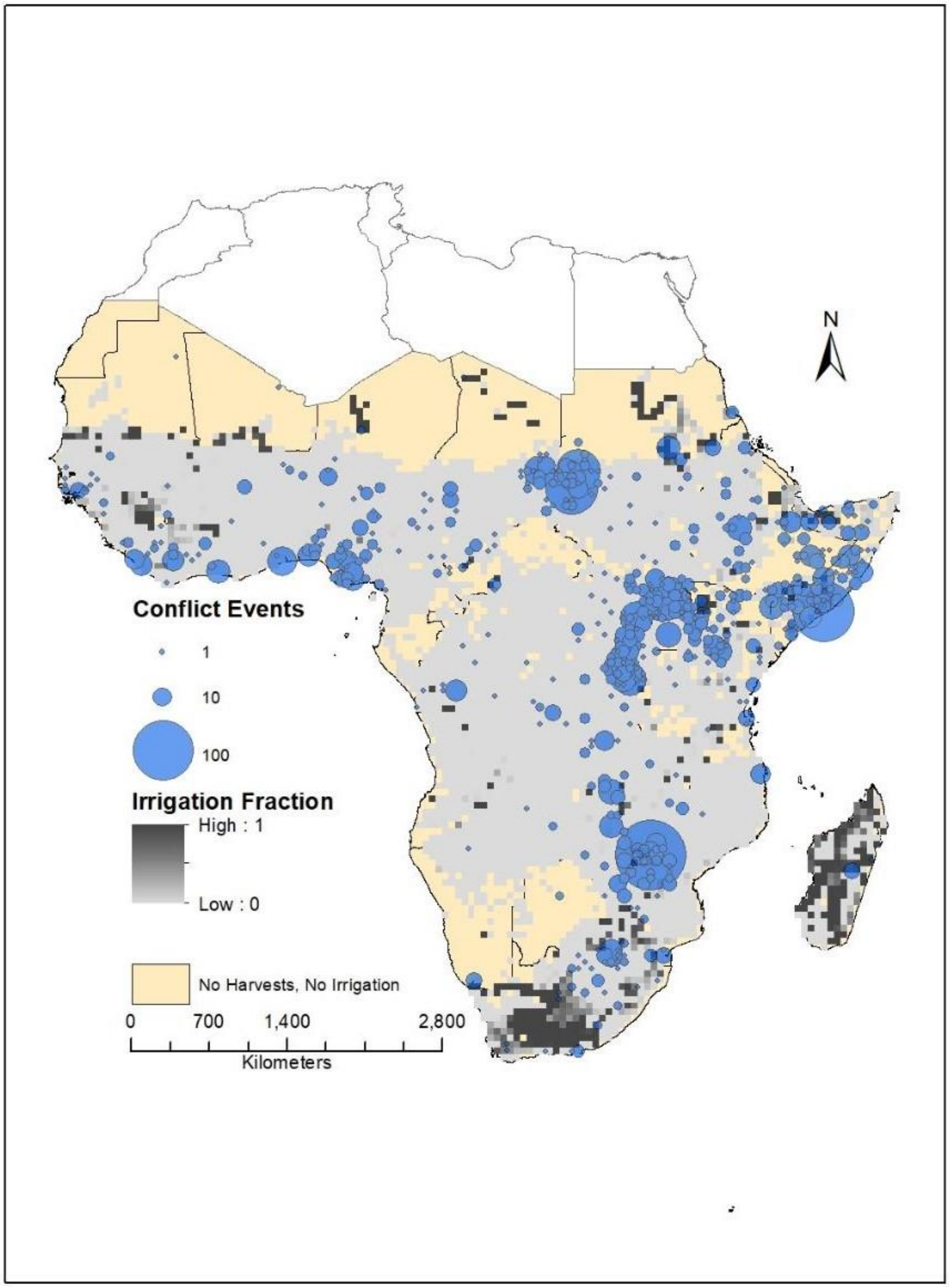


Figure 4.3: Distribution of conflict events and fraction of area equipped with irrigation across SSA, 2005

Source: Author's own based on ACLED and SPAM data

Figure 4.4 below shows the distribution of irrigation technology by region and by main crops. Two features stand out. First, very little of the agricultural land in SSA is equipped with irrigation technology. In fact, a quick analysis of the data shows that only 7.6% of the total harvested area in the sample is equipped for irrigation. For reference, around 6% of the continent of Africa is equipped with irrigation technology (You et al., 2011). Most of the irrigation efforts in the sample are concentrated in Madagascar and South Africa. This is consistent with the finding mentioned in section 4.2.3 that Madagascar, Nigeria, South Africa, and Sudan account for the vast majority of irrigated land in Africa (Burney& Naylor, 2012). Second, with the exception of Madagascar, rice growing regions are not necessarily equipped with irrigation technology at all. This is also consistent with the statistics mentioned earlier that only about a-sixth of the rice in Africa is produced using irrigation with the majority of that being in Madagascar (NRC, 1996). The use and adoption of irrigation technology in rice farming in Africa is in stark contrast to that of Asia where irrigation is the main source of water. This suggests that the pattern of low overlap between areas of conflict and irrigated regions shown in Figure 4.3 may not be due to water coordination efforts associated with rice cultivation practices in SSA.

While these figures provide a preliminary analysis and useful insight, the next section validates these relations with formal analyses. All Figures are revisited in the Results and Discussion section.

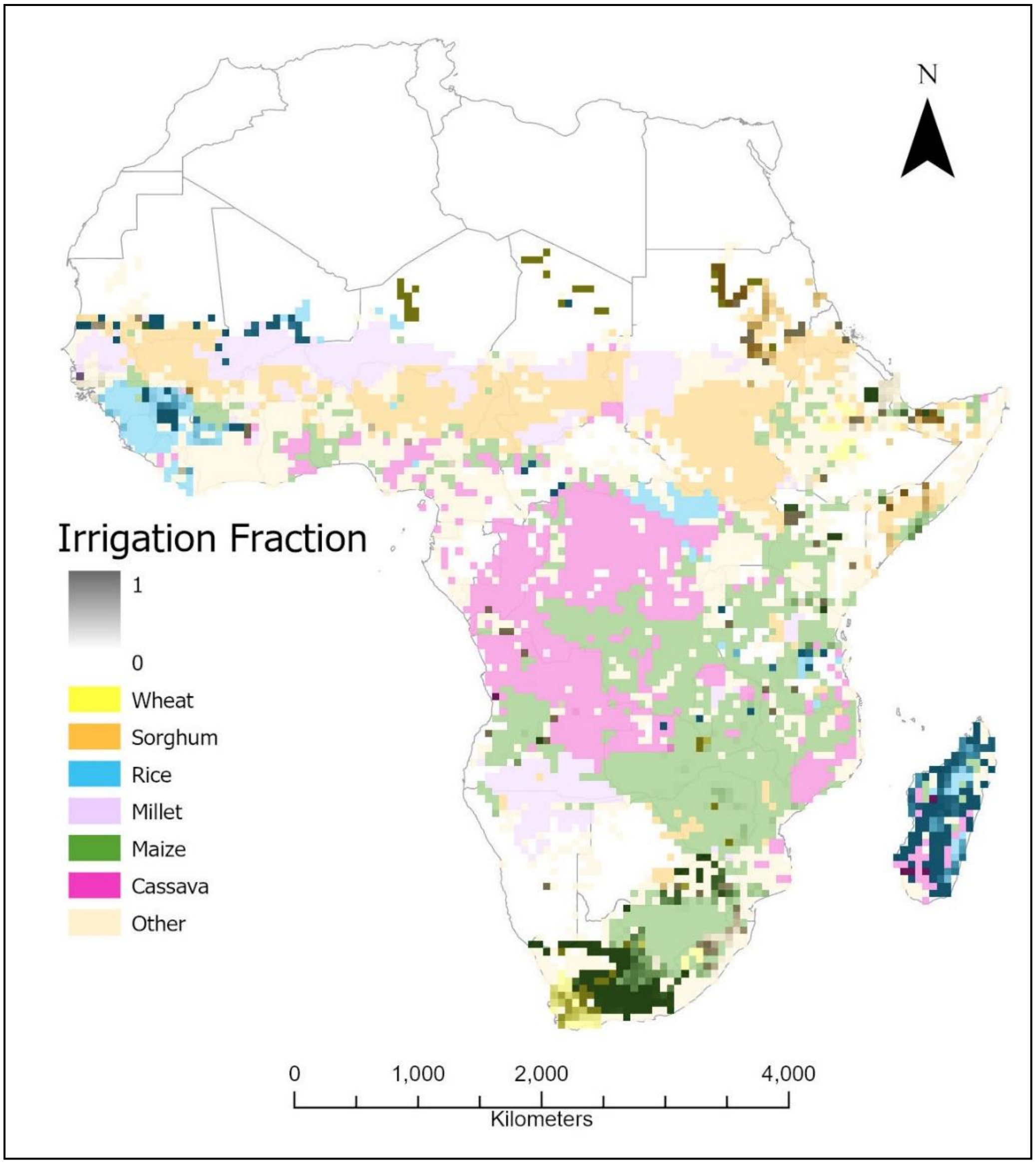


Figure 4.4: Distribution of share of land using irrigation and high input technology across crops, 2005

Source: Author's own representations based on ACLED and SPAM data

4.3.3 Control variables

I choose control variables for the study based on a thorough review of the literature. I include country; population density and nightlight density; ethnolinguistic and religious fractionalization index; indicators for democracy; distance to capital city, ports, and coast; ruggedness of terrain, and elevation level; natural resources such as industrial metals, diamond, and petroleum; agricultural controls; and past conflict as control variables. Since paddy rice is known to be affected by temperature and require significant amount of water (Talhelm et al., 2014), the controls also include malaria index to capture temperature, precipitation and weather shocks and the presence of rivers while the absolute latitude of the grid controls for different agro-ecological zones. Where data is not available from the SPAM dataset, I use data (at a resolution of 0.5X0.5 degree) from Zhang & Kibriya (2016) for control variables. The following sources are used: natural resources (precious metals, industrial metals, and diamonds) are drawn from the Mineral Resource Data System (U.S. Geological Survey, 1996) by the United States Geological Survey (USGS) and the PRIO Diamond Resources dataset (Gilmore, 2005); land data, including average elevation, and terrain ruggedness, is collected from the GTOPO30 global digital elevation model (U.S. Geological Survey, 1996); national income (GDP per capita) and democracy levels and indicators (Polity IV score) are obtained from Fearon & Laitin (2003); ethnolinguistic fractionalization index and religious fractionalization index are obtained from the Atlas Maradov Mira and CIA factbook respectively; and weather shock data is obtained from the Standardised Precipitation-Evapotranspiration Index (SPEI). Following Harari & La Ferrara (2018), weather shock is calculated as the fraction of months that have at best one standard deviation lower SPEI value than the sample period average.

4.4 Estimation Strategy

I first estimate the following equation,

$$y_i = \beta_0 + \beta_1 R_i + \boldsymbol{Z}_{\boldsymbol{i}}' \boldsymbol{\beta}_{\boldsymbol{2}} + \varepsilon_i \tag{1}$$

where y_i denotes one of the 10 measures of conflict for grid *i* in 2005 including conflict incidence, conflict intensity, conflict onset, battle incidence, battle intensity and the five conflict events specified in section 3.1. R_i is an indicator variable that equals 1 if rice is the main crop of grid i, and 0 otherwise. $\boldsymbol{Z}_{\boldsymbol{i}}'$ is a vector of grid level economic and geographic covariates including country fixed effects.

Hypothesis I. Rice cultivating units (grids) show lower prevalence of conflict on average.

H_0: $\beta_1 = 0$

H_A: $\beta_1 < 0$

Next, to test the effect of a history of rice farming on contemporary conflict, I estimate the following equation,

$$y_i = \beta_0 + \beta_1 R_i + \beta_2 H_i + \beta_3 (R_i\ X\ H_i) + \boldsymbol{Z}_{\boldsymbol{i}}' \beta_4 + \varepsilon_i \tag{2}$$

where y_i denotes one of the ten measures of conflict for grid i in 2005 and R_i is an indicator variable that equals 1 if rice is the main crop of grid i, and 0 otherwise. H_i is an indicator variable that equals 1 if the grid belongs to a country with a history of rice cultivation (of over a millennium) and 0 if not. $\boldsymbol{Z}_{\boldsymbol{i}}'$ is a vector of control variables. Thus, β_0 denotes the average present-day conflict in non-rice producing grids in regions without a history of rice farming. β_1 is the additional effect of rice cultivation on present day conflict in regions without a history of rice farming; or the mean difference in conflict between rice and non-rice producing areas in regions without long history of rice farming. This may be called the *crop effect.* β_2 shows the additional effect of other crops on conflict in regions with long history of rice farming or the average difference in conflict between non-rice producing grids in regions with and without long history of rice farming. I refer to this as the *history effect.* The main coefficient of interest, β_3, or the *interaction effect* signifies the mean difference in conflict between rice and non-rice producing grids in areas with and without a prior history of rice culture. Equation (2) essentially compares average conflict outcome under four scenarios: i) contemporary conflict in present-day rice farming units belonging to regions with history of rice cultivation; ii) contemporary conflict in present-day rice farming units belonging to regions without a prior history of rice; iii) contemporary conflict in present-day non-rice farming units belonging to regions that planted rice for centuries but have recently switched to other crops; and iv) contemporary conflict in present-day non-rice farming units from regions that did not traditionally farm rice. Thus, the interaction effect shows the average conflict in historically rice cultivating units that continue to grow rice at present compared to all other farming units of SSA. Since a culture of cooperation evolves over time, a history of rice cultivation should foster a stronger culture of peace and cooperation. Thus, grids with a history of rice farming that continue to farm rice at present are expected to be correlated with lower present-day conflict on average, resulting in an expected negative value of β_3.

Hypothesis II. Units (grids) with a history of rice farming that continue to farm rice at present show lower mean prevalence of present-day conflict.

H_0: $\beta_3 = 0$

H_A: $\beta_3 < 0$

Depending on whether the outcome is a count variable or a binary variable, estimations respectively use an Ordinary Least Squares (OLS) or a Linear Probability Model (LPM). Following studies of conflict risk, I choose the linear probability model over other binary response models for ease of interpretation (*for example,* Berman & Couttenier, 2012; Couttenier & Soubeyran, 2013; Harari & La Ferrara, 2018). The main advantage of the LPM is that the parameter estimates can be directly interpreted as the "mean marginal effect" of covariates on outcome thus making it a convenient approximation to the underlying response probability (Greene, 1993; Woolridge, 2002). To account for possible heteroskedasticity, all estimations report robust standard errors clustered at the grid level.

4.5 Results and Discussion

4.5.1 Comparative analysis of conflict in rice and non-rice growing regions

Table 4.3 shows the estimates of equation (1) or the difference in mean conflict between rice and non-rice producing grids of SSA. Results in Panel A shows that rice producing grids typically experience lower conflict than grids producing other crops. For example, grids that cultivate rice as the main crop have approximately 3 percentage-point lower risk of conflict than grids that cultivate some other crop as its main crop. All results shown include country fixed effects. Given that the average grid suffers only 8.8 percent risk of conflict, the results show an effect size of over 60 percent reduction in conflict risk in rice producing grids. Similarly, given that on average a grid experienced only 0.423 conflict events in 2005, the negative coefficient of 0.084 translates to about an 80 percent decline in the intensity of conflict experienced by grids cultivating rice versus other crops. The chance of conflict onset in a grid that did not experience conflict in the previous year is about 2.8 percentage point lower when the grid produces rice as the main crop. Turning to battle events, the effect size for battle incidence and intensity translate to reductions of about 17.5 and 34 percent respectively. Irrigation results will be discussed in Section 5.3. For further insight, Panel B disaggregates "all other crops" into the major staple crops of SSA. Thus, the results in Panel B compare the average conflict in rice producing grids with those growing one of the other major staple crops. The results show that for all cereal crops and all specifications of conflict, no grid shows lower conflict than grids where the main crop is rice. The results from the Table appear to reflect the visual representation of Figure 2 (in Section 3) which showed only a few overlaps between rice producing grids and conflict events.

Table 4.3: Conflict across grids growing rice and other crops

VARIABLES	(1) Conflict incidence	(2) Conflict intensity	(3) Conflict onset	(4) Battle incidence	(5) Battle intensity	(6) Incidence of Battle events where non-state actor overtakes territory
Panel A: Rice growing areas compared to areas growing any other crop						
Rice	-0.034**	-0.084*	-0.028**	-0.033**	-0.086**	-0.091***
	(0.0156)	(0.0440)	(0.0134)	(0.0156)	(0.0426)	(0.0306)
Share of irrigated land	0.006	-0.001	0.008	0.006	-0.005	0.003
	(0.0069)	(0.0195)	(0.0062)	(0.0068)	(0.0182)	(0.0141)
Observations	4,098	4,098	4,098	4,098	4,098	4,098
R-squared	0.172	0.119	0.074	0.172	0.121	0.120
Panel B: Rice growing areas compared to areas growing other major staple crops						
Wheat	0.003	-0.063	0.004	0.004	-0.054	-0.031
	(0.0171)	(0.0542)	(0.0146)	(0.0171)	(0.0525)	(0.0435)
Maize	0.026*	0.024	0.017	0.026*	0.028	0.046
	(0.0154)	(0.0402)	(0.0132)	(0.0154)	(0.0392)	(0.0282)
Pearl millet	0.042**	0.105*	0.018	0.041**	0.105*	0.107**
	(0.0184)	(0.0611)	(0.0156)	(0.0184)	(0.0597)	(0.0514)
Sorghum	0.055**	0.174	0.020	0.054**	0.171	0.183
	(0.0240)	(0.1684)	(0.0191)	(0.0240)	(0.1673)	(0.1648)
Cassava	0.030*	0.107*	0.017	0.029*	0.107*	0.108**
	(0.0161)	(0.0611)	(0.0138)	(0.0160)	(0.0585)	(0.0449)
All other crops	0.046***	0.132**	0.029**	0.046***	0.134**	0.126***
	(0.0174)	(0.0551)	(0.0148)	(0.0174)	(0.0542)	(0.0434)
Share of irrigated land	0.012	0.031	0.010	0.012	0.027	0.028
	(0.0073)	(0.0228)	(0.0063)	(0.0073)	(0.0216)	(0.0176)
Observations	4,098	4,098	4,098	4,098	4,098	4,098
R-squared	0.154	0.074	0.073	0.175	0.122	0.121
Controls	yes	yes	yes	yes	yes	yes
Country fixed effects	yes	yes	yes	yes	yes	yes

Note: *** p<0.01, ** p<0.05, * p<0.1. Grid-clustered standard errors in parentheses. The base category in all regressions in Panel A is “all other crops”. The base category in Panel B is rice. Robustness checks included the same regressions without irrigation controls, but the results remain unchanged. Other specifications included the other relevant conflict events (E1, E3, E6 and E7) from ACLED but the coefficients are not statistically significant for any of those individual events, possibly due to the small number of events in the data.

4.5.2 Comparative analysis of conflict in regions with differential history of rice

As previously discussed, West Africa has a long history of rice cultivation dating back over thousands of years (Linares, 2000; NRC, 1996), especially in seven countries: Ivory Coast, Liberia, Sierra Leone, Guinea, Guinea-Bissau, Gambia, and Senegal. 1,248 observations, about a sixth of the total, belong to one of the sixteen countries of West Africa included in the sample. Of these, 329 observations belong to one of these seven countries. Table 4 below shows the results of estimating equation (2) using two alternate measures for this millennial-old history of rice farming. In Panel A, history is measured as an indicator for whether the grid belongs to a West African country while in Panel B it is measured by an indicator for whether the grid belongs to one of the seven afore-mentioned rice growing West African nations. The results in both panels are similar, differing only slightly in magnitude of coefficients.

The coefficients in the Table provide several insights. First, as in the previous table, Rice producing grids show lower conflict on average than grids producing other crops in regions without long history of rice growing culture. This result holds for all specifications of conflict, except for the risk of conflict onset which is similar in rice and non-rice producing grids. Second, it appears that grids cultivating non-rice cereal crops have a higher risk of average conflict when grown in areas exposed to a history of rice farming. For example, grids growing crops other than rice in traditionally rice growing nations of West Africa have 6.6 percentage point higher conflict risk than when grown in areas with no prior history of rice farming. This contradicts expected results since a peaceful culture that evolved over time should have reflected in lower conflict irrespective of present-day crop choice and will be discussed shortly. Table 4 shows that rice producing grids in regions with a longer history of rice farming culture are correlated with additional reduction in conflict risk compared to grids producing other crops, as well as grids producing rice in regions without a history of rice farming. For example, grids with long history of rice farming that also grow rice at present show roughly between 3 and 6 percentage point (from panels A and B respectively) lower conflict risk when compared to grids that produce other crops as well as rice producing grids without a history of rice farming. Reconciling the results from the history and interaction effects in the Table, it appears that communities pursuing rice farming in traditionally rice grown regions experience lower conflict; but shifting away from rice in traditionally rice growing regions leads to conflict. A few explanations are possible. First, this conflict may have been driven by the shift in agriculture. The decision to surrender the crop of their forefathers, a crop that predominated the region for centuries, may have been followed by communal disagreements and conflict. It is highly unlikely that such a crucial decision in an agricultural community would be unanimous. Second, positive behavioral impacts of cooperative rice culture may have been negated by adverse effects of present-day, individualistic farming practices associated with other crops. It was previously discussed in Section 4.2 that farming crops such as wheat, maize or millet is different from rice. Moreover, fierce competition over scarce resources could have escalated to greater conflict in the absence of cooperative activities that communities engaged in through rice farming. Third, it is possible that cultural traits such as cooperation, harmony and conflict aversion wane over time unless continuously fostered which reflects in lower conflict when the community continues rice production but increases otherwise. This explanation fits the traditional, Neo-Classical growth model prediction that any external shock such as a transition in farming should only cause a temporary movement in the steady-state equilibrium. This would and not cause permanent changes as argued by scholars. However, an analysis of conflict in regions that produce crops other than rice is beyond the scope of this paper at this point but is certainly worth exploring in future endeavors.

Going back to the main estimation of this study, the coefficient of the interaction effect from equation (2), the results validate the preliminary visual probe in Figure 2 (in section 3) which shows little overlap between conflict events and rice producing grids in West African countries. Only eight rice producing grids belonging to the West African countries with a prior history of rice production, namely Guinea, Sierra-Leone and Liberia, have experienced present-day conflict. With the exception of one grid in Liberia which has had up to ten conflict events, the remaining have had one conflict at best in 2005. The map further reveals that within the seven historic rice producing nations of West Africa, the highest concentration of present-day conflict occurs within grids growing non-cereal crops at present. Thus, the main findings of the paper are reflected in the maps.

Table 4.4 provides further insight that while present day rice farming in areas with historical history of rice culture reduces the incidence of conflict, it does not necessarily reduce the intensity of conflict. This result appears plausible – history of a peaceful culture can shape community behavior and reduce the chance of communities engaging in conflict but may not be able to determine the number of conflict events experienced. Finally, it should be noted that though there is no significant difference in the risk of conflict onset between rice and non-rice producing grids in areas not exposed to rice cultivation history, long history of rice culture reduces the risk of conflict onset in rice producing grids.

Thus, based on the empirical evidence found in the paper, it may be concluded that historically rice cultivating regions that continue to grow rice as its main crop experience lower conflict risk than all other regions in SSA. This implies that the longer the history of a culture of rice cultivation, the greater the decline in conflict in rice growing areas. This supports the hypothesis that a culture of cooperation fostered by rice farming practices is not only associated with lower risk of conflict, but the reduction is further reinforced by longer history of such cultural practices. Building on the theory that inherent cooperative agricultural practices within the process of rice cultivation develops into a culture of peace over time, it is not surprising that the longer such practices prevail the stronger the peace promoting values which are in turn reflected in a decline in conflict.

Table 4.4: Conflict analysis in regions with and without rice history

VARIABLES	(1) Conflict incidence	(2) Conflict intensity	(3) Conflict onset	(4) Battle incidence	(5) Battle intensity	(6) Incidence of Battle events where non-state actor overtakes territory
Panel A: History measured as dummy for West Africa						
Rice	-0.031**	-0.211***	-0.002	-0.030**	-0.213***	-0.036***
	(0.0137)	(0.0696)	(0.0111)	(0.0137)	(0.0682)	(0.0112)
History (WA)	0.065***	0.124*	0.038***	0.064***	0.118*	0.056***
	(0.0144)	(0.0717)	(0.0115)	(0.0143)	(0.0699)	(0.0138)
Rice X History	-0.033**	0.027	-0.031**	-0.034**	0.029	-0.028*
	(0.0159)	(0.0520)	(0.0130)	(0.0159)	(0.0503)	(0.0143)
Irrigation	0.013*	0.057**	0.006	0.012*	0.052**	0.014**
	(0.0066)	(0.0257)	(0.0054)	(0.0066)	(0.0248)	(0.0061)
Observations	4,098	4,098	4,098	4,098	4,098	4,098
R-squared	0.175	0.119	0.038	0.175	0.121	0.170
Panel B: History measured by a dummy for rice growing nations of West Africa						
Rice	-0.028**	-0.204***	-0.001	-0.028**	-0.206***	-0.033***
	(0.0134)	(0.0688)	(0.0107)	(0.0133)	(0.0676)	(0.0109)
History (rice nations of WA)	0.066***	0.099	0.039**	0.065***	0.095	0.063***
	(0.0211)	(0.0700)	(0.0168)	(0.0211)	(0.0694)	(0.0211)
Rice X History	-0.055**	-0.002	-0.043**	-0.056**	0.000	-0.052**
	(0.0224)	(0.0663)	(0.0179)	(0.0224)	(0.0652)	(0.0213)
Irrigation	0.008	0.048*	0.004	0.008	0.044*	0.010*
	(0.0066)	(0.0259)	(0.0055)	(0.0066)	(0.0250)	(0.0061)
Observations	4,098	4,098	4,098	4,098	4,098	4,098
R-squared	0.079	0.031	0.020	0.079	0.031	0.076

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All specifications include controls. Robustness checks with different controls can be found in the Appendix. Since history is a region dummy, country fixed effects are excluded from all specifications shown. However, robustness checks with country fixed effects can be found in the Appendix. Including country fixed effects increases the magnitude of all coefficients slightly but the statistical significance remains the same as this table.

4.5.3 Relationship between irrigation, rice, and conflict

Contrary to the preliminary visual patterns observed in Figure 4.3, results from Tables 4.3 and 4.4 above show that irrigation is not negatively correlated with conflict. While Figure 4.3 indicated that areas such as Madagascar and parts of West and South Africa with the highest concentration of land equipped for irrigation showed little or no overlap with conflict, Table 4.4 shows that irrigation efforts are either uncorrelated or in some cases even positively correlated with conflict outcomes. This result holds in robustness checks (shown in the Appendix) with alternate measures of irrigation. Excluding irrigation controls from the specifications do not alter the main results either. In addition, Table 4.5 below shows that irrigation of staple crops, particularly rice, is correlated with higher conflict intensity across SSA. Given the limited availability of land equipped for irrigation in Africa, it is possible that irrigation leads to competition over the scarce resource thus increasing social friction and conflict. However, it should also be noted that a very small percentage of the grids are equipped for irrigation, as shown earlier in Figures 4.3 and 4.4, and an even smaller percentage of grids experience any conflict. Thus, the low number of data points make this result unreliable. Furthermore, a test for the joint significance of the indicator (irrigation) and interaction variables (crops) proves to be statistically insignificant.

Table 4.5: Irrigated crops and conflict

VARIABLES	(1) Incidence	(2) Intensity	(3) Onset	(4) Battle incidence	(5) Battle intensity	(6) E2
Rice	-0.052**	-0.187**	-0.039**	-0.051**	-0.184**	-0.172***
	(0.0209)	(0.0825)	(0.0176)	(0.0209)	(0.0816)	(0.0617)
Wheat	-0.060***	-0.298***	-0.039**	-0.059***	-0.289***	-0.249***
	(0.0179)	(0.0868)	(0.0172)	(0.0178)	(0.0847)	(0.0794)
Maize	-0.018*	-0.109**	-0.008	-0.017*	-0.107**	-0.078*
	(0.0090)	(0.0510)	(0.0082)	(0.0090)	(0.0509)	(0.0465)
Pearl millet	-0.005	-0.031	-0.003	-0.005	-0.034	-0.022
	(0.0114)	(0.0582)	(0.0108)	(0.0114)	(0.0579)	(0.0555)
Sorghum	0.010	0.045	0.000	0.010	0.040	0.061
	(0.0208)	(0.1884)	(0.0184)	(0.0208)	(0.1875)	(0.1872)
Cassava	-0.016	-0.032	-0.006	-0.016	-0.035	-0.023
	(0.0104)	(0.0564)	(0.0097)	(0.0104)	(0.0554)	(0.0486)
Irrigation	-0.009	-0.074*	0.001	-0.009	-0.072*	-0.049
	(0.0172)	(0.0411)	(0.0166)	(0.0172)	(0.0408)	(0.0370)
Rice X Irrigation	0.028	0.177**	0.016	0.026	0.163*	0.139**
	(0.0244)	(0.0856)	(0.0224)	(0.0244)	(0.0836)	(0.0658)
Wheat X Irrigation	0.028	0.177***	0.013	0.027	0.172***	0.142**
	(0.0232)	(0.0682)	(0.0224)	(0.0232)	(0.0664)	(0.0623)
Maize X Irrigation	-0.000	0.060	-0.008	-0.001	0.058	0.032
	(0.0184)	(0.0515)	(0.0177)	(0.0184)	(0.0511)	(0.0467)
Pearl millet X Irrigation	-0.006	0.100	-0.013	-0.007	0.091	0.048
	(0.0338)	(0.0778)	(0.0284)	(0.0338)	(0.0784)	(0.0795)
Sorghum X Irrigation	-0.016	-0.009	-0.015	-0.016	-0.008	-0.035
	(0.0262)	(0.1603)	(0.0243)	(0.0262)	(0.1595)	(0.1581)
Cassava X Irrigation	-0.000	0.091*	-0.016	0.000	0.092*	0.072
	(0.0214)	(0.0543)	(0.0208)	(0.0215)	(0.0541)	(0.0497)
Observations	4,098	4,098	4,098	4,098	4,098	4,098
R-squared	0.175	0.121	0.075	0.175	0.122	0.121

Note: *** p<0.01, ** p<0.05, * p<0.1. Grid-clustered standard errors in parentheses. In all specifications in this table irrigation is a dummy variable which equals one if the crop received any irrigation; and zero otherwise. However, the joint significance of the set of indicator and interaction variable is 0.446 and thus not statistically significant.

The results for the correlation between irrigated crops and conflict is a significant departure from studies linking irrigation to more peaceful cooperation in China (Talhelm et al., 2014), the Philippines (Tsusaka et al., 2015), and India (Von Carnap, 2017). However, low rates of adoption of irrigation technologies remain a big difference between rice produced in Africa with that of Asia. As previously discussed in Section 4.2 and shown in Figure 4.4, only a small fraction of the agricultural land in SSA is equipped for irrigation, and except for Madagascar, rice growing regions of Africa are not necessarily irrigated. While the potential for irrigation to increase crop production and improve livelihoods is nationally and internationally recognized (Burney & Naylor, 2012; Burney, Naylor & Postel, 2013; You et al., 2011), the failure of irrigation projects induced as part of structural adjustment programs in the 1970s and 1980s have discouraged irrigation efforts to continue (Moris, 2019; Moseley et al., 2010). Since much of the blame for this failure was directed at the formal structure of the irrigation schemes that were mostly controlled by the government, the development paradigm shifted and were geared towards a bottom-up or grassroots approach since the later part of 1980's. Subsequently, irrigation control shifted to the informal sector in the form of small-scale or smallholder irrigation.

The low rates of irrigation in SSA and the data hold two implications for the results. First, since irrigation is one of the channels through which rice farming fosters cooperation, the low adoption of irrigation may have resulted in a smaller effect size of rice farming on conflict in SSA than one would expect for a similar scenario in Asia. This is reassuring and makes a stronger case for rice farming and its associated negative impact on conflict. Second, it appears that the negative correlations between rice and conflict are more likely to have been driven by other cooperative practices of rice farming (discussed in Section 4.2) such as forming labor exchange cooperatives, and coordinating over seed transplanting and harvesting, and building and maintaining dikes as a communal activity.

4.5.4 Type of local conflict

Only certain categories of local conflict events appear to be correlated with rice cultivation. Though this entire analysis is based on those events only that could plausibly be affected by local norms and behaviors emanating from a culture of rice cultivation, additional estimations (Table A14 in Appendix) show that a few of these are uncorrelated with rice cultivation. While local battle events are negatively correlated with rice producing grids, disaggregated analysis shows that only battle events where a non-state actor gains territory (E2), is consistently correlated with rice growing regions while the other two are not related at all. In addition, riots/protests and violence against civilians are uncorrelated with any form of conflict. Two explanations can be offered for these findings. First, referring back to Table 1, the results may have been driven by the low number of cases for most events other than E2. Conflict itself is a rare event and disaggregating into categories leads to extremely low numbers in certain categories that may account for the statistical insignificance of the results. Second, it may well be the case that certain events in question are not affected by local norms or the behavior of local people. For example, while the analysis includes all three forms of local battle events, in retrospect, it is possible that only battles involving non-state actors reflect local people and their behaviors whereas the other two categories involve government forces and armed violent groups such as rebel groups. As such, these events are unaffected by local peace promoting behaviors and practices. Also, the perpetrators of violence against civilians may be rebel groups or militia which could explain why local norms and behaviors do not affect these categories of conflict. However, the statistically insignificant correlation between local cultures and local events of riots and/or protests is difficult to explain and perhaps, the low number of cases is the only plausible explanation.

To summarize, three main findings emerge from the paper. First, regions of SSA that produce rice as their main crop typically experience lower conflict. Second, traditionally rice farming regions that continue to grow rice at present experience lower conflict than all other regions in SSA. This implies that the longer the history of a culture of rice cultivation, the lower the expected conflict in rice growing areas. Third, this paper finds no evidence that irrigation reduces conflict in SSA. Thus, it is more likely that the lower prevalence of conflict in rice producing units is driven by cooperative practices other than water sharing efforts of the rice cultivation process.

4.6 Robustness Checks, Limitations and Future Work

4.6.1 Robustness checks and limitations

Alternate specifications using different measures of irrigation and conflict events were used to check the sensitivity of results (some of these are shown in the Appendix). For example, while all results shown in the main tables used the fraction of total harvested area equipped for irrigation, robustness checks included indicator variables for whether the production used irrigation or rainfed agriculture. Irrigation was also measured as the difference between total harvested area and the area using rainfed agriculture. However, the results remained robust. Sensitivity checks for each conflict event included using probabilities instead of total counts. Other checks included using weather shock for the entire year instead of weather shock to the growing season. Finally, specifications were checked using probit and multinomial probit instead of linear probability model. Marginal effects revealed similar results to the linear probability model.

Though this study used cross sectional analysis to uncover patterns and find robust correlations, some econometric concerns may arise with respect to the credibility of the correlations. These include the possibility that the correlations between rice cultivation and conflict may have been driven by reverse causality or bias from potentially omitted variables. For example, peaceful communities may choose to grow more rice; and human interventions such as political motivations may affect both agriculture and conflict. If this were the case, the statistical tests would overemphasize the results. While a randomized experiment would have allowed for better causal estimation, conducting such an experiment in a war zone would neither be feasible nor ethical. Given the difficulty of collecting data from conflict prone zones (*see* Blattman & Miguel, 2010), the only feasible option at hand may be to take analytical steps to mitigate potential confounders.

In this study I have tried to control for local confounders in the following ways. First, I include country fixed effects to capture regional/country level differences in political regime, climate, etc. Second, I include a rich set of economic, social, political and geographic covariates established in the greater conflict literature. Third, my estimation strategy takes on the form of a "difference-in-difference" style approach whereby I compare the conflict outcome of rice cultivating grids in areas with and without history of rice culture to non-rice producing grids in similar areas. This strategy can potentially cancel out all common trends and differences between units. Finally, my main interaction variables are indicators for rice being the dominant crop and for belonging to a region with rice farming legacy. Though rice production is usually associated with Asia, some countries of West Africa have since ancient times been as rice oriented as Asia (NRC, 1996). Most areas that grow rice in West Africa today have been growing it for over 2000 years (Linares, 2000). Thus, the long rice farming history referred to in this paper, dating back several millenniums, could not have been affected by present day conflict. As for the dominant crop for a grid, this is primarily driven by two factors: i) an ethnocultural preference, which develops over time, and ii) crop suitability of the land. Preference for staples develop over time and often last through generations. For example, in certain parts of Africa, traditional rituals are meaningless without the ancient grain – rice (NRC, 1996). Thus, it is unlikely that conflict outcomes have affected the choice to consume and therefore grow rice in such areas. Similarly, agricultural suitability partly determines whether or not an area cultivates rice. Suitability depends on factors such as climate and soil conditions relevant to agricultural production that change very slowly over time (Ang & Frederiksson, 2017; Nunn and Qian (2011). Thus, controlling for potential confounders, the predominant crop for a grid is unlikely to be affected by grid level conflict events. Nevertheless, it is not possible to make causal inferences between crop choice and conflict without the use of panel agricultural data, as proposed in the next steps.

4.6.2 Future Work

The findings from this study raise several questions and offer a few potential avenues to pursue further research. These include disentangling whether the relations uncovered are causal in nature, using a better measure for a legacy of rice farming, explicitly testing whether individuals from rice growing regions in sub-Saharan Africa have different beliefs and values, and testing whether culture affects conflict directly or via domestic institutions.

First, I discuss isolating causal effects for the correlations uncovered in this study. The challenges with obtaining and using conflict data make it difficult to arrive at causal relations. For example, a plausible attempt at causal exploration would involve using an instrumental variable that is strongly correlated with crop choice and uncorrelated with conflict or the disturbance term. However, valid instruments are difficult to find. A common instrument for studying the effect of rice farming legacy on other outcomes such as filing for individual scientific patents, family values, etc. is the "rice suitability index" or the "rice to wheat suitability ratio" (for example, Ang & Fredriksson, 2017; Talhelm et al., 2014; Zhu, Ang & Fredriksson, 2019). Crop suitability data is available from the Global Agro-Ecological Zones (GAEZ) database, developed by the United Nations Food and Agriculture Organization (UNFAO) and the International Institute for Applied Systems Analysis (IIASA). It provides information on suitability for various crops at a 5 X 5 arc minutes (about 10 km X 10 km) cell size. However, the biggest threat to validity with this approach is that the suitability index may be picking up factors which drives outcomes of interest as well. Additionally, I would argue that when the dependent of interest is conflict, suitability is not a good instrument since it is not uncorrelated with conflict in practice. War and violence can lead to long term environmental degradation and destroy the natural fertility of soil, thus affecting the suitability of land to specific crops. Hence, I am deterred from following this common route.

Another approach for arriving at credible correlations may be the use of a regression discontinuity design by exploiting border cells. Since the grid-cells are half decimal degree squares, comparing conflict on the borders of adjacent cells that cultivate rice and other crops may be insightful. One potential problem with this approach is spill over effects. Once a conflict erupts, it would not necessarily be contained within a locality and may spread to adjacent cells. A potential approach to circumvent this issue may be to use some measure of the distance from a bordering cell and compare conflict outcomes change as one moves away from the borders.

A more reliable causal inference estimation involves exploiting panel data. While ACLED has panel information on conflict events, new waves of the SPAM agricultural database may be used to obtain panel data on crops. The current paper uses SPAM version 2005. In recent years, SPAM has released two more versions – one for 2010 and the other for 2020. Exploiting the three waves of data and a switching model, it may be possible to test whether all else remaining equal, moving towards or away from rice cultivation changes the conflict experience of agricultural communities. One issue that can complicate this attempt at entangling causality is that the theory of rice culture implies that a long-standing culture of agricultural practices associated with rice cultivation is what induces cooperative behavior in local inhabitants. Thus, grids that have recently switched from cultivating rice to some other crop may have already developed a peaceful culture, and therefore a switching model may not reflect the actual effect of culture on conflict. One way to get around this issue may be to ensure that substantial time has elapsed since the switch. For example, it may be possible to use the three panels to identify grids that have switched from rice to other crops in 2010 and then examine the local conflict level experienced by the corresponding grids in 2020. This would still require the assumption that there were no further switches in grids after 2010.

Second, I discuss ways to refine the measure used in this study. The current analysis uses a dummy variable indicating whether an unit belongs to a historically rice farming region to measure a history of rice farming. While this can provide a reasonable indication for rice legacies being correlated with lower conflict given the available data, it is possible to improve on this estimator with refined data. Some possible improvements are outlined here. One, the quantity of rice produced at the grid level can be used as a measure of history to rice farming. However, while this data is available at a country level from AfricaRice, the current analysis is much more fine-grained and requires higher resolution data. Two, historical accounts of rice in West Africa suggest that it probably originated in the flood basin of the central Niger and prehistoric Africans carried it westward to Senegal, southward to the Guinea coast, and eastward as far as Lake Chad where the natives, through their hard work, developed it further (NRC, 1996). Therefore, an extension to this study is to examine whether these areas have historically experienced less conflict; although it might be difficult to obtain such early accounts of conflict. Three, if one can obtain a timeline for when rice was introduced into each country in each region, it is possible to test whether the order in which rice was introduced is correspondingly reflected in conflict trends across countries. For instance, even within rice growing nations of West Africa one could test whether rice grown in Cote d'Ivoire has a different impact on conflict than rice grown in Guinea and explore whether the difference is due to varying degrees of rice history in different countries. Four, this analysis could be extended to other regions in SSA. For example, AfricaRice documents that rice has been grown in many East and Southern African (ESA) countries for more than 500 years though it is only in the past two decades that consumption has increased significantly. A timeline can help establish how the level of history and subsequent trends in conflict compares between these countries and those of West Africa. Five, one can explore how rice import substitutions affected conflict over the years. In the last decade or so there has been increasing reliance on imported rice in Africa (Moseley, Carney & Becker, 2010). Subject to availability of reliable grid level data on rice import, it may be possible to analyze trends in conflict before and after rice farming has been substituted by import.

Third, it may be worthwhile to explicitly test whether historically rice cultivating communities of SSA show higher level of trust and cooperation at present. This can be tested using data from the Afrobarometer Survey - a pan-African series of national public attitude surveys. This subnational, geocoded dataset covers 6 rounds of surveys in 37 African countries between 1999 and 2015. It provides hyperlocal, time-varying information about the priorities, preferences, experiences, and opinions of more than 200,000 African citizens in 28,000 localities. The generalized trust question asks respondents whether most people can be trusted whereby respondents can have two answers: "you must be very careful in dealing with people"; "most people can be trusted". It also asks how much respondents trust their relatives, neighbors, and locally elected government council, as well as those in the same country from other ethnic groups, and those from the same ethnic group. The geo-location of the respondents allows for testing whether respondents from rice growing communities demonstrate higher degrees of trust and cooperative behavior.

Fourth, a final question that naturally arises from the literature on institutions is whether formal institutions shape or are shaped by informal ones. Therefore, another promising avenue of future research involves testing whether agricultural practices and a culture of cooperation reduces conflict directly, through the establishment of strong institutions in the past or present, or the interaction of the two. However, this requires further thought and much better data.

The current data do not allow analysis beyond what has been presented in this study. The steps outlined above are ways to improve the analysis that would require more time and geocoding before the data can be further analyzed or integrated. However, even then credibility or validity issues would remain. The validity of the instrument and any other cross-sectional analysis would lead to questionable causal claims. Historical data is difficult to obtain and proposed methods for panel data analysis is likely to suffer from issues discussed above. Therefore, it remains a challenge to make valid claims on the causal impact of generations of rice farming legacy on reducing contemporaneous conflict and robust correlations may be the best one can do.

4.7 Conclusion

This chapter has laid out theoretical arguments along with empirical evidence from multiple Social Sciences on the relationship and mechanisms through which a legacy of rice farming may influence culture and conflict. This relatively unexplored African heritage can inform studies on understanding historical legacies and deep cultural roots of conflict. Furthermore, this chapter has uncovered negative correlations between rice farming and contemporaneous sub-national conflict in sub-Saharan Africa. Based on the robust findings, it has proposed potential future avenues to disentangle causal mechanisms and relationships between a history of rice farming, the evolution of culture, domestic institutions, their interactions, and persistent effects on present day sub-national conflict.

The findings are consistent with a long tradition in the Social Sciences that view cultural values and identity as equal, if not more, important than formal institutions in determining sociopolitical outcomes. Recent studies of African historiography have emphasized a growing need to move beyond the formal-institutions analysis and incorporate culture and historical events to explain the deep roots of comparative development across Africa. Through this study I hope to have illustrated an overlooked channel by which yet another heritage of Africa may have shaped its culture and conflict. This chapter also informs the broader debate on the geographical and institutional pathways to contemporary conflict. Most scholars in the "geography matters versus history matters" debate agree that geographical factors have their strongest impact on economic and political outcomes through their effect on history. The findings in this chapter may be interpreted as further suggestive evidence for this hypothesis. Finally, this chapter contributes to the growing body of empirical evidence across multiple social sciences on the observed relationship between rice growing traditions and social behavior and outcomes.

References

1. Alesina, A., & Giuliano, P. (2015). Culture and institutions. *Journal of Economic Literature*, *53*(4), 898-944.
2. Ang, J. B., & Fredriksson, P. G. (2017). Wheat agriculture and family ties. *European Economic Review*, *100*, 236-256.
3. Ang, J. B., & Gupta, S. K. (2018). Agricultural yield and conflict. *Journal of Environmental Economics and Management*, *92*, 397-417. Aoki, M., OAKI, M. A., Greif, A., & Milgrom, P. (2001). *Toward a comparative institutional analysis*. MIT press.
4. Ang, J. B., Madsen, J. B., & Wang, W. (2021). Rice Farming, Culture and Democracy. *European Economic Review*, 103778.
5. Bellemare, M. F. (2015). Rising food prices, food price volatility, and social unrest. *American Journal of Agricultural Economics*, *97*(1), 1-21.
6. Bisin, A., & Verdier, T. (2001). The economics of cultural transmission and the dynamics of preferences. *Journal of Economic theory*, *97*(2), 298-319.
7. Blattman, C., & Miguel, E. (2010). Civil war. *Journal of Economic literature*, *48*(1), 3-57.
8. Bray, F. (1983). Patterns of evolution in rice-growing societies. *The Journal of Peasant Studies*, *11*(1), 3-33.
9. Burney, J. A., & Naylor, R. L. (2012). Smallholder irrigation as a poverty alleviation tool in sub-Saharan Africa. *World Development*, *40*(1), 110-123.
10. Burney, J. A., Naylor, R. L., & Postel, S. L. (2013). The case for distributed irrigation as a development priority in sub-Saharan Africa. *Proceedings of the National Academy of Sciences*, *110*(31), 12513-12517.

11. Collier, P., & Hoeffler, A. (1998). On economic causes of civil war. *Oxford economic papers*, *50*(4), 563-573.
12. Collier, P., & Hoeffler, A. (2002). On the incidence of civil war in Africa. *Journal of conflict resolution*, *46*(1), 13-28.
13. Collier, P., & Hoeffler, A. (2004). Greed and grievance in civil war. *Oxford economic papers*, *56*(4), 563-595.
14. Couttenier, M., & Soubeyran, R. (2013). Drought and civil war in sub-saharan africa. *The Economic Journal*, *124*(575), 201-244.
15. Elbadawi, I., & Sambanis, N. (2002). How much war will we see? Explaining the prevalence of civil war. *Journal of conflict resolution*, *46*(3), 307-334.
16. Esteban, J., Mayoral, L., & Ray, D. (2012). Ethnicity and conflict: An empirical study. *American Economic Review*, *102*(4), 1310-42.
17. Fearon, J. D., & Laitin, D. D. (2003). Ethnicity, insurgency, and civil war. *American political science review*, *97*(1), 75-90.
18. Guiso, L., Sapienza, P., & Zingales, L. (2006). Does culture affect economic outcomes?. *Journal of Economic perspectives*, *20*(2), 23-48.
19. Guiso, L., Sapienza, P., & Zingales, L. (2008). Social capital as good culture. *Journal of the European economic Association*, *6*(2-3), 295-320.
20. Harari, M., & Ferrara, E. L. (2018). Conflict, climate, and cells: a disaggregated analysis. *Review of Economics and Statistics*, *100*(4), 594-608.
21. Henrich, J. (2014). Rice, psychology, and innovation. *Science*, *344*(6184), 593-594.
22. Herbst, J. (2014). *States and power in Africa*. Princeton University Press.

23. Hodler, R., & Raschky, P. A. (2014). Economic shocks and civil conflict at the regional level. *Economics Letters*, *124*(3), 530-533.

24. http://dx.doi.org/10.1787/agr_outlook-2016-en

25. Linares, O. F. (2002). African rice (Oryza glaberrima): history and future potential. *Proceedings of the National Academy of Sciences*, *99*(25), 16360-16365.

26. Mach, K. J., et al. (2019). Climate as a risk factor for armed conflict. *Nature*, *571*(7764), 193-197.

27. Mach, K. J., Kraan, C. M., Adger, W. N., Buhaug, H., Burke, M., Fearon, J. D., ... & Roessler, P. (2019). Climate as a risk factor for armed conflict. *Nature*, 1.

28. MacLean, R., & Voss, J. (1996). Allocation of Water Resources in Africa: Potential for Moving Water. *Rached, E., Rathgebr, E. and Brooks, DB Water Management in Africa and the Middle East: Challenges and Opportunities. IDRC, Ottawa, Canada*, 39-49.

29. McGuirk, E., & Burke, M. (2017). *The economic origins of conflict in Africa* (No. w23056). National Bureau of Economic Research.

30. Michalopoulos, S., & Papaioannou, E. (2013). Pre-colonial ethnic institutions and contemporary African development. *Econometrica*, *81*(1), 113-152.

31. Michalopoulos, S., & Papaioannou, E. (2014). National institutions and subnational development in Africa. *The Quarterly journal of economics*, *129*(1), 151-213.

32. Michalopoulos, S., & Papaioannou, E. (2016). The long-run effects of the scramble for Africa. *American Economic Review*, *106*(7), 1802-48.

33. Michalopoulos, S., & Papaioannou, E. (2020). Historical legacies and African development. *Journal of Economic Literature*, *58*(1), 53-128.

34. Miguel, E., Satyanath, S., & Sergenti, E. (2004). Economic shocks and civil conflict: An instrumental variables approach. *Journal of political Economy*, *112*(4), 725-753.

35. Moris, J. R. (2019). *Irrigation development in Africa: lessons of experience*. Routledge.

36. Moseley, W. G., Carney, J., Becker, L., & Hanson, S. (2010). Neoliberal policy, rural livelihoods, and urban food security in West Africa: A comparative study of The Gambia, Côte d'Ivoire, and Mali. *Proceedings of the National Academy of Sciences*, *107*(13), 5774-5779.

37. Munshi, K. (2004). Social learning in a heterogeneous population: Technology diffusion in the Indian Green Revolution. Journal of Development Economics, 73(1), 185–213.

38. Mustafa, D., & Qazi, M. U. (2007). Transition from karez to tubewell irrigation: development, modernization, and social capital in Balochistan, Pakistan. *World Development*, *35*(10), 1796-1813.

39. National Research Council. (1996). *Lost crops of Africa: volume I: grains*. National Academies Press.

40. Nunn, N. (2009). The importance of history for economic development. *Annu. Rev. Econ.*, *1*(1), 65-92.

41. Nunn, N. (2012). Culture and the historical process. *Economic History of Developing Regions*, *27*(sup-1), 108-126.

42. OECD/FAO (2016). *OECD-FAO Agricultural Outlook 2016-2025*, OECD Publishing, Paris. *http://dx.doe.org/10.1787/agr_outlook-2016-en*

43. Raleigh, C., & Dowd, C. (2017). Armed Conflict Location and Event Data Project (ACLED) Codebook 2017.[sl]: ACLED. *Tilgjengelig fra http://www. acleddata. com/wpcontent/uploads/2017/01/ACLED_Codebook_2017. pdf [lastet ned 26.*

44. Raleigh, C., Linke, A., Hegre, H., & Karlsen, J. (2010). Introducing ACLED: an armed conflict location and event dataset: special data feature. *Journal of peace research*, *47*(5), 651-660.
45. Richards, P. (1986). *Coping with hunger: hazard and experiment in an African rice-farming system*.
46. Richards, P. (1996). Culture and community values in the selection and maintenance of African rice. *Valuing local knowledge: Indigenous people and intellectual property rights*, 209-229.
47. RUAN, J., XIE, Z. & ZHANG, X. (2015). 'Does Rice Farming Shape Individualism and Innovation', Food Policy 56, 51-58
48. Sarsons, H. (2015). Rainfall and conflict: A cautionary tale. *Journal of development Economics*, *115*, 62-72.
49. Talhelm, T. (2015). The rice theory of culture (Doctoral dissertation). *Charlottesville, VA: University of Virginia*.
50. Talhelm, T. (2019). Emerging Evidence of Cultural Differences Linked to Rice Versus Wheat Agriculture. *Current opinion in psychology*.
51. Talhelm, T., Zhang, X., Oishi, S., Shimin, C., Duan, D., Lan, X., & Kitayama, S. (2014). Large-scale psychological differences within China explained by rice versus wheat agriculture. *Science*, *344*(6184), 603-608.
52. Tsusaka, T. W., Kajisa, K., Pede, V. O., & Aoyagi, K. (2015). Neighborhood effects and social behavior: The case of irrigated and rainfed farmers in Bohol, the Philippines. *Journal of Economic Behavior & Organization*, *118*, 227-246.
53. Von Carnap, T. (2017). Irrigation as a Historical Determinant of Social Capital in India? A Large-Scale Survey Analysis. *World Development*, *95*, 316-333.

54. Wischnath, G., & Buhaug, H. (2014). Rice or riots: On food production and conflict severity across India. *Political Geography*, *43*, 6-15

55. Wood-Sichra, U., A.B. Joglekar and L. You. 2016. "Spatial Production Allocation Model (SPAM) 2005: Technical Documentation". *HarvestChoice Working Paper*. Washington, D.C.: International Food Policy Research Institute (IFPRI) and St. Paul: International Science and Technology Practice and Policy (InSTePP) Center, University of Minnesota

56. You, L., & Wood, S. (2005). Assessing the spatial distribution of crop areas using a cross-entropy method. *International Journal of Applied Earth Observation and Geoinformation*, *7*(4), 310-323.

57. You, L., Ringler, C., Wood-Sichra, U., Robertson, R., Wood, S., Zhu, T., ... & Sun, Y. (2011). What is the irrigation potential for Africa? A combined biophysical and socioeconomic approach. *Food Policy*, *36*(6), 770-782.

58. Zhang, Y., & Kibriya, S. (2016). *The impact of slave trade on current civil conflict in sub-Saharan Africa* (No. 333-2016-14542).

59. Zhang, Y., & Kibriya, S. (2018). Weather shock, Slave Trade, and Conflict: Evidence from Sub-Saharan Africa.

60. Zenna, N., Senthilkumar, K., & Sie, M. (2017). Rice production in Africa. In Rice production worldwide (pp. 117-135). Springer, Cham.

APPENDIX IV

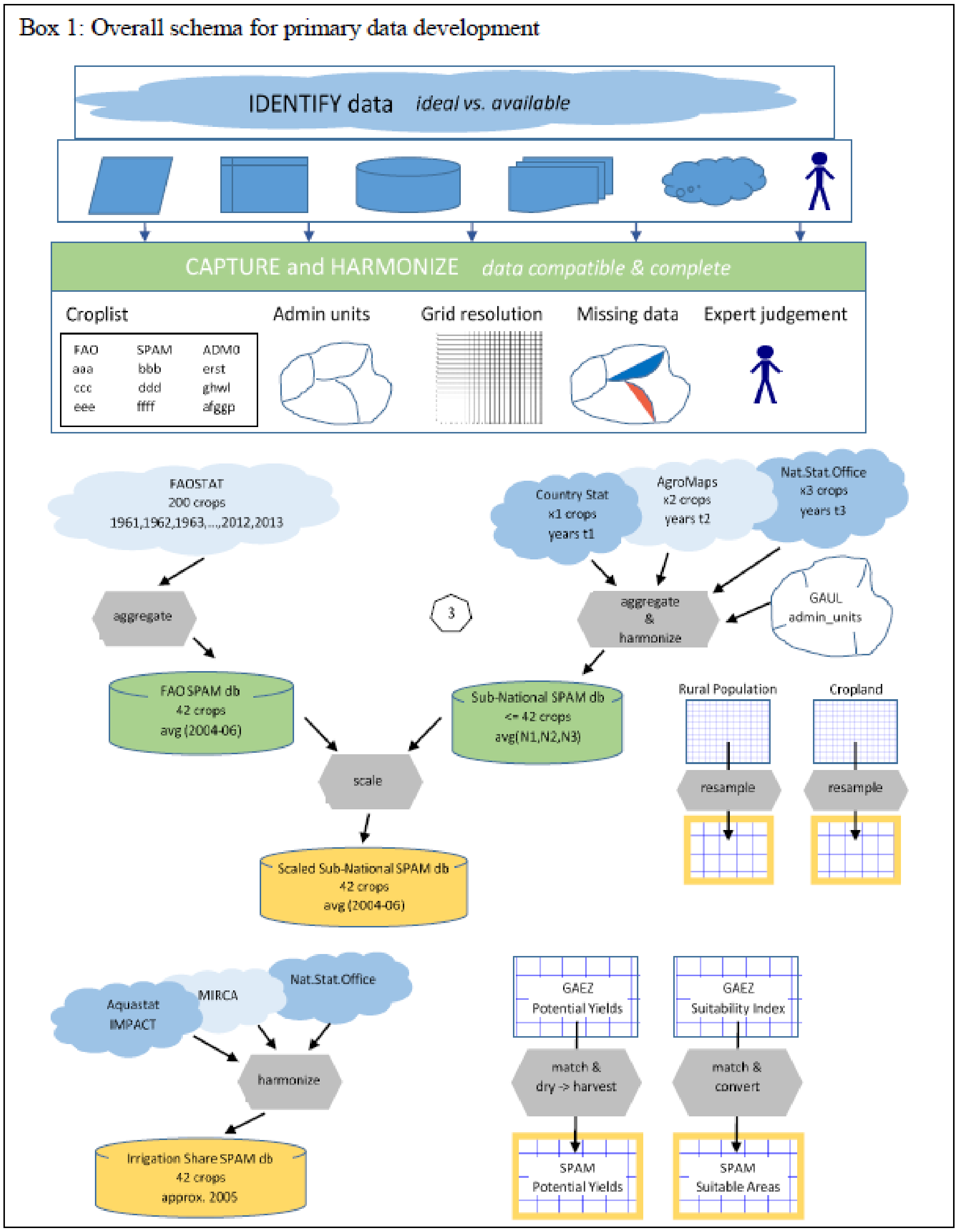


Figure 4.5: MAPSPAM 2005

Source: Wood-Sichra, Joglekar & You, 2016

Table 4.6: Effect of rice on conflict Incidence with control variables

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rice	-0.043***	0.001	0.006	0.006	0.001	-0.001	0.001	0.009
	(0.0118)	(0.0193)	(0.0234)	(0.0233)	(0.0227)	(0.0228)	(0.0227)	(0.0203)
Population density (log)			-0.079	-0.098**	-0.102**	-0.100**	-0.109**	-0.075**
			(0.0515)	(0.0489)	(0.0493)	(0.0491)	(0.0490)	(0.0338)
Nightlight density (log)			0.101***	0.100***	0.099***	0.098***	0.085***	0.057***
			(0.0121)	(0.0121)	(0.0121)	(0.0122)	(0.0117)	(0.0103)
GDP p.c. (log)			-0.032	0.003	0.012	0.019	0.026	0.045
			(0.0604)	(0.0559)	(0.0567)	(0.0570)	(0.0569)	(0.0526)
Polity score			-0.002	-0.002	-0.002	-0.002	-0.002	-0.001
			(0.0022)	(0.0022)	(0.0021)	(0.0021)	(0.0021)	(0.0018)
Ethnolinguistic fractionalization index				1.179*	1.122*	1.154*	1.307**	1.164*
				(0.6311)	(0.6139)	(0.6080)	(0.5964)	(0.6894)
Religious fractionalization index				-0.325	-0.240	-0.236	-0.319	-0.515
				(0.4531)	(0.4461)	(0.4412)	(0.4355)	(0.5119)
Malaria index					-0.004***	-0.003**	-0.003**	-0.002*
					(0.0013)	(0.0013)	(0.0013)	(0.0012)
Distance to capita					0.000***	0.000**	0.000***	0.000**
					(0.0000)	(0.0000)	(0.0000)	(0.0000)
Distance to port					-0.000	-0.000*	-0.000	-0.000
					(0.0000)	(0.0000)	(0.0000)	(0.0000)
Distance to coast					0.000	0.000	0.000	0.000
					(0.0000)	(0.0000)	(0.0000)	(0.0000)
Absolute latitude					-0.008***	-0.007***	-0.006***	-0.004***
					(0.0017)	(0.0016)	(0.0016)	(0.0015)
River						0.010	0.010	0.006
						(0.0085)	(0.0085)	(0.0078)
Ruggedness						-0.011	-0.011	-0.009
						(0.0107)	(0.0107)	(0.0100)
Elevation						0.000***	0.000***	0.000*
						(0.0001)	(0.0001)	(0.0001)
Industrial metals							0.051***	0.030*
							(0.0186)	(0.0161)
Diamond							0.044*	0.040*
							(0.0243)	(0.0228)
Petroleum							0.125***	0.113***
							(0.0357)	(0.0320)
Weather shock								0.035
								(0.0259)
Yield								-0.000
								(0.0000)
Past conflict								0.412***
								(0.0305)
Constant	0.091***	0.007*	0.951	-0.022	-0.005	-0.138	-0.323	-0.492
	(0.0039)	(0.0043)	(0.6272)	(0.6780)	(0.6809)	(0.6808)	(0.6856)	(0.6968)
Country fixed effects	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Observations	5,928	5,928	3,935	3,935	3,935	3,935	3,935	3,892
R-squared	0.001	0.100	0.120	0.122	0.133	0.138	0.147	0.290

Note: Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 4.7: Effect of rice on battle incidence with control variables

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rice	-0.031***	-0.020	-0.024*	-0.024*	-0.028**	-0.031**	-0.030**	-0.030**
	(0.0060)	(0.0125)	(0.0136)	(0.0134)	(0.0136)	(0.0132)	(0.0130)	(0.0128)
Population density (log)			-0.084	-0.085*	-0.093*	-0.090*	-0.094*	-0.076*
			(0.0520)	(0.0482)	(0.0483)	(0.0479)	(0.0478)	(0.0392)
Nightlight density (log)			0.022***	0.022***	0.021***	0.020***	0.016***	0.002
			(0.0059)	(0.0059)	(0.0058)	(0.0058)	(0.0055)	(0.0050)
GDP p.c. (log)			-0.074*	-0.060*	-0.054*	-0.047	-0.047	-0.036
			(0.0401)	(0.0307)	(0.0311)	(0.0308)	(0.0307)	(0.0260)
Polity score			-0.000	-0.000	-0.001	-0.001	-0.001	0.000
			(0.0015)	(0.0015)	(0.0015)	(0.0015)	(0.0015)	(0.0014)
Ethnolinguistic fractionalization index				0.114	0.044	0.095	0.149	0.079
				(0.2137)	(0.2286)	(0.2301)	(0.2355)	(0.2011)
Religious fractionalization index				0.247	0.336	0.329	0.294	0.198
				(0.2411)	(0.2472)	(0.2451)	(0.2476)	(0.2185)
Malaria index					-0.002**	-0.001	-0.001	-0.001
					(0.0010)	(0.0009)	(0.0009)	(0.0009)
Distance to capita					0.000***	0.000***	0.000***	0.000***
					(0.0000)	(0.0000)	(0.0000)	(0.0000)
Distance to port					-0.000	-0.000	-0.000	0.000
					(0.0000)	(0.0000)	(0.0000)	(0.0000)
Distance to coast					-0.000**	-0.000	-0.000	-0.000
					(0.0000)	(0.0000)	(0.0000)	(0.0000)
Absolute latitude					-0.004***	-0.003**	-0.002**	-0.001
					(0.0012)	(0.0011)	(0.0010)	(0.0010)
River						0.009*	0.009*	0.007
						(0.0055)	(0.0055)	(0.0052)

Ruggedness						-0.022***	-0.022***	-0.022***
						(0.0064)	(0.0064)	(0.0059)
Elevation						0.000***	0.000***	0.000***
						(0.0001)	(0.0001)	(0.0001)
Industrial metals							0.006	0.000
							(0.0100)	(0.0091)
Diamond							0.023*	0.017
							(0.0124)	(0.0118)
Petroleum							0.065**	0.055**
							(0.0275)	(0.0259)
Weather shock								0.013
								(0.0180)
Yield								0.000
								(0.0000)
Past conflict								0.209***
								(0.0239)
Constant	0.042***	0.002	0.970**	0.605*	0.609	0.474	0.426	0.328
	(0.0027)	(0.0025)	(0.4816)	(0.3639)	(0.3740)	(0.3657)	(0.3691)	(0.3489)
Country fixed effects								
Observations	5,928	5,928	3,935	3,935	3,935	3,935	3,935	3,892
R-squared	0.001	0.102	0.117	0.117	0.132	0.145	0.149	0.240

Table 4.8: Effect of rice on battle intensity with control variables

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rice	-0.117***	-0.036	-0.041	-0.041	-0.065	-0.073*	-0.071*	-0.079**
	(0.0211)	(0.0302)	(0.0381)	(0.0381)	(0.0408)	(0.0390)	(0.0389)	(0.0391)
Population density (log)			-0.083*	-0.086*	-0.120**	-0.107**	-0.118**	-0.052**
			(0.0492)	(0.0466)	(0.0483)	(0.0467)	(0.0475)	(0.0247)
Nightlight density (log)			0.041***	0.041***	0.042***	0.038***	0.023*	-0.033**
			(0.0120)	(0.0120)	(0.0116)	(0.0115)	(0.0126)	(0.0145)
GDP p.c. (log)			-0.099**	-0.086***	-0.067*	-0.041	-0.035	-0.009
			(0.0408)	(0.0330)	(0.0387)	(0.0398)	(0.0418)	(0.0326)
Polity score			-0.001	-0.001	-0.001	-0.001	-0.002	0.001
			(0.0061)	(0.0062)	(0.0060)	(0.0059)	(0.0059)	(0.0059)
Ethnolinguistic fractionalization index				0.216	-0.099	0.124	0.312	0.129
				(0.2427)	(0.3389)	(0.3258)	(0.3819)	(0.3459)
Religious fractionalization index				0.121	0.483	0.454	0.353	-0.082
				(0.2566)	(0.3055)	(0.2940)	(0.3218)	(0.3958)
Malaria index					-0.008	-0.005	-0.005	-0.004

					(0.0057)	(0.0053)	(0.0052)	(0.0051)
Distance to capita					0.000***	0.000***	0.000***	0.000***
					(0.0001)	(0.0001)	(0.0001)	(0.0001)
Distance to port					-0.000	-0.000	-0.000	0.000
					(0.0000)	(0.0000)	(0.0000)	(0.0000)
Distance to coast					-0.000	-0.000	-0.000	-0.000
					(0.0000)	(0.0000)	(0.0000)	(0.0000)
Absolute latitude					-0.009	-0.004	-0.003	-0.000
					(0.0070)	(0.0067)	(0.0067)	(0.0064)
River						0.054	0.054	0.046
						(0.0341)	(0.0342)	(0.0332)
Ruggedness						-0.102***	-0.100***	-0.104***
						(0.0342)	(0.0343)	(0.0345)
Elevation						0.001***	0.001***	0.001**
						(0.0004)	(0.0004)	(0.0004)
Industrial metals							0.045	0.020
							(0.0514)	(0.0504)
Diamond							0.044	0.030
							(0.0318)	(0.0309)
Petroleum							0.183**	0.144*
							(0.0820)	(0.0785)
Weather shock								0.232**
								(0.1077)
Yield								0.000
								(0.0000)
Past conflict								0.784***
								(0.1557)
Constant	0.139***	0.002	1.270***	0.917**	0.900**	0.367	0.157	-0.114
	(0.0156)	(0.0025)	(0.4757)	(0.3865)	(0.4545)	(0.4514)	(0.4891)	(0.5189)
Country fixed effects								
Observations	5,928	5,928	3,935	3,935	3,935	3,935	3,935	3,892
R-squared	0.001	0.083	0.090	0.090	0.096	0.101	0.102	0.139

Table 4.9: Effect of rice on event E2 with control variables

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rice	-0.117***	-0.044**	-0.043*	-0.043*	-0.062**	-0.068**	-0.066**	-0.074**
	(0.0175)	(0.0209)	(0.0263)	(0.0263)	(0.0293)	(0.0286)	(0.0285)	(0.0297)
Population density (log)			-0.085*	-0.087*	-0.116**	-0.107**	-0.115**	-0.053**
			(0.0500)	(0.0470)	(0.0481)	(0.0470)	(0.0473)	(0.0244)
Nightlight density (log)			0.035***	0.035***	0.037***	0.034***	0.023**	-0.030**
			(0.0111)	(0.0111)	(0.0108)	(0.0106)	(0.0110)	(0.0130)
GDP p.c. (log)			-0.096**	-0.082**	-0.067*	-0.049	-0.047	-0.023
			(0.0407)	(0.0326)	(0.0368)	(0.0376)	(0.0386)	(0.0290)
Polity score			-0.002	-0.002	-0.002	-0.003	-0.003	-0.000
			(0.0055)	(0.0055)	(0.0054)	(0.0054)	(0.0054)	(0.0053)
Ethnolinguistic fractionalization index				0.196	-0.073	0.087	0.222	0.061
				(0.2338)	(0.3047)	(0.2928)	(0.3376)	(0.3326)

Religious fractionalization index				0.159	0.462	0.439	0.356	-0.057
				(0.2520)	(0.2865)	(0.2747)	(0.2959)	(0.3729)
Malaria index					-0.005	-0.003	-0.002	-0.002
					(0.0050)	(0.0047)	(0.0046)	(0.0045)
Distance to capita					0.000***	0.000***	0.000***	0.000***
					(0.0001)	(0.0001)	(0.0001)	(0.0001)
Distance to port					-0.000	-0.000	-0.000	0.000
					(0.0000)	(0.0000)	(0.0000)	(0.0000)
Distance to coast					-0.000	-0.000	-0.000	-0.000
					(0.0000)	(0.0000)	(0.0000)	(0.0000)
Absolute latitude					-0.006	-0.003	-0.002	0.001
					(0.0062)	(0.0060)	(0.0060)	(0.0058)
River						0.038	0.038	0.030
						(0.0300)	(0.0300)	(0.0293)
Ruggedness						-0.074**	-0.072**	-0.075**
						(0.0295)	(0.0294)	(0.0298)
Elevation						0.001**	0.001***	0.001**
						(0.0003)	(0.0003)	(0.0003)
Industrial metals							0.020	-0.004
							(0.0397)	(0.0388)
Diamond							0.047	0.034
							(0.0316)	(0.0304)
Petroleum							0.156*	0.119
							(0.0815)	(0.0791)
Weather shock								0.238**
								(0.1067)
Yield								0.000
								(0.0000)
Past conflict								0.740***
								(0.1408)
Constant	0.130***	0.002	1.219**	0.850**	0.824*	0.451	0.321	0.065
	(0.0151)	(0.0025)	(0.4781)	(0.3849)	(0.4334)	(0.4287)	(0.4505)	(0.4783)
Country fixed effects								
Observations	5,928	5,928	3,935	3,935	3,935	3,935	3,935	3,892
R-squared	0.001	0.083	0.092	0.092	0.096	0.099	0.100	0.137

Table 4.10: Interaction effects of rice and irrigation with irrigation as a dummy variable

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Rice	0.016	-0.298*	0.046	-0.051**	-0.184**	-0.172***	-0.165**
	(0.0359)	(0.1531)	(0.0316)	(0.0209)	(0.0816)	(0.0617)	(0.0727)
Wheat	-0.098**	-0.326	-0.054**	-0.059***	-0.289***	-0.249***	0.090
	(0.0386)	(0.4182)	(0.0232)	(0.0178)	(0.0847)	(0.0794)	(0.2498)
Maize	-0.008	-0.219*	0.007	-0.017*	-0.107**	-0.078*	-0.160**
	(0.0133)	(0.1232)	(0.0102)	(0.0090)	(0.0509)	(0.0465)	(0.0686)
Pearl millet	0.009	0.009	0.009	-0.005	-0.034	-0.022	0.001
	(0.0161)	(0.1085)	(0.0127)	(0.0114)	(0.0579)	(0.0555)	(0.0417)

Sorghum	-0.004	-0.121	-0.000	0.010	0.040	0.061	-0.063
	(0.0246)	(0.2720)	(0.0169)	(0.0208)	(0.1875)	(0.1872)	(0.0693)
Cassava	-0.005	-0.156	0.002	-0.016	-0.035	-0.023	-0.101**
	(0.0142)	(0.1111)	(0.0107)	(0.0104)	(0.0554)	(0.0486)	(0.0492)
Irrigation (=1)	0.012	-0.137	0.013	-0.009	-0.072*	-0.049	-0.080
	(0.0335)	(0.1036)	(0.0292)	(0.0172)	(0.0408)	(0.0370)	(0.0517)
Rice X Irrigation	-0.032	0.331*	-0.057	0.026	0.163*	0.139**	0.174**
	(0.0465)	(0.1731)	(0.0398)	(0.0244)	(0.0836)	(0.0658)	(0.0806)
Wheat X Irrigation	0.077	0.318	0.058	0.027	0.172***	0.142**	-0.043
	(0.0598)	(0.3490)	(0.0468)	(0.0232)	(0.0664)	(0.0623)	(0.1726)
Maize X Irrigation	0.016	0.416	0.006	-0.001	0.058	0.032	0.339*
	(0.0370)	(0.2639)	(0.0316)	(0.0184)	(0.0511)	(0.0467)	(0.1937)
Pearl millet X Irrigation	-0.016	0.262*	-0.028	-0.007	0.091	0.048	0.147**
	(0.0427)	(0.1554)	(0.0350)	(0.0338)	(0.0784)	(0.0795)	(0.0641)
Sorghum X Irrigation	-0.001	0.363	-0.022	-0.016	-0.008	-0.035	-0.011
	(0.0518)	(0.4016)	(0.0320)	(0.0262)	(0.1595)	(0.1581)	(0.1902)
Cassava X Irrigation	-0.031	0.054	-0.008	0.000	0.092*	0.072	0.054
	(0.0367)	(0.1366)	(0.0306)	(0.0215)	(0.0541)	(0.0497)	(0.0710)
Constant	-0.472	0.095	-0.343	-0.224	-0.535**	-0.490**	0.385
	(0.3026)	(1.0191)	(0.2633)	(0.2143)	(0.2404)	(0.2314)	(0.6075)
Observations	4,098	4,098	4,098	4,098	4,098	4,098	4,098
R-squared	0.148	0.088	0.034	0.175	0.122	0.121	0.043

Table 4.11: Effect of rice on disaggregated conflict events

VARIABLES	(1) E1	(2) E2	(3) E3	(4) E4	(5) E5	(6) E6	(7) E7	(8) E8	(9) E9
Rice	0.003	-0.076**	-0.000	0.000	0.021	0.006	-0.003	-0.052	-0.062*
	(0.0175)	(0.0376)	(0.0005)	(0.0000)	(0.0238)	(0.0082)	(0.0018)	(0.0513)	(0.0357)
Constant	-0.010	-0.024	0.003	0.000	-0.008	0.008	0.005	0.534	0.614
	(0.0414)	(0.1463)	(0.0042)	(0.0000)	(0.0654)	(0.0063)	(0.0050)	(0.6262)	(0.7104)
Irrigation controls	yes	yes	yes	yes	yes	yes	yes	yes	yes
Other controls	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	3,451	3,451	3,451	3,451	3,451	3,451	3,451	3,451	3,451
R-squared	0.036	0.112	0.023		0.115	0.022	0.017	0.090	0.042

Table 4.12: Alternate specifications for E2

VARIABLES	(1)	(2)	(3)	(4)
Dep = E2				
Rice	-0.097***	-0.112***	-0.076**	-0.076**
	(0.0307)	(0.0335)	(0.0380)	(0.0376)
Whole year weather shock		0.275**		
		(0.1150)		
Growing season weather shock			0.145**	0.146**
			(0.0715)	(0.0715)
Yield	no	no	no	yes
Observations	4,142	4,098	3,451	3,451
R-squared	0.123	0.124	0.112	0.112

Table 4.13: Probability estimation using logit model

	Battle incidence	
VARIABLES	(1)	(2)
Rice	-1.383*	-1.893***
	(0.8145)	(0.6426)
Share of irrigated land		1.503**
		(0.6675)
Population density (log)	-15.979**	-16.484**
	(7.9558)	(8.1810)
Nightlight density (log)	0.395	0.397
	(0.3463)	(0.3444)
GDP p.c. (log)	4.025	4.247
	(5.5823)	(5.5996)
Polity score	0.010	0.011
	(0.0466)	(0.0467)
Past conflict	2.815***	2.821***
	(0.3048)	(0.3096)
Ethnolinguistic fractionalization index	441.057	455.784*
	(271.6732)	(276.2438)
Religious fractionalization index	-185.497*	-191.349*
	(100.1873)	(104.2409)
Malaria index	0.031	0.034
	(0.0343)	(0.0342)
Distance to capita	0.002***	0.003***
	(0.0007)	(0.0007)
Distance to port	0.000	0.000
	(0.0000)	(0.0000)
Distance to coast	-0.000	-0.000

	(0.0000)	(0.0000)
Absolute latitude	-0.045	-0.050
	(0.0546)	(0.0553)
River	0.055	0.018
	(0.2865)	(0.2875)
Ruggedness	-0.542	-0.542
	(0.4309)	(0.4264)
Elevation	0.004*	0.004*
	(0.0024)	(0.0024)
Industrial metals	0.590	0.637
	(0.4703)	(0.4704)
Diamond	0.931*	0.940*
	(0.5608)	(0.5674)
Petroleum	0.073	0.074
	(0.6156)	(0.6187)
Constant	-225.995	-234.505
	(187.6010)	(188.7248)
Observations	2,642	2,642

Table 4.14: Robustness checks for West Africa

VARIABLES	(1)	(2)	(3)	(4)	(6)
Rice	-0.010	-0.010	-0.004	0.015	-0.006
	(0.0178)	(0.0168)	(0.0153)	(0.0142)	(0.0164)
Wheat	-0.033***	-0.017*	-0.031***	-0.003	-0.012
	(0.0102)	(0.0104)	(0.0100)	(0.0108)	(0.0102)
Maize	-0.013***	-0.014***	-0.012***	-0.005	-0.005
	(0.0044)	(0.0050)	(0.0041)	(0.0050)	(0.0051)
Pearl millet	-0.005	-0.010*	-0.004	-0.008	-0.003
	(0.0054)	(0.0057)	(0.0054)	(0.0057)	(0.0078)
Sorghum	-0.015	-0.005	-0.014	-0.001	0.005
	(0.0109)	(0.0171)	(0.0108)	(0.0165)	(0.0151)
Cassava	-0.078	-0.160	-0.078	-0.164	0.003
	(0.0970)	(0.1380)	(0.0972)	(0.1411)	(0.0060)
West Africa	-0.062	-0.051	-0.063	-0.056	0.166
	(0.1553)	(0.1320)	(0.1553)	(0.1341)	(0.1048)
Rice X WA	-0.106***	-0.104***	-0.110***	-0.118***	-0.044
	(0.0397)	(0.0374)	(0.0390)	(0.0369)	(0.0280)
Maize X WA	-0.024	-0.021	-0.025	-0.029	-0.026
	(0.0212)	(0.0213)	(0.0211)	(0.0212)	(0.0240)
Pearl millet X WA	-0.019	-0.011	-0.019	-0.015	-0.008

	(0.0253)	(0.0246)	(0.0252)	(0.0247)	(0.0286)
Sorghum X WA	-0.023	-0.036	-0.024	-0.039	-0.037
	(0.0268)	(0.0291)	(0.0268)	(0.0288)	(0.0347)
Cassava X WA	0.066	0.135	0.066	0.137	-0.015
	(0.1042)	(0.1423)	(0.1044)	(0.1453)	(0.0383)
Past conflict		0.194***		0.195***	0.155***
		(0.0226)		(0.0227)	(0.0241)
Share of irrigated land	0.005	0.009			0.012
	(0.0084)	(0.0091)			(0.0105)
Irrigation dummy				-0.015***	
				(0.0047)	
Growing season weather shock					0.017
					(0.0147)
Constant	-0.089	-0.069	-0.090	-0.075	-0.080
	(0.0972)	(0.0842)	(0.0972)	(0.0848)	(0.1042)
Observations	4,098	4,098	4,098	4,098	3,451
R-squared	0.185	0.260	0.185	0.261	0.234
Controls	YES	YES	YES	YES	YES
Country fixed effects	YES	YES	YES	YES	YES

Table 4.15: History as WA dummy with country fixed effects included

VARIABLES	(1) Conflict incidence	(2) Conflict intensity	(3) onset	(4) Battle incidence	(5) Battle intensity	(6) E2 incidence
Rice	-0.004	-0.063	0.004	-0.003	-0.067	-0.010
	(0.0156)	(0.0591)	(0.0130)	(0.0156)	(0.0566)	(0.0124)
History=WA	0.062	0.287	0.082	0.059	0.276	0.049
	(0.0641)	(0.2454)	(0.0568)	(0.0641)	(0.2449)	(0.0649)
Rice X History	-0.103***	-0.070	-0.084**	-0.104***	-0.067	-0.097**
	(0.0388)	(0.0705)	(0.0346)	(0.0388)	(0.0679)	(0.0379)
Irrigation	0.002	-0.004	0.004	0.002	-0.007	0.004
	(0.0066)	(0.0202)	(0.0056)	(0.0066)	(0.0192)	(0.0060)
Constant	-0.213	-0.492**	-0.244	-0.214	-0.498**	-0.213
	(0.2035)	(0.2263)	(0.2059)	(0.2035)	(0.2245)	(0.2027)
Country f.e.	yes	yes	yes	yes	yes	yes
Observations	4,098	4,098	4,098	4,098	4,098	4,098
R-squared	0.175	0.119	0.038	0.175	0.121	0.170

Robust standard errors in parentheses

Connecting Text III

The previous three chapters together examined pathways through which the behavior and culture of households and communities and their operation within existing customary laws and traditions, determine the path of local conflict in sub-Saharan Africa. While understanding the root causes and dynamics of local conflict advances scholarly insights, effective tools to accurately predict the location of the next conflict may be a more pragmatic approach from a policy perspective. Timely conflict prediction allows policy options to identify conflict threats, divert resources efficiently and avoid economic and social costs. In this light, the next chapter evaluates the performance of conflict prediction models in order to determine the optimal policy option for predicting local conflict in sub-Saharan Africa.

Chapter 5

PREVENTION IS BETTER THAN CURE: MACHINE LEARNING APPROACH TO CONFLICT PREDICTION IN SUB-SAHARAN AFRICA

5.1 Introduction

"For many decision problems, it may be that prediction is of primary importance, and inference is at best of secondary importance. Even in cases where it is possible to do inference, it is important to keep on mind that the requirements that ensure this ability often come at the expense of predictive performance [1]."

From a policy perspective, predicting the occurrence of civil conflict accurately can reduce economic and social costs [2]. In 2017, the global economic impact of violence was estimated at 12.4 percent of world GDP ($14.6 trillion in purchasing power parity) [3]. Over the past six decades, the per capita growth has been three times higher in highly peaceful countries than in those with low levels of peace. This implies that allocating resources for predicting and preventing violence has significant positive economic benefits. In studies that evaluate the economic cost of conflict/violence two areas are analyzed: i) costs that arise as a consequence of violence and ii) the expenditures to prevent violence. While it appears that expenditures to prevent violence may be less expensive, policy makers are skeptical to implement prevention measures as they are viewed as riskier or even futile in the case of erroneous prediction models. However, we argue that using multiple classification algorithms to select the best predictive model can alleviate such confusion, reduce or eliminate costs due to conflict, and most importantly save human lives. Thus, our research offers empirical explorations and comparisons of logistic and machine learning algorithms that predict conflict and thereby minimize the economic and social costs of civil conflict.

Machine Learning approaches empirical analysis as algorithms that estimate and compare different alternate model scenarios. This fundamentally differs from econometric analysis, where the researcher chooses a specification based on principles and estimates the best model. Instead, Machine Learning is based on improving and changing models dynamically which is data driven. It enables researchers and policy makers to improve their predictions and analysis over time with evolving technology and data points. While research on conflict and violence in economic and policy sciences has progressed substantially in recent times, most of this literature focuses on identifying the possible drivers of conflict while ensuring the underlying assumptions of the data generating process are met [4,5]. Model selection in such econometric analysis of conflict is dominated by issues of identification, variable direction, as well as magnitude of estimates thereby imposing additional constraint on the response of policy makers. Most of the contemporary literature draws causal connection to conflict with sociopolitical and geographic variables [11, 12, 13, 14] whereby researchers identify variables and their associated, unbiased coefficients that cause conflict. The general goal is to maximize the t-statistic, thereby increasing estimation confidence and averting a type II error. However, we ask if from a policy perspective these investigations are effective in mitigating and more importantly preventing conflict? We argue that prevention is better than incurring the costs of conflict. Accordingly, we reformulate the research question to whether a region is susceptible to imminent conflict through a comparison of different econometric and machine learning models.

In order to predict incidences of conflict and examine how binary choice algorithms used in traditional econometrics compares to supervised machine learning algorithms, we commence with carefully collating and synthesizing data disaggregated information on violent conflict, geographical, socioeconomic, political and agricultural variables. We use data from the Armed Conflict Location and Event Data (ACLED) Project [9], which records violence prevalence information, as opposed to the Uppsala Conflict Data Program (UCDP) database, which codes civil war onset and fatalities information [5, 10]. Our rationale for this approach lies in our efforts to impact policy changes at a state or district level. Our approach of including a larger sphere of conflict events, especially less fatal ones, has the potential of generating effective policy pathways to prevent mass atrocities. Conflict prevention or mitigation goes beyond civil war deterrence and promotion of peace. Often smaller events are warning signs which, if detected, can prevent potential civil wars. Our combined variables are scaled to a 0.5 X 0.5 degree grid scale. The ACLED data provides information on the prevalence of conflict events including battles, violence against civilians, remote violence, rioting and protesting against a government, and non-violent conflict within the context of war. Next, we identify and describe a set of novel and standard prediction algorithms currently dominating the machine learning literature to predict the prevalence of conflict and examine model performance using out of sample test data. Machine learning provides a set of algorithms that can be used to improve prediction accuracy over classic predictive techniques such as logistic algorithm. Supervised (classification) machine learning (ML) approaches have gained popularity due to their ability to outperform simple algorithms such as logistic algorithm, providing more accurate out-of-sample predictions [6]. In as much as ML supervised algorithms have gained popularity, their use in conflict prediction is still limited and class imbalance is not explicitly addressed [7,8]. Imbalance occurs when the number of observations reporting conflict are not equally distributed and this might bias the algorithm's out-of-sample prediction towards the dominant class (observations or locations without conflict). In our study, we explicitly address classification imbalance. We argue that policy makers can obtain good prediction on incidences of conflict using existing data and applying ML algorithms that can assist in developing the best model for identifying the onset of conflict, thus reducing socioeconomic costs. Finally, we compare the performances of the logistic model (binary choice model) against four machine learning algorithms in predicting conflict events in sub-Saharan Africa: i) random forest model (combination of uncorrelated trees); ii) gradient boosting model (linear combination of trees); iii) multilayer perceptions (feed-forward artificial neural network); and iv) Support Vector Machines; and subsequently discuss the best path forward for policy makers.

Contemporary studies on conflict prediction suffer from a few drawbacks as discussed in detail in the following section. To begin with, most of these studies are limited in their geographical scope or use of high-resolution data. Furthermore, they often indicate low performance in accuracy. Finally, in cases where multiple models are compared the studies neither explicitly illustrate how addressing imbalance increases model performance nor compare predictive performance between using a reduced set of indicators and the full set of variables available to the researcher. In contrast, our initiative distinguishes and contributes to the novel yet emerging body of literature on conflict prediction in the following ways. First, in the realm of the Machine Learning and conflict mitigation literature our synthesized datasets are most exhaustive and fine combed. Our study spans over 40 countries in sub-Saharan Africa and examines conflict at a finer grid resolution within each country to predict conflict incidences at a sub-national level. From a policy perspective, our model selection is data driven whereby a large range of models are considered. However, we carefully select confidence intervals as well as document model selection processes. Second, while our intention is neither causal analysis nor impact evaluation, we list the top 20 predictors of conflict that may assist policymakers. We hope that our efforts can, in the least, enable policy makers to focus on a few key drivers of conflict, if the algorithm can identify them. Third, and most important, we compare performance across five classification algorithms while addressing issues of class imbalance that is pervasive in conflict analysis and provide information on the best call imbalance resampling methods that provides the best prediction. Additionally, we compare the differences on the possible efficacy and precision of these models. While the philosophical and scientific needs to investigate the drivers of conflict are undeniable, policy prescription needs to be pragmatic. In our model comparisons, we examine the gains obtained from using different ML models in comparison to the logistic model and discuss how performance metrics of the models may be of different value to different policy markers given the context. The trade-off between the models depends on the objective that the policy maker wants to achieve. Traditional causal analysis of conflict puts emphasis on in-sample prediction power and an extreme caution towards not accepting a false hypothesis. However, for successful policies to prevent random yet fatal conflict, out-of-sample predictions along with caution towards neglecting harbingers of catastrophic uprising, i.e. rejecting true hypotheses, may be more important. ML classification prediction problems are focused on a lower classification error and thereby provide a much better approach. The performance metrics shed light on the important policy conundrum of deciding whether institutions should prioritize preventing conflict or preserving resources for post-conflict development. From a research point of view, deciding on these performance measures are essential in terms of shifting resources in a preparedness campaign similar to relief efforts when a drought is predicted.

The rest of this article is structured as follows: the review section presents the current literature to identify variables and causal structures; the data and variable selection sections describes the source of the data, and provides context; while the results, discussion, and conclusion sections present the results and discuss their importance and contributions to the current scholarship.

5.2 Literature review

A third of the countries in sub-Saharan Africa were involved in some kind of civil war (defined as more than 1000 battle deaths per annum) or civil conflict in the mid-nineties. An estimated 750,000 to 1.1 million people were killed in conflict in Africa between 1989 and 2010 [15]. The potential costs of conflict can include substantial loss of lives, monetary losses, loss in investment and state capacity, forced displacement and mass migration, and an eventual breakdown of social cohesion, institutions and norms [15, 16, 17, 18, 19, 20, 21, 22].

Social scientists across disciplines agree that there are four main drivers of conflict: low socioeconomic development, low capabilities of the state, intergroup inequality, and recent history of violent conflict [12, 23, 24]. Socioeconomic development includes low per capita income, slow economic growth, high natural resource dependence, low levels of education, and high population [12, 14, 25, 26, 27, 28, 29]. Valid causal claims are usually argued either by exploiting exogenous sources of variation such as price shocks or natural resource endowment including diamonds and petroleum [29] or the use of instruments such as rainfall and rainfall growth [12, 25]. State institutional capabilities that often reflect how favorable or unfavorable conditions are to insurgents include lack of democracy, poor governance, fragile institutions, weak militaries, and rough terrains. The seminal works of both Collier & Hoeffler [4, 14] and Fearon & Laitin [28] establish that the roots of civil wars lie largely in opportunities for insurgency (through channels mentioned above such as fragile institutions and poor governance) and are not driven by political grievances. However, this stands in contradiction to proponents who argue that intergroup inequality and political grievances arising from ethnic and religious fractionalization across groups lead to civil conflict. Drivers of violent conflict can include both spatial, as well as temporal aspects, whereby a history of violent conflict in the past affects future conflict. In addition, agrarian societies, such as most of sub-Saharan Africa, are vulnerable to conflict arising from loss of agricultural income (a key livelihood activity), grievances from fluctuations in yield and productivity, price shocks as well as weather shocks.

Most existing studies on civil war onset have poor predictive power and place more importance on causal inference and statistical significance of individual variables. However, such statistical significance is sensitive to model specification as well as high quality data availability [5, 10, 30, 31]. For example, often cited logistic models used by Collier & Hoeffler [14] or Fearon & Laitin [28] tend to have less than 10 percent accuracy of prediction of war onset [5]. Using a multinomial logit model along with out-of-sample validation and a threshold of $p > 0.3$ for positive prediction, Hegre et al. [10] does a much better job of arriving at a true positive rate of 0.71, and a corresponding false positive rate of 0.085, seven to nine years into the future.

In contrast, statistical machine learning models appear to have superior performance when predicting conflict onset. For example, using data from Columbia and Indonesia Bazzi et al. [32] show that ML models are reliably able to identify persistent high-violence hot spots. Chadefaux [33] uses logit models to predict the onset of war with high confidence in relatively shorter time spans. Muchlinski et al. [31] finds that random forest algorithms can correctly predict 9 out of 20 civil wars, compared to zero correct prediction by three conventional logistic models (classic, Firth rare events, and L1-regularized). This result is supported by Mueller and Rauh [34] who use text mining to show that a random forest model performs extremely well in predicting conflict outbreaks, especially in previously peaceful countries. However, most conflict prediction models rely on national level aggregated statistics, which ignores within country variation and may not be representative of local contexts. Despite advances in both geospatial analysis and more sophisticated ML algorithms, conflict prediction models have rarely combined the two (granular geospatial data with ML algorithms). Only a handful of recent studies use sub-national data to predict conflict [35, 36, 37]. For example, van Weezel [37] uses sub-national data on Africa to analyze the predictive performance of a model of violent armed conflict with a range of geographic and socio-economic predictors commonly used in the causal literature. The study uses data from 2000 to 2009 to make out-of-sample predictions for 2010 to 2015 and provides an overview of the predictors of local level conflict in Africa. The author finds that conflict dynamics, determined by lagged values of spatial and temporal conflict incidences, is the strongest predictor of future conflict. Furthermore, persistence of conflict is easier to predict than the onset of new conflict events. A similar study is presented in Perry [36] with the added advantage of using two ML algorithms — a Naïve Bayes and Random Forest model — to predict district level conflict in Africa from 1997 to 2012. The study finds that selecting appropriate ML models can lead to substantial improvements in accuracy and performance and that using full models offer more predictive power than using prior outbreak of violence as the main indicator of current violence. Overall, it finds that Random Forest has an accuracy of 58.5% while the Naïve Bayes produces an accuracy of 24.6%. One drawback of the study is that it uses battle events as the only indicator of political violence though the author acknowledges that other categories of conflict events such as riots, protests, violence against civilians, remote violence, etc. should be included for more complete models of fragility and conflict prediction. Gill [35] compares the performance of Logistic model with Random Forest model in predicting political violence (defined as civil war battles and violence against civilians) in West Africa. District level data for the period 2015 – 2017 is used to train the model prior to making out-of-sample predictions for 2018. The study does not find a significant difference between the performance of the two models with an accuracy of about 74.7% for the logistic model and 78.4% for the Random Forest. The predictive power of Random Forest model with an F1 score of 0.58582 is similar to Perry [36], but the logistic regression model produces an F1 score of 0.61017 which indicates better performance than the F1 score of 0.257 in Weezel [37]. The variables in the study that contribute the most to predictive power in both models include conflict density, road density, area of the district, nighttime lights, and population density. Overall, the author finds that the Logistic regression model slightly outperforms the Random Forest model. One drawback of the study is that it is restricted to West Africa and uses district level data. In contrast, our study uses fine combed data from 40 countries across sub-Saharan Africa and compares between five classification algorithms to predict conflict while addressing class imbalance issues.

5.3 Data

To measure conflict, we use the Armed Conflict Location and Event Data Project (ACLED), version 7, developed by the International Peace Research Institute of Oslo, Norway, and the University of Uppsala, Sweden (referred to as PRIO/Uppsala) and obtained from Raleigh et al. (2010). This dataset documents nine different types of civil and communal conflict events in African states, from a period of 1997 to 2016, including battles, violence against civilians, remote violence, rioting and protesting against a government, and non-violent conflict within the context of the war. Most of the contemporary literature that use Uppsala data only considers battle related deaths that resulted in fatalities over 10,000. However, by essentially fine combing and considering the aforementioned smaller scale violence events, we expect to offer more robust models. We obtain crop data from the Spatial Production and Allocation Model (SPAM), 2005 version 1.0. This data was developed by the Food and Agricultural Organization (FAO), International Food Policy Research Institute (IFPRI), and the Center for Sustainability and the Global Environment (SAGE) at the University of Wisconsin-Madison. SPAM 2005 uses national and sub-national level crop statistics (averaged and standardized over a period of three years from 2004-2006 to FAO country levels) and provides detailed grid-level information such as crops cultivated, irrigation practices, inputs used, yield, production, etc. for 42 crops.

We combine large and publicly available datasets at a resolution of 0.5 degree latitude by 0.5 degree longitude (approximately 55km by 55 km at the equator). In cases where the data is originally available at a different resolution, it is resampled to fit our desired scale. For example, the crop data (originally 10X10km) is resampled to 55X55 km. Details of control variables included along with their corresponding data sources can be found in Tables 5.5 and 5.6 in the Appendix. Using (0.5X0.5 degree) grid-cells instead of administrative boundaries as the unit of analysis offers two advantages over most existing analysis. First, the fine-grained data allows for modelling conflict at a more localized level which can lead to better prediction models. Second, unlike administrative boundaries, grids are not likely to be governed by a particular political regimes. The final analysis spans 5928 grid-cells across 48 countries in sub-Saharan Africa. Conflict data is restricted to the years 2005 and 2006 to match with the socioeconomic and agricultural data available for the year 2005, while providing conflict predictions for the year 2006.

5.4 Variable selection and standardization

The objective of the study is to predict conflict at grid scale while comparing the logistic model to ML models for 2006 conflict incidences. In this section, we briefly describe the variables used. Conflict measures can differ depending on source, definition, and researcher decisions to aggregate and redefine incidents [22]. Most of these models use conflict prevalence or conflict onset as the main dependent variable (predicted variable). Prevalence is usually defined as the probability that the unit of observation experiences any conflict while onset indicates the start of a conflict in the given unit (e.g. cell-year). Since both these variables are binary, logistic model is the preferred model of choice in most cases.

Conflict is defined as a binary variable which equals one if the grid experienced any conflict, and zero otherwise. A broad set of predictor variables are selected based on previous literature and data availability. These are categorized as geographic, socioeconomic, political, and agricultural variables. Variables related to geography include location, geography of terrain and natural resources. Bordering grids spanning multiple countries are assigned to the country where the majority of the grid is located. Other location variables include the absolute value of the latitude of the grid; distances (km) to capital city, coastlines and ports; and a variable indicating the presence of rivers nearby. Ruggedness of terrain is measured by elevation, altitude and slope of terrain. Natural resources are accounted for by binary variables indicating the presence of diamonds, petroleum, and other industrial metals. Economic activity is measured by gross domestic product (GDP) per capita and nightlight density. Socio-political variables include population density, and ethnolinguistic and religious fractionalization indices. Governance indicators include the Polity (IV) score. Finally, past conflict is accounted for by an indicator variable which equals one if the grid experienced conflict in the past year, and zero otherwise. Agricultural variables include the dominant crop cultivated in a grid, total harvested area, crop yield, and irrigation practices as well as factors that may affect agricultural output such as crop price and weather shocks. SPAM 2005 provides detailed grid level statistics on physical area, harvested area, production, and yield for 42 crops. These are further organized by four productions systems: irrigated-high input, rainfed-high input, rainfed-low input, and rainfed subsistence production systems. We use the data to first define the dominant crop for each grid as the crop with the largest harvested area. Next, we use these statistics to generate variables indicating the fraction of total harvested area that is equipped with each of the four production systems for the dominant crop for each grid. Following Harari & La Ferrara [38] weather shocks are defined as the fraction of months per year that experience Standardized Precipitation and Evapotranspiration Index (SPEI) of one standard deviation or lower than the sample period means.

5.5 Methods, Materials and Classification techniques

We collate data on conflict incidences in 2006 and potential predictors the year prior (2005) and apply ML classification algorithms with the objective of examining whether there is 'added utility of use' compared to a logistic classification algorithm. This is a typical data science approach with the goal of evaluating model performance across accuracy, recall, and precision metrics. For algorithms that provide key predictor variables, this information is examined, and the trade-offs discussed.

5.5.1 Classification Algorithms

The study uses five classification algorithms: logistic, random forest, gradient boosting, support vector machines, and artificial neural networks. Predictors of conflict have historically relied on simple regression algorithms to provide additional insights on critical factors. But given the assumption and asymptotic properties of most of these algorithms (such as consistency, efficiency, etc.), there is a need to ensure a sample size where the number of observations are larger than the parameters (p) to be estimated. In the most applied causal or predictive algorithms such as linear regression or logistic algorithms, it might not be possible to include a large number of predictor variables. The need to select few variables or aggregate variables to obtain small number of response/predictor variables might mask important variables that correlate to the response variable. But supervised machine learning algorithms can provide tools for using these high dimension data without aggregation (either fat or thin) like the Support Vector Machines [39]. This paper uses the logistic algorithm as the base model due to its superior performance in terms of training speed, prediction, and interpretability. The logistic model can provide an output where most significant variables[15] are reported unlike artificial neural network and support vector machine algorithms at this time in their development. The other machine learning algorithms - random forest, gradient boosting, artificial neural networks, and support vector machines - might, in some cases, outperform logistic models in terms of predictive performance but at the expense of interpretability. As indicated, some of the algorithms such as support vector machines have higher performance on limited data that have large number of predictors and few observations. This diverse set of algorithms for evaluating and selecting the best predictive model is based on past studies allowing for wider application on diverse datasets [40, 41, 42]. Below we provide a brief description of each algorithm used.

[15] Since we wanted to keep the base model simpler we do not opt for a Lasso analysis and variable selection is not a key objective.

5.5.1.1 Logistic Regression

Logistic or logit regression algorithm is a probabilistic statistical classification model used to produce binary prediction model of a classification variable, dependent on multiple predictor variables [43]. Logistic model (LM), usually used after proper data transformation is applied to the initial data, produces quite a good performance compared to decision trees [44].

5.5.1.2 Decision Tress

Decision trees are non-linear models that can be used for classification and regression analysis. Decision trees require less data preparation and are easy to interpret. A tree has internal nodes and terminal nodes called leaves. Terminal nodes predict that each observation belongs to the most commonly occurring class of the training observation based on mean or mode. The Gini index or entropy measure are used to assess node purity and the class with the highest proportion of a given class at the terminal node is then selected. The number at the end of the leaf is the mean of the response of the variables. Because the predictor space is summarized as a tree using as set of splitting rules, the methods are referred to as decision trees [46]. We use two kinds of decision trees algorithms – random forests and gradient boosting.

Random forest is defined as a group of classification or regression trees that are trained on bootstrapped sample of the data. The random forest builds trees in a stepwise process that minimizes the prediction error. For categorical variables (classification problem), random forest algorithm assigns each observation to the most commonly observed class in its group. Random forest, an ensemble method that builds classifiers out of a large number of classifiers, has an advantage of reducing over fitting. For additional information on random forest conceptual model see [41, 46, and 47].

Gradient boosting is also an ensemble method that builds classifiers out of a large number of small classifiers. After the first tree is grown, the proceeding tree (second tree) is fit to the residuals from the predictions of the first tree with the goal of reducing the prediction error. Gradient boosting classifiers combine a large number of trees but does so sequentially, learning from previous estimates. Gradient boosting tends to build shorter trees than random forest and has been shown in some instances to provide better predictions than random forest. Gradient boosting has a draw-back compared to random forest in that it might take longer to train. For additional information on gradient boosting algorithms see [48] and for model specification and implementation see [41].

5.5.1.3 Support Vector Machines

Support Vector Machines (SVM) are kernel-based machine learning models that are used to analyze data to recognize patterns in both classification and regression analysis [41]. SVM is based on the structural risk minimization principle that minimizes the generalization error that is the true error on the test data [49, 50]. SVM classifier aims to split data into two groups using a hyperplane in high dimensional space. Among the strength of the SVM is that it can model non-linear relationships and is robust to outliers. SVM also uses a regularization parameter (‘C’) that is used to optimize the model by selecting the size of the margin of the hyperplane. Studies have shown SVM to outperform other ML algorithms in predictive performance, but it is difficult to obtain a ranking of which variables were important in predicting the predictor variable (in our case, conflict occurrence). SVM, at the moment, does not provide an output of variable importance. This makes it difficult to rank variables according to their contribution in improving prediction measures in the Scikit-Learn machine learning library. For additional information on model specification and implementation see [41, 50].

5.5.1.4 Artificial Neural Networks (Neural Nets)

Multi-Layer Perceptron (MLP) algorithms is one of the most widely used and popular neural networks. MLP is suitable for approximating a classifier function when the researcher is not familiar with the relationship between the input and output attributes. MLP is a feed forward artificial neural network model trained with a back-propagation algorithm which maps input data to a target variable [51]. The MLP is composed of an input layer, hidden layer(s), and an output layer. A multilayered neural network can have many hidden layers that hold the internal abstract representation of the training sample. These hidden layers allow the neural network to learn from complex datasets and improve performance on the test dataset. Multilayer perceptron algorithms have been used in classification work and in some instances outperform decision tree algorithms [44, 52]. For additional information on model implementation see [41].

5.5.2 Evaluating Model performance

To examine the accuracy of the different models in predicting occurrence of civil conflict, we split the data into a training and testing data set. There is debate on the best split between training and test datasets with most studies recommending the testing dataset range between 20 to 30 percent of the dataset [53]. We use the 80/20 split for the training and testing data. Classification algorithms in machine learning are prone to overfitting and to minimize this we use k-fold cross validation (cv) to optimize the hyper parameters of the model. Overfitting is based on the variance-bias trade-off; as an overfit model can have low classification error rate on the training dataset but high classification error rate on the test dataset (unseen data) [46]. We fit the model to the training dataset for each model and use the testing dataset to examine how the model performs on test (unseen) data.

As discussed above, the study uses five classification algorithms: logistic, random forest, gradient boosting, SVM, and MLP. Metrics of accuracy, precision, and recall are used to evaluate model performance and estimate the gains in prediction for each algorithm used. Accuracy is a measure of the overall prediction accuracy or the proportions of correctly predicted conflict outcomes. Precision (positive predictive values) is the percentage of results that are relevant (conflict). Recall (sensitivity) is the ratio of the true positive (TP) instances, in this case conflict, that are correctly detected by the model. Recall is a measure of completeness while precision is the measure of exactness. A model that has no false positive (FP) has a precision of 1 (100%) while a model with no false negative (FN) has a recall of 1 (100%).

$$Precision\ (PR) = {}^{TP}\!/_{(TP + FP)}$$

$$Recall\ or\ Sensitivity\ (RC) = {}^{TP}\!/_{(TP + FN)}$$

$$Accuracy = {}^{(TP + TN)}\!/_{(TP + FP + TN + FN)}$$

For most researchers, there is a debate between which of these two measures can be used to assess performance of the classification models: recall-precision, or the receiver operating characteristic (ROC) curve. Selecting between these two measures depends on the objective of the study. It is recommended that using recall-precision measure is better when the predicted class is imbalanced [54]. Since our predictor variable, incidences of conflict, is imbalanced we use the precision-recall measure.

It is important to note that there is a trade-off between recall and precision (Figure 5.1). Increasing model precision comes at the expense of recall. This implies that one has to find a balance between recall and precision. For this reason, F-score is also commonly used as a metric for evaluating a classifier performance. F-score is defined as the harmonic mean of precision and recall. A value closer to one implies that the classifier achieved a better combined precision and recall [44]. But it is not always that balance is required by the modeler or end user. For example, if the modeler is interested in detecting conflict areas, a model providing 30 percent precision but 70 percent recall would be acceptable. This implies that though there may be a few wrong alerts of conflict areas, the majority of conflict areas will be correctly identified [54].

		Predicted		
		No Conflict	Conflict	
Actual	No Conflict	True Negative (TN)	False Positive (FP)	
	Conflict	False Negative (FN)	True Positive (TP)	Precision
			Recall	

Figure 5.1: Illustration of recall-precision trade-off, adapted from [54]

Precision-recall curves are used to illustrate trade-offs between recall and precision. We summarize this information into a single value call the average precision score (APS) (see [42]). The higher the APS the better the score performs across these two measures. The APS is calculated from the precision-recall curve as the weighted mean precision achieved with an increase in recall at each threshold [42].

$$APS = \sum_{n} (RC_n - RC_{n-1})\, PR_n$$

Precision and recall are evaluated at the nth threshold along the curve.

5.5.2.1 Model imbalance

Model imbalance can affect the predictive performance of a classification model. Since our data exhibits class imbalance, with much fewer observations reporting conflict compared to ones with no conflict, there is a tendency of classification models to emphasize the dominant class in estimates (no conflict class). If accuracy is the only measure that is used for assessment, the model might have a high accuracy rate but a low recall rate. For example, if 90 percent of our observations did not report conflict, and the model predicted that none of our observations had conflict, we would have close to a 90 percent accuracy rate. But if the metric of interest is recall rate, we need to increase the percentage of correctly predicted conflict cases and reduce false negatives. Providing balanced (equal number of negative and positive instances) in the training data set can result in improved out-of-sample predictions on the test data [55].

Resampling is used to address the issue of imbalance in data sets. Over sampling and under sampling might have both advantages and disadvantages. Over sampling might lead to overfitting while under sampling might lead to discarding valuable information. In this study, we use techniques proven to overcome these disadvantages. There are three popular techniques: 1) Synthetic Minority Over-sampling Technique (SMOTE) that increases the minority class by introducing synthetic observations; 2) randomly under-sampling the majority class to match the minority class (NearMiss); and 3) randomly over-sampling the minority class till counts of both classes match [55, 56]. In our study, we examine the performance of two popular resampling techniques - SMOTE and the ‘NearMiss’ algorithms [55, 57]. The SMOTE draws a subset of the minority class, then generates additional similar synthetic observations, and finally, adds these observations to the original data [55]. The ‘NearMiss’ under sampling algorithm uses Euclidean distance to select samples from the majority class. In implementation of this algorithm in scikit-learn [41], there are three options: 1) selecting majority class example with minimum average distance to the three closest minority class examples; 2) selecting majority examples with minimum average distance to three further minority class observations; and 3) selecting majority class examples with minimum distance to each minority class observations. For this study, we use the techniques that selects observations based on minimum distance to the closest minority class observations [58].

5.5.2.2 Hyper Parameter Tuning

In order to select the best predictive model, tuning the model over a range of hyper parameters is important. We use the 5-fold cross-validation method where the training data is split into five folds, of which four are used to train the model and one is held back for testing in a recursive approach [43]. We discuss the parameters tuned and the ranges of the parameters for each algorithm in detail in the appendix.

Since classification machine learning algorithms might not perform well if the variables are not normalized, we use the in-built scikit-learn standard scaler to transform the continuous variables[16] [41]. The continuous variables are scaled, centered, and standardised with zero mean and unit variance. First, the algorithms are applied to transformed data without accounting for imbalance; next, imbalance is accounted for by oversampling the dominant class; and finally, imbalance is accounted for by under sampling the dominant class. This is especially important from a policy perspective because "normalized" variables are easier to estimate and predicated upon.

[16] Missing values are imputed based on author understanding of the variable and data.

5.6 Results

5.6.1 Descriptive Statistics

There are 5928 observations and 119 predictor variables in our data. Of the observations, 7.9 percent reported an incidence of conflict in 2006 while 92.1 percent did not. This implies an imbalance since the majority reported no conflict. Therefore, the no information rate for this imbalanced sample is 92.1 percent. Of the areas that reported conflict, over 55 percent had reported an incident of conflict in the previous year (2005). The average conflict area was close to the capital and had a lower gross domestic product per capita (Table 5.1).

Table 5.1: Descriptive statistics of conflict versus no conflict areas

Event	No Conflict (n=5461)	Conflict (n=467)
Past conflict	0.05 (0.22)	0.55 (0.49)
Presence of river	0.31 (0.46)	0.32 (0.46)
Industrial metals	0.08 (0.27)	0.19 (0.39)
Nightlight density	0.006 (0.04)	0.01 (0.04)
Area relying on rainfall and high inputs	4917.7 (10886)	8031.8 (17220)
Distance to capital city	542 (371)	487 (438)
Elevation level	89 (102)	148 (155)
GDP per capita	2310 (3238)	1310 (2249)
Area relying on rainfall and low inputs	4495 (11658)	7846 (15554)
Ruggedness of terrain	0.67 (0.89)	0.95 (1.12)
Maize	0.25 (0.43)	0.28 (0.45)
Malaria index	13.19 (9.4)	10.5 (8.2)

Source: Based on author’s calculations.

Note: Mean values are reported in the table with standard deviations reported in parentheses.

The best models that generate a list of important predictor variables are the random forest and gradient boosting model. These will be discussed shortly.

5.6.2 *Data Analysis*

The results are presented as follows: first we show the results of the model trained without adjusting for imbalance (Table 5.2); then we present the results after adjusting the conflict observations for imbalance using SMOTE and NearMiss under sampling (Table 5.3 and 5.4).

First, we are interested in exploring whether the prediction accuracy is above the no information rate without performing any adjustment for imbalance. As Table 5.2 shows, prior to adjustment all five models have a higher accuracy rate than the no information rate of 92.1 percent, with the gradient boosting algorithm having the highest rate. The Multilayer perceptron has the highest recall while the gradient boosting model the highest F1 score.

Table 5.2: Conflict prediction without adjusting for imbalance

Model	Accuracy	Precision	Recall	F1 Score	Average Precision Score
Logistic Regression	0.934	0.694	0.351	0.466	0.296
Random Forest	0.927	0.923	0.124	0.128	0.218
Gradient Boosting	0.940	0.759	0.423	0.543	0.368
Multilayer perceptron	0.930	0.609	0.433	0.506	0.310
Support Vector Machines	0.935	0.885	0.237	0.374	0.272
No information rate is 0.921					

After training the model on a balanced sample using SMOTE and Nearmiss algorithms, results indicate that on average SMOTE models have higher recall scores (Table 5.3 and 5.4).

Table 5.3: Conflict prediction with adjustment using SMOTE

Model	Accuracy	Precision	Recall	F1 Score	Average Precision Score
Logistic Regression	0.59	0.156	0.918	0.268	0.151
Random Forest	0.151	0.086	0.979	0.159	0.086
Gradient Boosting	0.082	0.082	1.00	0.151	0.082
Multilayer perceptron	0.79	0.224	0.619	0.329	0.17
Support Vector Machines	0.753	0.162	0.485	0.243	0.121

Table 5.4: Conflict prediction with adjustment using NearMiss technique

Model	Accuracy	Precision	Recall	F1 Score	Average Precision Score
Logistic Regression	0.418	0.106	0.825	0.188	0.102
Random Forest	0.19	0.085	0.918	0.156	0.085
Gradient Boosting	0.328	0.1	0.897	0.179	0.098
Multilayer perceptron	0.588	0.125	0.67	0.21	0.111
Support Vector Machines	0.428	0.109	0.835	0.193	0.105

Since models adjusted for imbalance using the SMOTE approach show better performance, we focus our discussing on these results. The gradient boosting model has the highest recall, followed by random forest and logistic model. The multilayer perceptron model has the highest F1 score, followed by logistic and Support Vector Machine models respectively. The Multi perceptron model also has the highest average precision score, which indicates the best trade-offs between precision and recall along the precision-recall curve.

The model with the highest recall score, gradient boosting algorithm, provides a ranking of the five most important predictors that contribute to predicting conflict - past conflict, location near a river, presence of industrial metals, density of night lights, and whether agriculture was rainfed or not. The list of 25 of the 119 variables in rank order are presented in Figure 5.2.

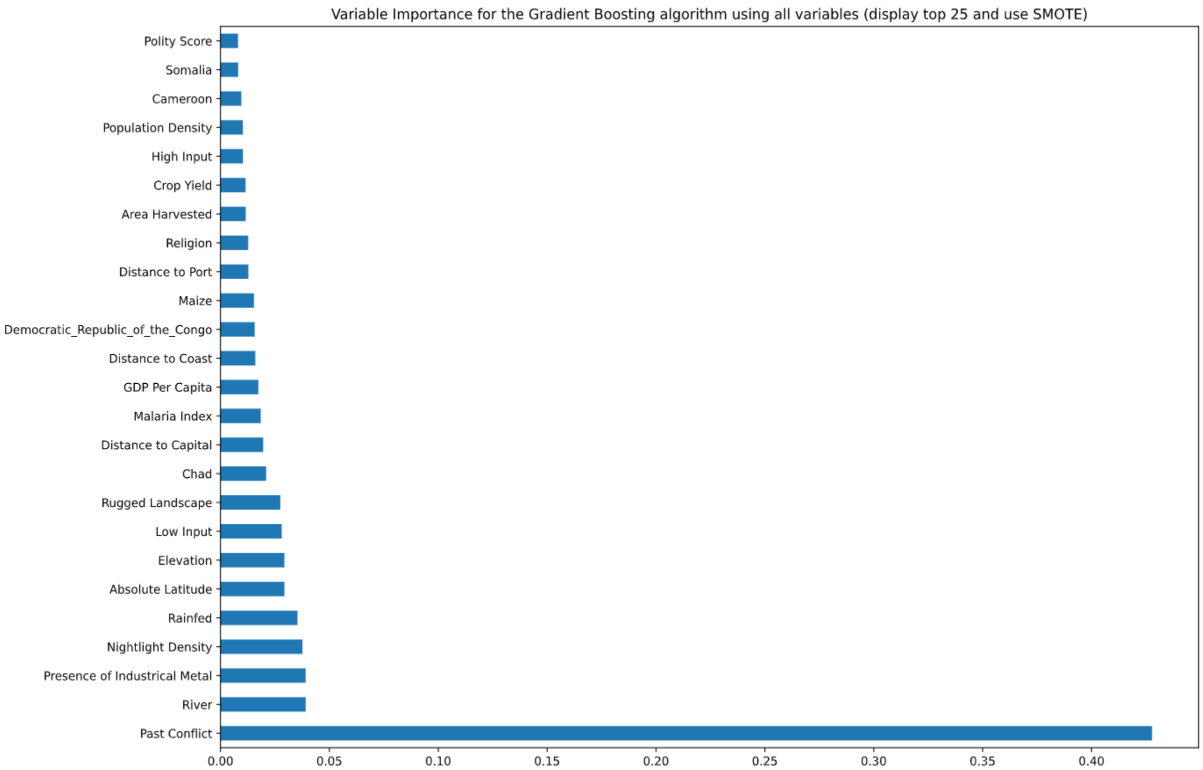


Figure 5.2: 25 of the most important variables used to predict conflict using the gradient boosting algorithm

5.7 Discussion

Our study compares the logistic regression algorithm to Machine Learning Algorithms in order to predict civil conflict in sub-Saharan Africa. This analysis creates a useful debate on which measure of prediction policy makers should prioritize - precision or recall - and how and when this choice would differ. In addition, this paper raises important questions as to why and when scientists should use ML approaches in the field of conflict assessment. Is it worthwhile to try and predict civil conflict with the new machine learning algorithms given the available data? As such, our effort improves on the existing literature by collating a higher resolution dataset, adding gradient boosting mechanisms to the analysis, and comprehensively comparing across several classification models while addressing issues of class imbalance.

The problem of class imbalance is ubiquitous in predicting conflict where we observe disproportionately lower cases of conflict compared to ones without conflict. Correcting for issues of imbalance improves performance of the algorithms, in particular the recall scores. Modeling conflict data that has an imbalanced class problem creates two issues: i) accuracy is not a good performance measure; and ii) training a model using an imbalanced class may not provide good out-of-sample predictions where recall is a key metric for model evaluation. Using the models trained on an imbalanced dataset, our results indicate a higher accuracy rate than those trained using techniques that adjust from imbalance, SMOTE and NearMiss, but lower recall rates. The models trained on an imbalanced dataset have a high accuracy rate which is higher than the no-information rate in three algorithms except the support vector machines and gradient boosting. In the latter two, we see minimal increases to the no-information rate. This indicates that the accuracy rate is lower than the rate of the largest class (no conflict) in three models. Therefore, accuracy might not be a good measure for imbalanced classes since it will be expected to have a high accuracy rate biased towards the majority class.

Our models are trained with future data to predict out of sample data. This approach mimics the real-world scenario of training current conflict to predict future data. Comparing performance across models, we find better predictions across all models for recall measures when imbalance is addressed compared to results where imbalance is not addressed. We investigate two types of class imbalance learning methods, that is, over-sampling technique (SMOTE) and under-sampling technique (NearMiss). We find that although the results in terms of which algorithm performs best on precision, recall, and F1 score are similar, overall SMOTE technique provides better results on precision and recall on average. Focusing on the SMOTE technique results, we find that when the metric of interest is recall, gradient boosting outperforms all other algorithms; but when the metric of interest is precision, Multilayer perceptron provides the best results. Since recall is a critical metric in predicting and preventing conflict, where we are more interested in identifying all potential conflict areas, gradient boosting appears to be the best model for policy makers[17]. Another advantage of gradient boosting model is that we can obtain a ranking of the most important variables (see figure 2 for variable importance in, predicting conflict). This not only improves our prediction of conflict but can also provide policy makers with guidance on which key predictors to collect data on in order to develop early warning systems. Not surprisingly, we observe that past conflict is a critical predictor of future conflict. In addition, areas located near a developed city, that can be proxied with high night light density, is a good predictor of conflict. Another key predictor was the presence of industrial metals. Extant literature concerning sub-Saharan Africa has previously shown that presence of minerals is a strong driver of conflict through extractive institutions.

From a policy perspective, our results emphasize the trade-offs between recall and precision in making appropriate decisions. On one hand, a high recall measure is of importance since we

[17] A key drawback of gradient boosting model is that it takes three times as much time to train on a 16GB RAM laptop than the other four algorithms. This implies that if time is an issue, random forest might be a second best option

prefer lower false negatives, i.e. identifying all areas with potential conflict. On the other hand, this implies that some areas that might be predicted as conflict zones will not observe conflict in reality and any resources diverted to such areas will come at the cost of other fragile areas. The question then becomes whether it is important for the policy makers be informed of all potential areas that might experience conflict, even though the true outcome may be that conflict does not occur in some of the identified areas. Or would they rather ensure that most areas where conflict might occur is correctly predicted with some areas being overlooked? This is the trade-off that policy makers need to assess depending on the existing infrastructure and available resources and on how much it might cost them to put in place preventative measures relative to the potential benefits.

Overall, we find that not all modern machine learning algorithms outperform the traditional logistic model. However, the ones that do so provide better performance across all metrics (recall and precision). When recall is the chosen measure, tree-based ML algorithms such as gradient boosting and random forest outperform the logistic regression algorithm; but if we rank the balance using the F1 score, the logit model ranks among the top half of the five algorithms. If the policy goal is to minimize this trade-off, policy makers can consider selecting a balanced model with a high F1 score. But if a modeler wants a balanced model with good precision and recall, then the multilayer perceptron algorithm provides the best model with a high F1 score. The trade-off with the multilayer perceptron that is similar with the SVM is that the model is like a black box. The model output provides predictions of performance metrics but does not indicate which variables are important in making that prediction. If understanding and examining the key predictor variables is of importance to the policy maker, then the best, balanced model with SMOTE is the gradient boosting model, followed by the logit model which also provides a ranking of variables by predictive power. Therefore, given the discussed trade-offs in performance, we present all recall and precision measures in our data that provide this information across all five models so that decision makers can choose the most appropriate model according to their goals and priorities.

5.8 Conclusion

In this paper we argue that given the availability of data, machine learning is the best way forward for policy makers to successfully predict and avert conflict. The alternative, causal inference models, is possible either by conducting Randomized Control Trials (RCT) or defining a set of assumptions. Creating conflict scenarios for an RCT or examining verifiable conditions is neither ethical nor possible. Further, data driven ML based policy approach will not have to suffer from threats to identification that many econometric models encounter.

Until the recent past, research on violent conflict primarily focused on generating correlations with sociopolitical and weather variables to imply causal inference through modeling finesse. However, as conflict datasets are getting more disaggregated over time and information on the relevant variables are getting richer, machine learning algorithms - both supervised and unsupervised - have been used in few of the conflict studies to forecast and predict the onset of both local violence and war (7,8.31,32,34,36, 59, 60). However, to our knowledge, this is among the first effort that takes advantage of continent-wide, big data and an even bigger set of carefully constructed variables of over 100 initial predictors, as well as comparing across multiple algorithms. We show a pathway to process intricate data with non-linear relationships and outperform maximum likelihood estimators in forecasting conflict onset.

Our pragmatic approaches neither emphasize *p*-values nor identifies deep rooted causes of conflict in Africa. From policy markers’ perspective, the associated coefficient or the average marginal affect of a variable is much less useful than the ability to predict the next conflict, along with some knowledge of its possible drivers. Hence, we plea to policy makers to maximize prediction performance by proposing empirical ways to make practical trade-off. Instead of minimizing only in-sample error, our techniques aim to maximize out of sample prediction power. We acknowledge the limitations of possible over fitting issues of machine algorithms. But we have also shed light on policy issues of whether to emphasize on issues of precision or recall. Furthermore, we encourage policy makers not to be hesitant to commit type II errors in order to react quickly to prevent potential conflict. In the least, our proposed approaches to the different machine learning algorithms as well logistic model presented can certainly be beneficial to prevent conflict events and assist in achieving sustainable development. In addition, investment in data collecting and real time access can enhance the application and delivery of good prediction models that can be obtained in a timely fashion and prove crucial to averting conflict. In recent times the collection of sub-national level conflict data has improved. New initiatives that stream real time data can also benefit from our research. Our models can also be updated at a real time to incorporate major early warning signs of conflict. Therefore, future studies should focus on how to develop and deploy algorithms that can provide timely recommendations. But this needs to be a multi-stakeholder initiative since data collection and curation are costly ventures that can only be undertaken if funds are available and the stakeholders can expect a net benefit.

References

1. Athey S, Imbens GW. Machine learning methods that economists should know about. Annual Review of Economics. 2019 Aug 2;11:685-725.
2. Kleinberg J, Ludwig J, Mullainathan S, Obermeyer Z. Prediction policy problems. American Economic Review. 2015 May;105(5):491-95.
3. The Economic Value of Peace (TEVP), 2018. https://reliefweb.int/sites/reliefweb.int/files/resources/Economic-Value-of-Peace-2018.pdf
4. Collier P, Hoeffler A. On economic causes of civil war. Oxford economic papers. 1998 Oct 1;50(4):563-73.
5. Ward MD, Greenhill BD, Bakke KM. The perils of policy by p-value: Predicting civil conflicts. Journal of peace research. 2010 Jul;47(4):363-75.
6. James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning. New York: springer; 2013 Feb 11.
7. Blair RA, Blattman C, Hartman A. Predicting local violence: Evidence from a panel survey in Liberia. Journal of Peace Research. 2017 Mar;54(2):298-312.
8. Gao C, Fei CJ, McCarl BA, Leatham DJ. Identifying Vulnerable Households Using Machine-Learning. Sustainability. 2020 Jan;12(15):6002.
9. Raleigh C, Linke A, Hegre H, Karlsen J. Introducing ACLED: an armed conflict location and event dataset: special data feature. Journal of peace research. 2010 Sep;47(5):651-60.
10. Hegre H, Karlsen J, Nygård HM, Strand H, Urdal H. Predicting armed conflict, 2010–2050. International Studies Quarterly. 2013 Jun 1;57(2):250-70.

11. Fearon JD. Primary commodity exports and civil war. Journal of conflict Resolution. 2005 Aug;49(4):483-507.
12. Miguel E, Satyanath S, Sergenti E. Economic shocks and civil conflict: An instrumental variables approach. Journal of political Economy. 2004 Aug;112(4):725-53.
13. Burke MB, Miguel E, Satyanath S, Dykema JA, Lobell DB. Warming increases the risk of civil war in Africa. Proceedings of the national Academy of sciences. 2009 Dec 8;106(49):20670-4.
14. Hoeffler A, Collier P. Greed and grievance in civil war. Oxford Economic Papers. 2004 Jun 22;56(4):563-95.
15. McGuirk E, Burke M. The economic origins of conflict in Africa.2017.
16. Autesserre S. Dangerous tales: Dominant narratives on the Congo and their unintended consequences. African Affairs. 2012 Apr 1;111(443):202-22.
17. Bellows J, Miguel E. War and local collective action in Sierra Leone. Journal of public Economics. 2009 Dec 1;93(11-12):1144-57.
18. Fatema N. Can land title reduce low-intensity interhousehold conflict incidences and associated damages in eastern DRC?. World Development. 2019 Nov 1;123:104612.
19. Justino P. Poverty and violent conflict: A micro-level perspective on the causes and duration of warfare. Journal of Peace Research. 2009 May;46(3):315-33.
20. Justino P. Violent conflict and human capital accumulation. IDS Working Papers. 2011 Nov;2011(379):1-7.
21. Voors MJ, Nillesen EE, Verwimp P, Bulte EH, Lensink R, Van Soest DP. Violent conflict and behavior: a field experiment in Burundi. American Economic Review. 2012 Apr;102(2):941-64.

22. The Economic Value of Peace (TEVP), 2018. https://reliefweb.int/sites/reliefweb.int/files/resources/Economic-Value-of-Peace-2018.pdf . Extracted June, 2020.
23. Blattman C, Miguel E. Civil war. Journal of Economic literature. 2010 Mar;48(1):3-57.
24. Mach KJ, Kraan CM, Adger WN, Buhaug H, Burke M, Fearon JD, Field CB, Hendrix CS, Maystadt JF, O'Loughlin J, Roessler P. Climate as a risk factor for armed conflict. Nature. 2019 Jul;571(7764):193-7.
25. Ciccone A. Transitory economic shocks and civil conflict. 2008.
26. Couttenier M, Soubeyran R. Drought and civil war in sub-saharan africa. The Economic Journal. 2014 Mar 1;124(575):201-44.
27. Elbadawi I, Sambanis N. How much war will we see? Explaining the prevalence of civil war. Journal of conflict resolution. 2002 Jun;46(3):307-34.
28. Fearon JD, Laitin DD. Ethnicity, insurgency, and civil war. American political science review. 2003 Feb 1:75-90.
29. Lujala P, Gleditsch NP, Gilmore E. A diamond curse? Civil war and a lootable resource. Journal of Conflict Resolution. 2005 Aug;49(4):538-62.
30. Cederman LE, Weidmann NB. Predicting armed conflict: Time to adjust our expectations?. Science. 2017 Feb 3;355(6324):474-6.
31. Muchlinski D, Siroky D, He J, Kocher M. Comparing random forest with logistic regression for predicting class-imbalanced civil war onset data. Political Analysis. 2016 Jan 1:87-103.

32. Bazzi S, Blair RA, Blattman C, Dube O, Gudgeon M, Peck RM. The promise and pitfalls of conflict prediction: evidence from Colombia and Indonesia. National Bureau of Economic Research; 2019 Jun 20.
33. Chadefaux T. Early warning signals for war in the news. Journal of Peace Research. 2014 Jan;51(1):5-18.
34. Mueller HF, Rauh C. The hard problem of prediction for conflict prevention.
35. Gill T. Comparing Random Forest with Generalized Linear Regression: Predicting Conflict Events in Western Africa. 2019.
36. Perry C. Machine learning and conflict prediction: a use case. Stability: International Journal of Security and Development. 2013 Oct 31;2(3):56.
37. Weezel SV. Estimating the effect of conflict on food supply at the national level. 2017.
38. Harari M, Ferrara EL. Conflict, climate, and cells: a disaggregated analysis. Review of Economics and Statistics. 2018 Oct 1;100(4):594-608.
39. McKenzie D, Sansone D. Predicting entrepreneurial success is hard: Evidence from a business plan competition in Nigeria. Journal of Development Economics. 2019 Nov 1;141:102369
40. Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers. InProceedings of the fifth annual workshop on Computational learning theory 1992 Jul 1 (pp. 144-152).
41. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J. Scikit-learn: Machine learning in Python. the Journal of machine Learning research. 2011 Nov 1;12:2825-30.

42. Mullainathan S, Spiess J. Machine learning: an applied econometric approach. Journal of Economic Perspectives. 2017 May;31(2):87-106.
43. Train KE. Discrete choice methods with simulation. Cambridge university press; 2009 Jun 30.
44. Vafeiadis T, Diamantaras KI, Sarigiannidis G, Chatzisavvas KC. A comparison of machine learning techniques for customer churn prediction. SiMution Modelling Practice and Theory. 2015 Jun 1;55:1-9.
45. Ruiz A, Villa N. Storms prediction: Logistic regression vs random forest for unbalanced data. arXiv preprint arXiv:0804.0650. 2008 Apr 4.
46. James G, Witten D, Hastie T, Tibshirani R. An introduction to statistical learning. New York: springer; 2013 Feb 11.
47. Breiman L. Random forests. Machine learning. 2001 Oct 1;45(1):5-32.
48. Friedman JH. Greedy function approximation: a gradient boosting machine. Annals of statistics. 2001 Oct 1:1189-232.
49. Elish KO, Elish MO. Predicting defect-prone software modules using support vector machines. Journal of Systems and Software. 2008 May 1;81(5):649-60.
50. Cortes C, Vapnik V. Support-vector networks. Machine learning. 1995 Sep 1;20(3):273-97
51. Ruby J, David K. Analysis of Influencing Factors in Predicting Students Performance Using MLP-A Comparative Study. International Journal of Innovative Research in Computer and Communication Engineering. 2015 Feb;3(2):1085-92.

52. Au WH, Chan KC, Yao X. A novel evolutionary data mining algorithm with applications to churn prediction. IEEE transactions on evolutionary computation. 2003 Dec;7(6):532-45.
53. Box GE, Meyer RD. An analysis for unreplicated fractional factorials. Technometrics. 1986 Feb 1;28(1):11-8.
54. Géron A. Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. O'Reilly Media; 2019 Sep 5.
55. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. Journal of artificial intelligence research. 2002 Jun 1;16:321-57.
56. Swamynathan M. Mastering machine learning with python in six steps: A practical implementation guide to predictive data analytics using python. Apress; 2019 Oct 1.
57. Mani I, Zhang I. kNN approach to unbalanced data distributions: a case study involving information extraction. InProceedings of workshop on learning from imbalanced datasets 2003 Aug 21 (Vol. 126).
58. Lemaître G, Nogueira F, Aridas CK. Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning. The Journal of Machine Learning Research. 2017 Jan 1;18(1):559-63.
59. Bessler DA, Kibriya S, Chen J, Price E. On forecasting conflict in the Sudan: 2009–2012. Journal of Forecasting. 2016 Mar;35(2):179-88.
60. Chen J, Kibriya S, Bessler D, Price E. The relationship between conflict events and commodity prices in Sudan. Journal of Policy Modeling. 2018 Jul 1;40(4):663-84.

APPENDIX V

A1. Hyper parameter tuning

Two parameters were set for the random forest technique; these are the number of variables used to grow the tree and number of trees. A range of [5,50,100, 150,200, 250] for number of trees has been assessed, as well as three different setting for the number of randomly selected variables per tree ([2,4,8,16,32,none[18]]). The number used to split the features are [3,5,10,20]. The random forests were trained using 5-fold cross-validation for tuning the parameters.

Three parameters were set for the gradient boosting technique; these are the number of variables used to grow the tree, number of trees, and the learning rate. A range of [5,50,100, 150,200,250] trees has been assessed, as well as different setting for the number of randomly selected variables per tree ([1,3,5,7,9,none]) and the learning rate [0.0001,0.001,0.01,0.1] . The gradient boosting were trained using 5-fold cross-validation for tuning the parameters.

For the SVM, the regularization parameter C, was set to a range of [0.1,1,10] and the kernel functions for selection was the radial basis function kernel (default) and the linear function kernel. Grid search using 5-fold cross-validation is used to tune the parameters. Three parameters in the MLP are tuned and these were the hidden layer sizes, the activation functions, and the learning rate. The hidden layer sizes with a range of [10,50.100], learning rate scheduled for weight updates area, constant learning rate, 'inscaling' learning rate that gradually decreases the learning rate at each time step, and the adaptive learning rate. The activation function for the hidden layers for selection are rectified linear unit function (relu), hyperbolic tan function ('tanh'), and the logistic sigmoid function ('logistic').

[18] Implies that the search can include all variables. "None" – no restriction.

Once the optimal model was tuned for each algorithm, it was saved. Then the saved model for each algorithm is tested on a single test (unseen) data set and the models are compared using the performance measures.

A2. Tables

Table 5.5. Variables and data sources

Variable	**Data Source**
Conflict	Armed Conflict Location and Event Data Project (ACLED), version 7
Agricultural data	Spatial Production and Allocation Model (SPAM), 2005 version 1.0
Natural resources	Mineral Resource Data System (U.S. Geological Survey, 1996)
Diamonds	PRIO Diamond Resources dataset (Gilmore, 2005)
Topographical data	GTOPO30 global digital elevation model (U.S. Geological Survey, 1996)
GDP and democracy indicators	Fearon & Laitin (2003)
Ethnolinguistic fractionalization index	Atlas Marodov Mira (Soviet ethnographic index)
Religious fractionalization index	CIA factbook
Weather shock	Standardized Precipitation-Evapotranspiration Index (SPEI)

Note: Data sources for natural resources, land, national income and democracy, ethnolinguistic and religious fractionalization, and weather shock are based on data collated by authors.

Table 5.6. Variables used in the prediction model

Geographical indicators	Political indicators	Main crops cultivated	
absolute latitude	corruption	cocoa	oil palm
country	government effectiveness	coconut	tropical fruit
distance to capital city	polity score	cotton	temperate fruit
distance to coast	political stability and absence of violence	cow pea	vegetables
distance to port	rule of law	groundnuts	other cereals
presence of railway	regulatory quality	banana	other fiber
presence of river	voice of accountability	barley	other pulses
elevation level		beans	other oil crops
malaria index	**Agricultural and climatic indicators**	sugarcane	all other crops
ruggedness of terrain	total harvested area for the main crop	sesame seed	
industrial metals	total yield from the main crop	coffee arabica	
diamond	area using irrigation and high inputs	coffee robusta	
petroleum	area relying on rainfall and high inputs	pearl millet	
	area relying on rainfall and low inputs	small millet	
Socioeconomic indicators	subsistence crop area relying on rainfall	wheat	
GDP per capita	indicator for irrigation technology	rice	
nightlight density	fraction of area equipped for irrigation	cassava	
population density	price of the main crop	maize	
ethnic index	rainfall shock in first growing season	sorghum	
religious index	rainfall shock in second growing season	plantain	
past conflict	rainfall shock for the whole year	potato	

Note: Ethnic and religious index refer to the ethnolinguistic and religious fractionalization indices respectively. Past conflict is an indicator variable for any conflict in the previous year. Polity score is an overall index for political indicators. Area using irrigation and high inputs refers to the total harvested area is equipped with full or partial control irrigation technology and using high technology inputs such as fertilizer and pesticides. Area relying on rainfall and low inputs refers to total harvested area that relies on rainfall for irrigation and uses little to no inputs such as fertilizer and pesticides. Subsistence crop area relying on rainfall refers to the total harvested area that relies on rainfall for irrigation and is harvested as a subsistence crop. Crops harvested twice a year may experience rainfall shocks in the first and/or the second growing seasons. Rainfall shock for the whole year is calculated as the number of months per year when the crop experiences deviations from expected monthly rainfall.

Chapter 6

CONCLUSION

This dissertation aimed to advance knowledge and scholarship on localized conflict in sub-Saharan Africa which are often neglected from both conflict analyses and new reports. Localized events are commonly rooted in the behavior of individuals, households, communities, and their interactions within social norms. Though such events often originate as innocuous exchanges between agents, they have the potential to intensify to wide scale unrest with severe consequences. This dissertation explored mechanisms through which household and community behavior, culture and local, informal institutions determine the path of local conflicts and evaluates the best policy options for predicting local conflict.

Based on the combined evidence, this dissertation reaches two main conclusions. First, it appears that a culture of cooperation and benevolence can potentially reduce local conflict between households and communities. Thus, development policy that incorporates awareness campaigns to educate households and communities on the virtues of cooperative behavior, such as helping one another in the field or with food can potentially help reduce local conflict. In addition, engaging communities in local team-building activities can foster a culture of cooperation that can not only avert conflict but also lay the foundation for conflict resolution or peaceful negotiations and settlements within the community should a conflict arise. Second, timely and accurate prediction of local conflict is crucial to reducing economic and social costs of conflict. Thus, investing in conflict early warning systems with real-time data can prove beneficial to conflict mitigation policies. A deeper understanding of the root causes of local conflict can inform the nature of data required for developing an effective early warning system. The overall evidence from this dissertation emphasizes the role of household and community behavior, values, and cultural norms in determining the nature and dynamics of localized conflict. Future prediction models can consider incorporating data on local traditions and culture.

Recent African historiography has criticized economists and political scientists for overemphasizing the role of formal institutions, especially those established through colonial rule, on explaining comparative development and conflict in Africa. In contrast, mechanisms involving informal institutions, such as ethno-cultural traits, customary laws, and Africa's precolonial history has been underemphasized. This dissertation tried to address this gap in the broader conflict and development literature by examining channels through which informal institutions, such as behavioral norms, cultural values, local traditions and precolonial history shaped the path of local conflict in contemporary sub-Saharan Africa. In addition, this dissertation has offered a comparative analysis of local conflict prediction models and proposed the optimal model based on the goals and priorities of policy makers. Better understanding of the root causes of local conflict, in conjunction with effective models for predicting local conflict can help avert local conflicts before they escalate to national crises, thus reducing social and economic costs and ultimately saving lives.

MASTER REFERENCE

1. Abadie, A., & Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica*, *74*(1), 235-267.
2. Albertus, M., & Kaplan, O. (2013). Land Reform as a Counterinsurgency Policy Evidence from Colombia. *Journal of Conflict Resolution*, *57*(2), 198-231.
3. Albertus, M., Brambor, T., & Ceneviva, R. (2018). Land inequality and rural unrest: theory and evidence from Brazil. *Journal of Conflict Resolution*, *62*(3), 557-596.
4. Alesina, A., & Giuliano, P. (2015). Culture and institutions. *Journal of Economic Literature*, *53*(4), 898-944.
5. Alston, L. J., & Mueller, B. (2010). *Property rights, land conflict and tenancy in Brazil* (No. w15771). National Bureau of Economic Research.
6. Ang, J. B., & Fredriksson, P. G. (2017). Wheat agriculture and family ties. *European Economic Review*, *100*, 236-256.
7. Ang, J. B., & Gupta, S. K. (2018). Agricultural yield and conflict. *Journal of Environmental Economics and Management*, *92*, 397-417. Aoki, M., OAKI, M. A., Greif, A., & Milgrom, P. (2001). *Toward a comparative institutional analysis*. MIT press.
8. Ang, J. B., Madsen, J. B., & Wang, W. (2021). Rice Farming, Culture and Democracy. *European Economic Review*, 103778.
9. Arezki, M. R., & Bruckner, M. (2011). *Food prices and political instability* (No. 11-62). International Monetary Fund.
10. Arruñada, B. (2003). Property enforcement as organized consent. *Journal of Law, Economics, and Organization*, *19*(2), 401-444.

11. Athey, S., & Imbens, G. W. (2019). Machine learning methods that economists should know about. *Annual Review of Economics*, *11*, 685-725.

12. Au, W. H., Chan, K. C., & Yao, X. (2003). A novel evolutionary data mining algorithm with applications to churn prediction. *IEEE transactions on evolutionary computation*, *7*(6), 532-545.

13. Autesserre, S. (2010). *The trouble with the Congo: Local violence and the failure of international peacebuilding* (Vol. 115). Cambridge University Press.

14. Autesserre, S. (2014). Going micro: Emerging and future peacekeeping research. *International Peacekeeping*, *21*(4), 492-500.

15. Autesserre, S. (2012). Dangerous tales: Dominant narratives on the Congo and their unintended consequences. *African Affairs*, *111*(443), 202-222.

16. Barnett, J., & Adger, W. N. (2007). Climate change, human security and violent conflict. *Political geography*, *26*(6), 639-655.

17. Barrett, C. B. (2001). Does food aid stabilize food availability? *Economic development and cultural change*, *49*(2), 335-349.

18. Barron, P., Kaiser, K., & Pradhan, M. (2009). Understanding variations in local conflict: Evidence and implications from Indonesia. *World Development*, *37*(3), 698-713.

19. Bazzi, S., Blair, R. A., Blattman, C., Dube, O., Gudgeon, M., & Peck, R. (2018). The promise and pitfalls of conflict prediction: evidence from Colombia and Indonesia. *The Review of Economics and Statistics*, 1-45.

20. Bellemare, M. F. (2015). Rising food prices, food price volatility, and social unrest. *American Journal of Agricultural Economics*, *97*(1), 1-21.

21. Bellows, J., & Miguel, E. (2009). War and local collective action in Sierra Leone. *Journal of Public Economics*, *93*(11), 1144-1157.

22. Berazneva, J., & Lee, D. R. (2013). Explaining the African food riots of 2007–2008: An empirical analysis. *Food Policy*, *39*, 28-39.

23. Besley, T. (1995). Property rights and investment incentives: Theory and evidence from Ghana. *journal of Political Economy*, 903-937.

24. Besley, T. J., & Ghatak, M. (2009). Improvement and extension of property rights. *Handbook of development economics*, *5*.

25. Bessler, D. A., Kibriya, S., Chen, J., & Price, E. (2016). On Forecasting Conflict in the Sudan: 2009–2012. *Journal of Forecasting*, *35*(2), 179-188.

26. Bisin, A., & Verdier, T. (2001). The economics of cultural transmission and the dynamics of preferences. *Journal of Economic theory*, *97*(2), 298-319.

27. Blair, R. A., Blattman, C., & Hartman, A. (2017). Predicting local violence: Evidence from a panel survey in Liberia. *Journal of Peace Research*, *54*(2), 298-312.

28. Blattman, C., & Miguel, E. (2010). Civil war. *Journal of Economic literature*, *48*(1), 3-57.

29. Blattman, C., Hartman, A. C., & Blair, R. A. (2014). How to promote order and property rights under weak rule of law? An experiment in changing dispute resolution behavior through community education. *American Political Science Review*, *108*(1), 100-120.

30. Bora, S., Ceccacci, I., Delgado, C., & Townsend, R. (2010). World Development Report 2011: Food sufficiency and Conflict. *Washington, DC: Agriculture and Rural Development Department, World Bank*.

31. Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992, July). A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory* (pp. 144-152).

32. Box, G. E., & Meyer, R. D. (1986). An analysis for unreplicated fractional factorials. *Technometrics*, *28*(1), 11-18.

33. Bray, F. (1983). Patterns of evolution in rice-growing societies. *The Journal of Peasant Studies*, *11*(1), 3-33.

34. Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5-32.

35. Brinkman, H. J., & Hendrix, C. S. (2011). *Food insufficiency and violent conflict: Causes, consequences, and addressing the challenges*. World Food Programme.

36. Broegaard, R. B., Vongvisouk, T., & Mertz, O. (2017). Contradictory land use plans and policies in Laos: tenure security and the threat of exclusion. *World Development*, *89*, 170-183.

37. Bromley, D. W. (2009). Formalising property relations in the developing world: The wrong prescription for the wrong malady. *Land Use Policy*, *26*(1), 20-27.

38. Brück, T., d'Errico, M., & Pietrelli, R. (2018). The effects of violent conflict on household resilience and food security: Evidence from the 2014 Gaza conflict. *World Development*

39. Burke, M. B., Miguel, E., Satyanath, S., Dykema, J. A., & Lobell, D. B. (2009). Warming increases the risk of civil war in Africa. *Proceedings of the national Academy of sciences*, *106*(49), 20670-20674.

40. Burney, J. A., & Naylor, R. L. (2012). Smallholder irrigation as a poverty alleviation tool in sub-Saharan Africa. *World Development*, *40*(1), 110-123.

41. Burney, J. A., Naylor, R. L., & Postel, S. L. (2013). The case for distributed irrigation as a development priority in sub-Saharan Africa. *Proceedings of the National Academy of Sciences*, *110*(31), 12513-12517.

42. Bush, R., & Martiniello, G. (2017). Food Riots and Protest: Agrarian Modernizations and Structural Crises. *World Development*, *91*, 193-207.

43. Bushman, B. J., DeWall, C. N., Pond, R. S., & Hanus, M. D. (2014). Low glucose relates to greater aggression in married couples. *Proceedings of the National Academy of Sciences*, *111*(17), 6254-6257.

44. Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys*, *22*(1), 31-72.

45. Carter, K. N., Kruse, K., Blakely, T., & Collings, S. (2011). The association of food security with psychological distress in New Zealand and any gender differences. *Social science & medicine*, *72*(9), 1463-1471.

46. Cederman, L. E., & Weidmann, N. B. (2017). Predicting armed conflict: Time to adjust our expectations?. *Science*, *355*(6324), 474-476.

47. Chadefaux, T. (2014). Early warning signals for war in the news. *Journal of Peace Research*, *51*(1), 5-18.

48. Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, *16*, 321-357.

49. Chen, J., Kibriya, S., Bessler, D., & Price, E. (2018). The relationship between conflict events and commodity prices in Sudan. *Journal of Policy Modeling*, *40*(4), 663-684.

50. Chilton, M., & Booth, S. (2007). Hunger of the body and hunger of the mind: African American women's perceptions of food insecurity, health and violence. *Journal of nutrition education and behavior*, *39*(3), 116-125.

51. Chimhowu, A., & Woodhouse, P. (2006). Customary vs private property rights? Dynamics and trajectories of vernacular land markets in Sub-Saharan Africa. *Journal of agrarian change*, *6*(3), 346-371.

52. Chiputwa, B., Spielman, D. J., & Qaim, M. (2015). Food standards, certification, and poverty among coffee farmers in Uganda. *World Development*, *66*, 400-412.

53. Ciccone, A. (2011). Economic shocks and civil conflict: A comment. *American Economic Journal: Applied Economics*, *3*(4), 215-27.

54. Cohen, M. J. (1993). Hunger 1994: transforming the politics of hunger. *Bread for the World Institute, Silver Spring, MD*.

55. Cohen, M. J., & Pinstrup-Andersen, P. (1999). Food sufficiency and conflict. *Social Research*, 375-416. Food sufficiency, Justice and Peace (2002). FAO.

56. Collier, P., & Hoeffler, A. (1998). On economic causes of civil war. *Oxford economic papers*, *50*(4), 563-573.

57. Collier, P., & Hoeffler, A. (2002). On the incidence of civil war in Africa. *Journal of conflict resolution*, *46*(1), 13-28.

58. Collier, P., & Hoeffler, A. (2004). Greed and grievance in civil war. *Oxford economic papers*, *56*(4), 563-595.

59. Cook, T. D., Shadish, W. R., & Wong, V. C. (2008). Three conditions under which experiments and observational studies produce comparable causal estimates: New findings from within-study comparisons. *Journal of policy analysis and management*, *27*(4), 724-750.

60. Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*(3), 273-297.

61. Couttenier, M., & Soubeyran, R. (2013). Drought and civil war in sub-saharan africa. *The Economic Journal*, *124*(575), 201-244.

62. Crump, R. K., Hotz, V. J., Imbens, G. W., & Mitnik, O. A. (2008). Nonparametric tests for treatment effect heterogeneity. *The Review of Economics and Statistics*, *90*(3), 389-405.

63. Dehejia, R. H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and statistics*, *84*(1), 151-161.

64. Deininger, K., & Feder, G. (2009). Land registration, governance, and development: Evidence and implications for policy. *The World Bank Research Observer*, *24*(2), 233-266.

65. 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.

66. Di Falco, S., Laurent-Lucchetti, J., & Veronesi, M. (2015). Property rights and conflicts: theory and evidence from the Highland of Ethiopia Preliminary and incomplete. Retrieved from: https://www.tse-fr.eu/sites/default/files/TSE/documents/conf/energy_climat/Papers/laurent-lucchetti.pdf

67. Elbadawi, I., & Sambanis, N. (2002). How much war will we see? Explaining the prevalence of civil war. *Journal of conflict resolution*, *46*(3), 307-334.

68. Elish, K. O., & Elish, M. O. (2008). Predicting defect-prone software modules using support vector machines. *Journal of Systems and Software*, *81*(5), 649-660.

69. Esteban, J., Mayoral, L., & Ray, D. (2012). Ethnicity and conflict: An empirical study. *American Economic Review*, *102*(4), 1310-42.

70. Esteban, J., & Ray, D. (2011). Linking conflict to inequality and polarization. *American Economic Review*, *101*(4), 1345-74.

71. Fang, X., Kothari, S., McLoughlin, C., & Yenice, M. (2020). The Economic Consequences of Conflict in Sub-Saharan Africa

72. Fatema, N. (2019). Can land title reduce low-intensity interhousehold conflict incidences and associated damages in eastern DRC? *World Development*, *123*, 104612.

73. Fatema, N., & Kibriya, S. (2018). *Givers of great dinners know few enemies: The impact of household food sufficiency and food sharing on low intensity interhousehold and community conflict in eastern Democratic Republic of Congo (*HiCN Working Paper 267). Retrieved from the Institute of Development Studies (IDS) at the University of Sussex, Households in Conflict Network website: http://www.hicn.org/wordpress/wp-content/uploads/2018/03/HiCN-WP-267.pdf.

74. Fearon, J. D. (1998). Bargaining, enforcement, and international cooperation. *International organization*, *52*(2), 269-305.

75. Fearon, J. D., & Laitin, D. D. (2003). Ethnicity, insurgency, and civil war. *American political science review*, *97*(1), 75-90.

76. Fearon, J. D., Humphreys, M., & Weinstein, J. M. (2009). Can development aid contribute to social cohesion after civil war? Evidence from a field experiment in post-conflict Liberia. *The American Economic Review*, *99*(2), 287-291.

77. Fearon, J. D. (2005). Primary commodity exports and civil war. *Journal of conflict Resolution*, *49*(4), 483-507.

78. Feder, G., & Nishio, A. (1998). The benefits of land registration and titling: economic and social perspectives. *Land use policy*, *15*(1), 25-43.

79. Field, E. (2005). Property rights and investment in urban slums. *Journal of the European Economic Association*, *3*(2-3), 279-290.

80. Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189-1232.

81. Galiani, S., & Schargrodsky, E. (2010). Property rights for the poor: Effects of land titling. *Journal of Public Economics*, *94*(9), 700-729.

82. Gao, C., Fei, C. J., McCarl, B. A., & Leatham, D. J. (2020). Identifying Vulnerable Households Using Machine Learning. *Sustainability*, *12*(15), 6002.

83. Garfinkel, M. R. (1994). Domestic Politics and International Conflict. *The American Economic Review*, *84*(5), 1294–1309.

84. Garfinkel, M. R., & Skaperdas, S. (2007). Economics of conflict: An overview. *Handbook of defense economics*, *2*, 649-709.

85. Géron, A., (2017). Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems. O'Reilly Media, Inc., Sebastopol, CA, USA, 2019.

86. Gill, T. M. (2019). *Comparing Random Forest with Generalized Linear Regression: Predicting Conflict Events in Western Africa* (Doctoral dissertation, The University of Arizona).

87. Global Humanitarian Report-An Overview. (2015). UN Office for the Coordination of Humanitarian Affairs

88. Goldstein, M., & Udry, C. (2008). The profits of power: Land rights and agricultural investment in Ghana. *Journal of political Economy*, *116*(6), 981-1022.

89. Grossman, H. I. (1991). A general equilibrium model of insurrections. *The American Economic Review*, 912-921.

90. Grossman, H. I. (1994). Production, appropriation, and land reform. *The American Economic Review*, *84*(3), 705-712.

91. Gulyani, S., & Talukdar, D. (2008). Slum real estate: The low-quality high-price puzzle in Nairobi's slum rental market and its implications for theory and practice. *World Development*, *36*(10), 1916-1937.

92. Guiso, L., Sapienza, P., & Zingales, L. (2006). Does culture affect economic outcomes?. *Journal of Economic perspectives*, *20*(2), 23-48.

93. Guiso, L., Sapienza, P., & Zingales, L. (2008). Social capital as good culture. *Journal of the European economic Association*, *6*(2-3), 295-320.

94. Haavelmo, T. (1954). *A study in the theory of economic evolution* (No. 330.1/H11s). Amsterdam: North-Holland.

95. Harari, M., & Ferrara, E. L. (2018). Conflict, climate, and cells: a disaggregated analysis. *Review of Economics and Statistics*, *100*(4), 594-608.

96. Heckman, J. J., Ichimura, H., & Todd, P. (1998). Matching as an econometric evaluation estimator. *The Review of Economic Studies*, *65*(2), 261-294.

97. Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The Review of Economic Studies*, *64*(4), 605-654.

98. Heflin, C. M., Siefert, K., & Williams, D. R. (2005). Food insufficiency and women's mental health: findings from a 3-year panel of welfare recipients. *Social Science & Medicine*, *61*(9), 1971-1982.

99. Hegre, H., Karlsen, J., Nygård, H. M., Strand, H., & Urdal, H. (2013). Predicting armed conflict, 2010–2050. *International Studies Quarterly*, *57*(2), 250-270.

100. Hendrix, C. S., & Salehyan, I. (2010). After the Rain: Rainfall Variability, Hydro-Meteorological Disasters, and Social Conflict in Africa.

101. Hendrix, C., & Brinkman, H. J. (2013). Food insufficiency and conflict dynamics: Causal linkages and complex feedbacks. *Stability: International Journal of Security and Development*, *2*(2).

102. Henrich, J. (2014). Rice, psychology, and innovation. *Science*, *344*(6184), 593-594.

103. Herbst, J. (2014). *States and power in Africa*. Princeton University Press.

104. Hirshleifer, J. (1991). The technology of conflict as an economic activity. *The American Economic Review*, *81*(2), 130-134.

105. Hirshleifer, J. (1995). Theorizing about conflict. *Handbook of defense economics*, *1*, 165-189.

106. Hodler, R., & Raschky, P. A. (2014). Economic shocks and civil conflict at the regional level. *Economics Letters*, *124*(3), 530-533.

107. Homer-Dixon, T. F. (1994). Environmental scarcities and violent conflict: evidence from cases. *International security*, *19*(1), 5-40.

108. Homer-Dixon, T. F. (1999). Environment, scarcity, and conflict. *Princeton University*.

109. Horowitz, D. L. (1985). *Ethnic groups in conflict*. Univ of California Press. http://dx.doi.org/10.1787/agr_outlook-2016-en

110. Humphreys, M., & Weinstein, J. M. (2008). Who fights? The determinants of participation in civil war. *American Journal of Political Science*, *52*(2), 436-455.

111. Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and statistics*, *86*(1), 4-29.

112. Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of economic literature*, *47*(1), 5-86.

113. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. (2013). *An Introduction to Statistical Learning*; Springer: Cham, Switzerland, Volume 112, p. 18.

114. Justino, P. (2009). Poverty and violent conflict: A micro-level perspective on the causes and duration of warfare. *Journal of Peace Research*, *46*(3), 315-333.

115. Justino, P. (2011). The impact of armed civil conflict on household welfare and policy responses. *Securing Peace: State-Building and Economic Development in Post-Conflict Countries*, *19*.

116. Justino, P. Violent conflict and human capital accumulation. *IDS Work. Pap.* 2011, *2011*, 1–17

117. Justino, P., Brück, T., & Verwimp, P. (Eds.). (2013). *A micro-level perspective on the dynamics of conflict, violence, and development*. Oxford University Press.

118. Kahl, C. H. (2006). *States, scarcity, and civil strife in the developing world*. Princeton University Press.

119. Kalyvas, S. N. (2006). *The logic of violence in civil war*. Cambridge University Press.

120. Kaplan, R. D. (1994). The coming anarchy. *Atlantic monthly*, *273*(2), 44-76.

121. Kibriya, S., Xu, Z. P., & Zhang, Y. (2017). The negative consequences of school bullying on academic performance and mitigation through female teacher participation: evidence from Ghana. *Applied Economics*, *49*(25), 2480-2490.

122. Kleinberg, J., Ludwig, J., Mullainathan, S. and Obermeyer, Z., 2015. Prediction policy problems. *American Economic Review*, 105(5), pp.491-95.

123. Lemaître, G., Nogueira, F., & Aridas, C. K. (2017). Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning. *The Journal of Machine Learning Research*, *18*(1), 559-563.

124. Leuveld, K. et al. (Forthcoming). The impact of agricultural extension and input subsidies on knowledge, input use and food security in Eastern DRC.

125. Li, M. (2013). Using the propensity score method to estimate causal effects: A review and practical guide. *Organizational Research Methods*, *16*(2), 188-226.

126. Linares, O. F. (2002). African rice (Oryza glaberrima): history and future potential. *Proceedings of the National Academy of Sciences*, *99*(25), 16360-16365.

127. Linkow, B. (2016). Causes and consequences of perceived land tenure insecurity: Survey evidence from Burkina Faso. Land Economics, 92(2), 308-327.

128. Lujala, P., Gleditsch, N. P., & Gilmore, E. (2005). A diamond curse? Civil war and a lootable resource. *Journal of Conflict Resolution*, *49*(4), 538-562.

129. Mach, K. J., et al. (2019). Climate as a risk factor for armed conflict. *Nature*, *571*(7764), 193-197.

130. MacLean, R., & Voss, J. (1996). Allocation of Water Resources in Africa: Potential for Moving Water. *Rached, E., Rathgebr, E. and Brooks, DB Water Management in Africa and the Middle East: Challenges and Opportunities. IDRC, Ottawa, Canada*, 39-49.

131. Macrae, J., & Zwi, A. B. (1992). Food as an instrument of war in contemporary African famines: a review of the evidence. *Disasters*, *16*(4), 299-321.

132. Malinowski, B. (1984) [1922]. *Argonauts of the Western Pacific: An account of native enterprise and adventure in the archipelagoes of Melanesian New Guinea*. Prospect Heights, Ill.: Waveland Press.

133. Mani, I., & Zhang, I. (2003). kNN approach to unbalanced data distributions: a case study involving information extraction. In *Proceedings of workshop on learning from imbalanced datasets* (Vol. 126). United States: ICML.

134. Marivoet, W., & De Herdt, T. (2019). What Happens with Household Assets during Economic Collapse? The Case of the Democratic Republic of the Congo (1975–2010). *The Journal of Development Studies*, *55*(4), 680-701.

135. Martin-Shields, C. P., & Stojetz, W. (2018). Food Security and Conflict: Empirical challenges and future opportunities for research and policy making on food security and conflict. *World Development.*

136. Mauss, M. (1970). *The Gift: Forms and Functions of Exchange in Archaic Societies*. London: Cohen & West.

137. Maxwell, D., Watkins, B., Wheeler, R., & Collins, G. (2003). The coping strategies index: A tool for rapidly measuring food security and the impact of food aid programs in emergencies. *Nairobi: CARE Eastern and Central Africa Regional Management Unit and the World Food Programme Vulnerability Assessment and Mapping Unit*.

138. **McCandless, E** (2012). *Peace Dividends and Beyond: Contributions of Administrative and Social Services to Peacebuilding*. New York: UN/PBSO.

139. McGuirk, E., & Burke, M. (2017). *The economic origins of conflict in Africa* (No. w23056). National Bureau of Economic Research

140. McGuirk, E., & Burke, M. (2020). The economic origins of conflict in Africa. *Journal of Political Economy*, *128*(10), 3940-3997.

141. McKenzie, D., & Sansone, D. (2019). Predicting entrepreneurial success is hard: Evidence from a business plan competition in Nigeria. *Journal of Development Economics*, *141*, 102369.

142. Mercier, M., Lionel Ngenzebuke, R. L. Verwimp, P. (2017) Violence exposure and deprivation: Evidence from the Burundi civil war.

143. Messer, E., Cohen, M. J., & d'Costa, J. (1998). *Food from peace: Breaking the links between conflict and hunger* (Vol. 24). Intl Food Policy Res Institute.

144. Michalopoulos, S., & Papaioannou, E. (2013). Pre-colonial ethnic institutions and contemporary African development. *Econometrica*, *81*(1), 113-152.

145. Michalopoulos, S., & Papaioannou, E. (2014). National institutions and subnational development in Africa. *The Quarterly journal of economics*, *129*(1), 151-213.

146. Michalopoulos, S., & Papaioannou, E. (2016). The long-run effects of the scramble for Africa. *American Economic Review, 106*(7), 1802-48.

147. Michalopoulos, S., & Papaioannou, E. (2020). Historical legacies and African development. *Journal of Economic Literature*, *58*(1), 53-128.

148. Miguel, E., Satyanath, S., & Sergenti, E. (2004). Economic shocks and civil conflict: An instrumental variables approach. *Journal of Political Economy*, *112*(4), 725-753.

149. Moris, J. R. (2019). *Irrigation development in Africa: lessons of experience*. Routledge.

150. Moseley, W. G., Carney, J., Becker, L., & Hanson, S. (2010). Neoliberal policy, rural livelihoods, and urban food security in West Africa: A comparative study of The Gambia, Côte d'Ivoire, and Mali. *Proceedings of the National Academy of Sciences*, *107*(13), 5774-5779.

151. Muchlinski, D., Siroky, D., He, J., & Kocher, M. (2016). Comparing random forest with logistic regression for predicting class-imbalanced civil war onset data. *Political Analysis*, *24*(1), 87-103.

152. Mueller, H.F.; Rauh, C. (2019). The hard problem of prediction for conflict prevention. *CEPR Discussion Paper, DP13748.*

153. Mullainathan, S., & Spiess, J. (2017). Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, *31*(2), 87-106.

154. Munshi, K. (2004). Social learning in a heterogeneous population: Technology diffusion in the Indian Green Revolution. Journal of Development Economics, 73(1), 185–213.

155. Musafiri, N. P. (2008). Land Rights and the Forest Peoples of Africa: Historical, Legal and Anthropological Perspectives.

156. Mustafa, D., & Qazi, M. U. (2007). Transition from karez to tubewell irrigation: development, modernization, and social capital in Balochistan, Pakistan. *World Development*, *35*(10), 1796-1813.

157. National Research Council. (1996). *Lost crops of Africa: volume I: grains*. National Academies Press.

158. Nunn, N. (2009). The importance of history for economic development. *Annu. Rev. Econ.*, *1*(1), 65-92.

159. Nunn, N. (2012). Culture and the historical process. *Economic History of Developing Regions*, *27*(sup-1), 108-126.

160. Nunn, N., & Qian, N. (2014). US food aid and civil conflict. *The American Economic Review*, *104*(6), 1630-1666.

161. O'brien, K. J., & Li, L. (2006). *Rightful resistance in rural China*. Cambridge University Press.

162. OECD/FAO (2016). *OECD-FAO Agricultural Outlook 2016-2025*, OECD Publishing, Paris. *http://dx.doe.org/10.1787/agr_outlook-2016-en*

163. Ogutu, S. O., Okello, J. J., & Otieno, D. J. (2014). Impact of information and communication technology-based market information services on smallholder farm input use and productivity: The case of Kenya. *World Development*, *64*, 311-321.

164. Østby, G. (2008). Polarization, horizontal inequalities and violent civil conflict. *Journal of Peace Research*, *45*(2), 143-162.

165. Paige, J.M. (1975). *Agrarian Revolutions.* New York: Free Press.

166. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, *12*, 2825-2830.

167. Perry, C. (2013). Machine learning and conflict prediction: a use case. *Stability: International Journal of Security and Development*, *2*(3), 56.

168. Pinstrup-Andersen, P., & Shimokawa, S. (2008). Do poverty and poor health and nutrition increase the risk of armed conflict onset? *Food Policy*, *33*(6), 513-520.

169. Poverty and household living conditions in the North-Kivu Province. (2009). UNDP Report.

170. Raeymaekers, T. (2008). Conflict and food security in Beni-Lubero: back to the future? *Alinovi, Hemrich y Russo (2008), págs*, 169-195.

171. **Raleigh, C** (2010). 'Political Marginalization, Climate Change, and Conflict in Sahelian States'. *International Studies Review* 12(1): 69–86.

172. Raleigh, C., & Dowd, C. (2017). Armed Conflict Location and Event Data Project (ACLED) Codebook 2017.[sl]: ACLED. *Tilgjengelig fra http://www. acleddata. com/wpcontent/uploads/2017/01/ACLED_Codebook_2017. pdf [lastet ned 26*.

173. Raleigh, C., Linke, A., Hegre, H., & Karlsen, J. (2010). Introducing ACLED: an armed conflict location and event dataset: special data feature. *Journal of peace research*, *47*(5), 651-660.

174. Richards, P. (1986). *Coping with hunger: hazard and experiment in an African rice-farming system*.

175. Richards, P. (1996). Culture and community values in the selection and maintenance of African rice. *Valuing local knowledge: Indigenous people and intellectual property rights*, 209-229.

176. Richards, P. (1998). *Fighting for the rain forest: war, youth & resources in Sierra Leone* (No. Reprinted Ed.). James Currey Ltd.

177. Rojas, M., & Guardiola, J. (2017). Hunger and the experience of being well: Absolute and relative concerns. *World Development*, *96*, 78-86.

178. Rosenbaum, P. R. (2002). Covariance adjustment in randomized experiments and observational studies. *Statistical Science*, *17*(3), 286-327.

179. Rosenbaum, P. R. (2002). Observational studies. In *Observational Studies* (pp. 1-17). Springer New York.

180. Rosenbaum, P. R. (2002). Sensitivity to hidden bias. In *Observational studies* (pp. 105-170). Springer, New York, NY.

181. Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41-55.

182. RUAN, J., XIE, Z. & ZHANG, X. (2015). 'Does Rice Farming Shape Individualism and Inno-vation', Food Policy 56, 51-58

183. Rubin, D. B. (2001). Using propensity scores to help design observational studies: application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, *2*(3-4), 169-188.

184. Ruby, J., & David, K. (2015). Analysis of Influencing Factors in Predicting Students Performance Using MLP-A Comparative Study. *International Journal of Innovative Research in Computer and Communication Engineering*, *3*(2), 1085-1092.

185. Ruiz, A., & Villa, N. (2008). Storms prediction: Logistic regression vs random forest for unbalanced data. *arXiv preprint arXiv:0804.0650*.

186. Salehyan, I. (2008). From climate change to conflict? No consensus yet. *Journal of Peace Research*, *45*(3), 315-326.

187. Sarsons, H. (2015). Rainfall and conflict: A cautionary tale. *Journal of development Economics*, *115*, 62-72.

188. Schomerus, M., & Allen, T. (2010). Southern Sudan at odds with itself: Dynamics of conflict and predicaments of peace.

189. Scott, J. C. (1977). *The moral economy of the peasant: Rebellion and subsistence in Southeast Asia*. Vol. 315. Yale University Press.

190. Skaperdas, S. (1992). Cooperation, conflict, and power in the absence of property rights. *The American Economic Review*, 720-739.

191. Stearns, J. (2012). *Dancing in the glory of monsters: The collapse of the Congo and the great war of Africa*. Public affairs.

192. Stearns, J. (2012). North Kivu: The background to conflict in North Kivu province of eastern Congo. Usalama Project, Rift Valley Institute, London, 2012.

193. Stewart, F. (1998). Food aid during conflict: can one reconcile its humanitarian, economic, and political economy effects. *American Journal of Agricultural Economics*, *80*(3), 560-565.

194. Stewart, F. (2011). Inequality in Political Power: A fundamental (and overlooked) dimension of inequality. *The European Journal of Development Research*, *23*(4), 541.

195. Swamynathan, M. (2019). *Mastering machine learning with python in six steps: A practical implementation guide to predictive data analytics using python*. Apress.

196. Taeb, M. (2004). Agriculture for Peace: Promoting agricultural development in support of peace.

197. Talhelm, T. (2015). The rice theory of culture (Doctoral dissertation). *Charlottesville, VA: University of Virginia*.

198. Talhelm, T. (2019). Emerging Evidence of Cultural Differences Linked to Rice Versus Wheat Agriculture. *Current opinion in psychology*.

199. Talhelm, T., Zhang, X., Oishi, S., Shimin, C., Duan, D., Lan, X., & Kitayama, S. (2014). Large-scale psychological differences within China explained by rice versus wheat agriculture. *Science*, *344*(6184), 603-608.

200. The Economic Value of Peace (TEVP). (2018). Available online: https://reliefweb.int/sites/reliefweb.int/files/resources/Economic-Value-of-Peace-2018.pdf.

201. The International Fund for Agricultural Development (IFAD), United Nations. IFAD in the Democratic Republic of the Congo.

202. Thompson, E. P. (1991). The Making of the English Working Class [1963].

203. Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

204. Tsusaka, T. W., Kajisa, K., Pede, V. O., & Aoyagi, K. (2015). Neighborhood effects and social behavior: The case of irrigated and rainfed farmers in Bohol, the Philippines. *Journal of Economic Behavior & Organization*, *118*, 227-246.

205. Turner, M. D., Ayantunde, A. A., Patterson, K. P., & Patterson III, E. D. (2011). Livelihood transitions and the changing nature of farmer–herder conflict in Sahelian West Africa. *The journal of development studies*, *47*(2), 183-206.

206. UNICEF, M. (2000). Multiple indicator cluster survey (MICS).

207. United Nations (2015). North Kivu Factsheet: MONUSCO. United Nations Organisation Stabilization Mission in the Democratic Republic of Congo.

208. Vafeiadis, T., Diamantaras, K. I., Sarigiannidis, G., & Chatzisavvas, K. C. (2015). A comparison of machine learning techniques for customer churn prediction. *Simulation Modelling Practice and Theory*, *55*, 1-9.

209. Van Acker, F. (2005). Where did all the land go? Enclosure & social struggle in Kivu (DR Congo). *Review of African Political Economy*, *32*(103), 79-98.

210. Van Weezel, S. (2017). Predicting Conflict Events in Africa at Subnational Level. *Available at SSRN 3019940*.

211. Van Weezel, S. (2017). Estimating the effect of conflict on food supply at the national level.

212. Verhofstadt, E., & Maertens, M. (2014). Can agricultural cooperatives reduce poverty? Heterogeneous impact of cooperative membership on farmers' welfare in Rwanda. *Applied Economic Perspectives and Policy*, ppu021.

213. Vlassenroot, K., & Huggins, C. (2004). Land, migration and conflict in Eastern DR Congo. *African Centre for Technology Studies*, *3*(4), 1-4.

214. Vlassenroot, K., & Huggins, C. (2005). Land, migration and conflict in eastern DRC. *From the ground up: land rights, conflict and peace in sub-Saharan Africa*, 115-195.

215. Vlassenroot, K., & Raeymaekers, T. (2008). Crisis and food security profile: The Democratic Republic of the Congo. *Beyond relief: food security in protracted crises*, 157-168.

216. Vlassenroot, K., & Raeymaekers, T. (Eds.). (2004). *Conflict and social transformation in Eastern DR Congo*. Academia Press.

217. Von Carnap, T. (2017). Irrigation as a Historical Determinant of Social Capital in India? A Large-Scale Survey Analysis. *World Development*, *95*, 316-333.

218. Voors, M. J., Nillesen, E. E., Verwimp, P., Bulte, E. H., Lensink, R., & Van Soest, D. P. (2012). Violent conflict and behavior: a field experiment in Burundi. *The American Economic Review*, *102*(2), 941-964.

219. Ward, M. D., Greenhill, B. D., & Bakke, K. M. (2010). The perils of policy by p-value: Predicting civil conflicts. *Journal of peace research*, *47*(4), 363-375.

220. Wendimu, M. A., Henningsen, A., & Gibbon, P. (2016). Sugarcane outgrowers in Ethiopia: “Forced” to remain poor? *World Development*, *83*, 84-97.

221. White, S. C., Fernandez, A., & Jha, S. (2016). Beyond the grumpy rich man and the happy peasant: mixed methods and the impact of food security on subjective dimensions of wellbeing in India. *Oxford Development Studies*, *44*(3), 332-348.

222. Wickham-Crowley, T. P. (1992). *Guerrillas and revolution in Latin America: A comparative study of insurgents and regimes since 1956*. Princeton University Press.

223. Winne, M. (2010). *Food Rebels, Guerrilla Gardeners, and Smart-Cookin' Mamas: Fighting Back in an Age of Industrial Agriculture*. Beacon Press.

224. Wischnath, G., & Buhaug, H. (2014). Rice or riots: On food production and conflict severity across India. *Political Geography*, *43*, 6-15

225. Wood-Sichra, U., A.B. Joglekar and L. You. (2016). "Spatial Production Allocation Model (SPAM) 2005: Technical Documentation". *HarvestChoice Working Paper*. Washington, D.C.: International Food Policy Research Institute (IFPRI) and St. Paul: International Science and Technology Practice and Policy (InSTePP) Center, University of Minnesota

226. Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

227. World Bank (2011). *World Development Report 2011: Conflict, Security, and Development*. Washington, DC: World Bank.

228. World Bank (2016). "Poverty and Shared Prosperity 2016: Taking on Inequality."

229. World Food Programme (WFP). 2014. Democratic Republic of Congo: Comprehensive Food Security and Vulnerability Analysis (CFSVA). Rome: WFP.

230. World Food Programme (WFP). 2016. Democratic Republic of Congo: Comprehensive Food sufficiency and Vulnerability Analysis (CFSVA).

231. World Food Programme (WFP). 2018. Democratic Republic of the Congo. Available online at http://www1.wfp.org/countries/democratic-republic-congo

232. Xie, Y., Brand, J. E., & Jann, B. (2012). Estimating heterogeneous treatment effects with observational data. *Sociological methodology*, *42*(1), 314-347.

233. You, L., & Wood, S. (2005). Assessing the spatial distribution of crop areas using a cross-entropy method. *International Journal of Applied Earth Observation and Geoinformation*, *7*(4), 310-323.

234. You, L., Ringler, C., Wood-Sichra, U., Robertson, R., Wood, S., Zhu, T., ... & Sun, Y. (2011). What is the irrigation potential for Africa? A combined biophysical and socioeconomic approach. *Food Policy*, *36*(6), 770-782.

235. Zhang, Y., & Kibriya, S. (2016). *The impact of slave trade on current civil conflict in sub-Saharan Africa* (No. 333-2016-14542).

236. Zhang, Y., Xu, Z. P., & Kibriya, S. (2021). The long-term effects of the slave trade on political violence in Sub-Saharan Africa. *Journal of Comparative Economics*.