Less Conflict During Harvest Season?

Analysing the Relationship between Satellite Data on Crop Season and Conflict on the Micro-Level in Kenya

Nicole Karnaus*& Sophia Seelos†

March, 2024

This study explores the relationship between harvest cycles and local conflict in Kenya from 2000 to 2019. Leveraging satellite imagery and vegetation software to analyze plant greenness (NDVI), we pinpoint the timing of harvest. Conducting a disaggregated analysis at the lowest available administrative level, we find that harvest has a negative impact on conflict. The decrease in conflict is predominantly driven by a reduction in protests and riots, shedding light on the type of confrontations affected by agricultural cycles. Our analysis indicates distinct temporal patterns in conflict dynamics throughout the crop cycle. The pre-harvest phases witness increased conflicts, while harvest and post-harvest periods experience a decline. These variations may be linked to liquidity constraints before harvest and increased funds post-harvest from output sales. This research offers valuable insights into the seasonality of conflicts in agriculturally dependent regions.

*Goethe University Frankfurt am Main, *karnaus@econ.uni-frankfurt.de*.
†University of Innsbruck,*sophia.seelos@uibk.ac.at*.

1. Introduction

Economic research on conflict traditionally revolves around the causes and consequences of conflict. A heavily debated subject in this context is the link between income and (civil) conflict. Scholars have repeatedly addressed the question of whether higher income leads to fewer conflicts. Initial cross-country studies, e.g. Collier & Hoeffler (2004); Fearon & Laitin (2003), have indeed indicated a negative correlation between per capita income and the incidence of conflicts (see Ray & Esteban, 2017, for an extensive review on empirical research about the relation between per capita income and conflict). A variety of more recent studies have used global commodity price shocks as proxies for income shocks to re-examine the causality. Results show that depending on the type of commodity, price shocks can increase or decrease the likelihood of local conflict, see, e.g., Bazzi & Blattman (2014); Brückner & Ciccone (2010); Dube & Vargas (2013). While exogenous international price shocks have proved to be a powerful tool for proving causality, they are specific to industries and economic sectors that are affected by the world market. This study contributes to the literature on income and conflict by focusing on the important but understudied sector of small-scale agriculture. Globally, the number of smallholder farms of less than 2 hectares is estimated at 525 million (Nagayets, 2005). In sub-Saharan Africa, a region that has been extensively used for conflict research, there are 33 million smallholder farmers (Nagayets, 2005). Most of them live in poverty - defined as living with less than $1.90 per person per day (Harris, 2019). If income is indeed linked to conflict, evidence on smallholder farmers may be particularly important, not only because of their current large number and low income but also because of their vulnerability to further deterioration. Climate change and the associated increase in the frequency of weather anomalies are expected to lead to more crop and income losses. Estimates of the total change in yields for the five most important (highest in calories, protein, and fat) rain-fed crops in the sub-Saharan region range from -8% to -22% over the next forty years (Schlenker & Lobell, 2010). The aim of our research is therefore to clarify whether changes in smallholder farmers' income from harvest have an impact on the incidence of conflict.

Contrary to other sectors or employment forms, smallholder farming does not provide income on a regular base (e.g. through wages) but is marked by seasonality, with harvest as the main income-generating event. Existing data on income in small-scale agriculture is often limited and restricted to surveys and case studies. We circumvent this limitation by using satellite data for the detection of the time of harvest. Our study focuses on Kenya, a country – like many others in the region – in which a large share of the population generates much of their income in small-scale agricultural production (FtMA, 2022). Smallholder farmers and their crops are still heavily dependent on rain and the time of harvest is mainly driven by the season's weather (FtMA, 2022). This dependence on seasonal weather and the related varying development of the crops provides some exogeneity in the timing of harvest. We elaborate on the assumption that the time of harvest is largely exogenous and investigate the relation between harvest season and the likelihood of local conflict with panel data at the ward (smallest administrative unit) level for Kenya. Conflict data is derived from ACLED, the Armed Conflict Location Events Data (Raleigh et al., 2010), which contains geocoded conflict events. To determine the harvest seasons for each ward, we use the Normalised Difference Vegetation Index (NDVI) derived from remote sensing (satellite) data Copernicus (2023). This index measures the changing greenness of plants and indicates crop development. Based on the NDVI values, the TIMESAT software from Eklundh & Jönsson (2012) is used to determine the time of harvest.[1] Combining the data, we end up with a database that provides harvest-conflict observations on a 10-day basis for the years 2000 to 2019. Controlling for ward-fixed effects and year-fixed effects enables us to elaborate on the relationship between harvest seasons and conflicts.

We argue that the underlying mechanism that relates income to conflict in our setting is the common argument of opportunity costs.[2] According to this argument, individuals weigh the benefits of rebellion (e.g., salaries, shares of appropriated resources, increased political power) against its costs (e.g., punishment, lost income, lost property). Given this trade-off, negative income shocks tend to divert productive resources, such as time, into conflict if income is sufficiently reduced. (see Becker, 1968; Grossman, 1991; Hirshleifer, 1991, for seminal models on the opportunity cost mechansim). Simply put, if the returns to aggression and war outweigh the returns to labor, the result will be more conflict. We argue that harvest increases the opportunity costs of conflict because it is associated with an increase in income. To further explore the impact of smallholders' seasonal income patterns and complement the analysis of harvest effects, this study also

[1]See Section 2.1 for more details on the detection of harvest and the modification of the data set.

[2]Other mechanisms that have been used to explain the relationship between income or commodity prices and conflict include the increased ability of the state to avoid conflict through higher tax revenues (negative relationship between income and conflict) and, on the other hand, increased incentives for violent resource appropriation due to increased value (positive relationship between income and conflict).

examines the general crop cycle and its relationship to conflict.

Our key finding is that at the time of harvest, there are significantly fewer conflicts. These negative effects are specifically persistent when analyzing protests and riots. Further, our additional analysis indicates that the crop cycle, in general, affects conflict likelihood. The number of conflicts increases during the start of the season and shortly before harvest, while we observe fewer conflicts post-harvest. These results may be related to liquidity constraints just before harvest and discretionary money from output sales in the post-harvest season.

This study most closely relates to the empirical literature on the relationship between commodity price shocks and conflict, as well as the relationship between weather shocks, agriculture and conflict. The findings on commodity price shocks are ambiguous. Studies such as Bazzi & Blattman (2014) disaggregate trade shocks (export price shocks) by product and find no significant positive relationship between natural resources and nationwide conflict. In contrast, Brückner & Ciccone (2010) show that negative export price shocks relate to more civil war outbreaks. Berman & Couttenier (2015) look at conflict patterns in 13 African countries and find that negative trade shocks are associated with more conflict outbreaks, but only sub-nationally. They state that such shocks may affect the location of violence but are not strong enough to influence the outbreak of national conflicts. The conclusion that the type of conflict, but also the type of commodity matter, is further corroborated by studies such as Angrist & Kugler (2008); Berman et al. (2017); Dube & Vargas (2013); Ubilava et al. (2023). Dube & Vargas (2013) find that positive exogenous price shocks for coffee in Colombia lead to fewer civil conflicts, while a price surge in oil leads to more civil conflicts. Angrist & Kugler (2008) show that Colombian municipalities experiencing an increase in coca prices also face a surge in violence. Berman et al. (2017) look at mines in Africa and use exogenous variations in world prices and mining data to show that positive price shocks in this sector lead to more local conflict. Ubilava et al. (2023) test the effect of a change in international cereal prices on conflict incidence in Africa. They find that a rise in international cereal prices increases the risk of conflict around harvest time and shortly after. Ubilava et al. (2023) provide important insights, but they base their analysis on general crop calendars that do not vary over time and assume that harvest always occurs at the same time.

The second branch of related literature, discussing the possible link between weather events and conflicts, also provides mixed evidence (see Koubi, 2019, for an extensive review on climate and conflict). Theisen & Holtermann (2011) find no direct link between drought and civil war. In contrast, other studies do find evidence for a weather-civil war relationship. Miguel et al. (2004) explore weather-induced income variation using rainfall variation as an instrumental variable for economic growth. They find a significant increase in conflicts when hit by negative weather shocks like droughts. Other authors provide similar evidence (Burke et al., 2009; Besley & Persson, 2011; Crost et al., 2018; Hsiang et al., 2011; Jia, 2014; Vesco et al., 2021). However, the robustness of some of these findings has been questioned (Buhaug, 2010).[3] Harari & Ferrara (2018) show that the timing of the negative weather shock within a season is crucial. Referring to their results, only negative weather shocks experienced during the growing season increase conflict likelihood in sub-Saharan Africa. McGuirk & Nunn (2021) examine the special case of herder conflicts in Africa. They focus on disputes between migrating pastoral groups and crop farmers and find evidence that due to droughts, pastoralists guide their herds on not-yet harvested land, which destroys harvest and leads to conflicts.

In contrast to previous studies, our analysis is not limited to commodities that are influenced by international price changes. Additionally, we do not depend on the incidence of extreme events, such as weather anomalies or price shocks. These may affect the likelihood of conflict through various channels (such as state capacity, commodity mechanism, grievances, etc.) with potentially persistent effects.[4] Our study provides insights into a seasonal but, in general, regular income pattern. We fall in line with a recently developing literature examining the seasonality of conflict, e.g., Linke & Ruether (2021); Hastings & Ubilava (2023). Linke & Ruether (2021) use precipitation data to proxy the progress of a season. They show that during the Syrian war, battles were more likely to occur in the growing season. Hastings & Ubilava (2023) study agriculture in South-East-Asia and find a higher likelihood of armed conflict during harvest season. Both studies relate their findings to the rapacity (or greed) mechanism, suggesting that perpetrators are more likely to engage in conflict when there is more at stake. In contrast to these studies, our novel approach uses harvest detection based on satellite data,

[3] Buhaug (2010) criticizes that, among other things, results might be biased if one does not account for ongoing conflicts. These events may not be driven by weather shocks but by the fact that there was already conflict before. The author further criticizes that studies like Burke et al. (2009) exclude conflicts with less than 1000 fatalities and look at especially conflict-prone periods in Africa, which may be driven by the Cold War. See Gleditsch (2012) for a critical review of early weather-conflict studies.

[4] Guardado & Pennings (2023) argue that commodity price shocks may have persistent effects on future conflict participation, for example, by changing the fighter's subjective value of grievance. The authors suggest that the opportunity cost mechanism may be even stronger than estimated by studies using price shocks. They recommend using seasonal and anticipated shocks to study the opportunity cost mechanism in the future.

allowing for a largely exogenous variation in harvest. Additonally, the analysis profits from high-resolution data and provides findings on different conflict types, especially on protests and riots. Agriculture remains the most important sector in developing countries. Our study focuses on this widespread form of employment, providing insights into the prevalent seasonal income derived from agriculture and its relationship to conflict.

This paper proceeds as follows. Section 2.1 provides an overview of the data and descriptive statistics. Section 3 explains the underlying empirical strategy, while Section 4 presents the results, followed by an additional analysis of harvest and crop cycle related income patterns in Section 5. Section 6.1 presents robustness checks and discusses some limitations of our study. The final section concludes.

2. Data and Descriptive Statistics

2.1. Data

Our empirical analysis is based on two main data sources. First, satellite images containing information about the normalized difference vegetation index (NDVI) retrieved from Copernicus (2023) which are employed in the TIMESAT software (Eklundh & Jönsson, 2012) to detect the time of harvest. Second, data from the Armed Conflict Location Events Data (ACLED) (Raleigh et al., 2010), containing information on the location of conflict events, the type of conflict, and the involved actors.

Our level of observation is wards in Kenya. Wards are the smallest administrative unit available (administrative unit 3), following constituencies (administrative unit 2) and counties (administrative unit 1). NDVI is an indicator of the greenness of biomass and is used, among others, for harvest detection. Copernicus (2023) currently provides worldwide information on NDVI starting from 1999 until June 2020. Data is available every 10 days (dekate) at a pixel size of 1 kilometer $\times$ 1 kilometer.[5] We aggregate the 10-day NDVI observations on the ward level, using the maximum NDVI value[6] for each 10-day observation. Combining the generated time series of NDVI values for each

[5]In contrast to other raw NDVI data, Copernicus (2023) data relies on images from the PROBA-V Collection 1, which corrects the atmospheric conditions of the input observations.

[6]To avoid measurement errors caused by cloud coverage that typically decreases the NDVI level, we consider the pixel with the maximum NDVI value in a given ward. This approach reduces the error of detecting lower NDVI values than are actually present.

ward, we then detect harvest seasons for the years 2000 until 2019 using the software TIMESAT (Eklundh & Jönsson, 2012). Based on our 10-day NDVI values, the program detects harvest seasons by smoothing the time series and detecting the maxima of the time series. As our observation level is wards, we do not detect the exact time of harvest per field or farmer but rather the time when the majority in the ward has started the harvest. A detailed explanation of the TIMESAT settings is provided in Appendix B. We additionally cross-validate the detection of harvest with survey data from farmers in Section 2.1.

In total, there are 1,442 wards in Kenya; however, given our approach of using detected harvest seasons in agricultural wards, we restrict our sample to wards that contain at least 80% of agricultural land.[7] This leaves us with 671 wards (median size of these wards is $59.44 km^2$). We argue that 80% of agricultural land is a good cutoff as these wards show fewer variations in harvest dates (see Section 2.2). This selection decreases the risk of incorrect harvest detection. The NDVI index measures all kinds of greenness, including trees and non-agricultural plants. The risk that other plant cycles affect and bias the harvest detection is non-neglectable in non-agricultural wards. Therefore, we restrict our sample to guarantee more accurate harvest season detection. We aim to include conflicts in the form of farmer-related unrest and protests, which are more likely to appear in surrounding major cities within a ward or a neighboring ward. To account for the possible participation of farmers in nearby cities, we extend the borders of each ward by a buffer of 20 kilometers. All conflicts which lie within the ward and the buffer are treated as conflicts for the respective ward.[8]

The data on conflicts are taken from the Armed Conflict Location and Event Dataset (ACLED) (Raleigh et al., 2010), which contains information on the date and the geolocation of conflict events in Kenya. We focus on the 2000–2019 period[9], which overlaps with our NDVI data. The information used by ACLED comes from various sources, including press accounts from regional and local news, humanitarian agencies, or research publications. The ACLED database contains information about the event type, the outcome as well as the characteristics of actors on both sides. Using ACLED data, we know

[7] The share of agricultural land has been calculated using FAO (2018)

[8] To account for the issue that, due to the buffers, the same conflict may be taken into account for more than one ward, we use spatial correction in our analysis. We allow for both cross-sectional spatial correlation and location-specific serial correlation by adjusting standard errors following the method of Conley (1999) and Hsiang et al. (2011).

[9] Our results do not include 1999 due to the low detection rate of harvest seasons. This is because TIMESAT needs to base its calculation of harvest season on the first year.

exactly whether the event was a battle, protest, or riot, which groups were involved, and how many fatalities were recorded. The precision estimate on time and geolocation is used to consider only the most precise data points in the analysis. Following Berman et al. (2017), we consider only locations that have the highest precision (indication of 1). For time, we include all observations which have at least a precision of 2. This means that the event occurred within the same week or on the exact date reported. The variation of one week is acceptable since our time unit of observations is on a 10-day basis. This approach leaves us with 77.6% of all conflicts. Additionally, we drop all duplicate conflict-ward events ending up with 3,799 conflict events in total appearing in agricultural wards and surrounding areas of 20 kilometers. It is important to notice that even though many researchers use the ACLED database for their studies and the data is cross-checked, potential biases and measurement errors are still possible.

2.2. Descriptive Statistics

Our final sample covers 671 wards for 20 years at a 10-day observation interval. Table 1 displays some descriptive statistics. There are some important features that are worth mentioning: Firstly, the unconditional probability of observing at least one conflict in a given cell and a given 10-day interval within a year is around 0.2% percent. In most wards, no event occurs over the entire period. The probability of observing a conflict during harvest seasons in a given ward is also 0.2%. Note that one year consists of, on average, 36 10-day observations.[10] Panel B differentiates between the types of conflict. Referring to ACLED, *Battles* are defined as armed clashes in which the government regains territory or a non-state actor overtakes territory, *Protests* denote peaceful protests. *Riots* include violent protests and mob violence. *Strategic developments* are non-violent activities of violent groups that may trigger or affect conflicts. *Violence against civilians*, finally, are attacks of militia and armed groups against civilians. Appendix A provides a more detailed explanation of the ACLED definitions of conflict, conflict type, and the involved actors. The most likely conflict types in our data are protests and riots, which make up 0.15% combined. Panel C shows that the higher the buffer size, the higher the likelihood of conflicts. This makes sense, given that conflicts such as protests and riots mainly appear in major surrounding cities rather than in rural villages. Considering a buffer of 20 kilometers, the probability of a conflict increases to 6.67%.

[10] For comparison: If we would observe on average one harvest per year per ward, the probability of observing harvest seasons in a given ward would be 0.027 %, which is very close to the actual data.

Table 1: Descriptive Statistics for Agricultural Wards

	Observations	Mean	Std. Dev.
Panel A: All Conflicts			
Conflict in wards	483,120	0.0020	0.045
Conflict if harvest = 1	24,152	0.0020	0.045
Conflict if harvest = 0	458,968	0.0020	0.045
Panel B: Different Types Conflicts			
Battles	483,120	0.0002	0.014
Viol. against. civilians	483,120	0.0004	0.019
Riots and protests	483,120	0.0015	0.039
Protests	483,120	0.0006	0.026
Riots	483,120	0.0009	0.030
Explosions	483,120	0	0
Strategic development	483,120	0.0004	0.007
Panel C: All Conflicts within Buffer			
Conflict in wards with 5km buffer	483,120	0.0132	0.114
Conflict in wards with 10km buffer	483,120	0.0269	0.162
Conflict in wards with 15km buffer	483,120	0.0454	0.208
Conflict in wards with 20km buffer	483,120	0.0667	0.250

Note: Authors' computations from ACLED and own generated harvest seasons. There are a total of 483,120 observations from 671 wards, spanning 20 years and 36 10-day observations per year. The mean describes the occurrence of at least one conflict in an agricultural ward within a 10-day interval. Panel A distinguishes between 10-day ward observations if a harvest season appears or not. Panel B distinguishes the different conflict types. Panel C summarizes the probability of conflict given an underlying buffer around the ward borders.

Figure 1 illustrates the variation of detected harvest depending on the agricultural area share of wards. Considering all wards (no cutoff) or a low cutoff of at least 50% agricultural area share, most wards show between 1-3 standard deviations, This corresponds to 1-3 10-day intervals (10-30 days). When changing the cutoff to at least 80% agricultural area share per ward, we observe less deviation, with most wards showing only 1-2 standard deviations in harvest timing. As the harvest detection software determines the time of harvest based on NDVI variation in the entire ward, the detection in low or no cutoff wards may be driven by other plants' greenness cycle (trees, grasses, etc.). Figure 2 plots the deviation in harvest for four particular wards containing 99%, 90%, 80%, and 70% agricultural land. We find similar timing for harvest seasons for agricultural wards containing at least 80% agricultural land. Harvests in these wards mostly take place from December to February (10-day intervals number 34 until 36, and 1 until 6) and from July to August (10-day intervals number 19 until 24).[11] This aligns with the general harvest seasons in Kenya (FAO, 2023b). Considering a ward with only 70% agricultural area provides a noisier detection of harvest months. Even though the harvesting months are similar to the higher agricultural wards, we find a less clear pattern of harvest months over the years. Therefore, the restriction to wards with at least 80% agricultural areas is reasonable.

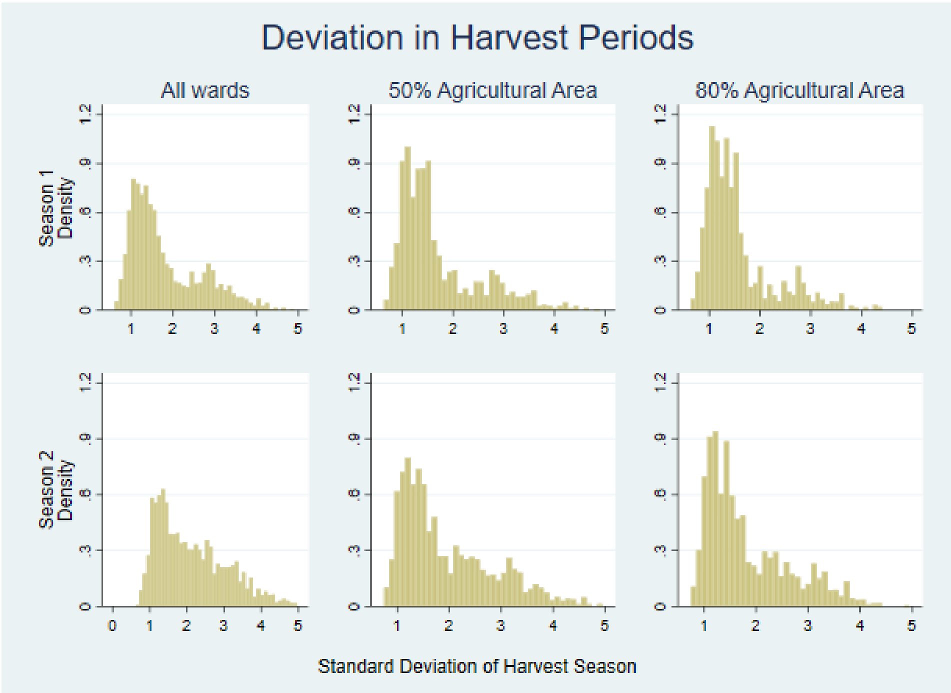

Figure 1: Deviation in Harvest Periods Depending on the Agricultural Area of Wards.

Note: This figure shows the deviation of harvest periods inside a ward. One standard deviation (on the x-axis) corresponds to a shift in harvest by 10 days (1 dekate). The variation in harvest periods is reduced when the sample is limited to wards with at least 50% agricultural area. Further reduction in variation is observed when the cutoff is increased to 80% of agricultural area in the wards.

To test the accuracy of our derived measure of harvest, we generate the same harvest variable for Tanzania and compare it with the most recent Tanzanian LSMS National Panel Survey from 2014[12]. In the survey, households indicate the month(s) of harvest

[11] Note that one year consists of 36 10-day intervals. The 10-day interval number 1 is, for example, January 1 until January 10.

[12] National Bureau of Statistics (NBS) [Tanzania]. 2015. National Panel Survey (NPS)-Wave 4, 2014-2015. Ref.TZA_2014_NPS-R4_v03_M. Downloaded from https://microdata.worldbank.org/index.php/catalog/2862 on September 9, 2023. Dar es Salaam, Tanzania: NBS. (www.nbs.go.tz)

for a plot in the given year. We take the median of the stated month(s) of harvest on the ward level (restricted to wards with at least 80% agricultural area as in the main analysis) and compare it with our detected time of harvest for the corresponding wards in 2014. For 81% of the 2163 observations, our detected harvest corresponds to the stated month(s) of harvest of the households; for 97%, it corresponds to the stated month(s) or a maximum deviation of one month. We are therefore confident that our detected harvest variable is a sound measure.[13]

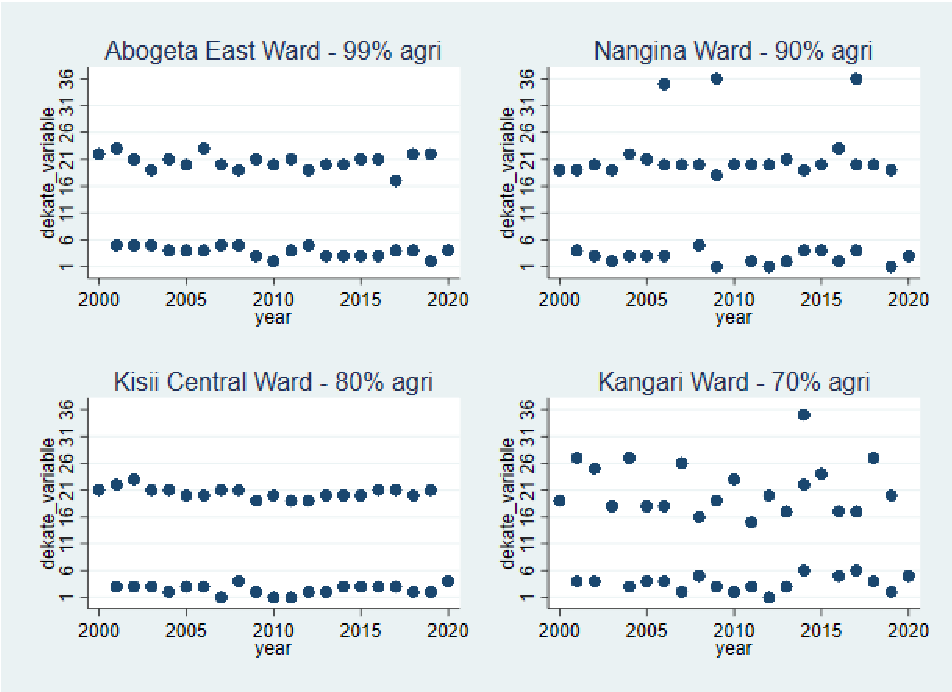

Figure 2: Harvest Season in Varying Degrees of Agricultural Wards

Note: This figure shows the detected time (dekate) of harvest by TIMESAT over years in a selection of wards with different proportions of agricultural land.

3. Empirical Strategy

In the following, we present our identification strategy and discuss some challenges.

Our empirical strategy is based on a large panel dataset including detected harvest seasons. Using a two-way fixed effect model, we assess whether harvest seasons affect

[13]We use Tanzanian data to test our measure, as Kenya is not part of the LSMS Survey.

violence in agricultural wards. The idea that harvest may affect conflict is based on the opportunity cost theory. During harvest time, farmers allocate time to generate output which then provides them with income from sales. We detect harvest on an aggregated ward level, thereby not accounting for the exact time of harvest per plot but the time when harvest has started on most plots. The detected dekate can be interpreted as main harvest season, which may affect the opportunity costs in two ways: First, farmers may allocate time to harvest (due to expected income). Second, they could have already started to generate income by spending time on selling the harvest output (see Section 5 for a discussion and analysis on income and crop cycle). In both cases, the opportunity costs of participating in a conflict increase. Therefore, we expect a decrease in conflicts at harvest time. We exploit the variations in harvest time and estimate a specification of the following form.

$$Conflict_{it} = \beta_1 Harvest_{it} + FE_i + FE_t + \varepsilon_{it} \tag{1}$$

where t denotes time and i ward. FE_i are ward fixed effects and FE_t are time fixed effects. The dependent variable $Conflict_{it}$ corresponds to the observation of conflict events at the 10-day ward level. Overall, the variable $Conflict_{it}$ is measured in terms of conflict intensity. Thus, all conflicts are summed up at the 10-day ward level.[14] We are aware that our dependent variable on conflicts is a count variable rather than a continuous variable with many zeros for areas not experiencing conflicts. Therefore, we provide further results in the robustness section using a Poisson pseudo-maximum-likelihood estimator accounting for the zero-inflated data. To account for the participation of smallholder farmers in conflicts, especially protests, in nearby major cities, we construct a buffer around each ward with a size of 20 kilometers. All conflicts that appear within a distance of 20 kilometers from the ward border will be counted as conflicts within the ward. Finally, the main explanatory variable $Harvest_{it}$ is a binary variable, turning 1 when harvest is detected at the ward-10-day interval level. Thus, β_1 is our main estimate of interest in equation 1 and measures the effect of the harvest season on the likelihood of conflicts.

[14] In contrast to our approach, Berman et al. (2017) rely on a specification of a simple incidence of conflicts. In other words, conflicts are not summed up. Instead, the appearance of one or more conflicts in a given area is used as a dummy variable. We verify that our estimates do not change significantly when considering the incidence definition of conflict similar to Berman et al. (2017). Results are reported in the robustness section using a linear probability model as well as a nonlinear conditional logit.

Establishing the causal impact of harvest on conflict involves some methodological challenges. To address causality, we focus on variations in harvest seasons. We argue that harvest appears largely exogenous. Due to (seasonal) weather variations, smallholder farmers do not know in advance when exactly harvest will take place. Many regions in sub-Saharan Africa, including Kenya, are mainly rainfed and not irrigated. Hence, the harvest timing highly depends on weather conditions. We show in the robustness section that our results are robust to restricting the sample to rainfed areas.[15] Kenya provides a good setting for our analysis because it shows multiple variations in harvest time. On the one hand, rainfall patterns and crops vary across Kenya (FAO, 2018, 2023a) and allow for the detection of different harvest seasons. Additionally, harvest often occurs on a biannual basis, further increasing the frequency and potential variation of harvest events. Besides general causality issues, reverse causation from local violence to harvest could affect our estimates. Regions that are affected by various conflicts might experience the destruction of harvest. Thus, no harvest is appearing even though there would have been a harvest period. We argue, however, that we should detect a deviation in harvest seasons in cases where agricultural products are destroyed before harvest. Assuming the growing season has started and a conflict destroying the agricultural products appears before harvest, we would detect a decrease in our NDVI measure. This decrease would be very abrupt and provide us with a harvest detection. We would then observe conflict during detected harvest, counteracting the expected effect in our models. This makes reverse causality a less critical issue as it would work in the opposite direction, only weakening the expected negative effect of harvest on conflict.

Another problem could be time-varying omitted variables that co-determine harvest season and violence in agricultural areas. The use of year-fixed effects and ward-fixed effects in the baseline specification controls for most of these problems. However, we cannot strictly rule out that unobservables affect our results.

Finally, spatial correlation can be an issue in our data, especially due to the buffers that allow one conflict to be assigned to multiple wards. Using geospatial data requires adjustments for spatial correlation since both conflict and harvest are clustered in space. Therefore, we estimate our results using a spatial correction allowing for both cross-sectional spatial correlation and location-specific serial correlation, applying the method developed by Conley (1999) and Hsiang et al. (2011). Our large panel database consisting

[15]The share of rainfed area is calculated similarly to the agricultural area using FAO Wapor data on land usage (FAO, 2018).

of 20 years and 671 wards makes the regression approach demanding. Therefore, we include standard errors corrected for spatial correction only for our main findings. Using Conley (1999) corrected HAC standard errors, we include a time dependence horizon of 1 year. For the spatial kernel, we adjust the radius to 40 kilometers which is also the median distance between agricultural county borders in Kenya. Considering the distance between agricultural ward borders, we receive a median distance of roughly 2 kilometers. Hence, our identification containing a 40-kilometer radius is rather conservative.

4. Results

Table 2 presents the baseline results for the main specification with various fixed effects. The dependent variable is the total number of conflicts in a ward (with a 20-kilometer buffer) within a 10-day period. The main explanatory variable is a binary variable, turning 1 when harvest is detected at the ward-dekate level.

The models report clustered standard errors at the ward level in parentheses. Clustered standard errors at the county level are shown in square brackets. The curly brackets show adjusted HAC standard errors following Conley (1999). We include ward-fixed effects in all models to control for time-invariant co-determinants of violence and harvest seasons at the smallest administrative level available. These co-determinants may include political instability, ethnic diversity, property rights, etc. Time-fixed effects control for variables that are constant across wards but vary over time. Examples of these variables include national elections, changes in national policies, and international agreements, which can lead to protests, riots, or conflicts and may affect harvest.

Model 1 accounts for ward-fixed effects and year-fixed effects. The coefficient of the harvest variable is negative and statistically significant at the 1%-level. The significance decreases to 5 % (10%) when clustering standard errors at the county level (controlling for spatial HAC-corrected standard errors). Model 2 does not account for year-fixed effects but introduces month-fixed effects instead. Clustering standard errors at the ward level, we get significant results at the 1%-level. The significance decreases to 5 % when accounting for clusters at the county level. The coefficient turns insignificant when accounting for spatial correlations. Given that many variations in conflict frequency appear on a yearly level (e.g., driven by election years), including year-fixed effects seems

Table 2: Identifying Main Specification of Harvest on Conflict

	Sum of Conflict (Buffer size: 20 km)			
	(1)	(2)	(3)	(4)
Harvest	-0.012	-0.009	-0.012	-0.012
	(0.002)***	(0.002)***	(0.002)***	(0.002)***
	[0.006]**	[0.005]*	[0.005]**	[0.005]*
	{0.007}*	{0.007}	{0.007}*	{0.007}*
Constant	0.039***	0.136***	0.086***	0.060***
	[0.006]	[0.009]	[0.010]	[0.000]
Ward FE	Yes	Yes	Yes	Yes
Year FE	Yes		Yes	
Month FE		Yes	Yes	
Year×County				Yes
N	483,120	483,120	483,120	483,120

Notes: Significance levels are defined as * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the ward level are provided in parentheses. Clustered standard errors at the county level are provided in brackets. Conley (1999) standard errors are shown in curly brackets, allowing for spatial correlation within a 40km radius and a 1-year serial correlation. The variable Harvest is a dummy variable that turns 1 when harvest is detected. The dependent variable is the sum of conflicts appearing within a ward and a 20-kilometer distance of the ward borders. There are a total of 483,120 observations from 671 wards, spanning 20 years and 36 10-day observations per year.

crucial. Model 3 includes both month and year-fixed effects. Neither size nor significance is affected compared to model 1, which includes year-fixed effects only. Finally, model 4 extends the specification of year-fixed effects by interacting them with counties. This controls for time-varying omitted variables co-determining harvest time and local violence in wards.[16] We obtain similar results in size and significance as compared to the specification with year-fixed effects only.

To shed more light on different kinds of conflict, we distinguish between protests, riots, strategic development, violence against civilians, and battles. Table 3 shows the results of this heterogeneity analysis.

Considering harvest season, we find statistically significant effects on protests, riots, and strategic development. The results on protests and riots are statistically significant at the

[16] Berman et al. (2017) use a similar approach by interacting year-fixed effects with country-fixed effects, using the country-level given that they analyze the whole continent of Africa.

Table 3: The Effect of Harvest on Different Types of Conflict

	(1) Protest	(2) Riots	(3) Strategic Dev.	(4) Violence	(5) Battle
Harvest	-0.005	-0.008	-0.001	0.003	-0.001
	(0.001)***	(0.001)***	(0.000)***	(0.001)***	(0.001)
	[0.003]*	[0.003]**	[0.000]***	[0.002]	[0.002]
	{0.003}*	{0.004}*	{0.000}***	{0.003}	{0.002}
Constant	0.005***	0.013***	0.000	0.011***	0.010***
	[0.002]	[0.003]	[0.001]	[0.002]	[0.002]
Ward FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
N	483,120	483,120	483,120	483,120	483,120

Notes: Significance levels are defined as * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the ward level are provided in parentheses. Clustered standard errors at the county level are provided in brackets. Conley (1999) standard errors are shown in curly brackets, allowing for spatial correlation within a 40km radius and a 1-year serial correlation. The variable Harvest is a dummy variable that turns 1 when harvest is detected. The dependent variable is the sum of the respective conflict type appearing within a ward and a 20-kilometer distance of the ward borders. *Protests* denote peaceful protests. *Riots* include violent protests and mob violence. *Strategic developments* are non-violent activities of violent groups that may trigger or affect conflicts. *Violence against civilians* are attacks of militia and armed groups against civilians. *Battles* are defined as armed clashes in which the government regains territory or a non-state actor overtakes territory. There are a total of 483,120 observations from 671 wards, spanning 20 years and 36 10-day observations per year.

1%-level when clustering at the ward level. Considering county clusters or spatial HAC corrected standard errors reduces the significance to 5-10 % for both types of conflicts. Model 3 analyses the effect of harvest on strategic development. The magnitude of strategic development is quite small. The coefficient is, however, statistically significant at the 1%-level even when accounting for HAC-corrected standard errors. Violence against civilians (Model 4), as well as battles (Model 5), do not show robust effects. The coefficient for battle is negative but not significant. The coefficient of violence against civilians is positive and significant at the 1%-level when clustering at the ward level, but turns insignificant when clustering at the county level or using HAC-corrected standard errors. Thus, the time of harvest seems to have no effect on conflicts such as battles and violence against civilians. The results on the type of conflict are in line with our conjectures. We expect farmers to be more likely to participate in conflicts in the absence of harvest period. The main actors for the conflict type violence against civilians and battles are the police, armed forces, and militia. In contrast, protests and riots mainly involve civilians, which includes farmers.

5. Additional Analysis: Crop Cycle and Income

The main analysis focuses on the effect of harvest on conflict and demonstrates that there is less conflict during this period. We argue that this is because harvest is the primary income-generating event for smallholder farmers, which increases the opportunity costs of participating in conflict. Drawing on the literature on crop cycles and farmers' behavior, the analysis is extended to examine the general income and conflict pattern over the crop cycle. Our aim is to provide a thorough insight into the crop cycle dynamics that complement our findings on the negative effects of harvesting on conflict.

Stephens & Barrett (2011) show that Kenyan farmers tend to sell their products at lower prices just after harvest instead of utilizing inter-temporal price arbitrage through storage. They present evidence that farmers face liquidity constraints before and around harvest, which makes it reasonable for them to sell even when prices are low. This, in turn, leads to higher consumption right after harvest but decreasing consumption afterward. Note that even though farmers may have non-agricultural income, which affects this consumption pattern, agricultural output sales still pose the largest source of income for small-holder farmers in Kenya (Karfakis et al., 2017). The results on liquidity constraints and the related sales and consumption pattern are in line with Mani et al. (2013) who investigate the cognitive function of Indian farmers over the planting cycle. The authors argue that farmers experience cycles of poverty because they do not (or cannot) smooth consumption: farmers are richest after harvest and poorest before harvest. Based on these findings, we argue that farmers' discretionary money from output sales (as well as consumption) peaks just after harvest and declines afterward. Thus, just before harvest, farmers have the smallest amount of discretionary funds and consume the least.[17] This leads us to anticipate less conflict during the harvest and post-harvest phases, but more conflict before the harvest period.

[17]Figure8 in the Appendix displays the distribution of conflicts and the likelihood of detected harvest across months. There is no discernible pattern indicating higher conflict rates during specific seasons, except for a tendency of more conflicts at the beginning of the year (which coincides with the first harvest) and towards the end of the year.

To examine this crop cycle-conflict pattern, we leverage two additional seasonality parameters detected by TIMESAT: the start of the season and the timing of the highest NDVI value within a season. The latter, occurring just a few weeks before harvest, serves as an excellent indicator of the pre-harvest period. The start of the season aligns approximately with the midpoint between two consecutive harvest periods when funds from previous harvest sales should have already decreased substantially (see Figure 3 in the Appendix for an illustration of the smoothed NDVI values for one season and the detected parameters). These seasonality parameters enable us to investigate the hypothesized relationship between the crop cycle and conflict dynamics in Table 4.

Model 1 in Table 4 includes dummy variables for both seasonality parameters. Both coefficients are positive and significant (at the 5%-level for start of season and at the 1%-level for the pre-harvest time). This shows that there are more conflicts between and before harvest, while harvest itself is still showing a negative effect on conflict frequency, significant at the 10%-level.

It is still not clear if and how the period after harvest affects conflict. Following our argument, conflict frequency should be lower after harvest when farmers have higher discretionary funds from harvest sales. To further explore this argument, we introduce a synthetic measure for discretionary money. The variable is normalized to 1 in the 10-day period just after harvest and declines afterwards – with the strongest decrease in the first few post-harvest periods. The following equation describes this process:

$$DiscretionaryMoney_{it} = \exp(-0.25(t - t_h - 1)) \tag{2}$$

where t is the dekate of interest and t_h is the dekate of the last detected harvest. The term $(t - t_h - 1)$ denotes the number of dekates between t and the last post-harvest dekate. Of course, the chosen exponential form of the discretionary money function is somewhat arbitrary and may take another form. The diminishing rate of decline captures, however, the implied pattern by Stephens & Barrett (2011) with the strongest decrease in funds just after harvest.[18]

We argue that the opportunity costs of conflict are higher when discretionary money allows the farmers to consume compared to the times when discretionary money is low

[18]With our specification, discretionary money is 1 in the first post-harvest dekate, has decreased by 40% after the second post-harvest and by 99% after the tenth post-harvest dekate. Note that the median number of dekates between two subsequent harvests is 18.

Table 4: Crop Cycle

	Sum of Conflict (Buffer size: 20 km)			
	(1)	(2)	(3)	(4)
Harvest	-0.010*	-0.017**	-0.017**	-0.015**
	(0.005)	(0.006)	(0.006)	(0.006)
Start of Season	0.014**			0.012**
	(0.006)			(0.006)
Pre-Harvest	0.017***			0.013***
	(0.004)			(0.005)
Discretionary Money		-0.022***	-0.022***	-0.019***
		(0.007)	(0.007)	(0.007)
Constant	0.038***	0.063***	-0.025	-0.027
	(0.006)	(0.017)	(0.046)	(0.046)
Max NDVI			Yes	Yes
Ward FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	483,120	468,094	468,094	468,094

Notes: Significance levels are defined as * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the county level are provided in parentheses. The variable Harvest is a dummy variable that turns 1 when harvest is detected. Start of Season and Pre-Harvest are dummy variables being equal to 1 if sowing time starts and NDVI reaches its peak, respectively. Discretionary money is normalized to one in the period after harvest and declines as specified in Equation 2. Max NDVI is the maximum NDVI value of the previous harvest season and serves as an indicator of the goodness of the previous harvest. The dependent variable is the sum of conflicts appearing within a ward and a 20-kilometer distance of the ward borders. There are a total of 483,120 observations from 671 wards, spanning 20 years and 36 10-day observations per year. The number of observations decreases when including discretionary money because the variable is not defined for observations before the first detected harvest in a ward.

and farmers face liquidity constraints.[19] We therefore expect conflict to decrease in discretionary money. Model 2 of Table 4 includes the synthetic measure for discretionary money. The coefficient is negative and significant at the 1%-level, indicating that higher discretionary money decreases conflict frequency. The discretionary money variable is normalized and does not account for the quality or output of the respective harvest,

[19] In addition to the opportunity mechanism, the suggested crop cycle-conflict pattern could also be explained by the grievance theory. This theory suggests that individuals rise up when their well-being deteriorates relative to others or to their past (Hendrix & Haggard, 2015; Winne & Peersman, 2021).

which may vary. Model 3, therefore, includes a measure of harvest quality. We use the maximum NDVI value of the previous harvest season to proxy the output of harvest. The coefficient of discretionary money in model 3 is robust to the inclusion of this control.[20] Model 4 includes all crop cycle variables that have been introduced. The results on harvest, start of season, pre-harvest, and discretionary money are confirmed. The coefficients are still significant at the 5%-level (1%-level for Pre-Harvest and Discretionary Money) and similar in size.

6. Robustness Checks and Limitations

This section applies further robustness checks to show that our main results are reliable. First, we distinguish between different specifications of our dependent variable in Table 5. Second, Table 6 varies the buffer size of the wards. Finally, Table 7 restricts the data to mainly rainfed areas to account for concerns about the exogeneity of the time of harvest. Following the robustness checks, we discuss some limitations of our study.

6.1. Robustness checks

Model 1 in Table 5 shows a specification with a binary dependent variable for conflicts. This shows the incidence of a conflict rather than the intensity. It equals 1 if the ward experiences a conflict within a 10-day time interval, otherwise, it equals 0. Considering a linear probability model and clustering at the ward level, we find statistically significant effects at the 1%-level. This implies that during harvest seasons, there is a decrease in the likelihood of conflict. However, when accounting for clusters at the county level, the significance level disappears. A similar pattern exists when considering a logit model. The size of the coefficient increases, implying that harvest season decreases the likelihood of conflicts by 8.4 percentage points. Nevertheless, this result is only significant when clustering at the ward level, which does not necessarily account for all spatial correlations. Model 3 shows a Poisson pseudo-maximum-likelihood estimator that accounts for zero-inflated observations. The wards experience no conflicting events during most of the periods. Additionally, we do not value conflicts as a continuous variable but rather as a count variable. The Poisson pseudo-maximum-likelihood estimator takes care of

[20] Appendix D includes all analyses of the main results with discretionary money instead of harvest and confirms that discretionary money has a negative effect on conflict frequency.

these drawbacks. Applying this method, we obtain a statistically significant effect at the 1%-level clustering at the ward level. The negative coefficient confirms, again, a decrease in conflicts. Clustering at the county level shows no significant results. Finally, we consider a transformation of our main dependent variable (model 4) and take the logarithm of the sum of conflicts. The logarithmic form helps to reduce the effect of outliers and the skewness of the data. The disadvantage of this approach is that it creates missing values if the variable contains values equal to zero. To overcome this problem, it is common to add a small constant. In our case, we add the value of 0.00001 to our sum of conflicts before taking the logarithm. Model 4 shows the result. Similar to the dependent dummy variable, we only find a significant effect when clustering at the ward level. Model 5 controls for potential biases that could result from transforming the independent variable in the way we did in Model 4. To overcome biases resulting from taking the logarithm of a variable by adding a small constant, the literature refers to the inverse hyperbolic sine transformation (IHST). This approach has the advantage that it approximates the natural logarithm of the variable, which is considered. The transformation proceeds as follows

$$IHST_i = log(y_i + (y_i^2 + 1)^{(\frac{1}{2})}) \tag{3}$$

where y_i denotes the value of a conflict. Model 5 shows the results using the inverse hyperbolic sine transformation. We see that when considering clusters at the county level, we still find a statistically significant negative effect at the 10%-level.

The robustness checks confirm our main results, showing that there is a negative effect of harvest season on conflict incidents and conflict intensity. While the magnitude of the coefficient varies depending on the specification, we can conclude that our estimates are robust in predicting a decrease in conflicts during harvest season.

We vary our buffers to show that significant levels remain, considering a smaller buffer. We construct buffers of the sizes 5 kilometers, 10 kilometers, 15 kilometers, and 20 kilometers and run a regression without any buffer. Results are shown in Table 6. There are no significant effects of harvest on conflict when considering no buffer or only a small buffer of 5 kilometers. However, results become statistically significant at the 10%-level when considering a 10-kilometer buffer or larger. Also, the size of the coefficient increases when expanding the buffer. The insignificant results for the coefficients with no buffer and a buffer of 5 kilometers may be explained by the fact that we are considering rural

Table 5: Model Specification Variation - Effect of Harvest on Conflict

	(1) LPM	(2) Logit	(3) Poisson ZIP	(4) log(Sum(Conflict))	(5) IHST
Harvest	-0.005	-0.084	-0.040	-0.058	-0.006
	(0.002)***	(0.028)***	(0.014)***	(0.015)***	(0.002)***
	[0.003]	[0.060]	[0.036]	[0.030]	[0.004]*
Constant	0.027***	-4.765***	0.297***	-11.193***	0.029***
	[0.004]	[0.193]	[0.043]	[0.047]	[0.004]
Ward FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
N	483,120	483,120	483,120	483,120	483,120

Notes: Significance levels are defined as * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the ward level are provided in parentheses. Clustered standard errors at the county level are provided in brackets. The variable Harvest is a dummy variable that turns 1 when harvest is detected. Model 1 uses the linear probability model and introduces a dummy variable as the dependent variable, turning 1 if a conflict appears within a ward and a 20-kilometer distance of the ward borders. Model 2 applies a logit probability model. Model 3 uses a pseudo maximum likelihood regression accounting for the large share of zero conflicts in the data. Model 4 applies a logarithm on the main dependent variable by adding a small constant (0.00001) to the zero values. Model 5 applies the inverse hyperbolic sine transformation on the main dependent variable. There are a total of 483,120 observations from 671 wards, spanning 20 years and 36 10-day observations per year.

wards and that the relevant conflicts for farmers seem to be protests and riots. Protests and riots appear more frequently in major cities such as the capital of the county. The median distance of wards to the next county capital city is 18 kilometers. This makes a buffer of 20 kilometers an adequate estimate to account for participation in protests in the nearest city. Additionally, the capital city of Nairobi experiences most of the conflicts in Kenya. In Appendix E, we elaborate more on the distribution of conflicts within districts.

Our analysis focuses on the frequency or incidence of conflicts but does not account for the possibility of ongoing disputes. Conflict observations may be driven by previous conflict in a ward. Table 8 in the Appendix accounts for previous conflict experience by introducing a dummy variable to the main analysis that turns 1 if the ward has experienced conflict in the previous dekate. Our results are robust to the inclusion of this control.

Finally, Table 7 restricts the sample to different cutoffs of rainfed area. In general, rainfed agricultural land is more dependent on the weather than irrigated. Thus, the

Table 6: Robustness Test: Buffer Variation - Effect of Harvest on Conflict

	(1) Conflict 0km	(2) Conflict 5km	(3) Conflict 10km	(4) Conflict 15km	(5) Conflict 20km
Harvest	-0.000	-0.002	-0.005	-0.008	-0.012
	(0.000)	(0.001)**	(0.002)***	(0.002)***	(0.002)***
	[0.000]	[0.001]	[0.002]**	[0.004]**	[0.006]**
	{0.000}	{0.002}	{0.003}*	{0.005}*	{0.007}*
Constant	0.001***	0.006***	0.012***	0.023***	0.039***
	[0.000]	[0.001]	[0.003]	[0.004]	[0.006]
Ward FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
N	483,120	483,120	483,120	483,120	483,120

Notes: Significance levels are defined as * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered standard errors at the ward level are reported in parentheses. Clustered standard errors at the county level are provided in brackets. The variable Harvest is a dummy variable that turns 1 when harvest is detected. The dependent variable changes in terms of buffer size. While model 1 reports only conflicts appearing inside the ward, the following models extend the area of interest by a 5 10, 15, and 20 kilometer buffer, respectively. There are a total of 483,120 observations from 671 wards, spanning 20 years and 36 10-day observations per year.

timing of harvest seasons in rainfed areas relies more on weather. In our main analysis, we do not distinguish between rainfed and irrigated land. However, to validate the exogeneity assumption of our harvest variable, we restrict our sample to different rainfed area cutoffs. Mainly varying cutoffs between 40% and 90%. Examining the rainfed area in Table 7, we observe no significant changes in terms of magnitude for the harvest season. In all our models we account for standard errors clustered at the county level. Varying the levels of rainfed area between 40% and 80% does not affect the significance nor the magnitude of the coefficient. However, model 6 reports insignificant results for areas with at least 90% rainfed, which may be attributed to the reduction of the sample size.

6.2. Limitations and Future Research

Our panel database, which has detected harvest and conflict frequency at the ward level for 20 years, provides insights into the effect of harvest seasons on conflict. Nonetheless, some limitations might need to be investigated further.

Firstly, smallholder farmers in sub-Saharan Africa own small plots. Therefore, detecting harvest seasons for an entire ward does not necessarily capture the exact point of harvest

Table 7: Robustness Test: Variation in Share of Rainfed Area

	Sum of Conflict (Buffer size: 20 km)					
	(1)	(2)	(3)	(4)	(5)	(6)
	40%	50%	60%	70%	80%	90%
Harvest	-0.012**	-0.012**	-0.012**	-0.012*	-0.011*	-0.011
	[0.006]	[0.006]	[0.006]	[0.006]	[0.006]	[0.007]
Constant	0.039***	0.039***	0.039***	0.040***	0.033***	0.029***
	[0.006]	[0.006]	[0.006]	[0.005]	[0.005]	[0.006]
Ward FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Cluster	County	County	County	County	County	County
N	479,520	475,920	464,400	444,240	406,800	330,480

Notes: Significance levels are defined as * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered standard errors at the county level are provided in parentheses. The variable Harvest is a dummy variable that turns 1 when harvest is detected. The dependent variable is the sum of conflicts appearing within a ward and a 20-kilometer distance of the ward borders. We vary the sample of wards by restricting to minimum of mean rainfed of 40%, 50%, 60%, 70%, 80%, and 90%, respectively.

for individual farmers. Thus, our harvest estimate needs to be interpreted as a proxy for the general harvest season for the entire ward. The low deviation of the harvest time for highly agricultural wards illustrated in Figure 1 (Section 2.1), however, indicates that our detection is sensible. The accuracy of our harvest detection is further supported by the additional cross-check with Tanzanian survey data on smallholder farmers. To further test the accuracy of our harvest detection, the location of crop-specific areas and their crop calendars should be taken into account. Detected harvest could then be compared to the crop-specific calendars. Furthermore, the timing of the annual heavy rain season could serve as an additional tool to test for correct harvest detection. Usually, farmers harvest before the heavy rain season; thus, harvest should be detected prior to the annual rain season.

Another critical point is the possible postponement of harvest. Farmers may harvest later to participate in a conflict. This would bias our estimates, showing fewer conflicts during harvest season, and might even explain why we find more conflict before harvest. Additional robustness would be to include crops whose harvest cannot be postponed. Analyzing these particular crops and running regressions, including only wards growing them, will further justify the results.

In constructing the equation for our discretionary money variable, we use a specific exponential function to capture the suggested diminishing decline in funds by Stephens & Barrett (2011). Various other specifications may, however, better reflect how farmers in rural Kenya spend their discretionary money. Therefore, different specifications should be considered in the future. In addition we cannot exclude the possibility that our synthetic measure picks up other factors that decrease in the post-harvest periods.

Finally, this project focuses on Kenya as one country. Focusing on the entire African continent could provide more insights and external validity. Moreover, the analysis would benefit from more spatial variation.

7. Conclusion

This paper provides an analysis of the impact of the harvest season on the likelihood and frequency of conflict in Kenya, Africa. Using satellite-based harvest detection and conflict data at the ward level, We find a significant negative impact of harvest season on conflict. Performing numerous robustness tests does not change these findings. We further investigate the general crop cycle and its effect on conflict. The results indicate that there are more conflicts just before harvest starts, as well as at the start of the season. This pattern may be due to seasonal liquidity constraints faced by farmers after they have spent their income and discretionary funds from harvest.

Our approach to identifying bi-annual ward-specific harvest seasons enables us to conduct an in-depth analysis that is not dependent on general crop calendars. Since harvests occur on at least an annual basis, we examine income changes that are relatively frequent and consistent, affecting a substantial portion of Kenya's population, rather than rare and extreme events such as export price shocks or weather anomalies. Our findings support the evidence that agricultural income and its timing have an important effect on participating in conflict.

This paper addresses the gap in research on the relationship between harvest seasons and conflicts. Many developing countries rely heavily on agriculture as a primary income source. More evidence on factors such as harvest periods and their resulting income influencing the timing of conflict can provide new insights into when and why conflicts emerge. Social unrest can harm social welfare. At the same time, it can also be socially productive, for example, by defending property rights, thereby strengthening incentives to engage in productive activity (Bates et al., 2002). Nonetheless, it is important to understand not only why social unrest appears but also when it is more or less likely to appear. These findings are of particular interest, given the higher likelihood of conflicts appearing in developing countries. Our empirical results suggest that individuals are less willing to join conflicts when income rises or when discretionary money (consumption) is relatively high. A steady income or discretionary funds to smooth consumption may, therefore, be especially important for smallholder farmers. This gains in importance in the context of climate change strongly affecting crop yields in the future. Policymakers may focus on several measures. Firstly, the government could invest in agricultural infrastructure, technology, and training for farmers to improve crop yields and increase income during the harvest season. Secondly, the government could promote a diversified income for farmers and encourage rural communities to diversify their income sources beyond agriculture. Finally, farmers should be supported in their efforts to smooth consumption over the planting cycle. Access to financial products or providing commitment devices could be possible approaches. In sum, providing a steady income for the rural population could lead to fewer conflicts, improving the overall situation of especially vulnerable populations.

References

Angrist, J. D. & Kugler, A. D. (2008). Rural windfall or a new resource curse? coca, income, and civil conflict in columbia. *The Review of Economics and Statistics*, 90(2), 191 – 215. https://doi.org/https://doi.org/10.1162/rest.90.2.191

Bates, R., Greif, A., & Singh, S. (2002). Organizing violence. *Journal of Conflict Resolution*, 46(5), 599–628. https://doi.org/10.1177/002200202236166

Bazzi, S. & Blattman, C. (2014). Economic shocks and conflict: Evidence from commodity prices. *American Economic Journal: Macroeconomics*, 6(4), 1–38. https://doi.org/10.1257/mac.6.4.1

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2), 169–217. https://doi.org/10.1002/9780470752135.ch25

Berman, N. & Couttenier, M. (2015). External shocks, internal shots: The geography of civil conflicts. *Review of Economics and Statistics*, 97(4), 758–776. https://doi.org/10.1162/REST_a_00521

Berman, N., Couttenier, M., Rohner, D., & Thoenig, M. (2017). This mine is mine! How minerals fuel conflicts in Africa. *American Economic Review*, 107(6), 1564–1610. https://doi.org/10.1257/aer.20150774

Besley, T. & Persson, T. (2011). The logic of political violence. *Quarterly Journal of Economics*, 126(3), 1411–1445. https://doi.org/10.1093/qje/qjr025

Brückner, M. & Ciccone, A. (2010). International commodity prices, growth and the outbreak of civil war in Sub-Saharan Africa. *Economic Journal*, 120(544), 519–534. https://doi.org/10.1111/j.1468-0297.2010.02353.x

Buhaug, H. (2010). Climate not to blame for African civil wars. *Proceedings of the National Academy of Sciences of the United States of America*, 107(38), 16477–16482. https://doi.org/10.1073/pnas.1005739107

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 of the United States of America*, 106(49), 20670–20674. https://doi.org/10.1073/pnas.0907998106

Collier, P. & Hoeffler, A. (2004). Greed and grievance in civil war. *Oxford Economic Papers*, 56(4), 563–595. https://doi.org/10.1093/oep/gpf064

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*, 92(1), 1–45. https://doi.org/10.1016/S0304-4076(98)00084-0

Copernicus (2023). European Union, Copernicus Land Monitoring Service. *European Environment Agency (EEA).*

Crost, B., Duquennois, C., Felter, J. H., & Rees, D. I. (2018). Climate change, agricultural production and civil conflict: Evidence from the philippines. *Journal of Environmental Economics and Management*, 88, 379–395. https://doi.org/https://doi.org/10.1016/j.jeem.2018.01.005

Dube, O. & Vargas, J. F. (2013). Commodity price shocks and civil conflict: Evidence from Colombia. *Review of Economic Studies*, 80(4), 1384–1421. https://doi.org/10.1093/restud/rdt009

Eklundh, L. & Jönsson, P. (2012). TIMESAT 3.1 Software Manual. *Lund University, Sweden*, 1–82. https://web.nateko.lu.se/TIMESAT/docs/timesat3_1_1_SoftwareManual.pdf

FAO (2018). WaPOR Database Methodology: Level 2. Remote Sensing for Water Productivity Technical Report: Methodology. *Rome, FAO. 72 pages. Licence: CC BY-NC-SA 3.0 IGO.* https://www.fao.org/3/I8225EN/i8225en.pdf

FAO (2023a). Earth observation country indicators kenya. https://www.fao.org/giews/earthobservation/country/index.jsp?code=KEN#. Accessed: July 26, 2023.

FAO (2023b). Food and agricultural organization of the united nations - crop calendar information tool for crop production. https://cropcalendar.apps.fao.org/#/home?id=KE&crops=0026,0027,0048,0065,0073,0075,0076,0087,0103,0113,0117,0127,0186,0226,0247,0259,0262,0272,0283,0285,0303,0318,0325,0327,0329,0334,0335,0341,0356,0370,0373. Accessed: July 26, 2023.

Fearon, J. D. & Laitin, D. D. (2003). Ethnicity, Insurgency, and Civil War. *American political science review*, 97(1), 75–90. https://doi.org/https://doi.org/10.1017/s0003055403000534

FtMA (2022). Making Markets Work Better for Smallholder Farmers - FtMA Kenya Country Brief. *Farm to Market Alliance*, https://ftma.org/wp-content/uploads/2022/02/FtMA_Kenya_CountryBrief_Dec-2022.pdf. Accessed: July 23, 2023.

Gleditsch, N. P. (2012). Whither the weather? climate change and conflict. *Journal of Peace Research*, 49(1), 3–9. https://doi.org/10.1177/0022343311431288

Grossman, H. I. (1991). A General Equilibrium Model of Insurrections. *American Economic Review*, 81(4), 912–921.

Guardado, J. & Pennings, S. (2023). Shock persistence and the study of armed conflict: Empirical biases and some remedies. *International Interactions*, 1–17. https://doi.org/https://doi.org/10.1080/03050629.2023.2266553

Harari, M. & Ferrara, E. L. (2018). Conflict, climate, and cells: a disaggregated analysis. *Review of Economics and Statistics*, 100(4), 594–608. https://doi.org/https://doi.org/10.1162/rest_a_00730

Harris, D. (2019). Intensification benefit index: how much can rural households benefit from agricultural intensification? *Experimental Agriculture*, 55(2), 273–287. https://doi.org/https://doi.org/10.1017/s0014479718000042

Hastings, J. & Ubilava, D. (2023). Agricultural shocks and social conflict in southeast asia. *arXiv preprint arXiv:2304.10027*. https://arxiv.org/ftp/arxiv/papers/2304/2304.10027.pdf

Hendrix, C. S. & Haggard, S. (2015). Global food prices, regime type, and urban unrest in the developing world. *Journal of Peace Research*, 52(2), 143–157. https://doi.org/https://doi.org/10.1177/0022343314561599

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

Hsiang, S. M., Meng, K. C., & Cane, M. A. (2011). Civil conflicts are associated with the global climate. *Nature*, 476(7361), 438–441. https://doi.org/10.1038/nature10311

Jia, R. (2014). Weather shocks, sweet potatoes and peasant revolts in historical china. *The Economic Journal*, 124(575), 92–118. https://doi.org/https://doi.org/10.1111/ecoj.12037

Karfakis, P., Ponzini, G., & Rapsomanikis, G. (2017). On the costs of being small: Case evidence from kenyan family farms. https://ageconsearch.umn.edu/record/288962/files/a-i7562e.pdf

Koubi, V. (2019). Climate change and conflict. *Annual Review of Political Science*, 22, 343–360. https://doi.org/https://doi.org/10.1146/annurev-polisci-050317-070830

Kumawat, M. & Khaparde, A. (2024). Land cover change detection using timesat software and machine learning algorithms near ujani dam: A case study. *Journal of Integrated Science and Technology*, 12(1), 717–717. https://pubs.thesciencein.org/journal/index.php/jist/article/view/a717

Linke, A. M. & Ruether, B. (2021). Weather, wheat, and war: Security implications of climate variability for conflict in syria. *Journal of Peace Research*, 58(1), 114–131. https://doi.org/https://doi.org/10.1177/0022343320973070

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

McGuirk, E. F. & Nunn, N. (2021). Transhumant pastoralism, climate change, and conflict in africa. *NBER Working Paper Series 28243*. https://doi.org/https://doi.org/10.2139/ssrn.3753152

Miguel, E., Satyanath, S., & Sergenti, E. (2004). Economic shocks and civil conflict: An instrumental variables approach. *Journal of Political Economy*, 112(4), 725–753. https://doi.org/10.1086/421174

Nagayets, O. (2005). Small farms: current status and key trends. *The future of small farms: Proceedings of a research workshop*, 355–367. https://doi.org/https://doi.org/10.2499/97808962976472020vp42

Raleigh, C., Linke, A., Hegre, H., & Karlsen, J. (2010). Introducing ACLED: An armed conflict location and event dataset. *Journal of Peace Research*, 47(5), 651–660. https://doi.org/10.1177/0022343310378914

Ray, D. & Esteban, J. (2017). Conflict and Development. *Annual Review of Economics*, 9, 1–29. https://doi.org/https://doi.org/10.1146/annurev-economics-061109-080205

Schlenker, W. & Lobell, D. B. (2010). Robust negative impacts of climate change on african agriculture. *Environmental Research Letters*, 5(1), 014010. https://doi.org/https://doi.org/10.1088/1748-9326/5/1/014010

Stephens, E. C. & Barrett, C. B. (2011). Incomplete credit markets and commodity marketing behaviour. *Journal of agricultural economics*, 62(1), 1–24. https://doi.org/https://doi.org/10.1111/j.1477-9552.2010.00274.x

Theisen, O. M. & Holtermann, H. (2011). Climate Wars? Assessing the Claim That Drought Breeds Conflict. *International Security*, 36(3), 79–106. https://doi.org/https://doi.org/10.1162/isec_a_00065

Ubilava, D., Hastings, J. V., & Atalay, K. (2023). Agricultural windfalls and the seasonality of political violence in africa. *American Journal of Agricultural Economics*, 105(5), 1309–1332. https://doi.org/https://doi.org/10.1111/ajae.12364

Vesco, P., Kovacic, M., Mistry, M., & Croicu, M. (2021). Climate variability, crop and conflict: Exploring the impacts of spatial concentration in agricultural production. *Journal of Peace Research*, 58(1), 98–113. https://doi.org/https://doi.org/10.1177/0022343320971020

Winne, J. D. & Peersman, G. (2021). The impact of food prices on conflict revisited. *Journal of Business & Economic Statistics*, 39(2), 547–560. https://doi.org/https://doi.org/10.1080/07350015.2019.1684301

A. Definition of Conflicts in ACLED

ACLED does not explicitly define conflict but aims to report all events of political violence and protests, where political violence is defined as the use of force by a group with a political purpose or motivation. ACLED data differentiates between six different types of conflict: Protests, riots, strategic development, violence against civilians, battles, and explosions and remote violence.

Protests are defined as a peaceful protest. The two actors participating in this kind of conflict are protesters and police forces. Riots are violent protests and mob violence. The main actors are rioters against civilians and police forces. Strategic development denotes non-violent activities of violent groups, e.g. recruitment drives, looting, peace talks and mass arrests or the arrests of high-ranking officials. The main actors are militia, armed groups, the government or police forces. Violence against civilians includes attacks by militia and armed groups. Hereby, the militia or armed groups fight against civilians. Finally, battles appear due to an armed clash, if the government regains territory or a non-state actor overtakes territory. The main actors in the sample of Kenya are mainly militias and to some extent military forces or police forces. Even though there exists one further conflict - explosions - we do not explain it in more detail, because the sample of agricultural wards does not contain any single event.

Please refer to the ACLED Code Book for the complete list of definitions:

`https://acleddata.com/acleddatanew/wp-content/uploads/2021/11/ACLED_Codebook_v1_January-2021.pdf`

B. TIMESAT Settings

Using the time series of NDVI values, we detect harvest seasons with the software TIMESAT (Eklundh & Jönsson, 2012). Based on NDVI values, the program detects harvest seasons by smoothing the time series. Figure 3 shows an example of one smoothed season with detected seasonality parameters in TIMESAT. For our data, the parameters for season start (point a) and season end (point b) are kept at the default setting of 0.5. Note that, in Figure 3, another setting (0.2) is used. With the season start and end default value of 0.5, the software detects the season end (harvest) as the point of time at which the NDVI value has decreased by 50% (between its maximum and minimum). Similarly, for the start of season, the software detects the point of time at which the NDVI value has increased by 50% (between its minimum and maximum). A recent study by Kumawat & Khaparde (2024) tests this parameter setting of 0.5 and compares the detected season start and end to actual field observations. They confirm that the detected end of season correctly corresponds with the observed time of harvest. The maximum NDVI is point e in Figure 3.

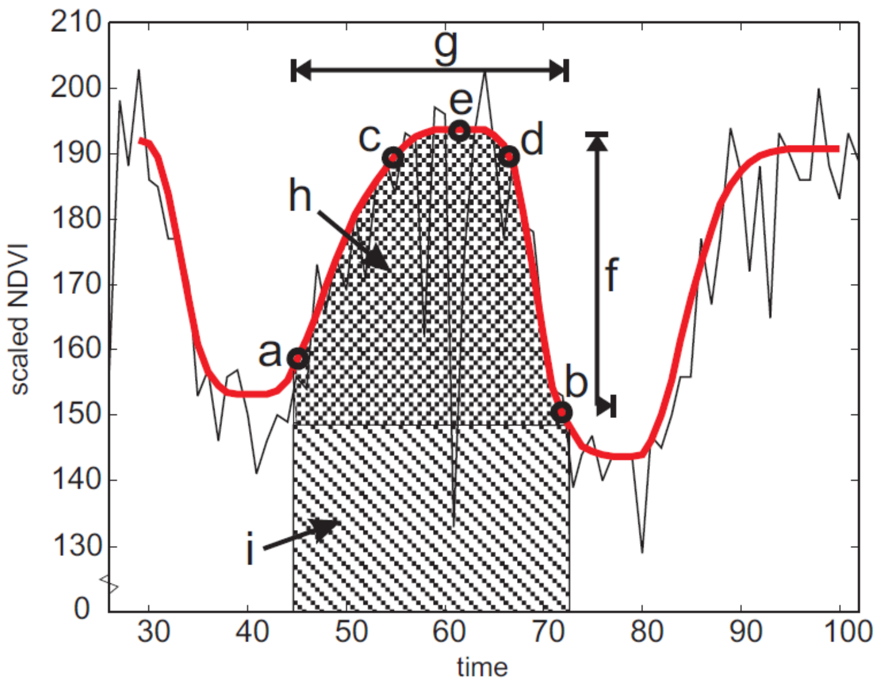

Figure 3: TIMESAT Seasonaliy Parameters

Notes: The figure is taken from Eklundh & Jönsson (2012). It shows an example for seasonality parameters detected by TIMESAT, where a and b mark the start and end of season, respectively; e is the maximum value obtained, c and d show the 80% levels. The area of i and h illustrate the cumulative effect of vegetation growth during the season and f and g the seasonal amplitude and length, respectively.

Besides the seasonality parameters, TIMESAT allows us to define the number of seasons within a year. If the seasonality parameter is set to 1 the program treats the data as if there is only one annual season; set to 0 the program will consider two annual seasons. Given that Kenya experiences two annual seasons, we set the parameter to 0. Out of the three fitting methods (adaptive Savitzky-Golay filter and fits to double logistic or asymmetric Gaussians functions), we choose the adaptive Savitzky–Golay filter, with 2 envelope iterations, and an adaptation strength of 2. We set the window size of the

Savitzky–Golay filter to 7 which corresponds in our data to 70 days. The choice of the fitting method and iterations is based on a variety of test runs. The Savitzky-Golay provided the most consistent detection of bi-annual harvest. We leave the spike parameter and spike method at default values. Figure 4 shows the user interface of TIMESAT and the chosen parameters for this study. See Eklundh & Jönsson (2012) for further descriptions of the filters and the software in general.

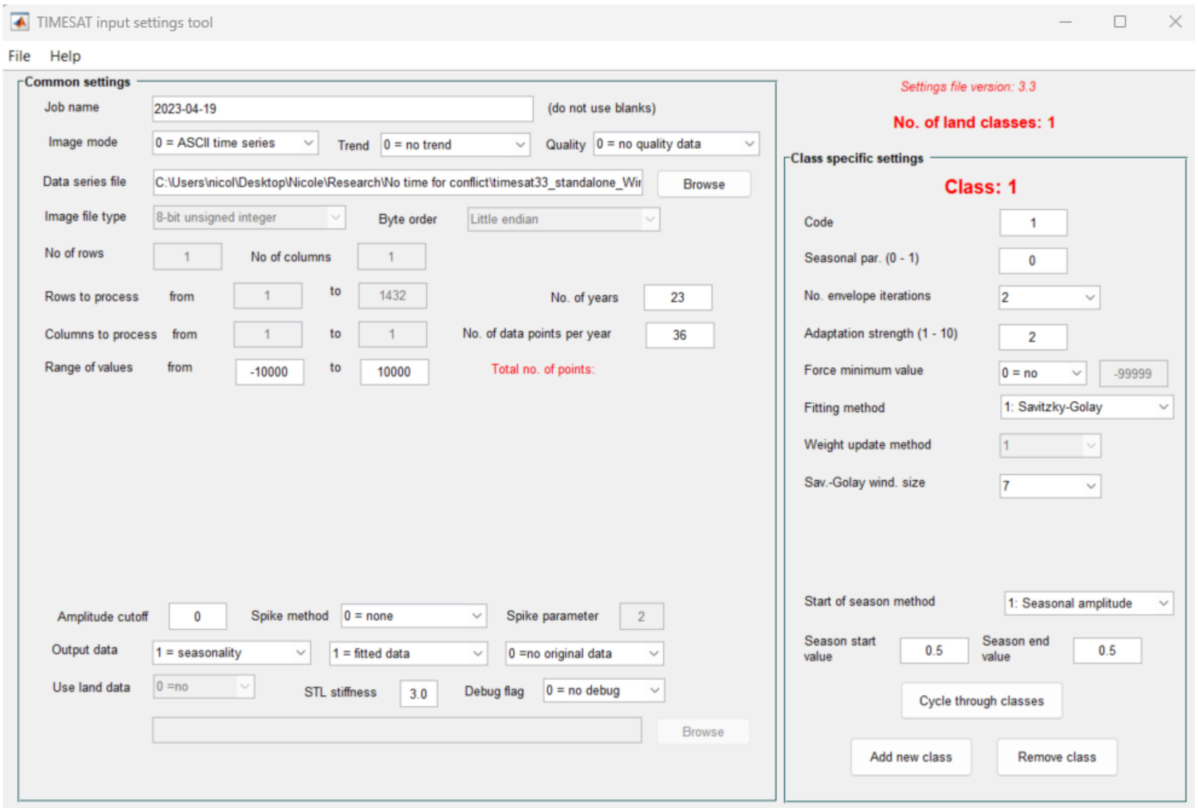

Figure 4: Settings TIMESAT

Figure 5 shows the smoothing (in orange) of one of our NDVI time series (in blue). The y-axis illustrates the NDVI values which in our case are in the range between 0 and 250, with 250 indicating the highest NDVI values. The x-axis illustrates the dates on a 10-day interval. The whole time series was calculated for the years 1999 to June 2020[21]. This example illustration contains information from January 2004 (illustrated as value 216) until the beginning of the year 2010 (illustrated as value 397). The dots between the spikes illustrate the start of the season (dot before a peak) and the end of the season

[21] Using the NDVI series up to June 2020 for harvest detection provides adequate harvest estimates for the crop cycle in the year 2019. We exclude the year 2020 in our econometric analysis given that it does not cover the entire year.

(dot after a peak), with the latter also indicating the times of harvest.

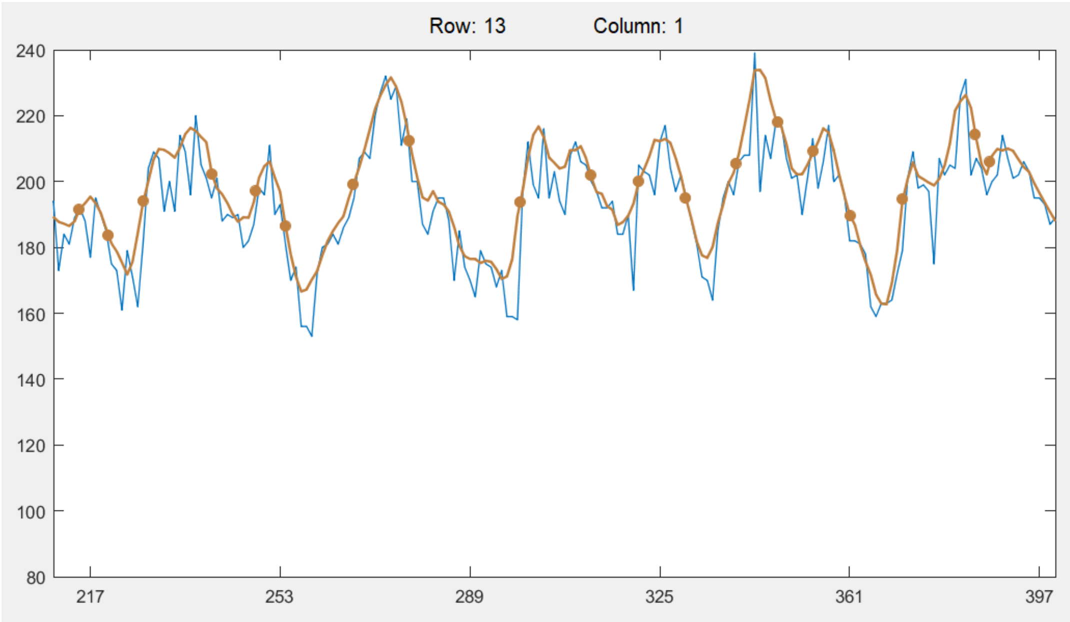

Figure 5: Example Harvest Detection

C. Additional Regression Results: Previous Conflict

Table 8 uses the same sample as the main analysis of this paper. We still consider conflicts within a buffer of 20 kilometers as dependent and harvest as independent variable. We additionally control for previous conflict experience as wards may face a higher risk of (continued) conflict when there was a conflict before. The models in Table 8 include a dummy variable indicating whether the ward experienced conflict in the previous 10-day interval. Our results on harvest are robust to the inclusion of this control. The coefficient of harvest is similar in size as in the main analysis and stays significant at the 5 %-level. The coefficient of the control variable for previous conflict experience is, as expected, significantly positive (1%-level).

Table 8: Previous Conflict Experience

	Sum of Conflict (Buffer size: 20 km)			
	(1)	(2)	(3)	(4)
Harvest	-0.0123**	-0.0122**	-0.0123**	-0.0122**
	(0.005)	(0.005)	(0.005)	(0.005)
Conflict$_{t-1}$	0.1561***	0.1541***	0.1559***	0.1540***
	(0.022)	(0.021)	(0.022)	(0.021)
Constant	0.0358***	0.0819***	-0.0419	0.0559
	(0.006)	(0.010)	(0.035)	(0.038)
Max NDVI			Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Month FE		Yes		Yes
N	482,449	482,449	482,449	482,449

Notes: Significance levels are defined as * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the county level are provided in parentheses. The variable Harvest is a dummy variable that turns 1 when harvest is detected. Conflict$_{t-1}$ is a dummy variable indicating whether the ward experienced a conflict in the previous 10-day interval. Max NDVI is the maximum NDVI value of the previous harvest season and serves as an indicator of the goodness of the previous harvest. The dependent variable is the sum of conflicts appearing within a ward and a 20-kilometer distance of the ward borders. We have in total 483,120 observations coming from 671 wards, 20 years, and 36 10-day observations within a year. The number of observations decreases by 671 when including Conflict$_{t-1}$ because the variable is not defined for the first dekate of a ward.

D. Additional Analysis on Discretionary Money

Table 9 employs the same models as the main table 2 but uses discretionary money as the independent variable instead of harvest. The coefficients for discretionary money are negative and significant at the 1%-level when clustering at ward level and including Ward and Year FE (Model 1) as well when including Ward and Year×County FE (Model 4). The significance decreases to 5% in both cases when clustering at the county level. We do not find any significant effects of discretionary money on conflict when including Month FE(Model 2 and Model 3).

Table 9: The Effect of Discretionary Money on the Main Specification

	Sum of Conflict (Buffer size: 20 km)			
	(1)	(2)	(3)	(4)
Discretionary Money	-0.025	0.004	0.003	-0.026
	(0.004)***	(0.004)	(0.004)	(0.004)***
	[0.011]**	[0.014]	[0.014]	[0.011]**
	{0.014}*	{0.012}	{0.012}	{0.014}*
Constant	0.073***	0.140***	0.096***	0.079***
	[0.014]	[0.011]	[0.018]	[0.007]
	{0.014}*	{0.012}	{0.012}	{0.014}*
Ward FE	Yes	Yes	Yes	Yes
Year FE	Yes		Yes	
Month FE		Yes	Yes	
Year×County				Yes
N	468,094	468,094	468,094	468,094

Notes: Significance levels are defined as * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the ward level are provided in parentheses. Clustered standard errors at the county level are provided in brackets. Discretionary money is normalized to one in the period after harvest and declines as specified in Equation 2. The dependent variable is the sum of the respective conflict type appearing within a ward and a 20-kilometer distance of the ward borders. There are a total of 483,120 observations from 671 wards, spanning 20 years and 36 10-day observations per year. The number of observations decreases when including discretionary money because the variable is not defined for observations before the first detected harvest in a ward.

Table 10 looks at different conflict types and uses the same models as Table 3 controlling for the decreasing discretionary money variable instead of harvest. When clustering standard errors on ward level, the effects of discretionary money on all types of conflict, except for violence against civilians, are negative and significant at 1%-level. The significance levels decrease to 5-10% for riots and strategic development when clustering at the county level. The results on protests turn insignificant.

Table 10: The Effect of Discretionary Money on Different Types of Conflict

	(1) Protest	(2) Riots	(3) Strategic Dev.	(4) Violence	(5) Battle
Discretionary Money	-0.002	-0.011	-0.002	-0.002	-0.002
	(0.001)***	(0.001)***	(0.000)***	(0.001)***	(0.002)***
	[0.002]	[0.004]***	[0.001]***	[0.002]	[0.001]
	{0.005}	{0.008}	{0.001}*	{0.006}	{0.004}
Constant	0.004	0.025***	0.001	0.017**	0.016**
	[0.002]	[0.006]	[0.001]	[0.007]	[0.007]
Ward FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
N	468,094	468,094	468,094	468,094	468,094

Notes: Significance levels are defined as * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the ward level are provided in parentheses. Clustered standard errors at the county level are provided in brackets. Discretionary money is normalized to one in the period after harvest and declines as specified in Equation 2. The dependent variable is the sum of the respective conflict type appearing within a ward and a 20-kilometer distance of the ward borders. *Protests* denote peaceful protests. *Riots* include violent protests and mob violence. *Strategic developments* are non-violent activities of violent groups that may trigger or affect conflicts. *Violence against civilians* are attacks of militia and armed groups against civilians. *Battles* are defined as armed clashes in which the government regains territory or a non-state actor overtakes territory. There are a total of 483,120 observations from 671 wards, spanning 20 years and 36 10-day observations per year. The number of observations decreases when including discretionary money because the variable is not defined for observations before the first detected harvest in a ward.

Again a similar analysis as for harvest has been considered for the discretionary money variable in Table 11 accounting for different specifications of the dependent variable. The negative coefficient of discretionary money indicates, again, a negative effect on conflict. The results are statistically significant at the 1% level when clustering standard errors at the ward level. Clustering at the county level makes the coefficients slightly less significant.

Table 12 tests the effect of discretionary money on conflict and varies the buffer sizes. Disrectionary money has a significant negative effect even when just considering conflicts within a ward (Model 1). The coefficients tend to increase in size and significance when increasing the buffer to 5, 10, and 20 kilometers.

Table 11: Model Specification Variation - Effect of Discretionary Money on Conflict

	(1) LPM	(2) Logit	(3) Poisson ZIP	(4) log(Sum(Conflict))	(5) IHST
Discretionary Money	-0.009	-0.152	-0.054	-0.106	-0.012
	(0.001)***	(0.025)***	(0.010)***	(0.016)***	(0.001)***
	[0.004]**	[0.081]*	[0.030]*	[0.050]**	[0.005]**
Constant	0.036***	-4.463***	0.418***	-11.085***	0.043***
	[0.006]	[0.228]	[0.104]	[0.069]	[0.009]
Ward FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
N	468,094	468,094	468,094	468,094	468,094

Notes: Significance levels are defined as * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the ward level are provided in parentheses. Clustered standard errors at the county level are provided in brackets. Discretionary money is normalized to one in the period after harvest and declines as specified in Equation 2. Model 1 uses the linear probability model and introduces a dummy variable as the dependent variable, turning 1 if a conflict appears within a ward and a 20-kilometer distance of the ward borders. Model 2 applies a logit probability model. Model 3 uses a pseudo maximum likelihood regression accounting for the large share of zero conflicts in the data. Model 4 applies a logarithm on the main dependent variable by adding a small constant (0.00001) to the zero values. Model 5 applies the inverse hyperbolic sine transformation on the main dependent variable. There are a total of 483,120 observations from 671 wards, spanning 20 years and 36 10-day observations per year. The number of observations decreases when including discretionary money because the variable is not defined for observations before the first detected harvest in a ward.

Table 12: Robustness Test: Buffer Variation - Effect of Discretionary Money on Conflict

	(1) Conflict 0km	(2) Conflict 5km	(3) Conflict 10km	(4) Conflict 15km	(5) Conflict 20km
Discretionary Money	-0.000	-0.003	-0.006	-0.013	-0.019
(exponential)	(0.000)	(0.001)***	(0.001)***	(0.002)***	(0.002)***
	[0.000]	[0.002]*	[0.003]*	[0.005]**	[0.006]***
	{0.000}*	{0.003}**	{0.006}**	{0.009}**	{0.014}*
Constant	0.001***	0.006***	0.016***	0.033***	0.062***
	[0.000]	[0.002]	[0.004]	[0.006]	[0.017]
Ward FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
N	468,094	468,094	468,094	468,094	468,094

Notes: Significance levels are defined as * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Clustered standard errors at the ward level are reported in parentheses. Clustered standard errors at the county level are provided in brackets. Discretionary money is normalized to one in the period after harvest and declines as specified in Equation 2. The dependent variable changes in terms of buffer size. While model 1 reports only conflicts appearing inside the ward, the following models extend the area of interest by a 5 10, 15, and 20 kilometer buffer, respectively. There are a total of 483,120 observations from 671 wards, spanning 20 years and 36 10-day observations per year. The number of observations decreases when including discretionary money because the variable is not defined for observations before the first detected harvest in a ward.

E. Overview Kenya Agriculture

Figure 6 shows the distribution of the agricultural areas in Kenya based on data by FAO (2018). We see that most of the agricultural land appears in West and Central Kenya. We also see that Nairobi - the capital city of Kenya - where most conflicts happen, is not far away from the agricultural wards. Similarly, Figure 7 shows the distance between ward centroids to the nearest capital city of a county. Looking at the distances and the map scale, 20 km is a rather conservative buffer. Most distances from ward centroid to the closest capital city of a county seem to be even larger.

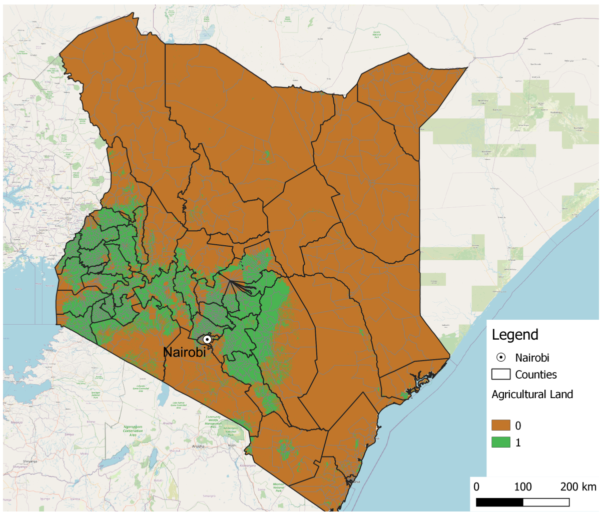

Figure 6: Overview Agricultural Area

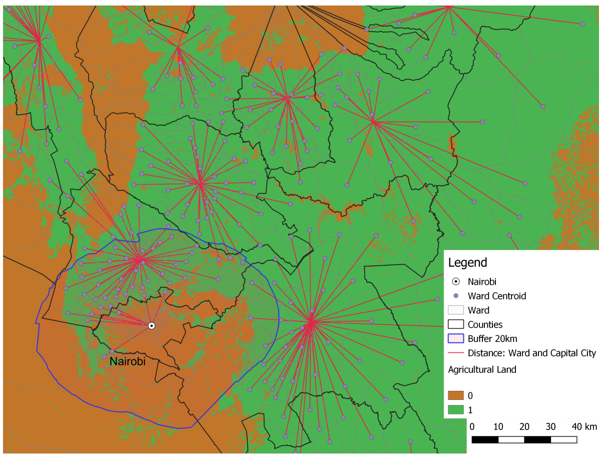

Figure 7: Overview Distance to nearest county capital cities from ward centroids

F. Overview distribution of Conflicts and Harvest within a Year

Figure 8 illustrates the probability of conflict and harvest across the years. The left panel shows wards without buffer; the right one includes the 20-km buffer. We see that most harvest appears in the months between December and February as well as between June and August. In contrast, the probability of conflicts does not seem to vary a lot throughout the year. Some increases are experienced in the months of October and December, as well as January.

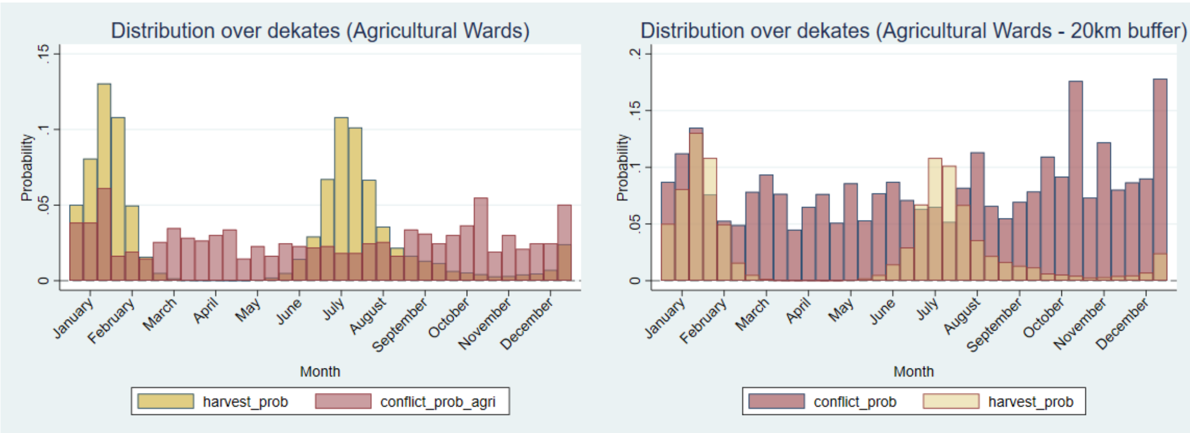

Figure 8: Conflict distribution in wards across months