Fertile Ground for Conflict*

Nicolas Berman† Mathieu Couttenier‡ Raphael Soubeyran§

This version: December 22, 2017

Abstract. We investigate how variations in soil productivity affect civil conflicts. We first present a model with heterogeneous land in which variations in input prices (fertilizers) affect appropriable rents and the opportunity costs of fighting. The theory predicts that spikes in input prices increase the likelihood of conflicts through their effect on income and inequality, and that this effect is magnified when soil fertility is naturally more heterogenous. We test these predictions using data on conflict events covering all Sub-Saharan African countries at a spatial resolution of 0.5 $\times$ 0.5 degree latitude and longitude over the 1997-2013 period. We combine information on soil characteristics and worldwide variations in fertilizer prices to identify local exogenous changes in input costs. As predicted, variations in soil productivity triggered by variations in fertilizer prices are positively associated with conflicts, especially in cells where land endowments are more heterogeneous. In addition, we find that the distribution of land fertility both within and across ethnic groups affects violence, and that the effect of between-group heterogeneity in soil quality is magnified in densely populated areas. Overall, our findings imply that inequality in access to fertile areas – an issue largely neglected in the literature dealing with the roots of Sub-Saharan African civil wars – constitutes a serious threat to peace at the local-level.

JEL classification: C23, D74, Q34

Keywords: conflict, land, fertility, inequality.

*We thank Catherine Benjamin, Johannes Buggle, Marshall Burke, Pascaline Dupas, Jérémy Laurent-Lucchetti, Robert Lifran, Manuel Oechslin, Nicolas Quérou, Marc Sangnier, Alexandre Sauquet, Petros Sekeris, Roberta Ziparo, and participants at various seminars and conferences. Bui Ha My provided outstanding research assistance. Nicolas Berman thanks the A*Midex for financial support (grant ANR-11-IDEX-0001-02 funded by the French government "Investissement d'Avenir" program). Mathieu Couttenier acknowledges financial support from the Swiss National Foundation (grant Ambizione-PZ00P1 161451). Raphael Soubeyran acknowledges financial support from the ANR GREEN-Econ research project (ANR-16-CE03-0005).

†Aix-Marseille University (Aix-Marseille School of Economics), CNRS, EHESS, Graduate Institute Geneva and CEPR. E-mail: nicolas.berman@univ-amu.fr.

‡University of Geneva. E-mail: mathieu.couttenier@unige.ch

§CEE-M, INRA, CNRS, SupAgro Montpellier, Univ. Montpellier, Montpellier. E-mail: raphael.soubeyran@inra.fr.

1 Introduction

Are African conflicts rooted in fertile soils? Over the last decades, unequal access to productive areas has often been mentioned as a key contributing factor to some of the deadliest wars on the continent. In Rwanda for instance, increasing pressure over land stemming from rapid population growth and soil depletion has most likely been one of the triggers of conflicts and genocide (André and Platteau, 1998). Disputes over arable land have also historically played their part in conflicts in Darfur (Faris, 2009).[1] In many countries however, the apparent "ethnic" dimension of these tensions has obscured the fact that they originated from rising land inequality and lack of access to fertile soils.[2] As a result, researchers have investigated the impact of ethnic divisions at great length, largely overlooking similar investigations about the role of soil fertility (Peters, 2004).

This paper studies theoretically and empirically how variations in soil productivity influence the occurrence of violence. We are interested both in the effect of changes in soil productivity over time and in the role of its distribution across space. We proceed in two steps. First, we present a dynamic theory of conflict over agricultural land in which two groups are endowed with units of land that differ in terms of natural fertility. At each period, soil productivity varies due to changes in fertilizer prices. We use this model to derive predictions that relate conflict likelihood to fertilizer prices variations and to the distribution of land quality across groups. These predictions are tested in the second part of the paper using detailed data on conflict events, soil characteristics and local fertilizer prices.

In our model, an increase in the price of fertilizer reduces agricultural rents for both groups and has two opposite effects on conflict probability: it at once lowers the value of what can be appropriated through violence and decreases the opportunity cost of fighting. However, due to heterogeneous land quality, the opportunity cost effect dominates: naturally fertile soils are less dependent upon fertilizer use, which implies that the rents generated from these lands are less sensitive to changes in fertilizer prices. As a consequence, when fertilizer prices rise, the group endowed with the less fertile soil (the land-poor group) has a greater incentive to fight: the decrease in rents from their own land (i.e. opportunity cost of fighting) is greater than the decrease in the potential gains that stand to be obtained from land-rich. In other words, spikes in fertilizer prices increase the risk of conflict through their effect on the level of rents and on the dispersion of rents across groups (i.e. rent inequality). We also show that the positive link between fertilizer price variations and conflict is magnified when natural soil quality is more heterogeneous.

In sum, our model generates two testable predictions: a) variations in fertilizer prices are positively associated with the likelihood of conflict, b) this association is stronger when natural soil fertility is more heterogeneous. We test these predictions using a dataset in which units of analysis are cells of 0.5 × 0.5 degree latitude and longitude (approx. 55km × 55km at the equator) covering all of Sub-Saharan Africa (SSA) from 1997 to 2013. Testing our theory requires data on (i) conflict events, (ii) soil fertility, and (iii) local variations in fertilizer prices.[3] First,

[1] The violent conflicts between Mauritania and Senegal and between Mali and Burkina Faso (the "wars between brethren") also originated from land disputes between pastoralists and herders (Van den Brink *et al.*, 1995).

[2] Ethnic diversity has been shown to be linked to land characteristics such as soil quality (Michalopoulos, 2012).

[3] Instead of fertilizer prices, one could have also used weather shocks to study (heterogeneous) changes in soil fertility across time and space. Weather shocks may however affect the likelihood of conflict through other channels than agricultural productivity, such as migration, infrastructure quality or competition over water – channels that

we combine original geo-localized datasets on soil characteristics with information on conflict occurrence. The Armed Conflict Location Events Data (ACLED) provides detailed information on the date, location and type of conflict events, and can be used to identify subsets of events that are more likely to be land-related. For each cell, we complement this conflict data with measures of the natural fertility of soils, defined as their inherent nutrient content. These measures come from the Harmonized World Soil Database and are available at a highly disaggregated level; this allows us to calculate our measure of land inequality, as the variance of natural soil fertility within each cell.

The final variable we construct is the local price of fertilizers. Direct information on time-varying prices at a spatially disaggregated level is not readily available. Furthermore, even if we were able to observe such prices, they would be endogenous to conflict. We circumvent this issue by computing a proxy for local fertilizer prices. The measure we build is informed by the fact that fertilizers are comprised of three main nutrients (nitrogen, phosphate and potassium) and that the ideal composition of fertilizers varies across crops. More precisely, for each cell we identify the main crop(s) produced – using the FAO's Global Agro-Ecological Zones (GAEZ) – and the required balance of nutrients for each of these crops. We then use data on the world prices of each nutrient to construct a cell-specific, time-varying indicator of fertilizer price. Because our identification strategy makes use of within-cell variations in fertilizer prices and conflicts over time, we are able to control for cell-fixed effects and unobserved common time shocks. Moreover, since we use the world prices of nutrients to compute local fertilizer prices, reverse causality from conflict to fertilizer prices is unlikely to drive our results. Using external data for a subset of countries and years, however, we show that such prices for specific nutrients are indeed transmitted to local fertilizer prices. We also provide indirect evidence that the mix of nutrients used by farmers in SSA is indeed correlated with the recommended mix for their particular crop.

We find empirical support for the predictions of the model. First, an increase in fertilizer price makes local conflicts more likely to occur.[4] Second, spikes in fertilizer prices are found to trigger more conflict in cells where soil fertility is naturally heterogeneous, i.e. in cells characterized by areas of both nutrient-rich as well as nutrient-poor soil. These effects are sizeable: for a cell with the average level of soil quality heterogeneity, a 10% increase in fertilizer price makes conflict 1.2 percentage points more likely. This figure roughly doubles in cells where soil fertility is one standard deviation more heterogeneous than the sample mean. These results hold across various measures of conflicts, soil fertility and nutrient mixes, and are robust to the use of alternative estimators and inference methods. Dropping the years during which commodities prices spiked (2008-2009) has little effect on our estimates. They also remain stable when we control for other co-determinants of violence that might be correlated with soil characteristics or with fertilizer

cannot be easily controlled for. See Sarsons (2015) for a discussion of the impact of rainfall on conflict through other channels than income.

[4]This result may seem surprising given the widespread belief that African farmers use little fertilizers compared to their Asian or Latin-American counterparts (Morris *et al.*, 2007). A recent World Bank report (Sheahan and Barrett, 2014) challenges this view using survey data for 6 SSA countries from the *Living Standards Measurement Study*. More than a third of the households are found to use inorganic fertilizers in this sample, and this share reaches 55% in Ethiopia and 77% in Malawi. Given that most of the labor in African agriculture is familial, labor costs are typically low and inputs such as fertilizers represent a large share of total variable costs. Moreover, even low levels of fertilizer use do not necessarily imply that African farmers are insulated from fertilizer prices variations. We indeed provide evidence, based on country-level data, that the imports and consumption of fertilizers of SSA countries are significantly impacted by fertilizer prices changes. This is consistent with Duflo *et al.* (2011), who find that fertilizer consumption in Africa is highly sensitive to price variations in the post-harvest season.

price variations (e.g. time-invariant or slow-moving characteristics such as geography, institutions, social cleavages, or time-varying determinants such as weather conditions, or the prices of produced or consumed commodities). Interestingly, we find quantitatively stronger results when we restrict our sample to events that are more likely to reflect conflict over land. The link between soil fertility variations, land inequality and conflict is also found to be more pronounced in countries with weak institutions and more insecure land tenure rights for indigenous communities.

As mentioned earlier, several dramatic conflict episodes, such as the war and genocide in Rwanda, were caused by a confluence of factors, namely population growth, scarcity of fertile land, and inherent ethnic tensions. In the final part of the paper, we find support for this account on a much larger scale. More precisely, combining our data on soil fertility with geocoded information on the contours of ethnic homelands, we split our cell-specific measure of soil fertility heterogeneity into two components: one component arising from heterogeneity in soil quality *across* ethnic groups, and the another component arising from differences in soil quality *within* ethnic groups. We find that both within- and between-group land inequality amplify the impact of fertilizer price variations on conflict. Inequality in soil quality between ethnic group tends however to matter quantitatively more, especially in densely populated areas.

Related literature. Our paper relates to several strands of the literature. The recent decade has seen a surge of empirical studies examining the roots of civil wars - first at the country-level, and more recently using spatially disaggregated data. The role of natural resource extraction, ethnic divisions, and income variations have received particular interest.[5] Our paper contributes in particular to the literature on agricultural income shocks, which typically associates conflict with changes in commodity prices or demand (e.g. Dube and Vargas, 2013, Bazzi and Blattman, 2014, Berman and Couttenier, 2015, McGuirke and Burke, 2017), weather conditions (e.g. Miguel *et al.*, 2004, Harari and Ferrara, 2012, Hsiang *et al.*, 2013, Couttenier and Soubeyran, 2014, Iyigun *et al.*, 2017, Adhvaryu *et al.*, 2017), or long-run changes in agricultural productivity (Iyigun *et al.*, 2015). We complement these works by demonstrating that variations in the price of an agricultural production technology (fertilizer) affect the likelihood of conflict, and that local heterogeneity in agricultural productivity plays a key role in this dynamic.

We also contribute more specifically to the relatively scarce empirical literature on land conflicts. Hidalgo *et al.* (2010) study the determinants of land invasions in Brazil, and emphasize the role of unequal landholding and land tenure systems. Di Falco *et al.* (2017) find that tenure security and climate affect the likelihood of observing land disputes in Ethiopia and Guardado (2016) finds that changes in crop prices trigger more violence in Peruvian districts where land ownership is primarily individual. Our paper differs from these studies in terms of geographic coverage (43 countries over a period of 17 years) as well as objectives. We focus on inequality in terms of natural soil fertility, and we identify periods during which this inequality rises using fertilizer price variations.

While we ultimately apply our theory to the case of agricultural production and fertilizer use, we begin by extending the conflict model developed by Chassang and Padró i Miquel (2009) to a situation in which rents are heterogeneous. Our paper therefore adds to existing models that

[5] On the role of natural resources, see, for example, Fearon and Laitin (2003), Ross (2004, 2006), Berman *et al.* (2017), and Sanchez de la Sierra (2017), and on ethnic fractionalization and polarization, see, Montalvo and Reynal-Querol (2005), Esteban *et al.* (2012), and Michalopoulos and Papaioannou (2016), among many others.

deal with the relationship between inequality among different individuals/groups and violence. Esteban and Ray (2011b) consider a rent-seeking game between groups of different sizes and show that the equilibrium level of conflict depends on inequality, fractionalization and polarization (see also Esteban and Ray, 1999). Fearon (2007), on the other hand, considers a rent-seeking model in which a rebel group enters into conflict with the government in order to appropriate the national tax revenue. In this framework, greater income inequality among citizens increases the number of relatively poor people and decreases the marginal cost of recruitment for both rebels and the government, which, in turn, leads to an increase in the intensity of the conflict. Compared with these papers, we use a model that more closely aligns with bargaining models, in which both peace and conflict are potential outcomes, rather than a rent-seeking game in which conflict is inherent. More importantly, while they find that the level of conflict increases with inequality in static frameworks, we consider a dynamic model that demonstrates that the probability of conflict increases *when* inequality in soil fertility is high, while the relationship between average inequality and the likelihood of conflict is ambiguous. The underlying mechanism comes from our agricultural production model: fertilizer prices impact inequality and conflict because they affect less the (appropriable) revenue of the rich than the opportunity cost of the poor. This in turn is due to the fact that fertile soil use less fertilizers, which makes their rents less sensitive to such price variations – a mechanism that we test empirically using household-level survey data.

While early cross-sectional empirical studies typically failed to find evidence of a positive link between income or wealth inequality and conflict (Lichbach, 1989), our results are consistent with more recent work. Macours (2011), for instance, finds that rebel recruitment was more intensive in Nepali districts where inequality between landlords and landless has previously increased. We also contribute to the debate on the effect of inequality within or across ethnic groups. While some studies argue that between-group inequality is a cause of conflict (Cederman *et al.*, 2011; Guariso and Rogall, 2017), other argue that it decreases conflict probability (Mitra and Ray, 2014).[6] Within-group inequality could matter as well (Huber and Mayoral, 2013), for instance if it made it easier for the rich to hire fighters within their own ethnic group (Esteban and Ray, 2008, Esteban and Ray, 2011a).[7] Our final set of results suggest that both between and within ethnic groups inequality shocks (in terms of soil fertility) have a conflict-inducing effect in SSA countries, between-group inequality being especially important in densely populated areas. This is in line with Mwesigye and Matsumoto (2016) who find, based on data from rural Uganda, that land conflicts are more likely to occur in areas characterized by high population growth and ethnically diverse communities. In general, our conclusions lend support to the view that land related violence must be analyzed in conjunction with wider processes of socio-ethnic divisions and discrimination, and with demographic changes. Our findings on the role of population density echo the recent cross-country analysis of Acemoglu *et al.* (2017).

Finally, this paper sheds light on the potential effects of the recent skyrocketing of fertilizer prices and contribute to the debate about technology adoption in SSA agriculture. Some literature has argued that access to modern inputs such as fertilizers were a key factor on the difference

[6] Guariso and Rogall (2017) provide cross-country evidence that economic inequality shocks (rainfall) between ethnic groups increase the likelihood of conflict.

[7] Using a theoretical model from a different perspective, De Luca and Sekeris (2012) show that the intensity of the fight between a rebel group (landless individuals) and landlords is greatest for intermediate values of land inequality between the landlords.

between the rapid growth of agricultural yields in Asia and the stagnation of yields in Africa (Morris *et al.*, 2007); this has led some authors to argue in favor of fertilizer subsidies for African countries (Duflo *et al.*, 2011). Our findings imply that, despite the relatively limited use of fertilizers in SSA countries, variations in nutrient prices do have a significant effect on agricultural yields, income, and political stability.

In the following section, we present our model of land heterogeneity and conflict. Section 3 describes the data and the methodology used to construct our main variables of interest. Section 4 contains our econometric approach, the baseline results, and a number of sensitivity analyses. In section 5 we study the impact of land inequality between and within ethnic groups. The last section concludes.

2 A Model of Land Heterogeneity and Conflict

In this section we develop a dynamic model of conflict among two groups that control an area of land characterized by heterogeneous soil fertility. We build the model in two parts. The first part explores the general dynamics of conflict between the two groups which enables us to make predictions about the relationship between economic shocks, inequality, and conflict. The second part focuses on agricultural output and links fertilizer price, soil fertility, and land rents. This enables us to deliver our two main testable predictions regarding the relationship between fertilizer price, the distribution of soil fertility and the likelihood of conflict.

2.1 Inequality Shocks and Conflict

This component of the model builds on Chassang and Padró i Miquel (2009), who assume that land productivity is homogenous and focus on the relationship between wealth and conflict.[8] Unlike Chassang and Padró i Miquel (2009), however, we consider land to be heterogeneous with respect to soil fertility. In this section, we extend their model to incorporate heterogeneity in land productivity in order to study the relationship between this measure of inequality and conflict.

Consider two groups $i \in \{1, 2\}$ that share a territory of size 2. Assume that each group controls 1 unit of land. The total rent from land i at time t is given by r_{it}. The amount of rent depends on fertilizer use and on soil fertility (we formalize these relationships later).

One group may decide to use violence in order to seize the land of the other group. For the sake of simplicity, we exclude the possibility of peaceful (costless) land transfers. We assume that if a group attacks, it has a first mover advantage and wins the conflict with probability $\mu > 1/2$.[9] If both groups decide to attack, then the probability of winning the conflict is 1/2 for both groups. When conflict occurs, there is a probability, denoted $d \in (0, 1]$, that all existing agricultural production is completely destroyed.[10]

[8]They show that the relationship between poverty and the likelihood of conflict is ambiguous while negative income shocks increase the likelihood of conflict.

[9]One could alternatively make the assumption that the richest group is also the strongest, i.e. it has a larger first mover advantage. Our results hold as long as inequalities are sufficiently large and/or the level of rents is small. A sufficient condition is $|r_{1t} - r_{2t}|/(r_{1t} + r_{2t}) > (\frac{\delta}{1-\delta} + 1 - d)|\mu_1 - \mu_2|$ for all t, where μ_i is the probability of winning for group $i = 1, 2$ if it decides to attack first. Our results are also unchanged if there is no first mover advantage ($\mu = 1/2$) as long as the rents are not symmetric, i.e. $r_{1t} \neq r_{2t}$ for all t. See Dow *et al.* (2017) for a conflict model in which the winning probability is endogenous and depends on agricultural income.

[10]One may also interpret d as the proportion of agricultural production that is destroyed.

In each period of time t, rents from agricultural lands vary, because *fertilizer prices and output prices vary.* Income shocks are captured by the fact that rent r_{it} is independently drawn from a cumulative distribution function $F_i(.)$ over $(0, +\infty)$. The expected value of r_{it} is denoted $\overline{r}_i$, with $\overline{r}_2 < \overline{r}_1$. Each group discounts the future by a factor δ.

The timing of the game is the following. First, r_{1t} and r_{2t} are observed by the two groups. If neither group chooses to attack, then each group produces and consumes their production, and the next period begins. If at least one group decides to attack, there is a decisive war, and the winner of the conflict uses all land to produce, consumes the yields produced, and controls all land forever. While this assumption is convenient, it can be relaxed: our results are qualitatively unchanged if the defeated group regains its land more than one period after the conflict occurred.[11]

Consider the choice of whether to attack for group $i \in \{1, 2\}$ at time t. If peace is reached at time t, then group i obtains the following expected payoff:

$$r_{it} + \delta V_i^P, \tag{1}$$

where V_i^P is the expected continuation value when peace is reached.

If group i decides to deviate and launch an attack, its expected payoff is given by:

$$\mu\left((r_{1t} + r_{2t})(1-d) + \delta V_i^V\right), \tag{2}$$

where V_i^V is the expected continuation value when war occurs and i is victorious.

This continuation value is:

$$V_i^V = E\left[\sum_{\tau=0,\ldots,+\infty} \delta^\tau r_{1\tau}\right] + E\left[\sum_{\tau=0,\ldots,+\infty} \delta^\tau r_{1\tau}\right], \tag{3}$$

or,

$$V_i^V = \frac{\overline{r}_1 + \overline{r}_2}{1-\delta} \equiv V^V. \tag{4}$$

Hence, peace is reached at time t only if:

$$r_{it} + \delta V_i^P > \mu\left((r_{1t} + r_{2t})(1-d) + \delta V^V\right), \tag{5}$$

for $i \in \{1, 2\}$.

To elaborate on this condition, we make the following assumption:

Assumption [no switching]: *The most profitable land (on average) is also the most profitable land at each point in time:* $\Delta r_t = r_{1t} - r_{2t} \geq 0$ *for all* t.

This assumption is sufficient to show that the group with less fertile soil (group 2) has a greater incentive to attack than the other group (group 1). Indeed, by this assumption, we can show that $V_1^P \geq V_2^P$, that is, the expected continuation value in the event of peace cannot be larger for the group who owns the less profitable land than for the group who owns the more

[11] See the appendix for the proof.

profitable land.[12] Hence, using condition (5), we deduce that peace occurs at time t only if group 2 has no incentive to launch an attack, that is:

$$(1-\mu(1-d))\, r_{2t} - \mu(1-d)\, r_{1t} > \delta\left(PV^V - V_2^P\right), \tag{6}$$

Condition (6) illustrates the trade-off between the *opportunity cost* of conflict and the *rapacity gain* (a benefit of conflict) in the current period. To see this, let Ψ_t denote the likelihood that a conflict occurs, independently of expected future play:

$$\Psi_t = \underbrace{\mu(1-d)\, r_{1t}}_{\text{Rapacity gain}} - \underbrace{(1-\mu(1-d))\, r_{2t}}_{\text{Opportunity cost}}. \tag{7}$$

Conflict is more likely (i.e. group 2 has a greater incentive to launch an attack) if the rent of the land-rich (group 1) is temporarily large or if the rent of the land-poor (group 2) is temporarily small. Indeed, independently from future play, group 2 faces a trade-off between the expected current loss from conflict, i.e. the opportunity cost, $(1-\mu(1-d))\, r_{2t}$, and the expected current gain from conflict, i.e. the rapacity gain, $\mu(1-d)\, r_{1t}$. The opportunity cost of conflict increases with the rent from its land, r_{2t}, and the rapacity gain increases with the rent from land controlled by group 1, r_{1t}.

Notice that the expected continuation value V_2^P is bounded above: $V_2^P \leq \frac{1}{2}\frac{\overline{r}_1+\overline{r}_2}{1-\delta}$. Hence, the right hand side in condition (6) is non negative. Thus, a necessary condition for peace is $1-2\mu(1-d) > 0$. This condition can be rewritten as $1-\mu(1-d) > \mu(1-d)$. This implies that the probability that a group loses its current rents is larger than the probability of winning the current rents of the other group. We maintain this assumption throughout the rest of the paper. This assumption excludes the trivial case in which conflict occurs immediately.

The likelihood of conflict defined in (7) can also be written in terms of average rents and the difference between rents:

$$\Psi_t = \frac{1}{2}\Delta r_t - [1-2\mu(1-d)]\, \overline{r}_t, \tag{8}$$

where $\overline{r}_t \equiv (r_{1t}+r_{2t})/2$ is the average rent at time t. This leads to the following result regarding the relationship between average rents, rent inequality, and conflict:

Proposition [Inequality and Conflict]: *Conflict occurs if current rents ($\overline{r}_t$) are low enough or if current rent inequality (Δr_t) is sufficiently large.*

The first part of this Proposition, i.e. conflict occurs if current rents are low enough, echoes Chassang and Padró i Miquel (2009). The second part of the Proposition, i.e. conflict occurs if current rent inequality is sufficiently large, however, differs from the results in the literature. Indeed, while existing studies argue that a higher average level of inequality increases the likelihood of conflict (Esteban and Ray, 2011b; Fearon, 2005), we find that a higher average level of inequality, measured by $\overline{r}_1 - \overline{r}_2$ (holding the average rent constant), may be associated with either more conflict (if V_2^P decreases) or less conflict (if V_2^P increases). Rather, our results suggest that, conflict is more likely to occur *when* current inequalities are high.

[12] More precisely, this holds for the subgame perfect equilibrium in which players use threshold strategies and launch an attack only when the realizations of the left-hand side in condition (6) are lower than a given constant threshold. We also show that such a subgame perfect equilibrium exists.

This general result cannot be easily tested because agricultural rents are difficult to measure accurately in developing countries and are moreover likely to be endogenous to conflict (or correlated with other co-determinants of conflicts). In the empirical section of the paper, we consider fertilizer price variation and the distribution of soil fertility as sources of exogenous and heterogeneous economic shocks. To derive our predictions, we complement our theory with a simple model of agricultural production.

2.2 Soil Fertility, Fertilizer and Conflict

We now specify how rents depend on fertilizer prices and soil fertility. For simplicity, we assume that agricultural production only depends on soil fertility and the quantity of fertilizer used (at the end of the present Section, we discuss why not considering other inputs such as land quantity is not crucial in our context). Let the rent from land i at time t be given by:

$$r_{it} \equiv \pi_t g\left(s_i, f_{it}\right) - c_t f_{it}, \tag{9}$$

where π_t is the output price, g is the production function, s_i is soil fertility, c_t is the price of fertilizer and f_{it} is the quantity of fertilizer used at time t. The production function is nondecreasing in both s_i and f_{it} and strictly concave in f_{it}, i.e. $g_s \geq 0, g_f \geq 0$ and $g_{ff} < 0$. Because fertilizers are comprised of nutrients (nitrate, phosphorus and potassium) and our empirical measure of soil fertility is the natural availability of nutrients in the soil (see next section), fertilizers and soil fertility are assumed to be substitutes, i.e. $g_{sf} < 0$. Existing empirical evidence supports this assumption as long as soil fertility is not extremely low.[13] Assuming that the price of agricultural output is π_t, group i's optimal choice of fertilizer is given by:

$$\text{Max}_{f_{it} \geq 0} \ \left\{\pi_t g\left(s_i, f_{it}\right) - c_t f_{it}\right\}. \tag{10}$$

The first order condition for an interior solution is given by:[14]

$$g_f\left(s_i, f_{it}\right) = \frac{c_t}{\pi_t}. \tag{11}$$

Hence, f_{it} is a decreasing function of fertilizer price, i.e. $\partial f_{it}/\partial c_t = 1/\left(\pi_t g_{ff}\right) < 0$. It is also a decreasing function of soil fertility:

$$\partial f_{it}/\partial s_i = -g_{sf}/g_{ff} < 0. \tag{12}$$

This condition implies that *farmers use lower amounts of fertilizer on fertile soils.*

Notice that, as expected, *rents from land decrease when the price of fertilizers increases.* Indeed, the effect of a marginal increase in the price of fertilizer on agricultural rents is given by:

$$\partial r_{it}/\partial c_t = -f_{it}. \tag{13}$$

[13] Marenya and Barrett (2009b), for instance, provide evidence of an S-shaped relationship between agricultural yields and fertilizer inputs and soil carbon stocks in Western Kenya. Therefore, as a robustness check in the empirical section, we drop the regions where soil fertility is below a certain threshold.

[14] Our results are not qualitatively affected if one take corner solutions into account.

The increase in the rent from land i generated by a marginal increase in the price of fertilizer equals minus the amount of fertilizer used.

We are now able to examine the effect of changes in fertilizer price on the likelihood of conflict. The derivative of the likelihood of conflict (7) with respect to fertilizer price is:

$$\frac{\partial \Psi_t}{\partial c_t} = \mu (1-d) \frac{\partial r_{1t}}{\partial c_t} - (1 - \mu (1-d)) \frac{\partial r_{2t}}{\partial c_t}. \tag{14}$$

Since an increase in the price of fertilizer decreases rents for both groups, the resulting effect on the likelihood of conflict is a priori ambiguous. However, the marginal decrease in the rents from the land with the most fertile soil is smaller than the marginal decrease in the rents from the land with the less fertile soil. Indeed, using (12) and (13), we obtain:

$$\frac{\partial r_{2t}}{\partial c_t} = -f_{2t} < \frac{\partial r_{1t}}{\partial c_t} = -f_{1t} < 0. \tag{15}$$

This condition implies that *the negative effect of fertilizer prices on agricultural rents is magnified for less fertile soils.*

Since the weight placed on rents from soils with lower fertility is larger than the weight placed on rents from soils with greater fertility, $1 - \mu(1-d) > \mu(1-d)$, the right-hand side in (14) is always positive. This result is summarized in the following proposition:

Proposition [Fertilizer Price and Conflict]: *An increase in fertilizer price makes conflict more likely. Formally,*

$$\frac{\partial \Psi_t}{\partial c_t} > 0. \tag{16}$$

This result can be interpreted in terms of current opportunity cost and rapacity gain. An increase in the price of fertilizer decreases the rents provided by the less fertile soil, hence the opportunity cost of fighting decreases. An increase in the price of fertilizer also decreases the rents provided by the soil with greater fertility, which decreases the incentive to seize this land. In other words, the rapacity gain also decreases. Since larger quantities of fertilizer are used on the less fertile soil, the rents from the poorer quality land more sensitive to an increase in fertilizer price. As a result, the former effect is stronger than the latter, which makes conflict more likely.

Next we explore how the impact of an increase in fertilizer price on conflict is affected by the initial distribution of soil fertility. Specifically, we consider the effect of a change in the difference in soil fertility (the effect of a change in average soil fertility level is relegated to the Appendix). To do so, let us denote $\overline{s}$ and Δs as the average and the difference of soil fertility between groups, respectively, with $\overline{s} = \frac{1}{2}(s_1 + s_2)$ and $\Delta s = s_1 - s_2$. Let us (re)define s_1 and s_2 as functions of $\overline{s}$ and Δs with $s_1 = \overline{s} + \frac{1}{2}\Delta s$ and $s_2 = \overline{s} - \frac{1}{2}\Delta s$. Condition (11) becomes:

$$g_f\left(\overline{s} + \frac{1}{2}\Delta s, f_{1t}\right) = \frac{c_t}{\pi_t} \text{ and } g_f\left(\overline{s} - \frac{1}{2}\Delta s, f_{2t}\right) = \frac{c_t}{\pi_t}. \tag{17}$$

These conditions characterize the amount of fertilizer used by each group as a function of the prices, the average soil fertility and the difference in soil fertility.

Now consider the effect of an increase in the difference in soil fertility on $\partial\Psi_t/\partial c_t$. Differentiating (14) with respect to the difference in soil fertility, we obtain:

$$\frac{\partial^2 \Psi_t}{\partial c_t \partial \Delta s} = \mu (1-d) \frac{\partial^2 r_{1t}}{\partial c_t \partial \Delta s} - (1 - \mu (1-d)) \frac{\partial^2 r_{2t}}{\partial c_t \partial \Delta s}. \tag{18}$$

The sign of the effect depends on the weighted difference between the cross derivatives of the rents with respect to the price of fertilizer and the difference in soil fertility. Using (17) and differentiating (13) with respect to the difference in soil fertility, these cross derivatives can be written as:

$$\frac{\partial^2 r_{1t}}{\partial c_t \partial \Delta s} = -\frac{1}{2} \frac{\partial f_{1t}}{\partial s_1} > 0 \text{ and } \frac{\partial^2 r_{2t}}{\partial c_t \partial \Delta s} = \frac{1}{2} \frac{\partial f_{2t}}{\partial s_2} < 0. \tag{19}$$

An increase in the difference of soil fertility dampens the negative effect of fertilizer price on the rents from the more fertile soil and it magnifies the negative effect of fertilizer price on the rent from the less fertile soil. This leads to the following result:

Proposition [Soil Fertility Heterogeneity and Conflict]: *The interaction effect of an increase in fertilizer price and in the difference in soil fertility on the likelihood of conflict is always positive. Formally,*

$$\frac{\partial^2 \Psi_t}{\partial c_t \partial \Delta s} > 0. \tag{20}$$

The intuition of this result is the following. An increase in the difference of soil fertility amplifies the decrease in opportunity cost and attenuates the decrease in rapacity gain. Hence, the cross-effect of increased fertilizer price and inequality in soil fertility on the likelihood of conflict is positive.

Notice that the last proposition is obtained holding the average level of soil fertility constant. How average soil fertility itself affects the relationship between fertilizer price and conflict is theoretically unclear. Section 7.4 of the appendix provides a detailed discussion of this issue. The general conclusion is that a change in average soil fertility may increase or decrease the likelihood of conflict following a fertilizer price shock depending on how soil fertility is measured. For instance, a nonlinear increasing transformation of the variable can reverse the sign of the relationship. Empirically, we indeed find that the sign of this effect varies depending on the indicator used to measure soil fertility.

Discussion. The two main theoretical predictions we investigate can be summarized as follows. First, an increase in fertilizer prices makes conflict more likely. This is because more costly inputs lead to a drop in the opportunity cost of engaging in conflict for the land-poor, which is greater in magnitude than the decrease in the rents that stand to be appropriated from the land-rich.

Second, the probability of conflict increases to an even greater extent if fertilizer prices increase in areas where soil fertility is more heterogeneous. When soil fertility is more heterogeneous, the decrease in the opportunity cost for the land-poor is magnified, and the drop in the rapacity gain is dampened.[15]

[15]Our main predictions therefore relate conflict likelihood to changes in rents triggered by variations in fertilizer prices. How soil fertility heterogeneity affects conflict independently of fertilizer price variations is theoretically

Our agricultural production model demonstrates that increases in fertilizer prices trigger conflict partly due to the fact that the most fertile soils require less fertilizer, which makes the rent from these lands less sensitive to variations in the price of this input. In section B of the online appendix, we provide direct evidence supporting these intermediary results using household data on fertilizer prices combined with information on land value. More precisely, we provide evidence supporting three implications of the agricultural production model that underlies our results, namely that: (a) farmers use lower amounts of fertilizers on fertile soils (an implication of equation 12), (b) rents from land decrease when the price of fertilizer increases (an implication of equation 13), (c) the negative effect of fertilizer prices on agricultural rents is larger for less fertile soils (an implication of equation 15).

We have assumed that both groups use fertilizers. In fact, if the price of fertilizers is sufficiently high, the land-rich (group 1) may stop using these inputs. This would not affect our two main predictions: in that case, the rents of group 1 become insensitive to fertilizer price variations. The rapacity gain would remain constant while the opportunity cost of group 2 would still vary with the price of fertilizer. In the online appendix D, we also consider an extension of the model in which the land-poor group (group 2) faces budget constraints, and as a result stops using fertilizers when the price goes beyond a certain level. This implies that the rents from the less fertile soil become insensitive to fertilizer price variations when the price of fertilizers is sufficiently high. Fertilizer prices now have a non-monotonic effect on conflict while the cross-effect of fertilizer price and soil fertility heterogeneity is still always positive. We provide some evidence of such non-monotonicity; we however find that in the data, the impact of fertilizer price changes on conflict is always positive, as predicted by our baseline model.

Finally, we have assumed that the only exogenous difference in agricultural productivity between the two groups is soil fertility. This leads us to conclude that the group with more fertile soil gets larger agricultural rents than the other group. This intermediate results is consistent with existing evidence (Marenya and Barrett, 2009a,b; Liverpool-Tasie *et al.*, 2017), and thus we believe that our simple agricultural production model is sufficient to capture the main intuitions. Notice that, from a theoretical point of view, exogenous characteristics that influence agricultural productivity beyond soil fertility (such as land quantity) may lead to a situation where the group with more fertile soil gets smaller rents than the other group (if the group with more fertile soil owns a much smaller land surface for instance). In this case, our first prediction could be reversed (to the extend that the weights $1-\mu(1-d)$ and $\mu(1-d)$ are of sufficiently similar magnitude) and our second prediction would be reversed. These alternative predictions are, however, based on an assumption that is not supported by existing empirical evidence and they are not consistent with our main empirical results.

3 Data

We now turn to the description of the data used and the construction of our main variables of

ambiguous. The online appendix C provides more discussion of this relationship. We show that, in contrast with fertilizer prices which only influence the tradeoff between current payoffs (left hand side in condition 6), soil fertility heterogeneity affects both the tradeoff between current payoffs and between future payoffs (right hand side in condition 6). The resulting effect is ambiguous. We however provide some suggestive evidence that, within our sample, soil fertility heterogeneity is positively correlated with conflicts across space.

interest. Testing the predictions of the model first requires defining a level of spatial aggregation. At this level of aggregation, we must then build measures of (i) conflict events, (ii) natural soil fertility, and (iii) variations in fertilizer prices. In this section we summarize the main variables used in our analysis. More information on the variables used in the paper is provided in the online appendix A.

3.1 Unit of analysis

Our units of analysis are cells of size 0.5×0.5 degrees latitude and longitude (around 55×55 kilometers at the equator), covering the entire set of SSA countries. Most of the data we use throughout the paper are available at a more disaggregated level. For this reason, we aggregate the data in order to generate a dataset at the *cell-year* level. We use this level of aggregation rather than administrative boundaries in order to ensure that our unit of observation is not endogenous to conflict events. We assign a country to each cell based on the end-of-period boundaries.

3.2 Conflict data

We use conflict event data from the *Armed Conflict Location and Event dataset* (Raleigh, Linke and Dowd, 2014) which contains information on the geo-location of conflict events in all African countries over the period from 1997 to 2013. These data have been widely used in recent conflict literature.[16] They contain information about the date, the location (longitude/latitude) of conflict events within each country, and the nature of the actors on both sides of the conflicts. Events are compiled from various sources, including press accounts from regional and local news, humanitarian agencies, and research publications. Geographic precision specifies at least the municipality level in more than 95% of cases, and is even finer (i.e. village level) for more than 80% of observations. For each data source, we aggregate the data by year and by 0.5×0.5 degree cell.[17] We focus on Sub-Saharan African countries only; North African countries possess significant reserves of phosphate (in particular in Morocco and Western Sahara), which would affect our identification strategy.

We construct a dummy variable equal to one if at least one conflict occurs in the cell during the year; we interpret this variable as cell-specific *conflict incidence*. As an alternative measure of conflict, we compute a variable containing the number of events observed in the cell during the year. Note that our results are also robust to modeling the onset and ending of cell-specific conflicts separately. Figure 1.A provides a visual representation of our data across cells (a larger version appears in the online appendix, figure A.2).[18] Events are observed over the entire Sub-Saharan African continent, with some clusters appearing in the Great Lakes Regions, Nigeria, and West Africa.

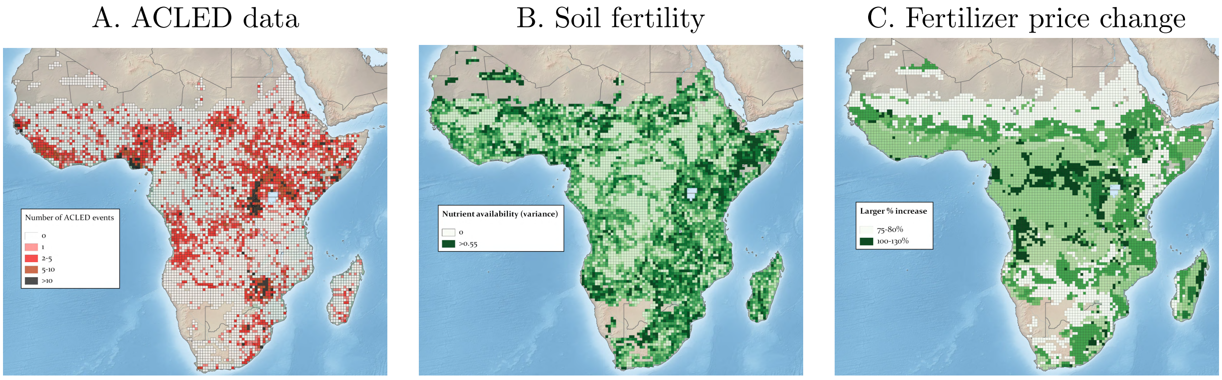
A unique feature of the ACLED dataset is that it includes information about the characteristics of the actors on both sides of the conflicts as well as – for a large subset of observations –

[16]See Besley and Reynal-Querol (2014), Michalopoulos and Papaioannou (2016), or Berman *et al.* (2017) for papers combining data that is structured similarly to our own with the same conflict data we use.

[17]We only keep events that are geolocalized at the finer precision level for our analysis. We also drop duplicated events, i.e. events for which all of the ACLED variable' content (precise date, location, actors, description, etc.) is the same for several observations – in these cases we retain only one observation for the event. This eliminates 1.7% of events. Finally, we drop from our baseline estimates events related to riots and protests in order to conform to the objectives of our model.

[18]Only the cells included in our final sample (i.e. the cells for which all of our main variables are non-missing) appear in these figures.

Figure 1: Data visualization



Note: Figure A: *Armed Conflict Location and Event dataset* (ACLED) events (all events). Figure B: cell-specific soil fertility variance. Data obtained from the International Institute for Applied System Analysis. Soils are ordered in five categories ranging from 1 ("no or slight constraints") to 5 ("very severe constraints"). Figure C: largest increase in the cell-specific fertilizer price index over the 1997-2013 period.

a description of the event. As an alternative measure of conflict incidence, we make use of the richness of the ACLED dataset in order to identify conflicts that are likely to be land-related as well as those that are likely to have occurred between local actors. We thus define "land conflicts" as events (i) involving communal militias (i.e. local armed groups, fighting for a local objective[19]) on both sides and/or (ii) whose description include specific keywords related to land.[20] Figure A.2.B in the online appendix shows the spatial distribution of our measure of land-related conflicts. These conflicts tend to be geographically concentrated, and mostly occur in East Africa, which illustrates a limitation of our measure of land-related conflicts. Specifically, the measure relies on the event description provided in the ACLED data, for which the methods and quality of coding vary widely across countries. In some countries, the description may be absent or very short, or may use a different terminology, which may have prevented us from identifying potential land-related events. For this reason, we use the broader measure of violent conflicts described above in our baseline, and to focus on land-related events for comparison purposes. Note that the ACLED dataset does not include information on who attacked in the first place.

Finally, as an alternative source of conflict data, we use the UCDP-Georeferenced Event dataset (UCDP-GED). This dataset only contains events pertaining to conflicts reaching at least 25 battle-related deaths per year. More importantly for our purposes, it does not include a description of the event, and therefore does not allow to identify – even imprecisely – land conflicts. As shown later, results on overall conflicts are however similar when using this alternative dataset.

3.3 Soil fertility

One of our objectives is to understand how the heterogeneity in soil fertility within regions affects the transmission of agricultural input price shocks to conflict. As an empirical counterpart

[19] ACLED refers to these groups as groups whose "goal [...] is often for the defense of localized territories, livelihoods, community wealth, etc."

[20] We include events whenever the descriptions include the following keywords: "land dispute", "dispute over land", "control of land", "over land", "clash over land", "land grab", "farm land", "land invaders", "land invasion", "land redistribution", "land battle", "over cattle and land", "invade land", "over disputed land", "over a piece of land".

to our model of soil fertility, we use the quantity of nutrients naturally available in the soil. When soil quality is poor, fertilizers can be used as substitutes.

Our baseline data source is the Harmonized World Soil Database (HWSD) built jointly by the FAO (Food and Agriculture Organization) and IIASA (International Institute for Applied System Analysis). This data uses models that incorporate location-specific soil attributes (texture, organic carbon, pH, and total exchangeable bases) to compute, among other measures, an index of nutrient availability.[21] The index assigns soils to one of five categories ranging from 1 ("no or slight constraints") to 5 ("very severe constraints"). Although, the data are available by spatial units of 5 arc-minutes (approx. 9km × 9km at the equator), while our units of observation are cells of 30 arc-minutes. For each cell, we thus compute the average value, the mode, as well as the standard deviation and the variance of this variable, which we use as a measure of the heterogeneity of soil fertility. We also compute the share of soil fertility in each category and define high fertility cells as those for which the share of soil with "no or slight constraints" is above 50%. Figure 1.B plots the spatial distribution of heterogeneity in soil fertility (see Figure A.3 in the online appendix, for larger maps and the distribution of average soil fertility). Figures A.7 to A.9 in the online appendix focus on four specific countries for which we plot the distribution of the mode, average and standard deviation of nutrient availability. In a country such as Ethiopia, soil fertility is relatively concentrated across cells, with most regions displaying a high level of average fertility (Figures A.7 and A.8); fertility within cells, in contrast, appears to be highly heterogeneous (Figure A.9). In Zimbabwe, however, nutrient availability is more concentrated, both across and within cells.

We also use two alternative datasets for robustness. First, the European Commission provides geocoded information on soils in Africa (Jones *et al.*, 2013). For each of the 31 different types of soils in Africa, the data describes their strengths and, weaknesses, as well as qualitative information regarding soil fertility. Based on this information, we rank soils in three categories: naturally fertile, medium level of fertility, and infertile.[22] As a measure of soil heterogeneity, we compute the Herfindahl index of the three types of soils by cell (see Figure A.10 for a visual representation of this index for selected countries). Finally, we incorporate information from the ORNL DAAC *Spatial Data Access tool* that estimates the nitrogen density of the soil at 0.08×0.08 degrees latitude and longitude (Wei *et al.*, 2009). Again, we compute both the average, the standard deviation and the variance of this variable by cell.[23]

3.4 Measuring local variations in fertilizer prices

The final piece of information required to test our predictions are cell-specific fertilizer prices. Direct information on local fertilizer prices is not widely available, and if it were, these prices would likely be endogenous to conflicts. To identify exogenous local variations in fertilizer prices,

[21] Each soil attribute is associated with a rating. The nutrient availability index is calculated as the average of the rating of the attribute with the smallest rating and the average rating of the three other attributes. See Fischer *et al.* (2008) for a detailed description.

[22] The following 8 soils are defined as naturally fertile: Andosols, Cambisols, Chernozems, Fluvisols, Gleysols, Kastanozems, Luvisols, and Phaeozems. See Table A.1 in the online appendix for more information on the different soils.

[23] Michalopoulos (2012) uses information on land quality for agriculture from Ramankutty *et al.* (2002), but this information is only available at a level of 0.5 by 0.5 degrees. This prevents us from using this alternative dataset to compute heterogeneity in land quality at the cell level. The measure is strongly correlated with our average measures of soil quality at the cell level.

we combine three types of data: (i) data on crop specialization (what crops are produced in each cell), (ii) information on crop-specific nutrients uptakes (what mix of nutrients should be included in the fertilizers used for each crop), and (iii) data on the annual price of each nutrient.

The general idea behind our measure is the following. Different cells are specialized in different crops, and the content of the fertilizers used to produce these crops should, in theory, differ (see the end of the present section for further discussion). Fertilizers typically contain a mix of nitrogen, phosphate and potassium (N-P-K), but this mix varies across different types of fertilizers in order to meet crop-specific needs.[24] Hence, combining data on local crop specialization with the international market price of each fertilizer component, we are able to construct a fertilizer price that varies across locations and over time.

First, we identify the main crop(s) produced by each cell using data from the FAO's Global Agro-Ecological Zones (GAEZ). This data is constructed from models that use location characteristics such as climate information (for instance, rainfall and temperature) and soil characteristics. This information is combined with crop' characteristics in order to generate a global GIS raster of the suitability of a grid cell for cultivating each crop. The main advantage of this data is that crop suitability is exogenous to conflicts, as it is not based on actual production. In our benchmark estimations, we use the main crop produced in the cell, which is defined as the crop with the highest suitability level.

Second, we gather information about the required N-P-K mix of nutrients (measured in kg/ha) for each crop. Our benchmark data are generated by the International Plant Nutrient Institute (IPNI) and include information on the nutrient requirements of 42 crops (Table A.2). We also verify that our results are not sensitive to the use of alternative data from the US Department of Agriculture (USDA).

Finally, we obtain the real international market prices of each nutrient from the World Bank Commodities Dataset. A graph of each price series over time is shown in the online appendix, Figure A.1. The three price spikes that occurred between 2008 and 2009 were due to a rise in demand triggered by US biofuel programs and to the imposition of a 135% Chinese export tariff on phosphate (see Schröder *et al.*, 2010).

With these data, we are now equipped to compute, for each cell, the international market price of a kilogram of local fertilizer based on the identified main crop and the required N-P-K mix for this crop:

$$P_{ct} = P_t^N \alpha_{i(c)}^N + P_t^P \alpha_{i(c)}^P + P_t^K \alpha_{i(c)}^K, \tag{21}$$

where the $\{P_t^N,\ P_t^P$, and $P_t^K\}$ represent the real international market prices of nitrogen, phosphate, and potassium, $i(c)$ is the crop with the highest suitability in the cell, and $\{\alpha_{i(c)}^N,\ \alpha_{i(c)}^P$, and $\alpha_{i(c)}^K\}$ are the required proportion (%) of the three nutrients for crop $i(c)$. These proportions are computed from the quantities of nutrients that are removed from the soil at the time of harvest (in kg/ha), with $\alpha_i^N + \alpha_i^P + \alpha_i^K = 1$, for each crop i. We use the main crop in our baseline estimations as re-weighting equation (21) across crops may introduce some noise. However, because multiple cropping practices may be prevalent in many regions, we demonstrate that our results

[24] Most fertilizers contain multiple nutrients. Each fertilizer has a N-P-K rating that consists of three numbers $(\alpha^N, \alpha^P, \alpha^K)$. The first number is the percentage of nitrogen (N), the second number is the percentage of phosphorus pentoxide (P_2O_5) , and the third number is the percentage of potassium oxide (K_20).

are robust to using the five most suitable crops in the computation of our price index.[25][26]

Figure 1.C above shows the largest yearly fertilizer price change observed for each cell over the period (see online appendix Figure A.3 for additional maps on fertilizer prices), which lies between 70% (for cells in which suitable crops require fertilizers with low amounts of phosphate) and almost 130% (for cells in which suitable crops require phosphate intensive fertilizers). As we will show show later, the average change in fertilizer price is positive but relatively low (around 6%), which implies that drops in prices are not uncommon.

Underlying assumptions. When interpreting P_{ct} as a measure of exogenous changes in local fertilizer prices, we make several implicit assumptions. The first is that the international market prices of nutrients are exogenous to conflict. Reverse causality could be an issue in our case only if the use of a specific mix of nutrients in conflict-affected cells in SSA impacts the international market price of this mix. This seems very unlikely: SSA countries are not large consumers nor producers of fertilizers. Consumption of fertilizer relies mostly on imports and almost all of the world production occurs in Europe, North America, and Asia (Hernandez and Torero, 2013).[27] Taken together, the countries in our dataset represent only 4% of world imports and 2% of world consumption of fertilizer.[28]

A second assumption is that fertilizer prices indeed affect fertilizer consumption. Section E.1 of the online appendix shows that this is the case. Using country-level data from FAO-Stats, we show that changes in the prices of nitrogen, phosphate and potassium fertilizers significantly reduce their corresponding consumption or imports by SSA countries.

Third, we assume that changes in the international market prices of nutrients are transmitted to local markets. We investigate the validity of this assumption using data on a subset of countries and nutrients. We gathered data for urea and phosphate, at the market level, for about 350 markets located in 17 countries over a 4-year period (2010-2013).[29] We regress local market-level prices on the international market price of the nutrient, both in logs. Controlling for various sets of fixed effects, we confirm that variations in international market prices do have a significant impact on local prices. Online appendix section E.2 provides additional details and discussion.[30]

A last identification assumption we make is that the fertilizers used by local farmers should indeed reflect, at least to some extent, the "ideal" mix of nutrients they should use given the crop(s) they grow. This assumption is weaker than it might appear at first: we do not need to assume that farmers systematically use the ideal mix. Our only requirement is that the fertilizers that are locally available partly reflect the specialization of the region – that is, that the nutrient

[25] We also use an alternative dataset, M3-Crops Monfreda *et al.* (2008), to identify the main crops. The maps of M3-Crops contain information about actual harvests for 137 different crops for the year 2000.

[26] When identifying the most suitable crop using GAEZ, we need to restrict our sample to the crops for which we observe the ideal NPK mix. For 78% of the cells, this has no incidence as the most suitable crop is one for which we have NPK data. Removing other 22% of the cells has little impact on our results (if anything these are marginally reinforced).

[27] See McArthur and McCord (2017) for a world map of nitrogen fertilizer production facilities.

[28] Authors' computations based on FAO-Stat data. Note also that, since fertilizer use is an investment, more conflict or anticipated conflicts (i.e. insecure property rights) should be associated with less fertilizer use (Jacoby *et al.*, 2002) and lower prices. We find the opposite: variations in international market prices and conflicts are positively correlated.

[29] We do not use this data directly in our conflict estimations because: (i) these local prices are likely to be endogenous to conflicts, and (ii) the time frame and geographical coverage are limited.

[30] We have also gathered country-level data from FAO-Stats to study how imports and consumption of fertilizers vary with world fertilizer prices. Reassuringly, we find that a higher price lowers both imports and consumption.

composition of fertilizers available in, say, a maize producing region should be closer to the mix of nutrients suitable for growing maize than the composition of fertilizers in other regions. In section E.3 of the online appendix, we provide evidence showing that this is indeed the case. We find that, at the country-level, our crop-specific fertilizer prices correlate positively with crop-specific producer prices. Given that we control for common time shocks in these estimations, our identification arises from differences in fertilizer prices across crops, which are themselves driven by different mixes of nutrients.

3.5 Other data

We complement our dataset with additional cell-specific information. In particular, we add cell-level information from PRIO-GRID v.2 (Tollefsen *et al.*, 2012): geographical characteristics, population, and weather. We also control for economic shocks such as variations in the demand for agricultural commodities (Berman and Couttenier, 2015) or in the price of locally produced minerals (Berman *et al.*, 2017). Finally, in the last part of the paper we make use of information on ethnic homelands from Murdock (1959) and Weidmann and Cederman (2010). More details about these variables and their sources are provided in the corresponding robustness section, and the full description appears in the online appendix, section A.5.

3.6 Final sample statistics

Table 1 provides descriptive statistics for our main variables (section A.6 of the online appendix provides additional statistics about the variables used in our robustness exercises). Our sample contains 6,565 cells belonging to 43 SSA countries, and covers the period from 1997 to 2013. Pr(conflict) equals 1 if at least one conflict event occurs within the cell-year.

At least one conflict event occurred in 7% of the 111,605 cell-year units. Unsurprisingly, the unconditional probability of land-related conflicts is much lower at 1%. Conditional on at least one conflict event being observed, the number of conflicts is 1.25 on average. The average level of soil fertility, computed from our baseline data on nutrient availability, is equal to 2.11, which is close to 2, the value defined as "moderate constraints" in our dataset (1 denotes "no or slight constraints", 5 denotes "very severe constraints"). The correlation between the mean and the variance of soil fertility is around 12%. Finally, the average fertilizer price in a cell is around \$210 per metric ton, with significant year-to-year variations observed during our time period: the third quartile of the log-change in this variable shows a 22% increase in price, and the largest observed increase is as high as 130%. Negative price shocks also occur often – the median of the variable is close to zero.

4 Econometric strategy and results

4.1 Empirical specification

We denote cells by c and years by t. Our first prediction states that variations in fertilizer prices have a positive impact on the likelihood of violence within a cell. We estimate the following specification:

$$\text{Conflict}_{ct} = \alpha_1 \ln P_{ct} + \mathbf{D'_{ct}}\beta + \eta_c + \mu_t + \varepsilon_{ct}, \tag{22}$$

Table 1: Descriptive statistics

	Obs.	Mean	S.D.	1^{st} Quartile	Median	3^{rd} Quartile
Pr(conflict> 0)	111605	0.07	0.26	0.00	0.00	0.00
Pr(conflict> 0) (land-related)	111605	0.01	0.08	0.00	0.00	0.00
# conflicts	111605	0.09	0.38	0.00	0.00	0.00
# conflicts (if > 0)	7945	1.25	0.75	0.69	1.10	1.61
Soil fertility (mean)	111588	2.11	0.79	1.42	2.00	2.81
Soil fertility (mode)	111588	2.07	0.92	1.00	2.00	3.00
Soil fertility (variance)	111588	0.38	0.29	0.16	0.42	0.50
Fertilizer price ($/ton)	111605	209.97	106.89	116.27	180.96	286.68
Δ ln fertilizer price	105040	0.06	0.24	-0.10	0.01	0.22
Share irrigated	111605	0.33	1.84	0.00	0.00	0.04
Share agriculture	111605	22.27	26.33	1.53	11.33	33.94

Source: Authors' computations. See main text and online appendix A for data sources. Pr(conflict¿0) is a dummy taking the value 1 if at least one conflict is observed in the cell that year. # conflicts is the number of conflict events observed. Soil fertility (mean), Soil fertility (mode) and Soil fertility (variance) are respectively the mean, mode and variance of nutrient availability within the cell. Fertilizer price ($/ton) is our measure of fertilizer price from equation 21.

where Conflict_{ct} is our conflict variable at the cell-year level, with conflicts being measured in terms of incidence (i.e. a binary variable coding for non-zero events) in our baseline specification, although we also estimate specifications using conflict intensity (number of events), onset and ending as alternative dependent variables. We also study the specific subset of events that we consider to be land-related. P_{ct} is our measure of cell-specific fertilizer price shocks from equation (21), which measures the price of the fertilizer for the main crop produced by the cell c at time t, or, in our robustness section, the price of the fertilizers for the 5 main crops produced by the cell. P_{ct} varies across cells and time because the fertilizers needed for different crops are characterized by different nutrient composition and the prices of these nutrients vary through time, as explained above. $\mathbf{D'_{ct}}$ is a set of potential time-varying cell-specific co-determinants of conflicts that we consider in our robustness checks. These include weather-related conditions, and other world demand or price shocks, related to mineral or agricultural output. Finally, η_c and μ_t are cell and year fixed effects, respectively. η_c accounts for any time-invariant or slow-moving cell characteristics, such as geography, institutions, or culture, that may affect conflict; μ_t captures common time shocks, in particular global commodity price variations, that might be correlated with fertilizer prices.

According to prediction 1, estimates of α_1 should be positive, e.g. spikes in fertilizer price make conflict more likely. In our theoretical model, this is due to the fact that, in an environment where soil fertility is unequally distributed, an increase in the price of fertilizer diminishes the opportunity cost of conflicts more than it decreases the rapacity gain, i.e. the production that can be appropriated through violence for groups who possess less fertile soils. Hence, conflict probability rises.

Our second prediction relates soil fertility heterogeneity to the effect of fertilizer price changes on conflict. We therefore augment equation (22) with an interaction term between fertilizer prices and a measure of soil fertility heterogeneity:

$$\text{Conflict}_{ct} = \alpha_1 \ln P_{ct} + \alpha_2 \ln P_{ct} \times \mathbb{V}(\text{Fertility}_c) + \alpha_3 \ln P_{ct} \times \overline{\text{Fertility}}_c + \mathbf{D'_{ct}}\beta + \eta_c + \mu_t + \varepsilon_{ct}, \quad (23)$$

where $\mathbb{V}(\text{Fertility}_c)$ is a measure of the heterogeneity of soil fertility in cell c. In our baseline estimations, this measure represents the variance of nutrient' availability within the cell. Note that we control for $\overline{\text{Fertility}}_c$, the average fertility level in the cell, as required by the model (see condition (20) and the paragraph before equation (17)). $\mathbf{D'_{ct}}$ includes interaction terms between $\ln P_{ct}$ and cell-specific variables that may be correlated with $\mathbb{V}(\text{Fertility}_c)$, and which we also consider in our robustness checks. These include geographical characteristics, as well as population density, ethnic fractionalization, and ethnic polarization.

Following the second prediction of the model, we expect α_2 to be positive, e.g. increases in fertilizer prices trigger more conflicts in cells in which natural soil fertility is more heterogeneous. This is because, keeping average fertility constant, in regions where soil fertility is more heterogeneous, the drop in the opportunity cost resulting from an increase in fertilizer prices is greater than in more homogeneous regions, and the decrease in the rapacity gain is lower. However, as discussed in the appendix, we note that the sign of α_3 is theoretically ambiguous and may depend on our measure of soil fertility.

Econometric issues. As a benchmark, equations (22) and (23) are estimated using a linear probability model (LPM), and include cell (η_c) and year (μ_t) fixed effects. A LPM more naturally handles multiple fixed effects and spatial correlation, and better deals with rare events than, for instance, a logit model (King and Zeng, 2001). We do, however, check the results we obtain using nonlinear estimators, specifically logit when the dependent variable is conflict incidence, or Poisson when the dependent variable is the number of events. We also consider a number of alternative specifications, adding in particular country-specific time trends or country$\times$year fixed effects to equations (22) and (23).

Given the high spatial resolution of the data, and because both conflicts and crop specialization are geographically clustered, we allow the error term to be spatially correlated, and autocorrelated in our baseline estimations. More precisely, we apply a spatial HAC correction to our standard errors, allowing for both cross-sectional spatial correlation and location-specific serial correlation, following the method developed by Conley (1999). As in Berman *et al.* (2017), we impose no constraint on the temporal decay for the Newey-West/Bartlett kernel that weights serial correlation across time periods. The horizon at which serial correlation is assumed to vanish can be infinite (i.e. 100,000 years). In the spatial dimension we retain a radius of 500km for the spatial kernel.[31] We also provide robustness results allowing for different radiuses and temporal decay.

The main potential threat to causal identification in equations (22) and (23) is arguably the existence of other shocks correlated with variations in fertilizer prices, or of cell characteristics correlated with soil fertility heterogeneity (in equation (23)). We therefore perform extensive robustness exercises, including time-varying variables such as variations in mineral prices, agricultural commodity demand, and climate, as well as time-invariant variables that reflect the geography and other characteristics of the cell, interacted with P_{ct}.

[31] We employ the STATA program from Berman *et al.* (2017), which is based on Hsiang *et al.* (2011) and Guimaraes and Portugal (2010).

4.2 Baseline results

Our baseline results appear in Table 2 below. Columns (1)-(2) consider the full set of conflict events, and columns (3)-(4) restrict the dependent variable to land-related conflicts. We find support for our two theoretical predictions. First, variations in local fertilizer price are positively and significantly correlated with conflict probability (columns (1) and (3)). Second, this effect is magnified in cells characterized by more heterogeneous soil fertility (columns (2) and (4)). The coefficient of the interaction between the average soil fertility level and fertilizer price is significant only in column (2).

Table 2: Baseline results

	(1)	(2)	(3)	(4)
Dep. var.	Conflict incidence			
Conflicts	—— All events ——		—— Land-related ——	
ln fertilizer price	0.119^a	0.156^a	0.018^b	0.019^b
	(0.034)	(0.036)	(0.008)	(0.009)
$\times$ $\mathbb{V}$(Fertility)		0.058^a		0.007^a
		(0.008)		(0.003)
$\times$ $\overline{\text{Fertility}}$		0.021^a		0.001
		(0.003)		(0.001)
Cell and Year FE	Yes			
Countries	42			
Period	1997-2013			
Observations	111605	111588	111605	111588
Average predicted conflict prob.	0.071	0.071	0.006	0.006
Increase in conflict prob. (%) after a 10% increase in fertilizer price[1]				
Average cell	17%	17%	28%	29%
1 s.d. more heterogeneous cell		27%		133%

[c] significant at 10%; [b] significant at 5%; [a] significant at 1%. Conley (1999) standard errors in parentheses, allowing for spatial correlation within a 500km radius and for infinite serial correlation. ln fertilizer price is our baseline fertilizer price shock, computed from using the required NPK mix (from IPNI) for the main crop produced by the cell (from GAEZ). $\overline{\text{Fertility}}$ is the mean of nutrient availability of the cell, i.e minus the categorial variable with values ranging from 1 to 5 (from HWSD). $\mathbb{V}$(Fertility) is the variance of the nutrient availability level within the cell (from HWSD). In columns (1) and (2) the dependent variable is a dummy taking the value 1 if at least a conflict event is observed in the cell during the year, 0 otherwise. In columns (3) and (4), the dependent variable is a dummy taking the value 1 if at least a land-related event is observed in the cell during the year, 0 otherwise. [1]: In columns (1) and (3), we compute the effect of a 10% increase in fertilizer prices, and take the ratio of this effect over the average predicted conflict probability in our sample. In columns (2) and (4), we compute the effect of a 10% increase in fertilizer prices, respectively for the average cell – a cell with the average value of $\overline{\text{Fertility}}$ and $\mathbb{V}$(Fertility) –, and then for a cell with the average value of $\overline{\text{Fertility}}$ and a value of 1 standard deviation above the mean of $\mathbb{V}$(Fertility); we take the ratio of these effects over the predicted conflict probability for the average cell.

When all conflicts events are considered, our estimates from column (1) imply that a 10% increase in fertilizer price raises conflict probability by 1.2 percentage points. Given than the predicted conflict probability is 0.07 on average in our sample, this number represents a 17% larger conflict probability. In cells where land is more heterogeneous, the effect is stronger. For instance, in cells where soil fertility is one standard deviation more heterogeneous than the sample average, the same 10% increase in fertilizer price is associated with a 27% higher likelihood of conflict. Interestingly, these figures are much larger in the case of land-related conflicts.[32]

These large figures might seem surprising at first given the widespread belief that fertilizers are not frequently used in Sub-Saharan Africa and therefore represent a small fraction of overall production costs in these countries. In fact, recent studies have shown that the use of fertilizers in this region is considerably higher than usually though. Sheahan and Barrett (2017), using data from the World Bank LSMS surveys on six SSA countries, find that a third of surveyed farmers use inorganic fertilizers, and that this share reaches 55% in Ethiopia and 77% in Malawi. The average amount of fertilizers used is 57 kg/ha, a number twice larger than estimates from previous studies. Given the average size of land in the LSMS data and the world price of a generic NPK fertilizer (using 2012 data from the World Bank), a back-of-the-envelop calculation implies an annual cost of fertilizers of $291 for the average household. Duflo *et al.* (2011) also report high fertilizer costs in their study in Kenya – around $20 for an harvest of maize of around $100. Importantly, given that most of the labor in African agriculture is familial, labor costs are typically low and inputs such as fertilizers represent a large share of total variable costs.

Underlying mechanisms. One of the reasons why increases in fertilizer prices trigger conflict in our model is that fertile soils use lower amounts of fertilizers, and therefore the rents generated by these lands are less sensitive to variations in fertilizer price. In section B of the online appendix, we provide evidence supporting these theoretical mechanisms. We make use of household information from the *Living Standards Measurement Study* (World Bank) from four African countries (Ethiopia, Malawi, Tanzania, Uganda). First, we find that being located in a fertile area decreases the average quantity of inorganic fertilizers used per acre. Second, we show that land value decreases when our measure of fertilizer prices increases, which supports the claim that rents decline when input prices rise. Third, the drop in land value following fertilizer price spikes is significant only in the least fertile areas. These results support the claim that the negative effect of fertilizer prices on agricultural rents is magnified for less fertile soils. While these results should be interpreted with caution due to the small number of countries and the likely presence of some degree of noise in the data, they are a favorable indication of our key theoretical mechanisms.

4.3 Sensitivity analysis

In this subsection we show that the baseline estimates (Table 2, columns (1) and (2)) are robust to a large battery of sensitivity checks. Most of the tables are relegated to the online appendix (section F), in which we discuss these results at length.

Measurement. We test the sensitivity of the results to alternative methods of measuring our main variables, i.e. (i) conflicts events, (ii) soil fertility, and (iii) fertilizer prices.

In Table A.12 in the online appendix, we use a number of different definitions of conflict. We begin by using an alternative dataset, UCDP-GED, that records only deadly events associated with civil wars, i.e. pertaining to conflicts associated with more than 25 conflict-related deaths

[32] The coefficients displayed in Table 2 are lower in columns (3) and (4), but this is only due to the very low probability of this type of event in our sample. Our quantification take into account these different probabilities.

in a given year. We also use a measure of conflict intensity (number of events), and measures of conflict onset or ending instead of incidence. Our main results hold when using this measure of conflict. Interestingly, we find that, following an increase in fertilizer price, the ending of conflict is less likely in cells where soil fertility is more heterogeneous.

Our baseline proxy for natural soil fertility is a measure of nutrient availability based on the soil texture, pH level, and amount of organic carbon present. This is the closest measure to that described in our model, in which natural soil fertility and fertilizer use are substitutes. We consider alternative measures in Table A.13 in the online appendix. These include: (i) a classification of soils from fertile to infertile from the EU Commission, (ii) the nitrogen density of the soil, and (iii) the percentage of irrigated land in the cell. More details are provided in the online appendix. Since soil fertility and fertilizer use are substitutes as long as natural soil fertility does not fall below a certain level (Marenya and Barrett, 2009b), we check whether our results are robust when we drop cells with very low average or maximum levels of soil fertility, as these are cells for which natural fertility and fertilizer may not be substitutes. Our main results hold under these conditions (see Table A.14).

Another set of measurement issues we consider relates to our fertilizer price variable. Table A.15 in the online appendix reports these results. First, the required nutrient mix by crop is experimentally measured and may vary from one data source to another (Halliday and Trenkel, 1992). We use the US Department of Agriculture as alternative source for the required nutrient mix. Second, in our baseline, we considered only the most suitable crop for the cell, hence assuming away the possibility of multiple cropping. As a robustness check, we identify the five main suitable crops for each cell in order to compute a fertilizer price index, weighted by the relative suitability of each crop within the cell. We also use an alternative data source to identify the main crops from actual harvested area (Monfreda *et al.*, 2008).

Estimation. First, we replicate the baseline estimates allowing for various levels of cross-sectional spatial correlation for a 100, 250, 750km or 1000km radius around the cell's centroid instead of 500km. The standard errors of our coefficients of interest remain statistically significant under these alternatives (Table A.16). Second, in Table A.17, we employ either a logit estimator (when the dependent variable is conflict incidence) or a PPML estimator (when considering the number of events).

Sensitivity to specific countries, years or crops. In the online appendix F.3, we perform a systematic sensitivity analysis and drop each country, year and main crop one by one from our sample. Our coefficients of interest remain remarkably stable. In particular, dropping the years during which commodities prices spiked (2008-2009) has little effect on our results – if anything it slightly increases our estimates. Across countries, Angola is found to contribute significantly to our estimates, although these remain highly significant when we drop it.

Omitted variables. Our final set of robustness checks controls for potential omitted variables (online appendix section A.5 provides the source of each variable, and online appendix section F.4 contains the tables). In Table A.18, we begin by adding country×year fixed effects (or country-specific time-trends) to filter out all time-varying characteristics that could be correlated to both the dynamics of conflicts and with global changes in fertilizer prices. Such fixed effects capture for instance the fact that in 2008, in the midst of the spike of fertilizers' and other commodities' prices, the stock exchange collapsed in South Africa and Nigeria. They can also capture country-specific abilities to subsidize fertilizers during bad times. It is reassuring that our results were not driven by unobserved heterogeneity across country×year. Columns (5) and (6) shows even more restrictive specification, where time-varying fixed effects are defined at the sub-national level (second administrative units). This rules out the possibility that our results are caused by regional characteristics correlated with conflicts and fertilizer prices (e.g., fertilizer subsidies).

Table 3 controls for cell-specific time-varying factors such as commodity prices and weather conditions. In recent literature conflict likelihood has been shown to be associated to fluctuations in the prices of agricultural commodities, produced (Dube and Vargas, 2013, Berman and Couttenier, 2015, McGuirke and Burke, 2017) or consumed (McGuirke and Burke, 2017), to changes in mineral prices (Berman *et al.*, 2017) or to rainfall variations (Harari and Ferrara, 2012, Guariso and Rogall, 2017, Adhvaryu *et al.*, 2017). These shocks could be correlated to our fertilizer prices. Similarly, soil quality might correlate with production and consumption patterns. We sequentially include agricultural price shocks, local rainfall variations, oil and mineral prices. The producer prices of agricultural goods are computed as in Berman and Couttenier (2015) and McGuirke and Burke (2017) from M3-crop data, as the sum of the world price of all crops produced in the cell, weighted by the share of each crop in cultivated land. Consistent with these papers, we find a negative and significant impact of such shocks (column 1). The crop consumption price index is computed as in McGuirke and Burke (2017) as the weighted sum of the world prices of consumed crops, with weights being defined at the country-level by nutritional intake shares from FAO Food Balance Sheets data. As in McGuirke and Burke (2017), it has a positive impact on conflict, although the coefficient is less !precisely estimated. In all instances our coefficients of interest remain statistically significant and of similar magnitude as in our benchmark specification. Similarly, including rainfall, the world price of oil for oil-producing cells or the price of the main mineral produced in the cell (as in Berman *et al.*, 2017) has little impact on our estimates.[33]

Finally, our measures of soil fertility could conceivably be correlated with a number of local characteristics affecting the response of the cell to economic shocks. These can be geographical (e.g. type of terrain) or socio-economic (e.g. income or, population density) in nature. Soil quality and its dispersion could affect ethnic diversity in particular, as found by Michalopoulos (2012). If this is the case, then our estimates would confound the magnifying effect of heterogeneous land endowments with that of ethnic divisions. Table 4 shows that indeed, some of these characteristics do affect the impact of fertilizer prices on conflicts. In columns (1) and (2), we therefore add to our baseline specification a set of interaction terms between fertilizer prices and geographical characteristics. Quite intuitively, we find that fertilizer price variations have a stronger impact in cells where agricultural area or harvested area represent a larger share of the total surface. Agricultural specialization could also correlate with cell-specific characteristics that affect the prevalence of malaria. In column (3), we find that the inclusion of the interaction between

[33]The number of observations is lower than in our baseline sample for several reasons. First, not all countries have consumption data in the FAO Food Balance Sheets, and so the consumer price index cannot be computed for our entire sample. Second, the data on minerals from Berman *et al.* (2017) stops in 2010.

Table 3: Additional time-varying controls

	(1)	(2)	(3)	(4)	(5)
Dep. var.	Conflict incidence				
ln fertilizer price	0.089^b	0.089^b	0.088^b	0.084^c	0.148^a
	(0.045)	(0.045)	(0.044)	(0.045)	(0.052)
$\times$ $\mathbb{V}$(Fertility)					0.040^a
					(0.015)
ln producer price index	-0.077^b	-0.077^b	-0.077^b	0.017	0.005
	(0.034)	(0.034)	(0.033)	(0.037)	(0.070)
$\times$ $\mathbb{V}$(Fertility)					-0.021
					(0.061)
ln consumer price index	0.063	0.064	0.068	-0.032	-0.045
	(0.045)	(0.045)	(0.045)	(0.047)	(0.073)
$\times$ $\mathbb{V}$(Fertility)					0.006
					(0.057)
Rainfall		-0.001	-0.001	-0.001	-0.041^a
		(0.006)	(0.006)	(0.006)	(0.012)
$\times$ $\mathbb{V}$(Fertility)					-0.013
					(0.010)
ln oil price $\times$ oil field			0.030^c	0.028	-0.076
			(0.018)	(0.017)	(0.050)
$\times$ $\mathbb{V}$(Fertility)					0.071^c
					(0.041)
ln main mineral price $\times$ mine				0.061^a	0.168^a
				(0.021)	(0.053)
$\times$ $\mathbb{V}$(Fertility)					0.091
					(0.068)
Cell and Year FE	Yes	Yes	Yes	Yes	Yes
Observations	78386	78386	78386	63686	63672

[c] significant at 10%; [b] significant at 5%; [a] significant at 1%. Conley (1999) standard errors in parentheses, allowing for spatial correlation within a 500km radius and for infinite serial correlation. ln producer price index is computed as in Berman and Couttenier (2015) and McGuirke and Burke (2017) from M3-crop data as the sum of the world price of all crops produced in the cell, weighted by the share of each crop in cultivated land. ln consumer price index is computed as in McGuirke and Burke (2017) as the weighted sum of the world prices of consumed crops, weights being defined at the country-level by nutritional intake shares from FAO Food Balance Sheets data. ln oil price equals the world price of oil interacted with a dummy denoting onshore petroleum deposits (from Prio-Grid v.2). ln price mineral is price of the main mineral produced by the cell during the period, and equals zero if no active mine is recorded in the cell over the period – see Berman *et al.* (2017). Finally, rainfall is the yearly total amount of precipitation (in millimeter) in the cell, based on monthly meteorological statistics from the Global Precipitation Climatology Centre (as appearing in Prio-Grid v.2).

fertilizer price and malaria suitability at the cell-level leaves our result unchanged.[34] In columns (4) to (9), we control for local socio-economic characteristics, specifically: population density, nighttime lights (which could serve as a proxy for either GDP or population density), and ethnic polarization or fractionalization at the cell level.[35] We find that conflict is more responsive to

[34]Cervellati *et al.* (2016) show that suitable conditions for malaria increase the incidence of civil violence. We employ the *Malaria Ecology Index* used by Kiszewski *et al.* (2004) and developed by Gordon McCord as a measure of malaria at the cell level.

[35]These data were obtained from PRIO-GRID, the Global Land Survey dataset, Murdock (1959) and Weidmann and Cederman (2010), respectively. Demographic and economic variables are measured before the start of the

Table 4: Additional time-invariant controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. var.	Conflict incidence								
Controls	—— Geography ——				——— Socio-economic ———				
Ethnic groups data						Murdock		GREG	
ln fertilizer price	0.154^a	0.145^a	0.148^a	0.149^a	0.147^a	0.152^a	0.152^a	0.153^a	0.152^a
	(0.036)	(0.038)	(0.035)	(0.036)	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)
× $\mathbb{V}$(Fertility)	0.054^a	0.055^a	0.059^a	0.056^a	0.054^a	0.058^a	0.058^a	0.059^a	0.059^a
	(0.008)	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
× % agriculture	0.025^c								
	(0.013)								
× % forest	-0.027^a	-0.038^a							
	(0.010)	(0.010)							
× % barren	-0.015^c	-0.017							
	(0.009)	(0.016)							
× % water	-0.033	-0.034							
	(0.037)	(0.038)							
× % harvested		0.139^b							
		(0.062)							
× Malaria index			0.000						
			(0.000)						
× Density pop.				0.000^b					
				(0.000)					
× Lights					0.069^a				
					(0.011)				
× Eth. Pol.						0.035^c		0.034	
						(0.021)		(0.025)	
× Eth. Frac.							0.015		0.017
							(0.010)		(0.011)
Cell and Year FE					Yes				
Observations	111486	96951	111588	111588	111588	111588	111588	110857	110857

[c] significant at 10%; [b] significant at 5%; [a] significant at 1%. Conley (1999) standard errors in parentheses, allowing for spatial correlation within a 500km radius and for infinite serial correlation. % agriculture, % forest, % barren, % water and % harvested are the percentage of the cell's area covered respectively by agricultural land, forest, barren land, water and the percentage harvested (all from PRIO-GRID). The malaria index is a measure of the incidence of malaria at the cell-level from Gordon McCord. Density pop. is the log of population density in 1990 from LandScan. lights is the log of nighttime luminosity in 1995 from PRIO-GRID. Ethnic frac. and ethnic pol. are measures of ethnic fractionalization computed from Murdock (1959) (in cols 6-7), or from the Geo-referencing of Ethnic Groups (Weidmann and Cederman, 2010) (cols. 8-9). ln fertilizer price × $\overline{\text{Fertility}}$ is included as a control variable but its coefficient is not reported.

fertilizer price changes in denser cells, as measured by population or nighttime luminosity, as well as in more ethnically diverse locations (although the estimates are less significant in the latter case). However, in all estimations our coefficients of interest remain remarkably stable.

period. A complete description of the variables is provided in section A.5 of the online appendix.

4.4 Country characteristics and impact of fertilizer price

In this section, we explore whether countries are heterogeneous in the way they react to variations in fertility triggered by fertilizer prices. We begin by examining the role of fertilizer use, and then turning to institutional characteristics.

Fertilizer use. African countries use a lower amount of fertilizers than other developing countries (Morris *et al.*, 2007; Duflo *et al.*, 2011). We expect the effect of fertilizer price and its interaction with heterogeneity in soil fertility to be magnified in countries that use more fertilizers. This is indeed demonstrated in Table 5, columns (1) and (2), where we use data on the intensity of fertilizer use for a subset of 30 countries from FAO-Stats. The effect of fertilizer price changes is clearly increasing in country-level fertilizer use (column (1)). Quantitatively, the direct effect of fertilizer prices on conflict doubles when we compare the country that uses the least fertilizer (Central African Republic) with the country that uses the most (South Africa). Similarly, the magnifying effect of heterogeneity in soil fertility (column (2)) is stronger in countries using more fertilizers (the p-value is 0.106 for the interaction term and 0.16 for the triple interaction term).

Institutional quality. Institutions often play a major role in the way in which the abundance of natural resources affects economic growth and the risk of conflict (see e.g. van der Ploeg, 2011). In our context, fertilizer price variations, through their effects on income and inequality, may trigger conflicts more easily in countries with weak institutions. In columns (3)-(6) of Table 5, we provide evidence demonstrating that, conflict triggered by variations in soil productivity is indeed more likely to occur in countries with weak institutions, especially in areas where soil fertility is more unequally distributed. We use two measures of institutional quality: i) the standard and synthetic measure of institutional quality (International Country Risk Guide) and ii) the standard democracy score (Polity IV). We use pre-sample scores to mitigate endogeneity concerns (year 1996). For each of the two measures of institutional quality, we begin with our baseline specification, which we progressively augment by adding an interaction term between fertilizer price and institutional quality and with a triple interaction term between fertilizer price, soil fertility dispersion and institutional quality. We find that the effect of fertilizer price variations is dampened in countries with institutions of good quality (columns (3) and (5)). Interestingly, this effect is only observed in countries where soil fertility is unevenly distributed (columns (4) and (6)). Finally, we make use of information on the legal security of land tenure. Specifically, we rely on estimates of the extent to which the security of land tenure for indigenous peoples and communities is legally codified in national laws (see section A.5 for more information). We find a smaller effect of fertilizer price on conflict in countries where land tenure is better secured (columns (7) and (8)).

5 Ethnic divisions and population pressure

5.1 Land inequality within and between ethnic groups

Thus far, in the model as well as in our empirical analysis, we have abstracted from the ethnic dimension of conflicts. The issues of heterogeneous access to fertile soils and of ethnic diversity

Table 5: Country characteristics

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var.	Conflict incidence							
Country characteristic	Fertilizer Use		ICRG		Polity IV		Legal Security	
ln fertilizer price	0.016	0.106^b	0.153^a	0.162^a	0.125^a	0.159^a	0.124^b	0.169^a
	(0.041)	(0.049)	(0.048)	(0.049)	(0.042)	(0.043)	(0.054)	(0.055)
$\times$ $\mathbb{V}$(Fertility)		-0.042		0.154^a		0.063^a		0.041
		(0.059)		(0.027)		(0.009)		(0.030)
$\times$ Country characteristic	0.008^a	0.003	-0.068^b	-0.000	-0.001^c	-0.000	-0.009^c	-0.013^b
	(0.002)	(0.002)	(0.027)	(0.030)	(0.001)	(0.001)	(0.005)	(0.005)
$\times$ $\mathbb{V}$(Fertility) $\times$ Country characteristic		0.006		-0.177^a		-0.003^b		-0.008
		(0.004)		(0.047)		(0.001)		(0.013)
Cell and Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	101609	101609	87074	87074	98668	98651	75140	75123

LPM estimations. ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. Conley (1999) standard errors in parentheses, allowing for spatial correlation within a 500km radius and for infinite serial correlation. Fertilizer use intensity is obtained from FAO-Stats, ICRG from the International Country Risk Guide, the standard democracy score from Polity IV, and data on the legal security of land tenure from *LandMark: The Global Platform of Indigenous and Community Lands*. ln fertilizer price $\times$ $\overline{\text{Fertility}}$ is included as a control variable but its coefficient is not reported.

are however likely to be intertwined. First, regions where the soil quality is more heterogeneous are on average more ethnically diverse (Michalopoulos, 2012). Second and more importantly, pre-existing ethnic tensions may exacerbate the effect of rising land inequality. The civil war and subsequent genocide in Rwanda illustrate this possibility, as they have been at least partly triggered by the combination of historical ethnic divisions between Tutsi and Hutu (divisions which are themselves linked to their agricultural practices), population growth and soil depletion (André and Platteau, 1998).

Our previous results show that variations in land inequality over time (triggered by fertilizer price variations that have heterogeneous effects across soil types) increase the likelihood of violence. Changes in land inequality could, however, occur either across or within the homelands of ethnic groups, both of which could in theory affect the likelihood of conflict. Larger between-group inequality could exacerbate grievances, frustrations, and feelings of relative deprivation (Cederman *et al.*, 2011; Guariso and Rogall, 2017), and these issues could coincide with the rapacity gain/opportunity cost channels described in our model. Within-group inequality could matter, as well (Esteban and Ray, 2008, 2011a; Huber and Mayoral, 2013). In Esteban and Ray (2011a), for instance, more inequality within ethnic groups makes conflicts more likely because it makes it easier for the rich to finance them by hiring fighters. This type of effect falls beyond the scope of our theory, but we can nonetheless assess its empirical relevance.

Our methodology is the following. Combining maps of ethnic homelands with our baseline information on nutrient availability, we construct measures of within- and between-ethnic groups soil heterogeneity using a simple variance decomposition.[36] Specifically, we decompose the total

[36] The empirical strategy we use in this section is in the same spirit than Guariso and Rogall (2017). They compute a measure of inequality using rainfall on ethnic homelands and aggregate this information in order to provide cross-country evidence that economic inequality shocks (in terms of rainfall) between ethnic groups increase the likelihood of conflict. However, they do not find any significant effect of economic inequality shocks (rainfall) within ethnic groups on the likelihood of conflict.

variance of soil quality observed in a given cell over the territories of identified ethnic groups into the variance within groups and the variance across groups:[37]

$$\mathbb{V}(\text{Fertility}) = \mathbb{V}_W(\text{Fertility}) + \mathbb{V}_B(\text{Fertility})$$

Or:

$$\frac{1}{A}\sum_{e\in E}\sum_{j\in T(e)}(s_j - s)^2 = \frac{1}{A}\sum_{e\in E}\sum_{j\in T(e)}(s_j - s_e)^2 + \frac{1}{A}\sum_{e\in E}A_e(s_e - s)^2 \tag{24}$$

where e is an ethnic-group, E is the set of ethnic groups observed in the cell, j is our geographical unit of observation of nutrient availability (pixels of 5 arc-minutes), and $T(e)$ is the set of geographical units covered by the homeland of ethnic group e. s_j, s_e, and s denote nutrient availability in area j, average nutrient availability over the area in which group e is observed, and average nutrient availability across all the pixels of the cell, respectively. Finally, A and A_e are, respectively, the sizes (in number of pixels) of the area covered by all ethnic groups and by ethnic group e. In other words, the overall variance of soil fertility in the cell (over the areas covered by at least one ethnic group) is the sum of the average variance within groups and of the variance of the average fertility across groups.

We compute two versions of $\{\mathbb{V}(\text{Fertility}), \mathbb{V}_W(\text{Fertility}), \mathbb{V}_B(\text{Fertility})\}$ based on two alternative datasets that map the borders of ethnic homelands. The first version is based on Murdock (1959); the second version is based on the Geo-referencing of Ethnic Groups (GREG), which is drawn from the Soviet Atlas Narodov Mira (Weidmann and Cederman, 2010). Given our level of spatial disaggregation, our preferred source is Murdock (1959), which contains many more groups than GREG (835 versus 250 over the geographical area we consider). This is an important consideration in order to be able to identify sufficient variation in between-group variance within our cells. For completeness, we present the results obtained using both measures.

Unsurprisingly, the contribution of the within-group component is very large in all cases, although it is smaller when using the Murdock (1959) data (92% versus 96% using GREG).[38] In our baseline specification, we replace the interaction term between fertilizer price and $\mathbb{V}(\text{Fertility})$ by two interaction terms between fertilizer price and each component of the variance of soil fertility. The between-group variance is set to zero in cells where only one ethnic group is present (in a robustness check, we control for interactions with the number of ethnic groups).

Table 6 contains the results. Columns (1) to (4) show the estimates obtained using our baseline data on ethnicity from Murdock (1959). Columns (5) to (8) use GREG data instead. In each case, we begin by reporting the results on the effect of the overall variance (cols. 1 and 5),[39] and then split the variable into its within and between-group components (cols. 2 and 6). The remaining of regressions add controls for ethnic polarization (cols. 3 and 7) and the number of ethnic groups in the cell (cols. 4 and 8). Within-group land inequality is found to be a significant predictor in all columns, and quantitatively the effect is relatively stable across different ethnic group datasets.

[37]See for instance Helpman *et al.* (forthcoming) for an application of such a decomposition to within and between sector *wage* inequality.

[38]Figure A.6 in the online appendix plots the variances, within and across groups, obtained using the Murdock (1959) data.

[39]Note that the estimated coefficients are not the same as in our baseline table. This is because the variance of soil fertility is computed over the territories wherein at least one ethnic group is identified, which do not necessarily cover the entire cell.

Table 6: Soil fertility heterogeneity within and between ethnic groups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var.	Conflict incidence							
Ethnic groups data	Murdock				GREG			
ln fertilizer price	0.152^a	0.151^a	0.147^a	0.152^a	0.156^a	0.156^a	0.152^a	0.136^a
	(0.036)	(0.036)	(0.035)	(0.035)	(0.036)	(0.036)	(0.035)	(0.035)
$\times$ $\mathbb{V}$(Fertility)	0.047^a				0.060^a			
	(0.007)				(0.008)			
$\times$ $\mathbb{V}_B$(Fertility)		0.090^a	0.072^b	0.081^b		0.054	0.026	0.010
		(0.030)	(0.031)	(0.033)		(0.040)	(0.039)	(0.038)
$\times$ $\mathbb{V}_W$(Fertility)		0.044^a	0.046^a	0.045^a		0.060^a	0.063^a	0.064^a
		(0.007)	(0.007)	(0.007)		(0.009)	(0.009)	(0.009)
$\times$ Ethn. Pol.			0.037^c	0.060^b			0.039	-0.042
			(0.022)	(0.029)			(0.026)	(0.036)
$\times$ # groups				-0.004				0.013^a
				(0.004)				(0.005)
Cell and Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	111316	111316	111316	111316	110755	110755	110755	110755

LPM estimations. ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. Conley (1999) standard errors in parentheses, allowing for spatial correlation within a 500km radius and for infinite serial correlation. ln fertilizer price $\times$ $\overline{\text{Fertility}}$ is included as a control variable. $\mathbb{V}_W$(Fertility) is the variance of soil fertility (nutrient availability) within ethnic groups, computed by cell. $\mathbb{V}_B$(Fertility) is the variance of soil fertility between ethnic groups, computed by cell. $\mathbb{V}$(Fertility) is the overall variance of soil fertility, computed by cell. Ethn. Pol. is an index of ethnic polarization in the cell. # groups is the number of ethnic groups present in the cell. Data on ethnic groups are obtained from Murdock (1959) in cols 1-4, and from the Geo-referencing of Ethnic Groups (Weidmann and Cederman, 2010) in cols. 5-8.

Between-group inequality, on the other hand, only magnifies the effect of fertilizer price variations in columns (2) to (4), when using Murdock's data. In columns (5) to (8) the coefficient estimate is positive but statistically insignificant at conventional levels. A reason that might explain this discrepancy is the highest variability of the measure in the first four columns due to the largest number of groups present in the Murdock dataset. Controlling for ethnic polarization or the number of ethnic groups has little impact on our coefficients of interest.

5.2 The role of population pressure

As mentioned above, anecdotal evidence suggests that competition for fertile lands is more likely to trigger conflict in densely populated areas. Returning to the example of Rwanda, while population pressure has surely been one of the war's causes, several scholars have even argued that this in fact served as the main motivation for a genocide that was purposefully planned by a small Hutu elite (e.g. Diamond, 2005). More recently, Acemoglu *et al.* (2017) used cross-country data to show that exogenous changes in population growth were positively associated with civil wars.

In Table 7, we investigate the links between population pressure, soil fertility shocks, and conflicts. In columns (1) and (3), we add interaction terms between the log of population density in the cell in 1990 (to mitigate endogeneity concerns) and fertilizer prices. Fertilizer price shocks

Table 7: Ethnic inequality and population pressure

	(1)	(2)	(3)	(4)
Dep. var.	Conflict incidence			
Ethnic groups data	—— Murdock ——		—— GREG ——	
ln fertilizer price	0.140^a	0.144^a	0.145^a	0.149^a
	(0.038)	(0.038)	(0.038)	(0.038)
$\times$ $\mathbb{V}_B$(Fertility)	0.069^b	0.002	0.031	-0.018
	(0.034)	(0.032)	(0.047)	(0.048)
$\times$ $\mathbb{V}_W$(Fertility)	0.051^a	0.044^a	0.071^a	0.063^a
	(0.008)	(0.007)	(0.010)	(0.010)
$\times$ Density pop.	0.003^b	-0.000	0.003^b	0.000
	(0.001)	(0.002)	(0.001)	(0.001)
$\times$ Ethn. Pol.	0.037	0.040^c	0.026	0.027
	(0.024)	(0.024)	(0.028)	(0.028)
$\times$ $\mathbb{V}_B$(Fertility) $\times$ Density pop.		0.030^b		0.026
		(0.013)		(0.017)
$\times$ $\mathbb{V}_W$(Fertility) $\times$ Density pop.		0.007^a		0.010^a
		(0.003)		(0.003)
Cell and Year FE	Yes	Yes	Yes	Yes
Observations	97580	97580	97410	97410

LPM estimations. ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. Conley (1999) standard errors in parentheses, allowing for spatial correlation within a 500km radius and for infinite serial correlation. ln fertilizer price $\times$ $\overline{\text{Fertility}}$ is included as a control variable. $\mathbb{V}_W$(Fertility) is the variance of soil fertility (nutrient availability) within ethnic groups, computed by cell. $\mathbb{V}_B$(Fertility) is the variance of soil fertility between ethnic groups, computed by cell. Density pop. is the log of population density in 1990. Ethn. Pol. is an index of ethnic polarization in the cell. Data on ethnic groups come from Murdock (1959) in cols. 1-2, and from the Geo-referencing of Ethnic Groups (Weidmann and Cederman, 2010) in cols. 3-4. The F-test (and p-value) reports whether the estimates of $\times$ $\mathbb{V}_B$(Fertility) $\times$ Density pop. and $\times$ $\mathbb{V}_W$(Fertility) $\times$ Density pop. are statistically significantly different.

indeed have a stronger effect in densely populated areas. What is more, this effect is only observed in cells where land is unequally distributed. When the variance of soil quality is zero, population density no longer has an effect (cols. 2 and 4). On the other hand, in denser areas, the impact of both within- and between-group inequality on conflict are stronger.[40] Although less precisely estimated, the interaction between population density and between-group inequality has a much larger coefficient than the interaction with inequality within groups.[41]

In a nutshell, the results presented in Tables 6 and 7 imply that violence is more likely to occur when inequality rises, whether between or within groups, and especially in densely populated areas.[42] Between-group inequality appears to contribute more than within-group inequality when population density is high. Yet, changes in within-group inequality also matter significantly: a larger part of our sample is composed of cells in which a single ethnic group is observed (hence all of the variance is within-group); in these cells, fertilizer prices do have a significant impact on

[40] The p-value of the coefficient on the triple interaction with between-group inequality on column (4) is 0.12.

[41] The null hypothesis of the F-test, i.e. the equality of the coefficient estimates on the $\times$ $\mathbb{V}_B$(Fertility) $\times$ Density pop. and $\times$ $\mathbb{V}_W$(Fertility) $\times$ Density pop., is rejected in column (2) but (4).

[42] Section F.5 of the online appendix contains a number of additional robustness checks of these results (controlling for ethnic fractionalization, and for country $\times$ year fixed effects).

conflict, especially when soil fertility is heterogeneous. More generally, the findings in this section clearly point to the need for more research focusing on the nexus between inequality, population growth, and conflicts.

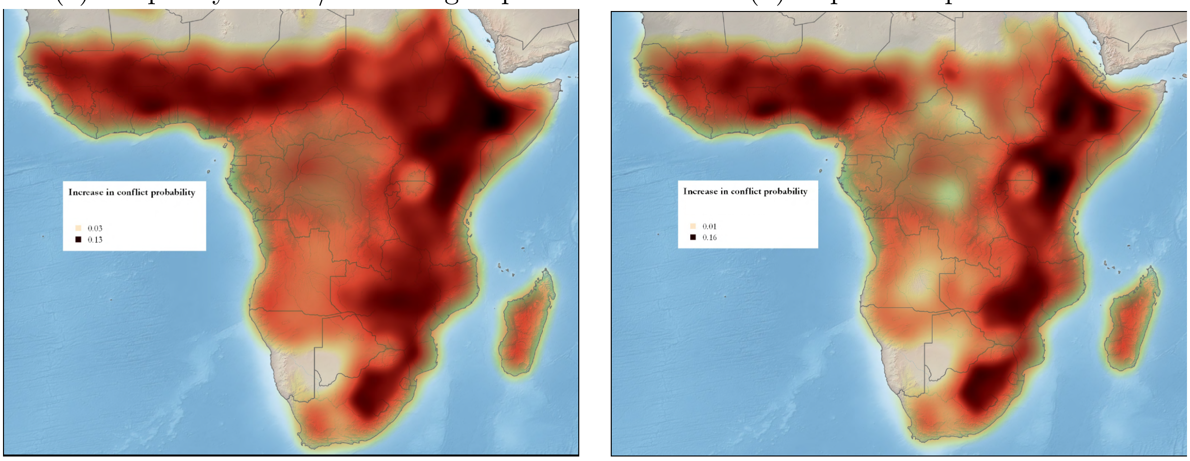
5.3 The geography of soil fertility and conflicts

Which areas are more likely to be affected by variations in soil productivity? Figures 2.a and 2.b provide a visual representation of the predicted impact of a one standard deviation increase in fertilizer prices on conflict across Sub-Saharan countries. Figure 2.a is based on column (3) of Table 6: the heterogeneity of the effect arises primarily from differences across regions in terms of land inequality (either between or within ethnic groups). Figure 2.b is constructed from column (2) of Table 7, i.e. we consider also variations in terms of population density. Note that Figures A.5 and A.6 in the online appendix show the spatial distribution of each of these variables.

Figure 2: Effect of a one standard deviation increase in fertilizer prices on conflict

(a) Inequality within/between groups

(b) Population pressure



Note: These figures represent the predicted increase in the probability of conflict occurring, following a one standard deviation increase in fertilizer prices. Figure (a) plots the predictions obtained from column (3) of Table 6; Figure (b) adds an interaction with population density, and displays the predictions from column (2) of Table 7.

When considering only the variance in soil fertility, a stronger effect is found along a diagonal that begins in Northern Ethiopia, and continues to South Africa. The estimated impact becomes more heterogeneous when population density is taken into account. In particular, the increase in conflict probability becomes much stronger around Nigeria, Rwanda and more generally in the Great Lakes region. It remains high in most parts of Ethiopia, Kenya, and in Eastern South Africa, regions characterized by large populations and heterogeneous land endowments.

These estimates correspond to conflict narratives in these regions. Peters (2004), for instance, writes that:

> *"[...] some of the most intense competition and conflict over land resources and patterns of exclusion are found in the most densely populated areas such as Rwanda and Burundi, where commentators have included land conflicts in the complex causes of 'ethnic' hostility and civil war, and in the Kenyan Highlands and Hausa areas of Northern Nigeria where land sales and landlessness have long been common. Even where overall population density may not be high, intense competition has developed over valued resources, such as wetlands and river valleys in*

semi-arid regions or in areas with a single annual rainy season. Many of these, dubbed 'key resources' by ecologists working in Southern Africa [...], are coming under intensified use, and generating increased social competition and conflict among farmers [...]" (Peters, 2004, p.293).

6 Conclusion

In this paper, we provide an analysis of the effect of variations in soil productivity on violence at the local level. From a theoretical perspective, we show that changes in land inequality, defined as the level of geographical dispersion of natural agricultural soil fertility, positively affects the likelihood of conflict. Changes in fertilizer prices, through their effect on income and inequality, also increase the likelihood of conflict, especially in regions characterized by high levels of initial inequality. Combining data on local agricultural specialization, soil fertility and conflict events over SSA countries with information on international market prices of fertilizer, we find support for these predictions. We conduct a variety of robustness exercises, controlling for potential omitted factors, changes in estimation techniques, and alternative methods of measuring our key variables. We also find that country characteristics play a role: specifically, the impact of changes in fertilizer price – and hence, the effect of increases in soil fertility heterogeneity – is magnified in countries with weak institutions, or more insecure land tenure arrangements for local communities.

In the last part of the paper we incorporate an ethnic dimension to the analysis, and find that the distribution of soil fertility both within and across ethnic groups matters. Between-group inequality matters especially in densely populated areas, a result that accords well with case-specific evidence on several well-known civil war episodes throughout Africa.

Our work has a number of implications and indicates several avenues for future research. With respect to policy, our results suggest that fertilizer price fluctuations have a significant effect on the occurrence of violence because they generate agricultural yields fluctuations. Therefore, policies aiming at limiting such fluctuations and reducing fertilizer prices could be considered as tools to reduce conflict. Land redistribution – an issue often neglected in empirical papers dealing with the roots of civil wars in Africa – should also be a key consideration of strategies to reduce conflict.

In general, our findings imply that inequality in access to fertile lands, both within and across ethnic groups, must be considered as a serious threat to peace at the local-level. The results presented in the last section suggest that complex interactions exist between land inequality both between and across ethnic groups, population density, soil fertility, and conflict. The model presented in this paper features only some of these elements. One specific direction for future research would be to extend our theory to include more than two groups. Such a model could be used to shed light on these complex interactions, as well as to study the emergence and the diffusion of conflicts over space as a result of unevenly distributed economic shocks.

References

Acemoglu, D., Fergusson, L. and Johnson, S. (2017), "Population and Civil War", Working Paper 23322, National Bureau of Economic Research.

Adhvaryu, A., Fenske, J., Khanna, G. and Nyshadham, A. (2017), "Resources, conflict and economic development in Africa", Unpublished manuscript.

André, C. and Platteau, J.-P. (1998), "Land relations under unbearable stress: Rwanda caught in the Malthusian trap", *Journal of Economic Behavior & Organization*, vol. 34 n° 1: pp. 1 – 47.

Bazzi, S. and Blattman, C. (2014), "Economic Shocks and Conflict: Evidence from Commodity Prices", *American Economic Journal: Macroeconomics*, vol. 6 n° 4: pp. 1–38.

Berman, N. and Couttenier, M. (2015), "External Shocks, Internal Shots: The Geography of Civil Conflicts", *Review of Economics and Statistics*, vol. 97 n° 4: pp. 758–776.

Berman, N., Couttenier, M., Rohner, D. and Thoenig, M. (2017), "This Mine is Mine! How Minerals Fuel Conflicts in Africa", *American Economic Review*, vol. forthcoming.

Besley, T. and Reynal-Querol, M. (2014), "The Legacy of Historical Conflicts: Evidence from Africa", *American Political Science Review*, vol. 108 n° 2: pp. 319–336.

Van den Brink, R., Bromley, D. W. and Chavas, J.-P. (1995), "The economics of Cain and Abel: Agro-pastoral property rights in the Sahel", *Journal of Development Studies*, vol. 31 n° 3: pp. 373–399.

Cederman, L.-E., Weidmann, N. and Gleditsch, K. (2011), "Horizontal Inequalities and Ethnonationalist Civil War: A Global Comparison", *American Political Science Review*, vol. 105 n° 3: pp. 478–495.

Cervellati, M., Esposito, E., Sunde, U. and Valmori, S. (2016), "Malaria Risk and Civil Violence", CEPR Discussion Papers 11496, C.E.P.R. Discussion Papers.

Chassang, S. and Padró i Miquel, G. (2009), "Economic Shocks and civil war", *Quarterly Journal of Political Science*, vol. 4 n° 3: pp. 211–28.

Conley, T. (1999), "GMM Estimation with Cross Sectional Dependence", *Journal of Econometrics*, vol. 92: pp. 1–45.

Couttenier, M. and Soubeyran, R. (2014), "Drought and Civil War In Sub-Saharan Africa", *Economic Journal*, vol. 124 n° 575: pp. 201–244.

De Luca, G. and Sekeris, P. G. (2012), "Land inequality and conflict intensity", *Public Choice*, vol. 150 n° 1: pp. 119–135.

Di Falco, S., Laurent-Lucchetti, J., Veronesi, M. and Kohlin, G. (2017), "Property Rights, Water Scarcity and Land Disputes", Technical report, Mimeo.

Diamond, J. (2005), *Collapse: how societies choose to fail or succeed*, New York, Viking.

Dow, G. K., Mitchell, L. and Reed, C. G. (2017), "The economics of early warfare over land", *Journal of Development Economics*, vol. 127 n° Supplement C: pp. 297 – 305.

Dube, O. and Vargas, J. F. (2013), "Commodity Price Shocks and Civil Conflict: Evidence from Colombia", *Review of Economic Studies*, vol. 80 n° 4: pp. 1384–1421.

Duflo, E., Kremer, M. and Robinson, J. (2011), "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya", *American Economic Review*, vol. 101 nº 6: pp. 2350–90.

Esteban, J. and Ray, D. (1999), "Conflict and Distribution", *Journal of Economic Theory*, vol. 87 nº 2: pp. 379 – 415.

Esteban, J. and Ray, D. (2008), "On the Salience of Ethnic Conflict", *American Economic Review*, vol. 98 nº 5: pp. 2185–2202.

Esteban, J. and Ray, D. (2011a), "A Model Of Ethnic Conflict", *Journal of the European Economic Association*, vol. 9 nº 3: pp. 496–521.

Esteban, J. and Ray, D. (2011b), "Linking Conflict to Inequality and Polarization", *American Economic Review*, vol. 101 nº 4: pp. 1345–74.

Esteban, J., Mayoral, L. and Ray, D. (2012), "Ethnicity and Conflict: An Empirical Study", *American Economic Review*, vol. 102 nº 4: pp. 1310–42.

Faris, S. (2009), *Forecast: The Surprising and Immediate Consequences of Climate Change*, New York: Holt and Company.

Fearon, J. (2005), "Primary commodity exports and civil war", *Journal of Conflict Resolution*, vol. 49 nº 4: pp. 483–507.

Fearon, J. (2007), "Economic Development, Insurgency, and Civil War", in Helpman and Elhanan (editors), *Institutions and Economic Performance*, Earthscan, London.

Fearon, J. and Laitin, D. (2003), "Ethnicity, Insurgency and Civil War", *American Policial Science Review*, vol. 97 nº 1: pp. 75–90.

Guardado, J. (2016), "Land Tenure, Price Shocks, and Insurgency: Evidence from Peru and Colombia", *Unpublished Manuscript*.

Guariso, A. and Rogall, T. (2017), "Rainfall Inequality, Political Power, and Ethnic Conflict in Africa", Discussion Paper 391/2017, LICOS Discussion Paper Series.

Guimaraes, P. and Portugal, P. (2010), "A simple feasible procedure to fit models with high-dimensional fixed effects", *Stata Journal*, vol. 10 nº 4: pp. 628–649.

Halliday, D. J. and Trenkel, M. E. (1992), *IFA world fertilizer use manual*, International Fertilizer Industry Association, Paris.

Harari, M. and Ferrara, E. L. (2012), "Conflict, Climate and Cells: A Disaggregated Analysis", Working Papers 461, IGIER, Bocconi University.

Helpman, E., Itskhoki, O., Muendler, M.-A. and Redding, S. (forthcoming), "Trade and Inequality: From Theory to Estimation", *Review of Economic Studies*.

Hernandez, M. A. and Torero, M. (2013), "Market Concentration and Pricing Behavior in the Fertilizer Industry: A Global Approach", *Agricultural Economics*, vol. 44 nº 6: pp. 723–734.

Hidalgo, F. D., Naidu, S., Nichter, S. and Richardson, N. (2010), "Economic Determinants of Land Invasions", *Review of Economics and Statistics*, vol. 92 nº 3: pp. 505–523.

Hsiang, S., Meng, K. and Cane, M. (2011), "Civil Conflicts are Associated with the Global Climate", *Nature*, vol. 476: pp. 438–441.

Hsiang, S. M., Burke, M. and Miguel, E. (2013), "Quantifying the Influence of Climate on Human Conflict", *Science*, vol. 341 n° 6151.

Huber, J. D. and Mayoral, L. (2013), "Inequality, Ethnicity and Civil Conflict", Working Papers 744, Barcelona Graduate School of Economics.

Iyigun, M., Nunn, N. and Qian, N. (2015), "Agricultural Productivity, Conflict and State Size: Evidence from Potatoes, 1400-1900", *Mimeo.*

Iyigun, M., Nunn, N. and Qian, N. (2017), "Winter is Coming: The Long-Run Effects of Climate Change on Conflict, 1400-1900", Working Paper 23033, National Bureau of Economic Research.

Jacoby, H. G., Li, G. and Rozelle, S. (2002), "Hazards of Expropriation: Tenure Insecurity and Investment in Rural China", *American Economic Review*, vol. 92 n° 5: pp. 1420–1447.

Jones, A. *et al.* (2013), "Soil Atlas of Africa", Technical report, Publications Office of the European Union, Luxembourg. 176 pp.

King, G. and Zeng, L. (2001), "Logistic Regression in Rare Events Data", *Political Analysis*, vol. 9: pp. 137–163.

Kiszewski, A., Mellinger, A., Spielman, A., Malaney, P., Sachs, S. and Sachs, J. (2004), "A global index representing the stability of malaria transmission", *American Journal of Tropical Medecine and Hygiene*, vol. 70 n° 5: pp. 486–498.

Lichbach, M. I. (1989), "An Evaluation of "Does Economic Inequality Breed Political Conflict?" Studies", *World Politics*, vol. 41 n° 4: pp. 431–470.

Liverpool-Tasie, L. S. O., Omonona, B. T., Sanou, A. and Ogunleye, W. O. (2017), "Is increasing inorganic fertilizer use for maize production in SSA a profitable proposition? Evidence from Nigeria", *Food Policy*, vol. 67 n° Supplement C: pp. 41 – 51.

Macours, K. (2011), "Increasing inequality and civil conflict in Nepal", *Oxford Economic Papers*, vol. 63 n° 1: p. 1.

Marenya, P. P. and Barrett, C. B. (2009a), "Soil quality and fertilizer use rates among smallholder farmers in western Kenya", *Agricultural Economics*, vol. 40 n° 5: pp. 561–572.

Marenya, P. P. and Barrett, C. B. (2009b), "State-conditional Fertilizer Yield Response on Western Kenyan Farms", *American Journal of Agricultural Economics*, vol. 91 n° 4: pp. 991–1006.

McArthur, J. W. and McCord, G. C. (2017), "Fertilizing growth: Agricultural inputs and their effects in economic development", *Journal of Development Economics*, vol. 127: pp. 133 – 152.

McGuirke, E. and Burke, M. (2017), "The economic origin of conflict in Africa", NBER Working Paper 23056.

Michalopoulos, S. (2012), "The Origins of Ethnolinguistic Diversity", *American Economic Review*, vol. 102 n° 4: pp. 1508–39.

Michalopoulos, S. and Papaioannou, E. (2016), "The Long-Run Effects of the Scramble for Africa", *American Economic Review*, vol. 106 n° 7: pp. 1802–1848.

Miguel, E., Satyanath, S. and Sergenti, E. (2004), "Economic Shocks and Civil Conflict: An Instrumental Variables Approach", *Journal of Political Economy*, vol. 112 n° 4: pp. 725–753.

MITRA, A. and RAY, D. (2014), "Implications of an Economic Theory of Conflict: Hindu-Muslim Violence in India", *Journal of Political Economy*, vol. 122 n° 4: pp. 719–765.

MONFREDA, C., RAMANKUTTY, N. and FOLEY, J. A. (2008), "Farming the planet: Geo-graphic Distribution of Crop Areas, Yields, Physiological Types, and Net Primary Production in the Year 2000", *Global Biochemical Cycles.*

MONTALVO, J. G. and REYNAL-QUEROL, M. (2005), "Ethnic Polarization, Potential Conflict, and Civil Wars", *American Economic Review*, vol. 95 n° 3: pp. 796–816.

MORRIS, M., KELLY, V. A., KOPICKI, R. J. and BYERLEE, D. (2007), "Fertilizer Use in African Agriculture : Lessons Learned and Good Practice Guidelines", Technical report, ARD Directions in Development Washington, DC: World Bank.

MURDOCK, G. P. (1959), *Africa: Its Peoples and their Culture History*, McGraw-Hill.

MWESIGYE, F. and MATSUMOTO, T. (2016), "The Effect of Population Pressure and Internal Migration on Land Conflicts: Implications for Agricultural Productivity in Uganda", *World Development*, vol. 79 n° C: pp. 25–39.

PETERS, P. E. (2004), "Inequality and social conflict over land in Africa", *Journal of Agrarian Change*, vol. 4 n° 3: pp. 269–314.

VAN DER PLOEG, F. (2011), "Natural Resources: Curse or Blessing?", *Journal of Economic Literature*, vol. 49 n° 2: pp. 366–420.

RALEIGH, A. L., CLIONADH and DOWD, C. (2014), "Armed Conflict Location and Event Data Project (ACLED) Codebook 3", Centre for the study of civil war, international peace research institute, oslo (prio).

RAMANKUTTY, N., FOLEY, J., NORMAN, J. and MCSWEENEY, K. (2002), "Global distribution of cultivable lands: current patterns and sensitivity to possible climate change", *Global Ecology and Biogeography*, vol. 11 n° 5: pp. 377–392.

ROSS, M. (2006), "A Closer Look at Oil, Diamonds, and Civil War", *Annual Review of Political Science*, vol. 9 n° 1: pp. 265–300.

ROSS, M. L. (2004), "What Do We Know about Natural Resources and Civil War?", *Journal of Peace Research*, vol. 41 n° 3: pp. 337–356.

SARSONS, H. (2015), "Rainfall and conflict: A cautionary tale", *Journal of Development Economics*, vol. 115: pp. 62 – 72.

SCHRÖDER, J., SMIT, A., CORDELL, D. and ROSEMARIN, A. (2010), "Sustainable Use of Phosphorus", Technical report, Plant Research International, part of Wageningen UR Business Unit Agrosystems Report 357.

SHEAHAN, M. and BARRETT, C. (2017), "Ten striking facts about agricultural input use in Sub-Saharan Africa", *Food policy*, vol. 67: pp. 12–25.

SHEAHAN, M. and BARRETT, C. B. (2014), "Understanding the agricultural input landscape in Sub-Saharan Africa : recent plot, household, and community-level evidence", Technical report, Policy Research working paper ; no. WPS 7014. Washington, DC: World Bank Group.

SANCHEZ DE LA SIERRA, R. (2017), "On the Origns of the State: Stationary Bandits and Taxation in Eastern Congo", Technical report, Mimeo UC Berkeley.

Tollefsen, A. F., Strand, H. and Buhaug, H. (2012), "PRIO-GRID: A unified spatial data structure", *Journal of Peace Research*, vol. 49 n° 2: pp. 363–374.

Wei, Y., Santhana-Vannan, S.-K. and Cook, R. (2009), "Discover, visualize, and deliver geospatial data through OGC standards-based WebGIS system", *Geoinformatics, 17th International Conference*, pp. 1–6.

Weidmann, J. K. R., Nils B. and Cederman, L.-E. (2010), "Representing Ethnic Groups in Space: A New Dataset", *Journal of Peace Research*, vol. 47 n° 4: pp. 491–99.

7 Appendix: Theory

7.1 The continuation values in the event of peace are such that $V_1^P \geq V_2^P$

In the body of the paper, we state that, in the event of peace, the continuation value of group 1 is larger than the continuation value of group 2. This condition is quite intuitive and holds if we consider, for example, threshold strategies following Chassang and Padró i Miquel (2009). We focus on the most efficient subgame perfect equilibrium that is characterized by threshold strategies such that groups play peace if $(1-\mu(1-d))\,r_{2t} - \mu(1-d)\,r_{1t} > \widetilde{\Omega}$, and play conflict otherwise. Let H denote the cumulative distribution of $\Omega_t = (1-\mu(1-d))\,r_{2t} - \mu(1-d)\,r_{1t}$.

As in the aforementioned paper, we focus on the equilibrium with the lowest possible threshold. The optimal threshold $\widetilde{\Omega}$ is the lowest value of Ω_t that satisfies condition (6) with equality. The corresponding continuation value in the event of peace, $\widetilde{V}_i^P$, is the solution to the following equation:

$$\begin{aligned}\widetilde{V}_i^P &= H\left(\widetilde{\Omega}\right)\frac{1}{2}\left[E\left(r_{it}+r_{jt}|\,\Omega_t<\widetilde{\Omega}\right)(1-d)+\delta V^V\right] \\ &\quad + \left(1-H\left(\widetilde{\Omega}\right)\right)\left[E\left(r_{it}|\,\Omega_t>\widetilde{\Omega}\right)+\delta\widetilde{V}_i^P\right],\end{aligned} \tag{25}$$

for $i,j \in \{1,2\}, i \neq j$. Using our assumption that $r_{2t} \leq r_{1t}$ for all t, we deduce that $\widetilde{V}_1^P \geq \widetilde{V}_2^P$.

7.2 The subgame perfect equilibrium with threshold strategies exists

We must show that a threshold $\widetilde{\Omega} \in (0,+\infty)$ exists. Substituting (4) into (25), solving for $\widetilde{V}_2^P$, and rearranging, we find

$$\begin{aligned}\widetilde{V}_2^P &= \frac{1}{2}\frac{\overline{r}_1+\overline{r}_2}{1-\delta} - \frac{1}{2}d\frac{H(\Omega_t)\,E\left(r_{1t}+r_{2t}|\,\Omega_t<\widetilde{\Omega}\right)}{1-\delta\left(1-H\left(\widetilde{\Omega}\right)\right)} \\ &\quad -\frac{1}{2}\frac{\left(1-H\left(\widetilde{\Omega}\right)\right)\left(E\left(r_{1t}-r_{2t}|\,\Omega_t>\widetilde{\Omega}\right)\right)}{1-\delta\left(1-H\left(\widetilde{\Omega}\right)\right)},\end{aligned} \tag{26}$$

for $i,j \in \{1,2\}, i \neq j$.

The optimal threshold $\widetilde{\Omega}$ is the lowest value of Ω_t that satisfies condition (6) with equality. Rearranging, we find

$$\widetilde{\Omega} = \delta\left(\mu\frac{\overline{r}_1+\overline{r}_2}{1-\delta} - \widetilde{V}_2^P\right), \tag{27}$$

Condition (26) implies that $\widetilde{V}_2^P \leq \frac{1}{2}\frac{\overline{r}_1+\overline{r}_2}{1-\delta}$. Hence, the right hand side in (27) is always strictly positive. The right hand side is bounded, an upper bound being $\delta\mu\frac{\overline{r}_1+\overline{r}_2}{1-\delta}$. The right hand side is also continuous in $\widetilde{\Omega}$. The left hand side in (27) can take any value in $(-\infty,+\infty)$. Hence condition (27) has a solution, $\widetilde{\Omega}^* > 0$.

7.3 Our results are qualitatively unchanged if the losing group regains its land $n > 1$ periods after its defeat

In this section, we show that, if the losing group regains its land $n > 1$ periods after its defeat, then the results are qualitatively unaffected. If the losing group can return, then the continuation value in the event of a defeat will no longer be zero. Moreover, since land is heterogeneous, the continuation values of the loser and of the winner will not be equal.

Condition (25) must be replaced by the following condition:

$$\begin{aligned}\widetilde{V}_i^P &= H\left(\widetilde{\Omega}\right)\left[\frac{1}{2}\left[E\left(r_{it}+r_{jt}|\,\Omega_t<\widetilde{\Omega}\right)(1-d)+\delta\widetilde{V}_i^V\right]+\frac{1}{2}\delta\widetilde{V}_i^L\right] \\ &\quad+\left(1-H\left(\widetilde{\Omega}\right)\right)\left[E\left(r_{it}|\,\Omega_t>\widetilde{\Omega}\right)+\delta\widetilde{V}_i^P\right],\end{aligned} \tag{28}$$

for $i,j\in\{1,2\}, i\neq j$, where $\widetilde{V}_i^V$ denotes the continuation value of group i when it wins the conflict and $\widetilde{V}_i^L$ the continuation value of group i when it loses the conflict. Given that the losing group regains its land $n>1$ periods after its defeat, we have:

$$\widetilde{V}_i^V=\sum_{\tau=0}^{n-2}\delta^\tau E(r_{i\tau}+r_{j\tau})+\delta^{n-1}\widetilde{V}_i^P=\frac{1-\delta^{n-1}}{1-\delta}(\overline{r}_1+\overline{r}_2)+\delta^{n-1}\widetilde{V}_i^P \tag{29}$$

and,

$$\widetilde{V}_i^L=\delta^{n-1}\widetilde{V}_i^P, \tag{30}$$

Substituting (29) and (30) into (28) and rearranging, we find:

$$\begin{aligned}\widetilde{V}_i^P &= \frac{1}{2}\frac{\overline{r}_i+\overline{r}_j}{1-\delta}-\frac{1}{2}d\frac{H\left(\widetilde{\Omega}\right)E\left(r_{it}+r_{jt}|\,\Omega_t<\widetilde{\Omega}\right)}{1-\delta+\delta(1-\delta^{n-1})H\left(\widetilde{\Omega}\right)} \\ &\quad-\frac{1}{2}\frac{\left(1-H\left(\widetilde{\Omega}\right)\right)\left(E\left(r_{1t}-r_{2t}|\,\Omega_t>\widetilde{\Omega}\right)\right)}{1-\delta+\delta(1-\delta^{n-1})H\left(\widetilde{\Omega}\right)},\end{aligned} \tag{31}$$

Using our assumption that $r_{2t}\leq r_{1t}$ for all t, we deduce that $\widetilde{V}_1^P\geq\widetilde{V}_2^P$.

Using (29), we find that condition (27) must be replaced by:

$$\widetilde{\Omega}=\delta\left(\mu\frac{1-\delta^{n-1}}{1-\delta}(\overline{r}_1+\overline{r}_2)-(1-\delta^{n-1}\mu)\widetilde{V}_2^P\right), \tag{32}$$

Since $\mu\leq 1$, we use the same reasoning as in the case in which the loser can never return to show that equation (32) has a solution, $\widetilde{\Omega}^*>0$, and thus that an equilibrium with threshold strategies exists.

7.4 Effect of a change in average soil fertility

Consider the effect of an increase in the average soil fertility. The marginal increase in the average soil fertility on $\partial\Psi_t/\partial c_t$ is given by

$$\frac{\partial^2\Psi_t}{\partial c_t\partial\overline{s}}=\mu\,(1-d)\,\frac{\partial^2 r_{1t}}{\partial c_t\partial\overline{s}}-(1-\mu\,(1-d))\,\frac{\partial^2 r_{2t}}{\partial c_t\partial\overline{s}}. \tag{33}$$

The sign of the effect depends on the weighted difference between the cross derivatives of the rents with respect to the price of fertilizer and the average soil fertility. Using (17) and differentiating (13) with respect to the average soil fertility, we find:

$$\frac{\partial^2 r_{1t}}{\partial c_t \partial \overline{s}} = -\frac{\partial f_{1t}}{\partial s_1} > 0 \text{ and } \frac{\partial^2 r_{2t}}{\partial c_t \partial \overline{s}} = -\frac{\partial f_{2t}}{\partial s_2} > 0. \tag{34}$$

Since the two cross-derivatives have the same sign, the resulting cross-effect, $\frac{\partial_t^2 \Psi}{\partial c_t \partial \overline{s}}$, has an a priori ambiguous sign. The following Proposition shows that the effect can indeed be either positive or negative:

Proposition [Soil Fertility and Conflict]: *The cross-effect of an increase in fertilizer price and average soil fertility on the likelihood of conflict is positive if and only if the effect of an increase in soil fertility on fertilizer use is sufficiently larger for more fertile soil compared to less fertile soil. Formally,*

$$\frac{\partial^2 \Psi_t}{\partial c_t \partial \overline{s}} > 0 \Leftrightarrow \left| \frac{\partial f_{1t}/\partial s_1}{\partial f_{2t}/\partial s_2} \right| > \frac{1-\mu\left(1-d\right)}{\mu\left(1-d\right)}. \tag{35}$$

An increase in the average soil fertility reduces both the decrease in the opportunity cost of fighting and the decrease in the rapacity gain. The resulting effect is ambiguous and notably depends on the specific form of the production function (specifically, it depends on the sign of the third derivatives of g). A necessary condition for the cross-effect to be positive is that fertilizer use is more sensitive to soil fertility when soil fertility increases. Indeed, since $\frac{1-\mu(1-d)}{\mu(1-d)} > 1$, the cross-effect is positive only if $\left| \frac{\partial f_{1t}/\partial s_1}{\partial f_{2t}/\partial s_2} \right| > 1$.

Moreover, the following Corollary states that the sign of the effect is sensitive to the measurement of soil fertility.

Corollary [Sensitivity to Soil Fertility Measurement]: *Assume that the production function is such that $g\left(s_i, f_{it}\right) \equiv v\left(f_{it} + m\left(s_i\right)\right)$ with $v' > 0$, $v'' < 0$, $m' > 0$. Hence, if $m'' \leq 0$, then $\frac{\partial^2 \Psi_t}{\partial c_t \partial \overline{s}} < 0$ and, if $m'' > 0$,*

$$\frac{\partial^2 \Psi_t}{\partial c_t \partial \overline{s}} > 0 \Leftrightarrow \frac{m'\left(s_1\right)}{m'\left(s_2\right)} > \frac{1-\mu\left(1-d\right)}{\mu\left(1-d\right)}. \tag{36}$$

Proof of Corollary: *First order condition (11) becomes $f_{it} = h\left(\frac{c_t}{\pi_t}\right) - m\left(s_i\right)$ where $h \equiv v'^{-1}$. Hence, if $m'' \leq 0$, we have $\left| \frac{\partial f_{1t}/\partial s_1}{\partial f_{2t}/\partial s_2} \right| = \frac{m'(s_1)}{m'(s_2)} < 1 < \frac{1-\mu(1-d)}{\mu(1-d)}$.*

This Corollary states that the sign of the cross-effect of an increase in fertilizer price and average soil fertility is sensitive to a monotonic increasing transformation of soil fertility. In other words, it is sensitive to the measurement of soil fertility.

Fertile ground for conflict

Online Appendix

December 22, 2017

Nicolas Berman[1] Mathieu Couttenier[2] Raphael Soubeyran[3]

[1]Aix-Marseille University (Aix-Marseille School of Economics), CNRS, EHESS, Graduate Institute Geneva and CEPR. E-mail: nicolas.berman@univ-amu.fr.

[2]University of Geneva. E-mail: mathieu.couttenier@unige.ch

[3]LAMETA, Univ. Montpellier, CNRS, INRA, Montpellier SupAgro, Univ. Paul Valéry. E-mail: raphael.soubeyran@inra.fr.

Contents

A Additional data description **ii**
A.1 Structure of the dataset . . . ii
A.2 Conflict data . . . ii
A.3 Soil fertility . . . iii
A.4 Fertilizer prices . . . vi
A.5 Other data . . . x
A.6 Extended sample statistics . . . xii
A.7 Maps . . . xiv
A.8 Soil fertility: kernel density . . . xx
A.9 Effect of a standard deviation increase in fertilizer prices on conflict . . . xxii

B How fertilizer price variations affect fertilizer use and rents? **xxv**

C Soil fertility heterogeneity and conflict: theory and cross-sectional evidence **xxvii**

D Budget constraints and non monotonic effects **xxix**

E Local fertilizer prices: identifying assumptions **xxxi**
E.1 Do international market prices of fertilizer affect fertilizer consumption? . . . xxxi
E.2 Are international market prices of fertilizer transmitted to local prices? . . . xxxii
E.3 Do farmers use the appropriate mix of nutrients? . . . xxxiii

F Additional results **xxxiv**
F.1 Robustness: Measurement . . . xxxiv
F.1.1 Conflict definition . . . xxxiv
F.1.2 Soil fertility definition . . . xxxv
F.1.3 Fertilizer prices computation . . . xxxvii
F.2 Robustness: Estimation . . . xxxviii
F.2.1 Standard errors . . . xxxviii
F.2.2 Alternative estimators . . . xxxix
F.3 Sensitivity to specific countries, years or crops . . . xl
F.4 Robustness: Omitted variables . . . xlv
F.5 Robustness: Land inequality across and within ethnic groups . . . xlvi

G Additional references **xlvii**

A Additional data description

A.1 Structure of the dataset

The structure of our main dataset is a full grid of Sub-Saharan African countries divided in sub-national units of 0.5×0.5 degrees latitude and longitude (i.e. roughly 55×55 kilometers at the equator). This is the exact same level of aggregation as the one used in the PRIO-GRID version 2 (Tollefsen *et al.*, 2012), which allows us to easily include cell-specific information from this dataset. We use this level of aggregation rather than administrative boundaries to ensure that our unit of observation is not endogenous to conflict events. To each cell we assign a country based on the end-of-the-period boundaries. The country which represents the largest share of the cell's area is assigned to this cell.

A.2 Conflict data

We use the Armed Conflict Location and Event dataset ACLED (Raleigh and Dowd, 2014) which contains information on the geo-location of conflict events in all African countries from 1997 to 2013. We use ACLED version 4 as it contains cells identifiers which allows to directly match the data with the dataset PRIO-GRID v.2.[4] In the ACLED dataset, the unit of observation is the "event". We have information about the date (precise day most of the time), longitude and latitude of conflict events within each country. ACLED dataset contains information on the accuracy of the geo-referencing of the events: at least the municipality level in more than 95% of the cases, and is even finer (village) for more than 80% of the observations. We keep only events which are geolocalized with the finer precision level for our analysis. We also drop duplicated events, i.e. events for which all the ACLED variables' content (precise date, location, actors, description, etc.) is the same for several observations – in this case we keep only one observation for the event.

These events are obtained from various sources, including press accounts from regional and local news, humanitarian agencies or research publications. ACLED records all political violence, including violence against civilians, rioting and protesting within and outside a civil conflict, without specifying a battle-related deaths threshold. A unique feature of the ACLED dataset is that it contains some information about the characteristics of the events that allows us to identify local conflict that are likely to be land-related. We make use of this information in the following way. We consider an event as being land-related and to occur between local actors if (i) it involves communal militias on both sides of the conflict (i.e. local armed group, fighting for a local objective), and/or (ii) its description includes at least one of the following words / combinations

[4] http://www.acleddata.com/wp-content/uploads/2014/acled_with_prio.zip

of words: "land dispute", "dispute over land", "control of land", "over land", "clash over land", "land grab", "farm land", "land invaders", "land invasion", "land redistribution", "land battle", "over cattle and land", "invade land", "over disputed land", "over a piece of land".

As alternative measure of conflict, we use the UCDP-GED dataset that records only events pertaining to conflicts reaching at least 25 battle-related deaths per year, but does not include information about the characteristics of the events.[5]

A.3 Soil fertility

Baseline measure. To construct our baseline measure of soil fertility we use information on nutrient availability from the Harmonized World Soil Database,[6] which is the outcome of a joint work of the Food and Agriculture Organization (FAO) and of the International Institute for Applied System Analysis (IIASA). This data is constructed from models that use location specific soil attributes (soil texture, soil organic carbon, soil pH and total exchangeable bases) to compute an index of nutrient availability. Ratings are associated to soil attributes. The index is the average of the rating of the attribute with the smallest rating and the average rating of the three other attributes. The raw data is in the form of raster file containing a rating of nutrient availability from 1 to 7 at the 0.08$\times$0.08 degrees latitude and longitude level (approx. 9km $\times$ 9km at the equator). The categories are the following: 1/ No or slight constraints; 2/ Moderate constraints; 3/ severe constraints; 4/ very severe constraints; 5/ Mainly non-soil; 6/ Permafrost; 7/ Water. We consider only categories 1 to 5. We aggregate the data by 0.5 $\times$ 0.5 degree cell. For each cell we compute the mode, average, and standard deviation of nutrient availability. The mode and the standard deviation are our baseline measures of local soil fertility level and dispersion, respectively.

Alternative measure: EU commission. The European Commission (Jones *et al.*, 2013) provides geolocalized information on soils in Africa; which we use as an alternative fertility measure for robustness. For each of the 31 different types of soils in Africa recorded by the European Commission, they gather information on the strengths, weaknesses, and opportunities for each type of soil. Based on this information, we rank soils in three categories: naturally fertile, medium level of fertility and infertile. The following 8 soils over the 31 are defined as naturally fertile: Andosols, Cambisols, Chernozems, Fluvisols, Gleysols, Kastanozems, Luvisols, Phaeozems. Table A.1 displays some information on the main characteristics of each of the 31 types of soils as well as our final coding of each type of soil. Once soils have been classified into our three categories,

[5]See for instance Berman and Couttenier (2015) for a discussion of the two datasets.

[6]http://webarchive.iiasa.ac.at/Research/LUC/External-World-soil-database/HTML/SoilQuality.html?sb=10

we compute two statistics that we use as measures of fertility level and dispersion: the share of fertile soil in the cell and the Herfindahl index of the three types of soils in the cell. The latter is computed as the sum of the square of the share of each soil in the cell. It is an index of soil concentration (i.e. an inverse index of dispersion).

Table A.1: Classification of soils using data from the European Commission Joint Research Center

Soil type (1)	Strengths (2)	Weaknesses (3)	Opportunities (4)	"Naturally fertile" (5)
Acrisols	Can support some agriculture	With dry seasons, Acrisols may become very hard	Productive if fertilizers applied	0
Alisols	Fertile once acidity is reduced	Very acid	Shallow-rooting crops	0
Andosols	High level of nutrients. Easy to cultivate	High level of phosphate fertilizer that has to be fixed by iron	Suitable for a wide variety of crops	1
Anthrosols	Physical properties favorable for cultivation	Has to be maintained	Cultivation opportunities in unfavorable conditions - support local produce	0
Arenosols	Easy to work	Only a small amount of organic matter, nutrients and water	Land use limited to extensive grazing	0
Calcisols	Cultivation easy with irrigation	High pH and need to fertilizers	Irrigated winter wheat, melons and cotton	0
Cambisols	Among the best agricultural soil in Africa	Strongly weathered, it contains limited amounts of nutrients.	Water-holding can be high	1
Chernozems	Very fertile soils	Timely cultivation and careful irrigation to preserve the fertility	High agricultural potential	1
Cryosols	Carbon sink	Limited extend in Africa	High ecological and environmental value	0
Durisols	Excellent for vine-growing	Fairly restricted once a pan has developed	Need to breaking up the pan for successful	0
Ferralsols	Sustain limited cultivation with fertilizers	Require specific soil management	Liming to correct the actual amount of exchangeable aluminium	0
Fluvisols	Fertile with a regular supply of nutrients	Flood control or drainage may be needed	High agricultural potential	1
Gleysols	Very suitable for wetland rice	Most crops need to be drained	Arable cropping and dairy farming if drained	1
Gypsisols	Support extensive grazing	Do not hold much water	Limited	0
Histosols	Contains more than 20% organic matter	Density and bearing capacity are low	High ecological and environmental value	0
Kastanozems	High fertile soil	Nutrient imbalance, irrigation is necessary if hot summer season	Winter wheat during rainy season	1
Leptosols	Solid foundation for construction	Unsuitable for growing crops	Grazing for cattle	0
Lixisols	Fairly productive. Low inherent fertility	Do not hold much organic matter	Irrigation needed to grow crops during dry season	0
Luvisols	Productive soil	Poor in organic matter	Productive soils if managed appropriately	1
Nitisols	Suitable for a wide range of crops	For annual cropping, fertilizer application is needed	Intensive liming to overcome aluminium toxicity	0
Phaeozems	Highly productive soil	Water-holding capacity is limited	Good agricultural potential	1
Planosols	Grassland and wetland rice	Use is limited because of waterlogging	Grassland with irrigation in the dry season is good land	0
Plinthosols	Building stone	Issue with water holding	Limited for agriculture	0
Podzols	Sustain a light forest cover	Low nutrients level, available moisture and pH. Unattractive for farming	Limited. Extensive grazing	0
Regosols	Well supplied with nutrients	Water-holding capacity is limited	Shrub and tree cultivation	0
Solonchaks	Support natural habitats. Grazing	Salty	Cultivation only after salt have been flushed	0
Solonetz	Support natural habitats. Grazing	High pH	Improved if pH lowered	0
Stagnosols	Fertile owing to their moderate degree of leaching	Limited because of oxygen deficiency	Be improved by deep ploughing	0
Technosols	Solid foundation for construction	Sealed by artificial surface	Limited	0
Umbrisols	Suitable for woodland	Acidity	Require heavy investment to make them productive	0
Vertisols	Productive under right measures	Heavy to work when wet	Raised beds made out of the surface layer	0

Columns 1 to 4 display data from the European Commission Joint Research Center (Jones *et al.*, 2013). Column (5) provides our coding of each type of soil based on the description provided in columns 1 to 4: a dummy which is coded 1 if the soil is considered to be "Naturally fertile" and 0 otherwise.

A.4 Fertilizer prices

As explained in the manuscript, to identify local variations in fertilizer prices, we combine data on (a) crop specialization, (b) crop-specific nutrients uptakes (which mix of nutrients should be included in the fertilizers used for each crop), and (c) annual prices of each nutrient.

Agricultural specialization. We compute our baseline measure of agricultural specialization from the FAO's Global Agro-Ecological Zones (GAEZ).[7] It contains the suitability of each location for 45 different cultivating crops. This dataset is constructed from models that use location characteristics such as climate information (rainfall and temperature for instance) and soil characteristics. The climate information is based on the average information over the period 1961-1990. This information is combined with crops' characteristics (in terms of growing requirements) to generate a global GIS raster of the suitability of a grid cell for each crop. Suitability is then defined as the percentage of the maximum yield that can be attained in each grid cell. As several suitabilities are computed based on different scenarii, we consider the one where crop production has been considered with intermediate input level conditions. Based on these levels of suitability, we select the most suitable crops, or the five most suitable crops that we next use in our fertilizer price index calculation.[8] Table A.2 shows the share of cells for which each crop is the most suitable.

As an alternative measure of agricultural specialization, we use the M3-CROPS dataset from Monfreda *et al.* (2008), which contains information on the harvested area in hectares for 137 different crops for grid-cells of 5 arc minutes×5 arc minutes resolution for the year 2000.

Crop-specific nutrients mix. Data of nutrients removal by crop from the International Plant Nutrient Institute is shown in Table A.2.[9] For each crop, it provides the percentage of nitrogen (N), phosphorus (P) and potassium (K) that is removed from the ground at the harvest (for instance, maize removes 48% of nitrogen, 16% of phosphorus and 36% of potassium). We also use alternative nutrients removal by crop data from the U.S. Department of Agriculture (USDA).[10] Table A.3 contains this alternative data.

International market prices. Data on the real international market prices of each nutrient come from the World Bank Commodities Dataset.[11] Figure A.1 displays the time variations

[7] http://gaez.fao.org/Main.html
[8] See Nunn and Qian, 2011, for more details and a discussion on the FAO-GAEZ data.
[9] http://www.ipni.net/
[10] http://plants.usda.gov/npk/NutrientSources
[11] http://databank.worldbank.org/data/databases/commodity-price-data

of each series.

Table A.2: IPNI Nutrient Removal Share by Crop (%) and crop share in cells

Product	N	P2O5	K20	% cells most suitable
Alfalfa	46	10	44	2.3
Barley	58	23	19	0.2
Buckwheat	64	19	17	0.7
Cabbage	46	11	43	0.1
Chick Pea	56	28	16	0.0
Citrus	41	8	51	2.2
Cotton	49	22	29	1.9
Cowpea	64	17	19	4.6
Dry Pea	64	17	19	0.0
Dryland Rice	56	29	15	16.7
Flax	65	19	16	0.02
Green Gram	64	17	19	0.7
Groundnut	72	11	17	0.0
Maize	48	16	36	3.4
Oats	62	23	15	0.03
Pearl Millet	64	18	18	18.9
Pigeon Pea	56	28	16	7.0
Rape	57	29	14	1.3
Reed Canary Grass	43	19	38	17.7
Wetland Rice	56	29	15	0.4
Rye	65	21	14	0.1
Sorghum	50	30	20	15.86
Soy Bean	63	14	23	0.7
Sugar Cane	29	20	51	2.7
Sugar Beet	28.5	16.5	55	0.0
Sun Flower	59	21	20	1.0
Sweet Potato	30	13	57	0.4
Switch Grass	24	13	64	0.0
Tobacco	35	9	56	0.5
Tomato	27.5	10.5	62	0.2
Wheat	58.5	26.5	15	0.5
White potato	27	14	59	0.1

Notes: This table shows nutrients (N, P, K) exports share for each crop, computed thanks to the quantity of nutrient removed from the field at crop harvest, in kg/ha. For instance, the second raw second column value is $\alpha^N_{Alfalfa} = q^N_{Alfalfa}/(q^N_{Alfalfa} + q^P_{Alfalfa} + q^K_{Alfalfa})$, where $q^N_{Alfalfa}$, $q^P_{Alfalfa}$ and $q^K_{Alfalfa}$ are the quantities of N, $P205$ and $K20$ removed from the field when Alfalfa is harvested, in kg/ha. The data comes from the International Plant Nutrition Institute (IPNI). These represent what IPNI scientists believe to be the best estimates of typical values to date of nutrient uptake for different crops grown in different countries of the world. See http://www.ipni.net/article/IPNI-3296. % cells most suitable is the share of cells in our final dataset for which this crop is the most suitable one (from GAEZ).

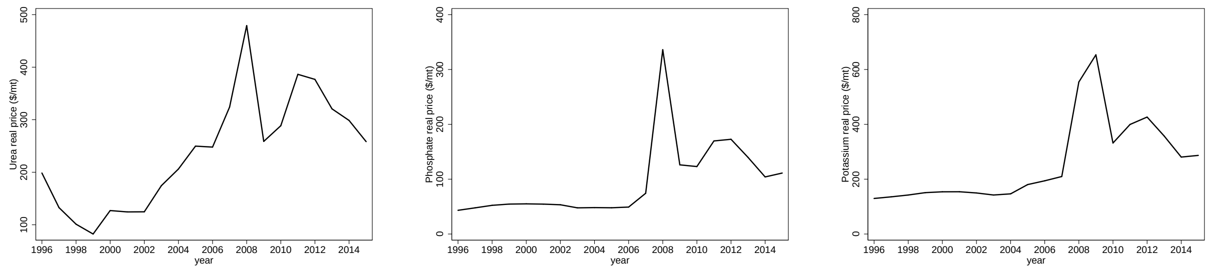
Table A.3: USDA Nutrient Removal Share by Crop (%)

Product	N	P2O5	K20
Alfalfa	40	9.5	50.5
Banana	24	6.5	69.5
Barley	62.5	21.5	16
Buckwheat	57.5	24.5	18
Cabbage	45.5	11	43.5
Carrot	29	14.5	56.5
Chick Pea	54	18.5	27.5
Citrus	47.5	8.5	44
Coconut	43.5	21	35.5
Cotton	55	24.5	20.5
Cowpea	54	18.5	27.5
Dry Pea	68.5	15.5	16
Dryland Rice	53.5	25.5	21
Flax	61.5	21.5	17
Green Gram	62	14	24
Groundnut	74	14	12
Maize	59	26	15
Oats	59.5	24.5	16
Onion	43	18	39
Pearl Millet	62.5	24	13.5
Pigeon Pea	58	15.5	26.5
Rape	60	22	18
Reed Canary Grass	32	10	58
Rice	53.5	25.5	21
Rye	59	24	17
Sorghum	60.5	24.5	15
Soy Bean	57.5	12.5	30
Sugar Cane	31.5	10.5	58
Sugar Beet	36	15	49
Sun Flower	57.5	26	16.5
Sweet Potato	36	14	50
Switch Grass	29.5	9	61.5
Tobacco	41.5	12	46.5
Tomato	29	12	59
Wheat	61	24	15
White potato	35	12.5	52.5
Yams	24.5	13	62.5

Notes: This table shows nutrients (N, P, K) exports share for each crop, computed thanks to the quantity of nutrient removed from the field at crop harvest, in kg/ha. For instance, the second raw second column value is $\alpha^N_{Alfalfa} = q^N_{Alfalfa}/(q^N_{Alfalfa} + q^P_{Alfalfa} + q^K_{Alfalfa})$, where $q^N_{Alfalfa}$, $q^P_{Alfalfa}$ and $q^K_{Alfalfa}$ are the quantities of N, $P205$ and $K20$ removed from the field when Alfalfa is harvested, in kg/ha. The data comes from the U.S. Department of Agriculture (USDA). These are the estimates of typical values of nutrient removal for crops grown in different countries of the World, according to the USDA. See http://plants.usda.gov/npk/NutrientSources

Figure A.1: N-P-K prices

(a) Urea (b) Phosphate (c) Potassium



Source: World Bank.

A.5 Other data

Cell-specific variables (time-invariant). In Table 4 in the manuscript, we include interaction terms using cell level data compiled in PRIO-GRID version 2 (Tollefsen *et al.*, 2012).[12] Specifically, we use the share of cropland (*% agriculture*) in 1990, the share of forest area (*% forest*) in 1990, and the share of water area (*% water*) in 1990 from the ISAM-HYDE land-cover and land-use data (Meiyappan *et al.*, 2012). We use the share of harvested area for the cell's main crop (*% harvested*) in 2000 from Portmann *et al.* (2010) and average nighttime lights gathered in 1992-93 from the DMSP-OLS Nighttime Lights Time Series version 4.[13] We also use the *Malaria Ecology Index* from Kiszewski *et al.* (2004) and developed by Gordon McCord as a measure of malaria incidence at cell level.[14] We use population density data from the Global Land Survey 1990 gathered by all eight Landsat sensors.[15]

Cell-specific variables (time-varying). In Table 3, we use cell level data on drought (SPEI) coded as the average value of the Standardized Precipitation Evaporation Index (developed by Vicente-Serrano *et al.*, 2010) in the cell-year.[16] Data on agricultural commodities demand come from Berman and Couttenier (2015). The variable, labeled *ln agr. demand* represents the (log of the) sum of the world import value of commodities (net of the import value of the country) weighted by the share of the agricultural commodity in the cell. Data on mineral produced in the cell come from Berman *et al.* (2017). We use the interaction variable *ln price mineral* $\times$ *mines* > 0. *ln price mineral* is the (log of the) is the average international market price of the minerals

[12] http://grid.prio.org/
[13] https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html
[14] https://sites.google.com/site/gordoncmccord//datasets
[15] http://cidportal.jrc.ec.europa.eu/ftp/jrc-opendata/GHSL/
[16] http://spei.csic.es/

produced in the cell with weights equal to the share of each mineral in total production value over the period. $mines > 0$ is a dummy taking the value 1 if at least 1 mine is recorded in the cell at any point over the period.

Ethnicity. In Tables 4, 6 and 7 in the body of the paper, we use measures of ethnic fractionalization and polarization. We use two different data sources. First, we make use of the digitalized version of the Ethnographic Atlas from Murdock (1959) that divides Africa into 843 regions and is provided in Reynal-Querol (2014). Second, we use the Geo-referencing of Ethnic Groups (GREG) that is drawn from the Soviet Atlas Narodov Mira (Weidmann and Cederman, 2010). For each cell and datasets, we compute the share of each ethnic groups in order to compute a cell-specific measure of ethnic fractionalization and polarization.

Country specific variables. The data on country-level fertilizer use intensity, used in Table 5 in the body of the paper, is from Morris *et al.* (2007) and covers a subset of 30 countries. We use two alternative measures of institutions i) the indicator of quality of government from International Country Risk Guide (ICRG, 2013) and ii) the democracy score of Polity IV (2013). We use pre-sample scores to mitigate endogeneity concerns (year 1996). We also make use of information on legal security of indigenous and community lands in Africa from *LandMark: The Global Platform of Indigenous and Community Lands*.[17] For both indigenous peoples and communities, through ten different indicators, they compute an index that estimates to what extend the security of land tenure is established in national laws.

[17] http://www.landmarkmap.org/data/

A.6 Extended sample statistics

Table A.4: Descriptive statistics (extended)

	Obs.	Mean	S.D.	1^{st} Quartile	Median	3^{rd} Quartile
Pr(conflict>0)	111605	0.07	0.26	0.00	0.00	0.00
Pr(conflict>0) (land-related)	111605	0.01	0.08	0.00	0.00	0.00
# conflicts	111605	0.09	0.38	0.00	0.00	0.00
# conflicts (if >0)	7945	1.25	0.75	0.69	1.10	1.61
Pr(conflict) (UCDP-GED)	91910	0.03	0.18	0.00	0.00	0.00
Soil fertility (mean)	111588	0.38	0.29	0.16	0.42	0.50
Soil fertility (mode)	111588	2.11	0.79	1.42	2.00	2.81
Soil fertility (variance)	111588	0.23	0.30	0.03	0.17	0.25
Percentage rich soil (EU commission)	111605	11.14	21.01	0.00	0.00	12.70
Herfindalh index soils	111061	0.81	0.21	0.60	0.89	1.00
Nitrogen density (mean)	111605	918.58	352.67	686.49	908.45	1117.81
Nitrogen density (standard dev.)	111605	168.18	179.26	32.74	124.99	242.64
% irrigated	111605	0.33	1.84	0.00	0.00	0.04
Fertilizer price ($/ton)	111605	209.97	106.89	116.27	180.96	286.68
Δ ln fertilizer price	105040	0.06	0.24	-0.10	0.01	0.22
% agriculture	111605	22.27	26.33	1.53	11.33	33.94
% urban	111605	0.08	0.93	0.00	0.00	0.00
% harvested	97053	3.01	5.10	0.29	1.07	3.42
% forest	111605	38.24	32.00	7.56	31.76	66.62
Rainfall	111605	977.82	542.67	541.50	970.82	1362.28
SPEI	88278	-0.11	0.33	-0.32	-0.08	0.12
Ethnic Fractionalization (GREG)	110874	0.17	0.21	0.00	0.00	0.36
Ethnic Polarization (GREG)	110874	0.08	0.09	0.00	0.00	0.18
Ethnic Fractionalization (MURDOCK)	111605	0.20	0.22	0.00	0.09	0.42
Ethnic Polarization (MURDOCK)	111605	0.09	0.10	0.00	0.05	0.20
Malaria index	111605	12.52	9.52	4.26	11.63	19.70
Population in 1990	111605	70648.61	184630.40	1231.60	14811.55	67151.21
Nighttime lights	91910	0.23	1.63	0.00	0.00	0.03

Source: See text for details about data sources and variables' computations.

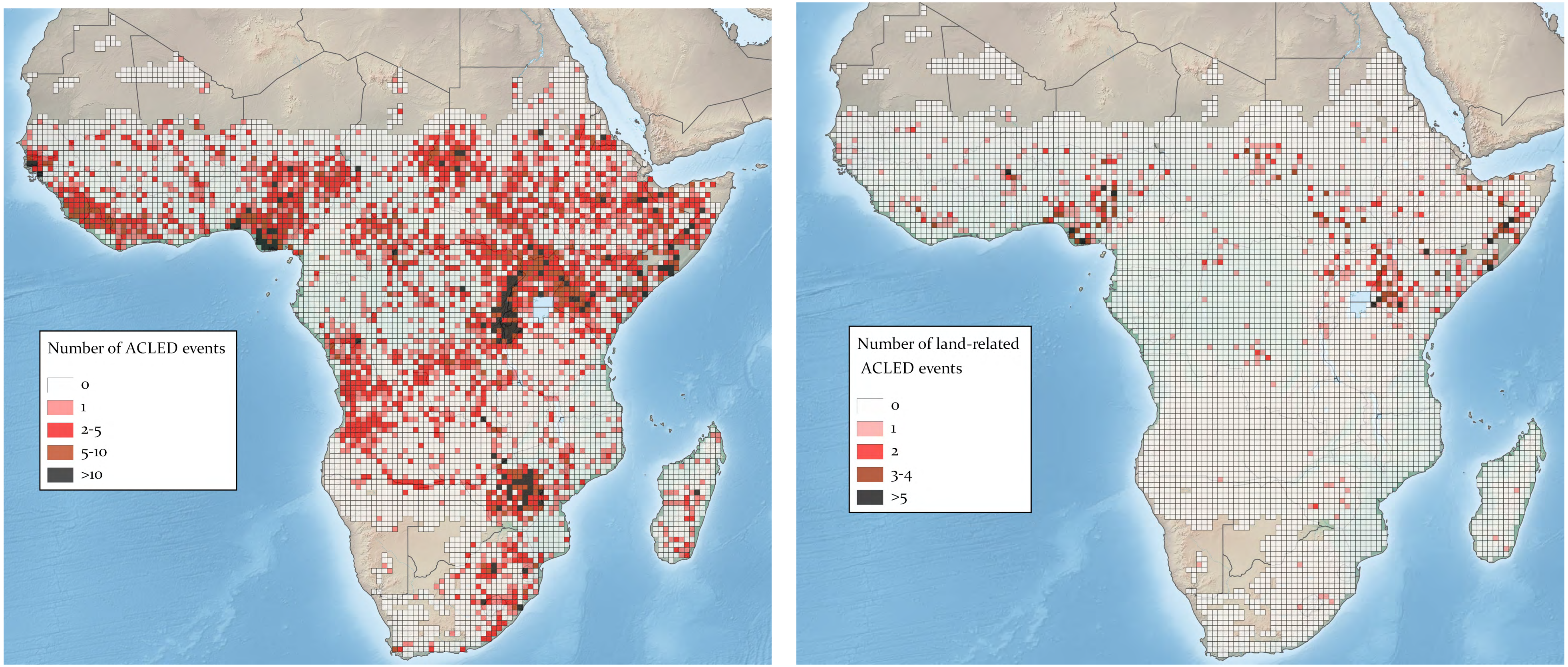
A.7 Maps

Maps A.2 to A.6 show the spatial distribution of conflict events, soil fertility (average and dispersion), fertilizer prices (average and largest yearly increase), population density in 1990, ethnic polarization and the between/within variance of nutrient availability by ethnic groups.

Figure A.2: ACLED data

A- ACLED data

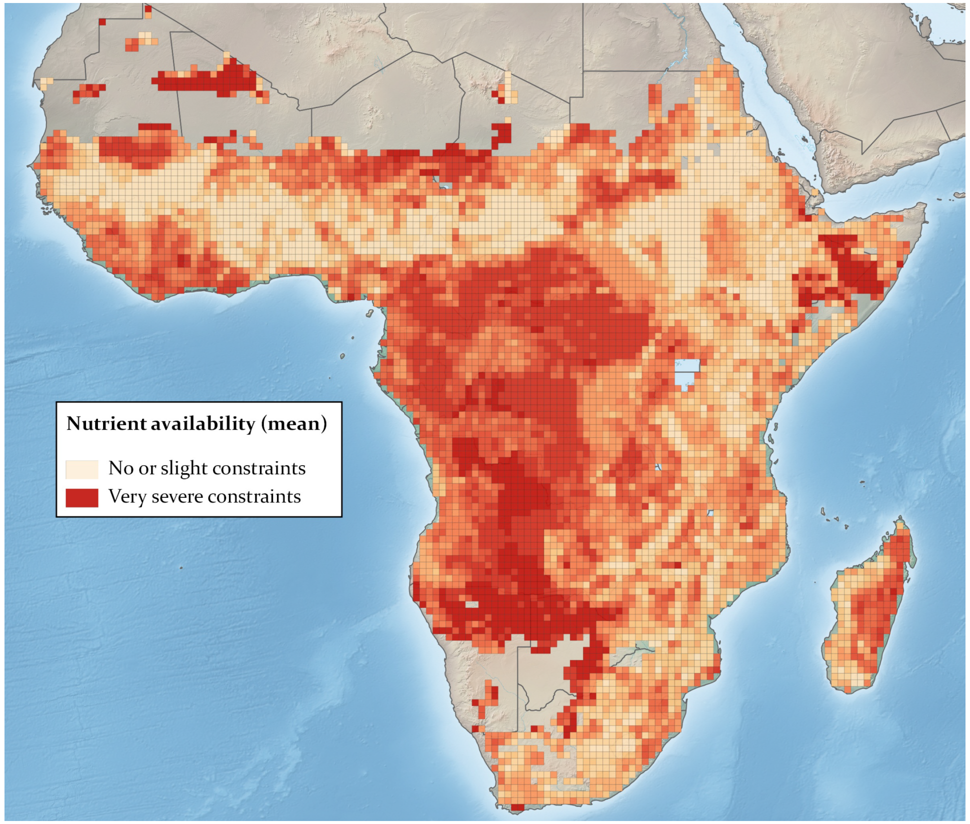
B - Land conflicts



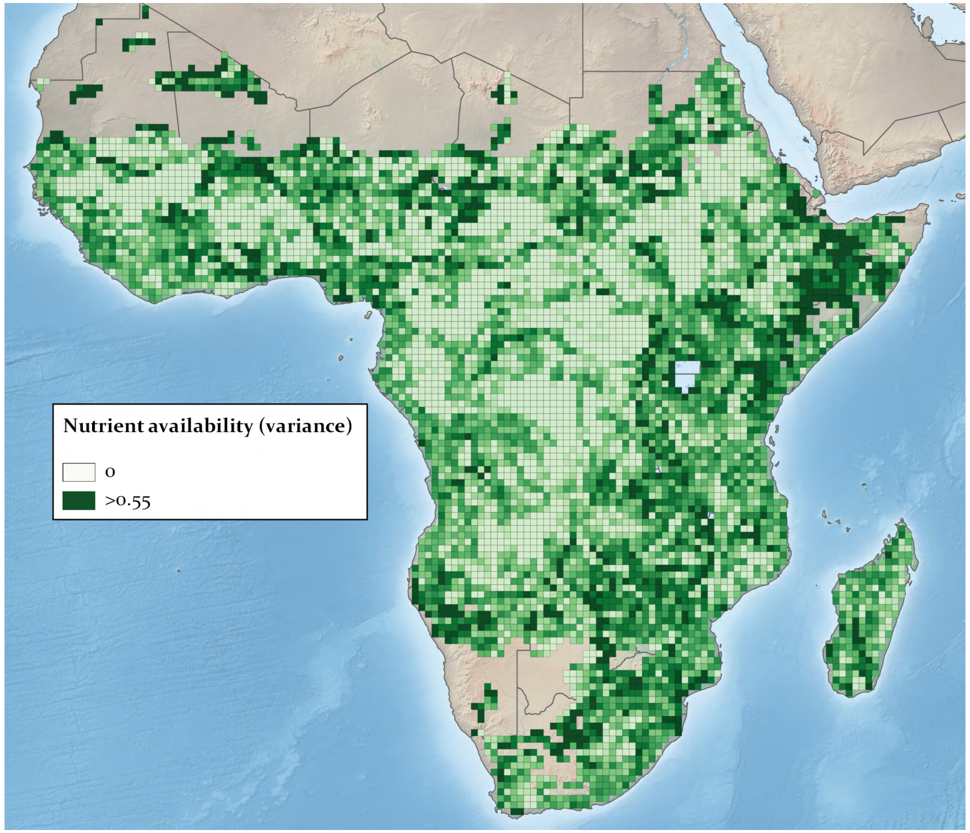
Source: Figure A: All ACLED events. Figure B: Only land-related events defined as events (i) involving communal militias on both sides and/or (ii) whose description include specific keywords related to land including “land dispute”, “dispute over land”, “control of land”, “over land”, “clash over land”, “land grab”, “farm land”, “land invaders”, “land invasion”, “land redistribution”, “land battle”, “over cattle and land”, “invade land”, “over disputed land”, and “over a piece of land”.

Figure A.3: Soil fertility

A- Soil fertility (average)



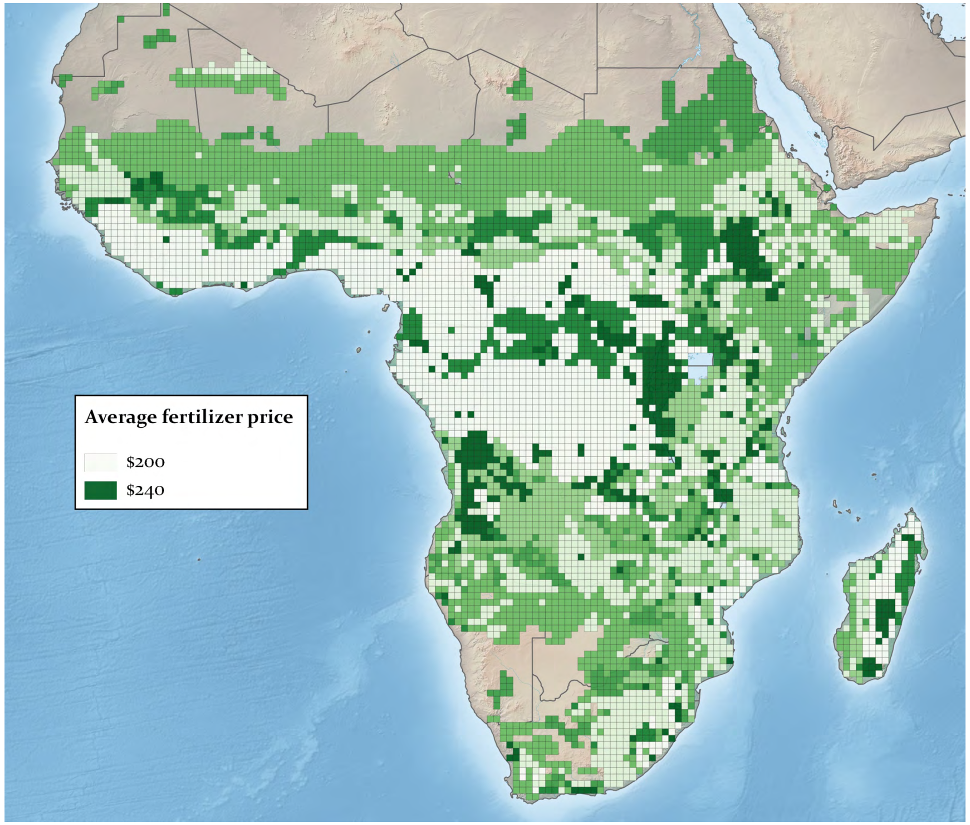
B- Soil fertility (dispersion)



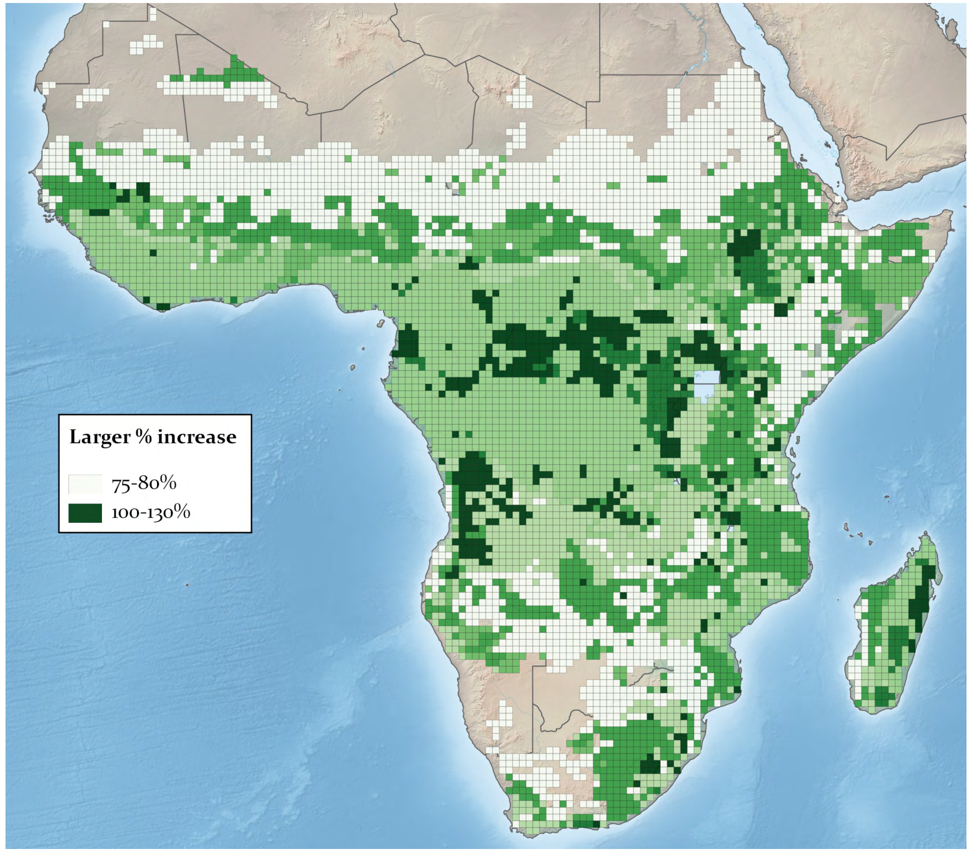
Source: Figure A: Data from the Harmonized World Soil Database. Soils are ordered in five categories ranging from 1 ("no or slight constraints") to 5 ("very severe constraints"). For each cell, we compute the average value of these categories. Figure B: cell-specific standard deviation instead of average value.

Figure A.4: Fertilizer prices

A- Fertilizer price (average)



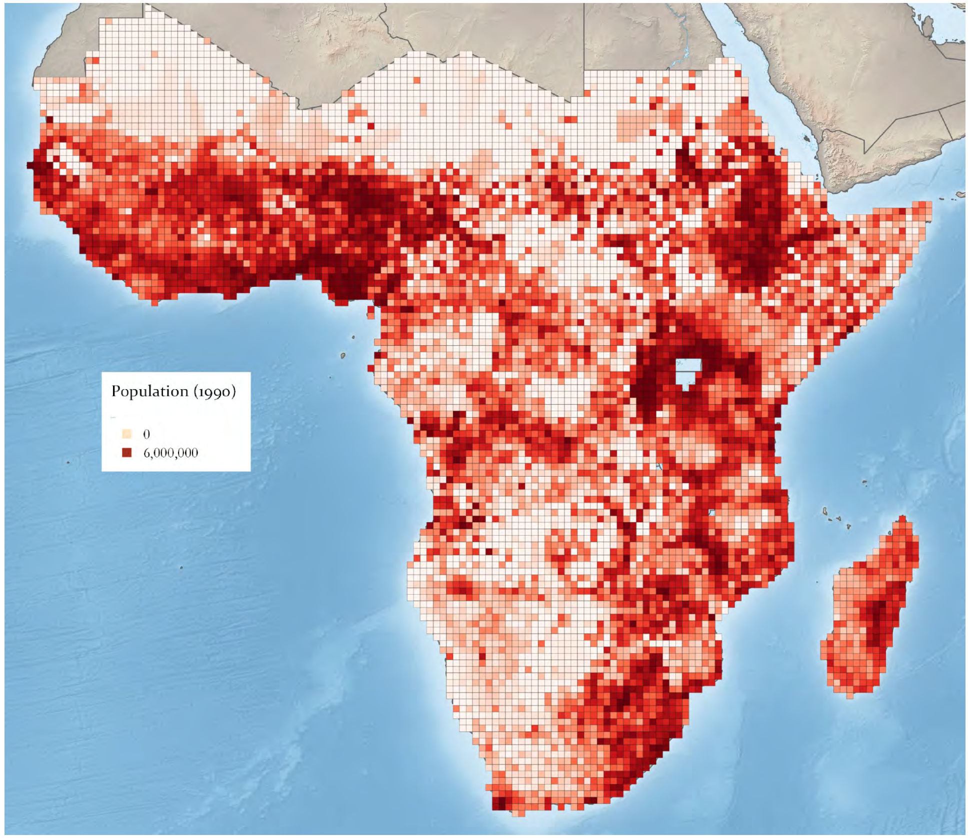
B- Fertility price (largest yearly increase)



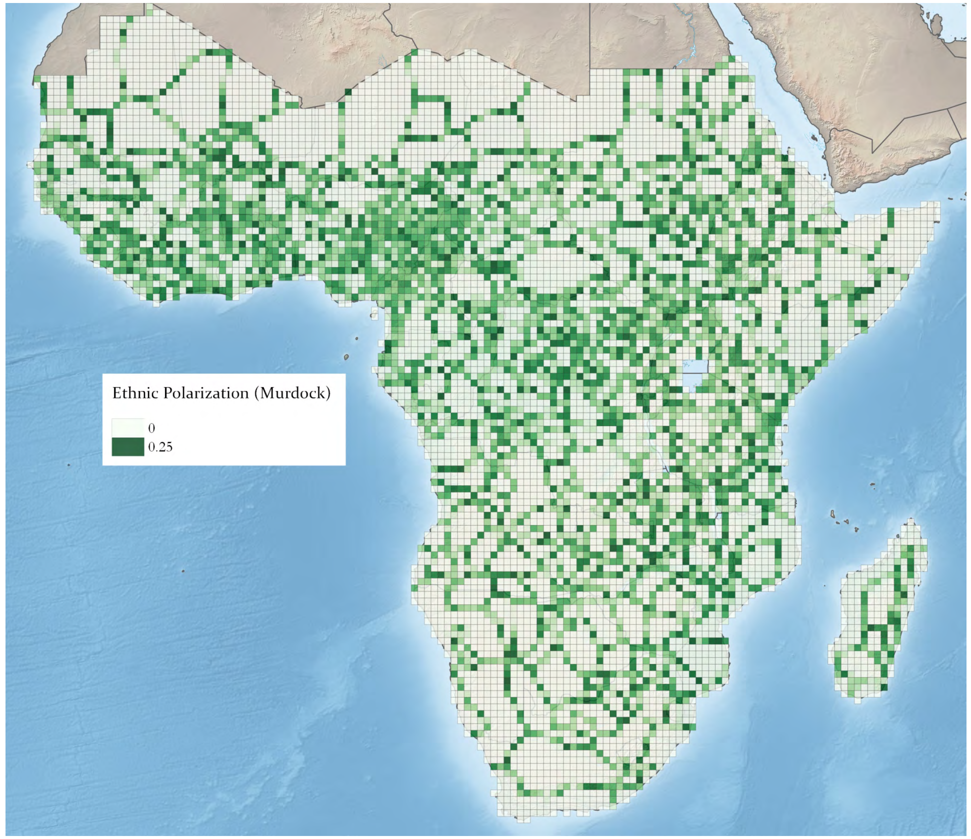
Note: Figure A: Average value of our fertilizer price index over the period 1997-2013, by cell. B: largest increase in the fertilizer price index over the period 1997-2013.

Figure A.5: Population and ethnic polarization

A- Population in 1990



B- Ethnic Polarization (Murdock)

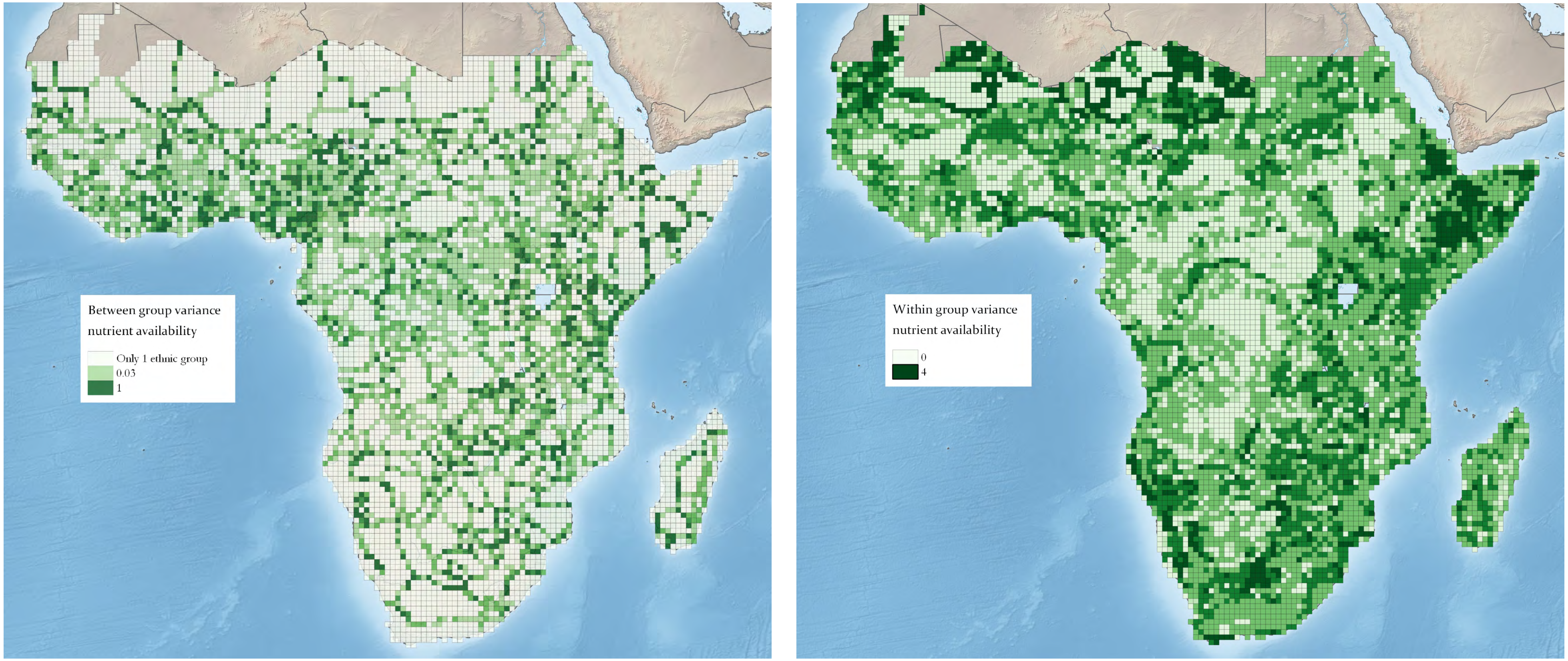


Source: Figure A: Population in 1990, from Landsat Global Land Survey. Figure B: Ethnic polarization index, computed from Murdock (1959). We use the standard formula: $POL = \sum_{i=1}^{N} g_i^2(1 - g_i)$ where g_i is the share of the cell's area occupied by ethnic group i, and N is the total number of groups in the cell.

Figure A.6: Between/within variance of nutrient availability by ethnic group

A- Between group nutrient availability variance

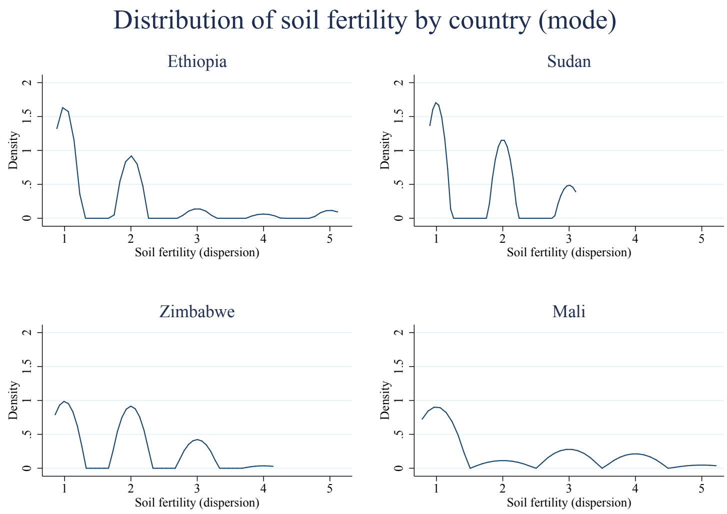
B- Within group nutrient availability variance



Source: Figure A represents the variance of nutrient availability between ethnic groups in the cell, $\mathbb{V}_B(\text{Fertility}) = \frac{1}{A}\sum_{e\in E} A_e(s_e - s)^2$. Figure B shows the variance within ethnic group for each cell, i.e. $\mathbb{V}_W(\text{Fertility}) = \frac{1}{A}\sum_{e\in E}\sum_{j\in T(e)}(s_j - s_e)^2$ (see main text for details). These variances are computed from Murdock (1959) and from the Harmonized World Soil Database.

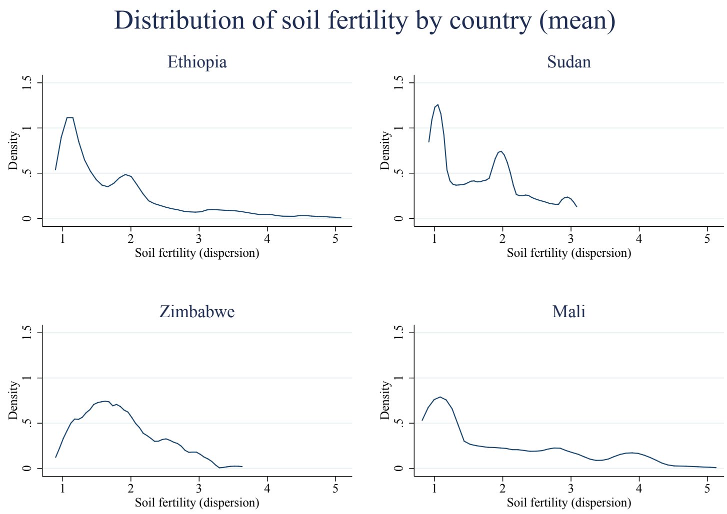
A.8 Soil fertility: kernel density

Figure A.7: Data visualization



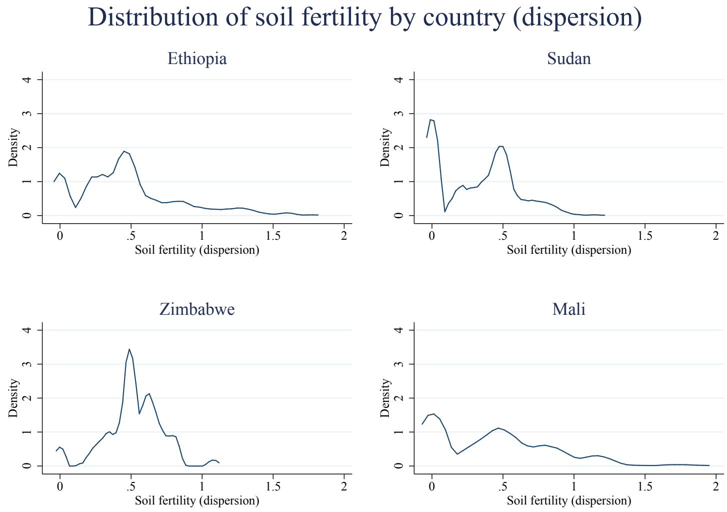
Density estimates presented are weighted by the Epanechnikov kernel. Data from the Harmonized World Soil Database. The figures represent the distribution of soil fertility, measured in terms of the cell-specific mode of nutrients availability. 1 denotes no or slight constraints, 5 very severe constraints.

Figure A.8: Data visualization



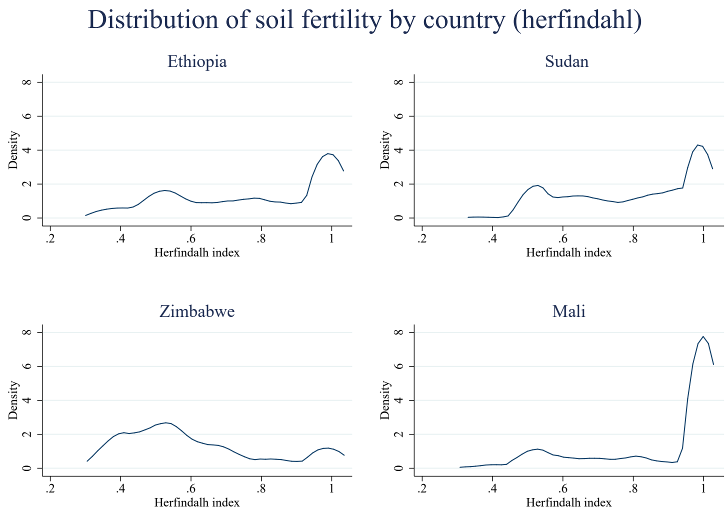
Density estimates presented are weighted by the Epanechnikov kernel. Data from the Harmonized World Soil Database. The figures represent the distribution of soil fertility, measured in terms of the cell-specific mean of nutrients availability. 1 denotes no or slight constraints, 5 very severe constraints.

Figure A.9: Data visualization



Density estimates presented are weighted by the Epanechnikov kernel. Data from the Harmonized World Soil Database. The figures represent the distribution of soil fertility, measured in terms of the cell-specific standard deviation of nutrients availability.

Figure A.10: Data visualization

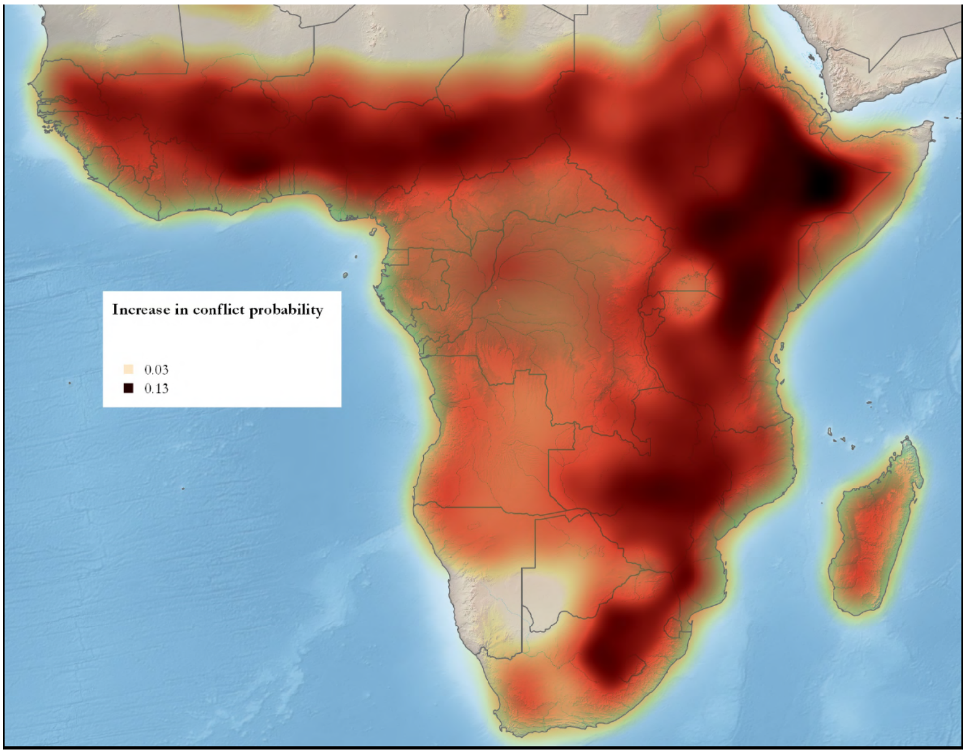


All density estimates presented are weighted by the Epanechnikov kernel. Data from the European Commission. Soils are ranked in three categories: naturally fertile, medium level of soil fertility and infertile. For each cell, we compute an Herfindalh index based on the share of the three types of soils.

A.9 Effect of a standard deviation increase in fertilizer prices on conflict

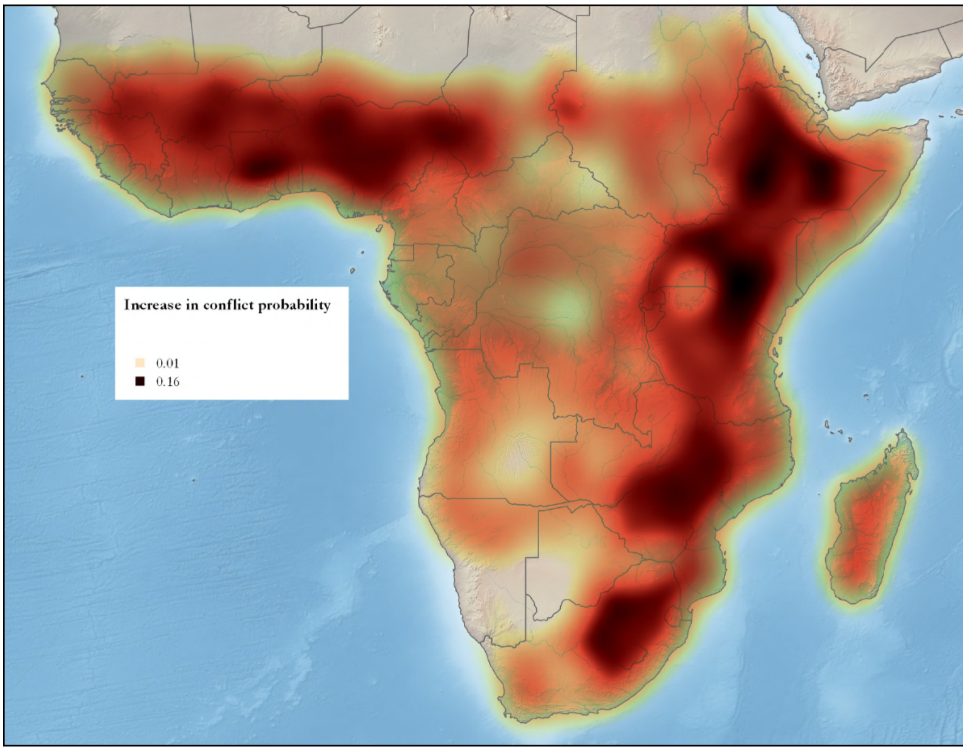
The next two figures represent the increase in conflict probability, in percentage points, that is predicted to occur following a standard deviation increase of fertilizer prices. Figure A.11 plots the predictions obtained from column (3) of Table 6 and Figure A.12 adds interaction with population density, and shows the predictions from column (2) of Table 7.

Figure A.11: The geography of soil fertility and conflicts: within and between ethnic groups inequality



This figure represents the increase in conflict probability, in percentage points, that is predicted to occur following a standard deviation increase of fertilizer prices. It plots the predictions obtained from column (3) of Table 6.

Figure A.12: The geography of soil fertility and conflicts: within and between ethnic groups inequality and population pressure



This figure represents the increase in conflict probability, in percentage points, that is predicted to occur following a standard deviation increase of fertilizer prices. It plots the predictions obtained from column (2) of Table 7.

B How fertilizer price variations affect fertilizer use and rents?

In section 2, we consider a stylized agricultural production model and use implications of this model to derive our main predictions on the link between fertilizer price variations, soil fertility heterogeneity, and the likelihood of conflict. Here we provide evidence supporting the three main implications from the agricultural model, namely: (a) farmers use lower levels of fertilizer on fertile soils (an implication of equation 12 in section 2), (b) rents from land decrease when the price of fertilizer increases (an implication of equation 13 in section 2), and (c) the negative effect of fertilizer prices increase on agricultural rents is magnified for less fertile soils (an implication of equation 15 in section 2).

We use household data from the World Bank Living Standard Measurement Surveys (LSMS) to provide several pieces of evidence supporting these three implications from the agricultural production model. We restrict our analysis to the rounds of surveys covering an African country and for which (i) the GPS coordinates of the household / plot is available, and (ii) at least two years of data are available. For each of these surveys, we gather data on (i) the use of inorganic fertilizers, and (ii) land value. This leaves us with the five countries: Ethiopia, Malawi, Nigeria, Tanzania, Uganda (panel data is available for a few other countries, but the surveys do not include our variables of interest). We then use the geo-location of the households to assign to each household a value of soil fertility based on our maps of nutrient availability (see section 3.3 of the main text). We categorize households location into two categories: fertile soils ("no or slight constraints" in terms of nutrient availability); infertile soils (from "moderate" to "very severe" constraints). Fertile soils represent 48% of the households.

Our results can be summarized as follows.

a) Farmers use lower levels of fertilizer on fertile soils. In order to test this implication, we compute the quantity of inorganic fertilizers used per acre for each household in the surveys. The availability of such information is limited but we are nevertheless able to construct this variable for 4,512 households located in four countries (Ethiopia, Malawi, Tanzania, Uganda). Table A.5 shows that being located in a fertile area decreases significantly the quantity of fertilizer used on average.

b) Rents decrease when the price of fertilizer increases. In order to test this implication, we make use of information on land value from the LSMS surveys. The households who are land-owners are asked the following question: "If you were to sell this parcel of land today, how much could you sell it for?". Using information on the area of each plot, we construct a land

Table A.5: Fertilizer use and soil fertility

	(1)	(2)	(3)
Dep. var.	Fertilizer quantity (kg/ac)		
Fertile soil	-157.099	-195.089[b]	-216.126[b]
	(131.553)	(90.029)	(85.979)
Fixed effects	-	Year	Country × year
Observations	9118	9118	9118

[c] significant at 10%; [b] significant at 5%; [a] significant at 1%. Conley (1999) standard errors in parentheses, allowing for spatial correlation within a 500km radius and for infinite serial correlation. OLS estimations. The unit of observation is the household-plot-year. Fertile soil is a dummy taking the value 1 if the household if located in a fertile area (i.e. in an area with no or moderate nutrient availability constraints). Fertilizer use is the amount of inorganic fertilizer per acre used by the household for each plot.

value per acre in USD, as a proxy for rents. We then study how it varies with our fertilizer price. Table A.6 displays the results. Column (1) shows the results of an estimation performed at the individual level, while in columns (4) and (5) we aggregate the data up to the cell-year level. In the latter case we compute either the median or the average land value. We find that an increase of cell-specific fertilizer price reduces the value of land – although the coefficient is not always significant when land value is aggregated at the cell-level.

c) The negative effect of fertilizer prices on agricultural rents is magnified for less fertile soils. In order to test this implication, we perform separate estimations for nutrient-rich and for nutrient-poor areas. The results provided in Table A.6 show that, while land value drops significantly with fertilizer prices in nutrient-poor areas (columns (2), (6), (8)), the effect is insignificant (and sometimes positive) for fertile zones (columns (3), (7), (9)).

Table A.6: Land value and fertilizer prices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. var.	ln land value								
Aggregation level	Household-plot-year			Cell-year					
Statistic				Mean	Median	Mean		Median	
Soil fertility	All	Poor	Rich	All	All	Poor	Rich	Poor	Rich
ln fertilizer price	-0.706[a]	-0.736[a]	-0.449	-0.630	-0.443[c]	-1.158[b]	0.997	-0.693[a]	0.197
	(0.190)	(0.200)	(0.549)	(0.553)	(0.263)	(0.496)	(1.148)	(0.228)	(0.717)
Individual FE	Household-plot			Cell					
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	26143	15879	10264	756	756	560	337	560	337

[c] significant at 10%; [b] significant at 5%; [a] significant at 1%. Conley (1999) standard errors in parentheses, allowing for spatial correlation within a 500km radius and for infinite serial correlation. OLS estimations. From columns (1) to (3) the unit of observation is household-plot-year level. From columns (4) to (9) the unit of observation is the cell-year level, and land value represents either the median or the average land value.

C Soil fertility heterogeneity and conflict: theory and cross-sectional evidence

Theory. We show theoretically that the effect of soil fertility heterogeneity on conflict is ambiguous. Let us consider the effect of an increase in the heterogeneity of soil fertility (Δs), holding average soil fertility ($\overline{s}$) constant. It affects both the tradeoff between current payoffs (left hand side in condition 6) and between future payoffs (right hand side in condition 6). The left hand side of condition (6) decreases when Δs increases and then conflict becomes more likely. However, we can show that the effect of an increase in soil fertility heterogeneity on the right hand side is à priori ambiguous. Indeed, let us consider the effect on the continuation value in case of victory, V^V. Using the envelope theorem, we find

$$\frac{\partial V^V}{\partial \Delta s} = E\left(\sum_{\tau=t+1,\dots,+\infty} \delta^\tau p_\tau \left(g_s(s_1, f_{1\tau}) - g_s(s_2, f_{2\tau})\right)\right), \tag{37}$$

where $f_{i\tau}$ for $i = 1, 2$ is characterized by condition (11).

The right hand side in condition (37) depends on $(g_s(s_1, f_{1\tau}) - g_s(s_2, f_{2\tau}))$. Its sign is negative if the derivative of $g_s(s_i, f_{i\tau})$ with respect to s_i is negative, or:

$$g_{ss} + g_{sf}\frac{\partial f_{it}}{\partial s_i} = g_{ss} - (g_{sf})^2/g_{ff} < 0. \tag{38}$$

Condition (38) holds when g_{ss} is negative and sufficiently large (it is positive otherwise). Hence, if g_{ss} is negative and large, we have $\frac{\partial V^V}{\partial \Delta s} < 0$, i.e. the continuation value in case of victory decreases with the heterogeneity of soil fertility. This leads us to the following conclusion:

Result [Soil fertility heterogeneity and conflict]: *An increase in soil fertility heterogeneity (Δs) can decrease the continuation value of victory (V^V), and thus has an ambiguous effect on the risk of conflict.*

Cross-sectional evidence. The relationship between soil fertility heterogeneity and conflict is difficult to test convincingly because of its cross-sectional nature: soil characteristics may be correlated with many unobserved local characteristics affecting violence, such as geography or socio-economic factors (on the other hand, the cross effect of variations in fertilizer price and natural soil fertility is more difficult to attribute to such unobserved variables). In this section we however provide some evidence showing that indeed, conflict tends to occur more in cells where soil fertility heterogeneity is high.

We regress average conflict incidence at the cell-level on the variance of nutrient availability, controlling for fertility level and for country fixed effects. Standard errors are clustered over space as in our baseline specifications. The results are provided in Table A.7: the effect of soil fertility heterogeneity is positive and statistically significant at the 5% level. In columns (2) to (4), we add sequentially a number of geographical and socio-economic factors; the coefficient of $\mathbb{V}$(Fertility) remains statistically significant and is quantitatively reinforced.

Table A.7: Land inequality and conflict: cross-sectional evidence

	(1)	(2)	(3)	(4)
Dep. var.		Conflict incidence (mean)		
$\mathbb{V}$(Fertility)	0.008^b	0.019^b	0.020^b	0.020^b
	(0.004)	(0.009)	(0.008)	(0.008)
$\overline{\text{Fertility}}$	-0.008^a	-0.010	-0.014^b	-0.013^b
	(0.003)	(0.006)	(0.006)	(0.005)
Country Fixed effects	Yes	Yes	Yes	Yes
Observations	8267	6220	5946	5941
Additional controls	-	Geography	+ Socio-Economic	+ Ethnic

^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. Robust standard errors in parentheses. OLS estimations. Conley (1999) standard errors in parentheses, allowing for spatial correlation within a 500km radius. $\overline{\text{Fertility}}$ is the mean of nutrient availability of the cell. $\mathbb{V}$(Fertility) is the variance of the nutrient availability level within the cell (from HWSD). In column (2) we add controls for geography-related cell-characteristics: the share of urban, agricultural, barren, forest, and water area. In column (3) we further add socio-economic characteristics: GDP per capita and density of population. In column (4) we further add an indicator of ethnic polarization.

D Budget constraints and non monotonic effects

Theory. Poor African farmers generally lack access to credit and thus face budget constraints that are not taken into account in our theoretical model. As a consequence, the quantity of fertilizer they purchase may be constrained when the price of fertilizer is high. Assume, for simplicity, that the budgets of the two groups are equal. Hence, when the price of fertilizer increases, the budget constraint of the land-poor (group 2) will become saturated first (because they use larger quantities of fertilizer). In such a case, the quantity of fertilizer group 2 purchases will become insensitive to fertilizer price variations. As a consequence, the rents from the less fertile soil (i.e. the opportunity cost of group 2) become insensitive while the rents from the most fertile soil (i.e. the rapacity gain of group 2) remain sensitive to fertilizer price variations. Hence, condition (14) becomes:

$$\frac{\partial \Psi_t}{\partial c_t} = P\left(1-d\right)\frac{\partial r_{1t}}{\partial c_t} < 0. \tag{39}$$

Thus, our first prediction as to be restated as follows:

Proposition [Fertilizer Price and Conflict with Budget Constraints]: *Assume that the two groups face a budget constraint. When the price of fertilizer price increases, the likelihood of conflict first increases and then decreases for sufficiently high levels of fertilizer price.*

Now, consider an increase in the heterogeneity of soil fertility. Condition (18) becomes:

$$\frac{\partial^2 \Psi_t}{\partial c_t \partial \Delta s} = P\left(1-d\right)\frac{\partial^2 r_{1t}}{\partial c_t \partial \Delta s} > 0. \tag{40}$$

An increase in the difference of soil fertility dampens the negative effect of fertilizer price on the rents from the most fertile soil. Hence, our second prediction remains unchanged:

Proposition [Soil Fertility Heterogeneity and Conflict with Budget Constraints]: *Assume that the two groups face a budget constraint. The cross-effect of an increase in fertilizer price and in the difference in soil fertility on the likelihood of conflict is still always positive.*

Table A.8 below tests these modified predictions. We add to our baseline specification the square of fertilizer prices in column (1), and its interaction with the variance of soil fertility. We find that the marginal effect of fertilizer price variations is lower for large spikes of prices. However, the overall effect (and the cross-effect of fertilizer prices and the variance of soil fertility) is always positive (i.e. the level of price for which the effect would reach zero given the estimated coefficient is largely unrealistic).

Table A.8: Fertilizer prices and conflict: non-monotonicity

	(1)	(2)
Dep. var.	Conflict incidence	
ln fertilizer price	0.631^{b}	1.093^{a}
	(0.281)	(0.317)
$\times$ $\mathbb{V}$(Fertility)		0.532^{a}
		(0.142)
ln fertilizer price^2	-0.047^{c}	-0.088^{a}
	(0.025)	(0.028)
$\times$ $\mathbb{V}$(Fertility)		-0.045^{a}
		(0.013)
Cell and Year Fixed effects	Yes	Yes
Observations	111605	111588

^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. Conley (1999) standard errors in parentheses, allowing for spatial correlation within a 500km radius and for infinite serial correlation. ln fertilizer price is our baseline fertilizer price shock, computed from using the required NPK mix (from IPNI) for the main crop produced by the cell (from GAEZ). $\mathbb{V}$(Fertility) is the variance of the nutrient availability level within the cell (from HWSD). Column (2) controls for the average level of soil fertility, interacted with fertilizer prices and its square.

E Local fertilizer prices: identifying assumptions

E.1 Do international market prices of fertilizer affect fertilizer consumption?

A widespread belief is that Sub-Saharan African farmers use little fertilizers, which may make them insensitive to international market price variations in these inputs. In this section we estimate the effect of fluctuations in fertilizer prices on consumption and imports of fertilizers at the country level. The data come from FAO-Stat. We consider three items which correspond to the three N-P-K nutrients that we use in our main empirical exercises: phosphate, nitrogen and potash fertilizers. For each item, we gather data on consumption, nutrient use and imports from FAO-Stat. The data varies by country and year. We then regress the log of imports, consumption and nutrient use on the log of the corresponding international market price of fertilizer. The results are displayed in Table A.9 below. We control in all estimations for country, item (N, P, K) and year dummies. In all cases, variations in fertilizer prices are found to negatively affect imports or consumption.

Table A.9: Impact of international market prices on consumption and imports

	(1)	(2)	(3)
Dep. var.	ln imports	ln consumption	ln nutrient use
ln international market prices	-0.156^b	-0.190^b	-0.119^b
	(0.076)	(0.081)	(0.048)
Fixed effects		Country, Item, Year	
Observations	1142	1147	758

^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. Estimations are run at the country-item-year level. Standard errors clustered by item-year in parentheses. OLS estimations. ln international market prices is the ln international market price of the fertilizer or nutrient (phosphate, potash or nitrogen fertilizer).

E.2 Are international market prices of fertilizer transmitted to local prices?

In this section we provide piece of evidence that variations in the international market prices of nutrients or fertilizers are indeed transmitted to local markets in Sub-Saharan Africa. Data on local fertilizer prices is not widely available. However, we were able to gather data on the prices of three inputs: (i) urea (nitrogen), (ii) phosphate, and (iii) Diammonium phosphate (DAP) fertilizers, for a subset of local markets throughout Africa. The data is provided by the Africa Fertilizer initiative[18] and covers over 300 markets in 17 countries.[19] We pool available data on the three inputs and regress the local price in logs on the international market price, controlling for various fixed effects. The results are provided in Table A.10 below. Column (1) controls for input fixed effects and year dummies; columns (2) to (4) progressively sequentially add market × input, country × year, and market × year fixed effects. In column (5) we estimate a model in first difference. In all specifications, international market and local price series are strongly correlated, with all coefficients being statistically significant at the 1% level.

Table A.10: Impact of international market prices on local prices

	(1)	(2)	(3)	(4)	(5)
Dep. var.			ln local price		Δln local price
ln international market price	3.166^a (1.117)	6.374^a (1.659)	6.819^a (1.609)	7.125^a (1.843)	
Δln international market price					6.439^a (1.959)
Fixed effects	Input Year	Market-input Year	Market-input Country-year	Market-input Market-year	- Year
Observations	970	970	713	348	445

^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. Robust standard errors in parentheses. OLS estimations. ln international market price is the world price of the fertilizer or nutrient (DAP fertilizer, urea, or phosphate). Local price is its market-level counterpart. Estimations are run at the market-input-year level.

[18] AfricaFertilizer.org

[19] Benin, Burkina Faso, Burundi, Ethiopia, Ivory Coast, Ghana, Kenya, Malawi, Mali, Mozambique, Niger, Nigeria, Senegal, Tanzania, Togo, Uganda and Zambia.

E.3 Do farmers use the appropriate mix of nutrients?

When computing our local fertilizer prices, we assumed implicitly that the fertilizers used by farmers locally should indeed reflect, at least to some extent, the ideal mix of nutrients they should use given the most suitable crop(s) in their region. Put differently, we make the assumption that the fertilizers available locally partly reflect the specialization of the region – that is, the nutrients' composition of the fertilizers available, say, a rice producing region should be closer to the mix of nutrients suitable for growing rice than the composition of fertilizers in other regions. In this section, we provide evidence supporting this assumption.

Our strategy is the following. Our fertilizer prices are crop-specific. Hence, if the local mix of nutrients used for each crop is indeed correlated with the required one, we should find a positive correlation between crop-specific producer prices and our fertilizer prices (controlling for common time shocks). We have compiled country-level data on crop specific producer prices (in USD) from FAO-STAT.[20] In Table A.11 we regress producer prices on fertilizer prices at the country × crop × year level. More precisely, we estimate whether our measure of fertilizer price correlates with the producer prices of the crops produced by each country. We control for time dummies in all specifications, and for country or crop × country fixed effects. The last columns additionally includes controls for weather conditions (rainfall and temperature). Despite the limited number of observations and the limited variations in the data, we find a positive and significant (at the 10 or 5% level) correlation between these two prices.

Table A.11: Impact of fertilizer prices on producer prices

	(1)	(2)	(3)	(4)
Dep. var.	ln producer prices			
ln fertilizer price	0.716^b	0.738^b	0.170^c	0.170^b
	(0.311)	(0.318)	(0.086)	(0.086)
Fixed effects	-	Country	Country × crop	Country × crop
	Year	Year	Year	Year
Weather-related controls	No	No	No	Yes
Observations	3923	3923	3906	3906
R^2	0.075	0.288	0.772	0.772

^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. Standard errors clustered by product × year in parentheses. OLS estimations. OLS estimations. producer price is the producer price of the crop. Fertilizer price is the average fertilizer price for the country (based on equation 21). The unit of observation is the crop-country-year.

[20] http://www.fao.org/faostat/en/home

F Additional results

F.1 Robustness: Measurement

F.1.1 Conflict definition

Table A.12: Definition of a conflict

Dep. var.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Conflict incidence (UCDP-GED)		# events (ACLED)		Conflict onset (ACLED)		Conflict ending (ACLED)	
ln fertilizer price	0.083^a	0.088^a	0.776^c	1.119^b	0.086^a	0.114^a	-0.143	-0.181
	(0.030)	(0.031)	(0.407)	(0.450)	(0.025)	(0.026)	(0.183)	(0.187)
$\times$ $\mathbb{V}$(Fertility)		0.036^a		0.650^a		0.035^a		-0.179^a
		(0.006)		(0.241)		(0.006)		(0.034)
$\times$ $\overline{\text{Fertility}}$		0.005^c		0.205^a		0.015^a		-0.040^b
		(0.003)		(0.067)		(0.003)		(0.019)
Cell and Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	91910	91896	111605	111588	104570	104553	13600	13599

^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. Conley (1999) standard errors in parentheses, allowing for spatial correlation within a 500km radius and for infinite serial correlation. Conflict onset is a dummy that takes the value 1 if a conflict is observed in the cell in t but none in $t-1$, and is coded as missing if a conflict is observed in $t-1$. Conflict ending is a dummy that takes the value 1 if no conflict is observed in the cell in t but a conflict is observed in $t-1$, and is coded as missing if no conflict is observed in $t-1$.

F.1.2 Soil fertility definition

In this subsection, we estimate the sensitivity of our baseline estimates to alternative definitions of soil fertility. Column (1), Table A.13 replicates our baseline estimates for an ease of comparison. In column (2), we use our baseline measure of nutrient availability, but include the range of nutrient availability values (maximum-minimum value) as measure of soil fertility heterogeneity, instead of the variance. From column (3) to column (5), we estimate the sensitivity of our results to the use of alternative data sources. First we use information from the European Commission on the quality of 31 soil types in Africa that we rank in three different categories according to the level of natural soil fertility that are coded on the basis of the description of the specific properties of the 31 soil types (see section A for more details). The variable *Concentration Fertility* is the Herfindahl index of this variable (column 3). In column (4), we make use of information on the nitrogen density in the soil and in column (5) on the percentage of irrigated land in the cell. Soil quality heterogeneity is found to magnify the effect of fertilizer price variations in all columns, the only insignificant coefficient being the one in column (4) where Nitrogen density is used as a proxy (the p-value is 0.2 in this case).

As mentioned in the body of the paper, soil fertility and fertilizer use are substitutes (an assumption we make in the theoretical model) as long as natural soil fertility does not fall below a certain level. This suggests that we should remove low fertility cells from the empirical analysis in order to be consistent with the assumption of the theoretical model. In Table A.14, we estimate the sensitivity of our baseline estimate (with the nutrient availability data source from the *Harmonized World Soil Database* - see section A for more details). We alternately remove i) cells for which the *average* constraints is "severe" or more (columns (1) and (2)), ii) cells for which the *maximum* constraints is "severe" or more (columns (3) and (4)). Our results are robust to these changes in the sample.

Table A.13: Alternative fertility measure

	(1)	(2)	(3)	(4)	(5)
Dep. var.	Conflict incidence				
Fertility measure	Nutrient availability (baseline)	Nutrient availability (range)	Rich soils (EU commission)	N density (ORNL-DAAC)	Irrigated land (PRIO-GRID)
ln fertilizer price	0.156^a	0.145^a	0.138^a	0.207^a	0.117^a
	(0.036)	(0.036)	(0.038)	(0.055)	(0.034)
$\times$ $\mathbb{V}$(Fertility)	0.058^a	0.018^a		0.003	0.171^b
	(0.008)	(0.003)		(0.002)	(0.069)
$\times$ Concentration Fertility			-0.028^b		
			(0.014)		
$\times$ $\overline{\text{Fertility}}$	0.021^a	0.020^a	0.010^c	-0.015^a	-0.001
	(0.003)	(0.003)	(0.006)	(0.005)	(0.002)
Cell and Year FE	Yes	Yes	Yes	Yes	Yes
Observations	111588	111588	111061	111571	111605

LPM estimations. ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. Conley (1999) standard errors in parentheses, allowing for spatial correlation within a 500km radius and for infinite serial correlation. In col. 1 we replicate our baseline estimate. In column 2, we use our baseline measure of nutrient availability, but include the range (maximum-minimum value) as measure of land heterogeneity, instead of the variance. In col. 3, concentration fertility is the Herfindahl index of types of soils computed from EU Commission data. In col. 4, $\mathbb{V}$(Fertility) is the log of the standard deviation of the Nitrogen Density variable from ORNL-DAAC (http://webmap.ornl.gov/). In col. 5, $\mathbb{V}$(Fertility) is the standard deviation of the share of irrigated land from PRIO-GRID. $\overline{\text{Fertility}}$ denote: in cols. 1-2, the average level of nutrient availability; in col. (3), the share of the cell covered by "rich soils"; in col. (4): the log of the level of Nitrogen Density; in col. (5): the average share irrigated.

Table A.14: Removing low fertility cells

	(1)	(2)	(3)	(4)
Dep. var.	Conflict incidence			
Robustness	Drop if *average* constraints are "severe" or more		Drop if *maximum* constraints are "severe" or more	
ln fertilizer price	0.105^a	0.140^a	0.109^a	0.141^a
	(0.032)	(0.034)	(0.033)	(0.035)
$\times$ $\mathbb{V}$(Fertility)		0.058^a		0.079^a
		(0.009)		(0.013)
$\times$ $\overline{\text{Fertility}}$		0.020^a		0.020^a
		(0.004)		(0.004)
Year FE	Yes	Yes	Yes	Yes
Cell FE	Yes	Yes	Yes	Yes
Observations	103326	103326	96679	96679

LPM estimations. ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. Conley (1999) standard errors in parentheses, allowing for spatial correlation within a 500km radius and for infinite serial correlation. Cols. 1 and 2: we drop all cells for which the average nutrient availability constraints are measured as "severe" or more. Cols. 3 and 4: we drop all cells for which the maximum nutrient availability constraint is measured as "severe" or more.

F.1.3 Fertilizer prices computation

Table A.15 displays some sensitivity estimates of our baseline results using alternative fertilizer price shocks. First, we use an alternative dataset on the required nutrient mix by crop from the US Department of Agriculture (columns (1) and (2)). In our baseline estimates, we consider the most suitable crop in each cell. In columns (3) and (4), we replicate our baseline estimates using the five main suitable crops in each cell. More precisely, we compute a fertilizer price index weighted by the relative suitability of each crop within cell. Last, in columns (5) and (6), we use an alternative measure of crop specialization (M3-CROP) build on the actual harvested area (year 2000). Our results are unaffected.

Table A.15: Alternative fertilizer price shocks

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var.	Conflict incidence					
Robustness	Alternative nutrient mix		Multiple crops		M3-crop data	
ln fertilizer price	0.099^a	0.121^a	0.167^a	0.189^a	0.074	0.108^b
	(0.030)	(0.031)	(0.051)	(0.050)	(0.052)	(0.051)
$\times$ $\mathbb{V}$(Fertility)		0.053^a		0.054^a		0.061^a
		(0.008)		(0.008)		(0.009)
$\times$ $\overline{\text{Fertility}}$		0.014^a		0.015^a		0.027^a
		(0.003)		(0.003)		(0.004)
Cell and Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	111605	111588	111605	111588	88842	88842

LPM estimations. ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. Conley (1999) standard errors in parentheses, allowing for spatial correlation within a 500km radius and for infinite serial correlation. Cols. 1 and 2: alternative optimal nutrient mix from USDA used in the computation of local fertilizer prices. Cols. 3 and 4: we include the 5 main crops (weighted by their suitability) instead of the main crop in the computation of local fertilizer prices. Cols. 5-6: main crop computed from M3-CROP data instead of GAEZ.

F.2 Robustness: Estimation

F.2.1 Standard errors

Table A.16: Alternative levels of clustering

	(1)	(2)
Dep. var.	Conflict incidence	
ln fertilizer price	0.119	0.156
distance: 100	$(0.024)^a$	$(0.025)^a$
distance: 250	$(0.029)^a$	$(0.031)^a$
distance: 750	$(0.037)^a$	$(0.039)^a$
distance: 1000	$(0.037)^a$	$(0.040)^a$
$\times$ $\mathbb{V}$(Fertility)		0.058
distance: 100		$(0.007)^a$
distance: 250		$(0.007)^a$
distance: 750		$(0.008)^a$
distance: 1000		$(0.009)^a$
$\times$ $\overline{\text{Fertility}}$		0.021
distance: 100		$(0.002)^a$
distance: 250		$(0.003)^a$
distance: 750		$(0.004)^a$
distance: 1000		$(0.004)^a$
Year FE	Yes	Yes
Cell FE	Yes	Yes
Observations	111605	111588
Countries	43	43
Period	1997-2013	1997-2013

LPM estimations. ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. Conley (1999) standard errors in parentheses, allowing for spatial correlation within a 500km radius and for infinite serial correlation. ln fertilizer price is our baseline fertilizer price shock, computed from using the optimal NPK mix (from IPNI) for the main crop produced by the cell (from GAEZ). Level fertility is the average level of nutrient availability of the cell, ranging from 1 to 5 (from HWSD). $\mathbb{V}$(Fertility) is the variance of the nutrient availability level within the cell (from HWSD).

F.2.2 Alternative estimators

Table A.17: Nonlinear estimators

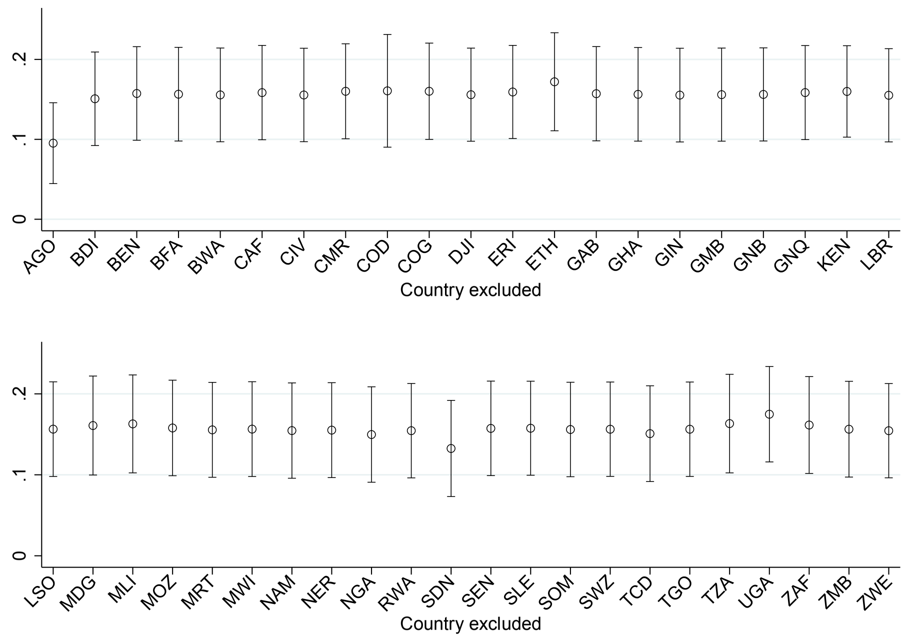
	(1)	(2)	(3)	(4)
Dep. var.	Conflict incidence		# events	
ln fertilizer price	1.730^{c}	2.553^{c}	1.443	2.559^{c}
	(0.935)	(1.324)	(1.155)	(1.452)
$\times$ $\mathbb{V}$(Fertility)		1.202^{a}		1.293^{b}
		(0.332)		(0.591)
$\times$ $\overline{\text{Fertility}}$		0.482^{c}		0.714^{a}
		(0.258)		(0.217)
Cell and Year FE	Yes	Yes	Yes	Yes
Estimator	FE Logit		PPML	
Observations	42772		42908	

^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. Standard errors clustered by country in parentheses. ln fertilizer price is our baseline fertilizer price shock, computed from using the required NPK mix (from IPNI) for the main crop produced by the cell (from GAEZ). Level fertility is the average level of nutrient availability of the cell, ranging from 1 to 5 (from HWSD). $\mathbb{V}$(Fertility) is the variance of the nutrient availability level within the cell (from HWSD). In columns (1) and (2) the dependent variable is a dummy taking the value 1 if at least a conflict event is observed in the cell during the year, 0 otherwise. In columns (3) and (4), the dependent variable is the number of conflict events observed in the cell during the year.

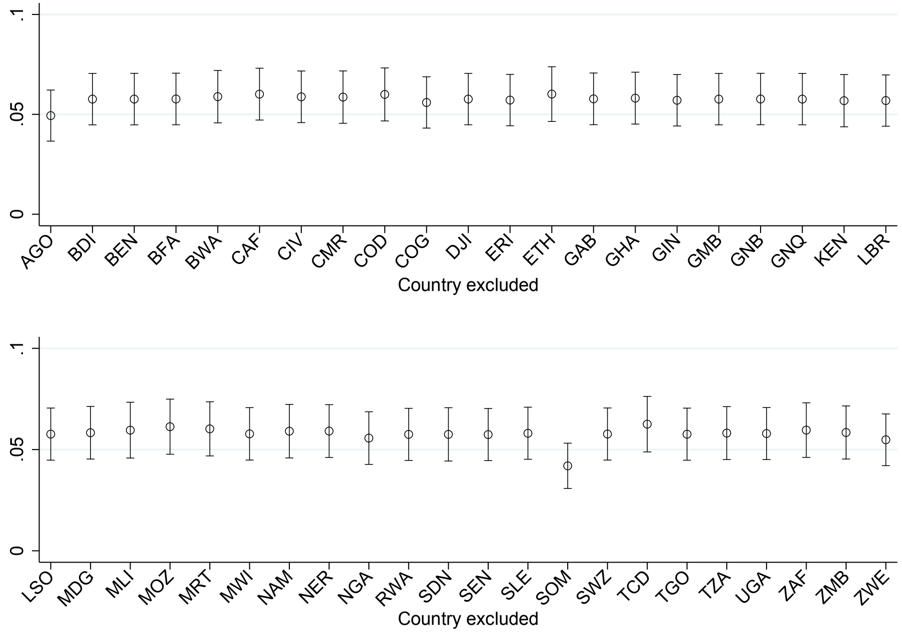
F.3 Sensitivity to specific countries, years or crops

Figure A.13: Exclusion of country one by one

A- Estimate of *ln fertilizer price*



B- Estimate of *ln fertilizer price*$\times$ $\mathbb{V}$(Fertility)

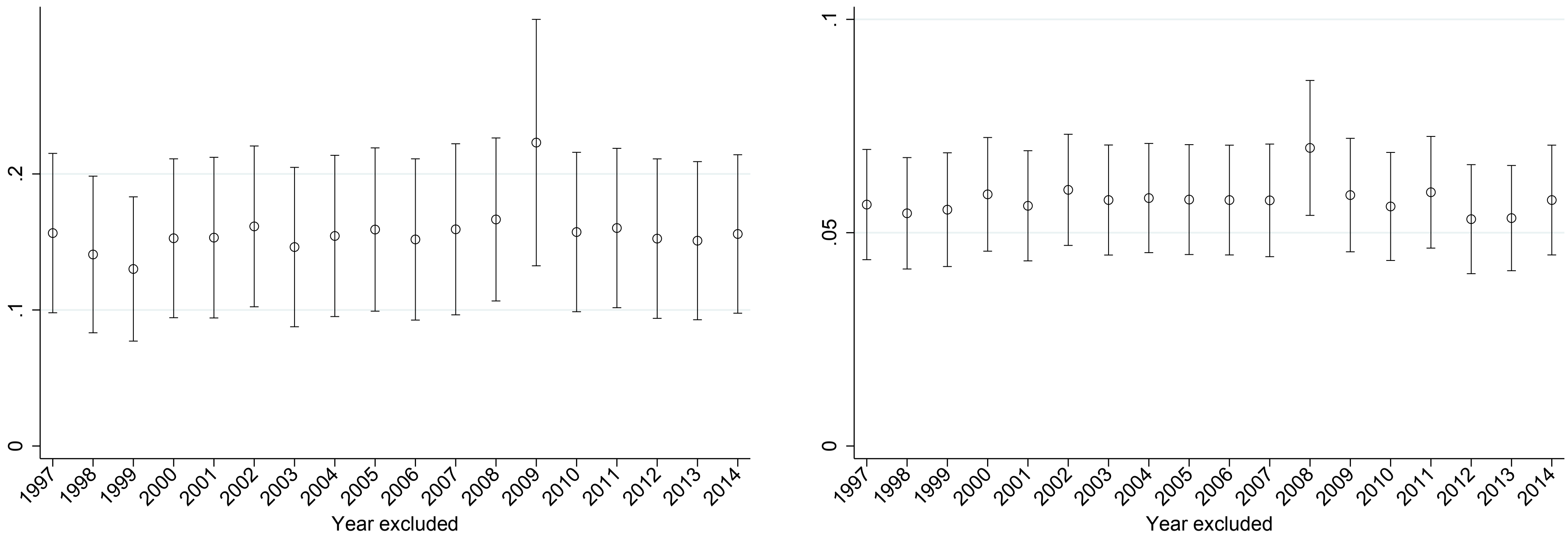


Note: Figure A and B report estimates of Table 2, column (2), when each country is excluded alternatively. Estimated coefficients and confidence interval at 90% are reported. Each estimated coefficients and confidence interval come from a single estimation.

Figure A.14: Exclusion of year one by one

A- Estimate of *ln fertilizer price*

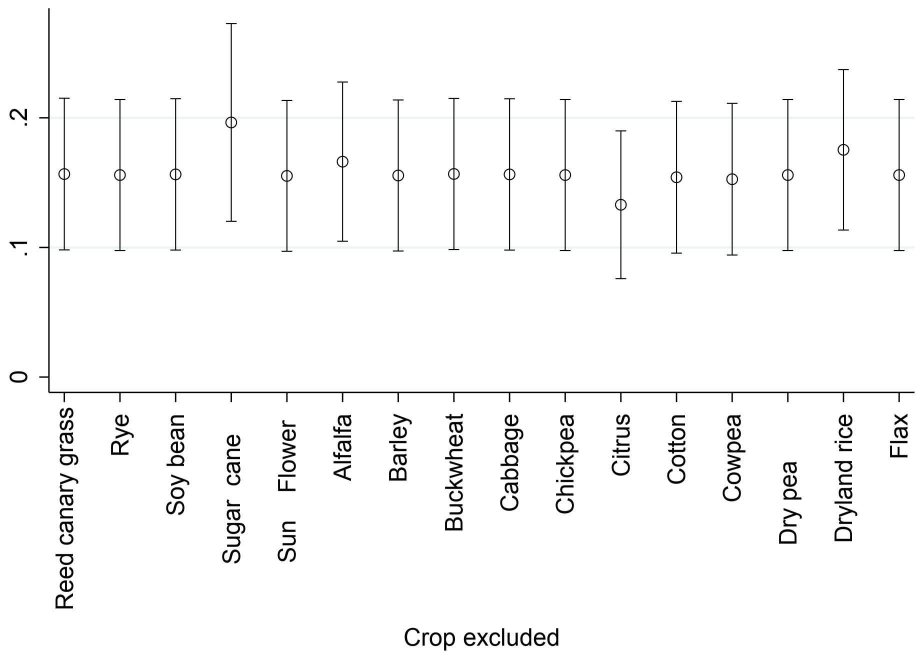
B- Estimate of *ln fertilizer price* $\times$ $\mathbb{V}$(Fertility)



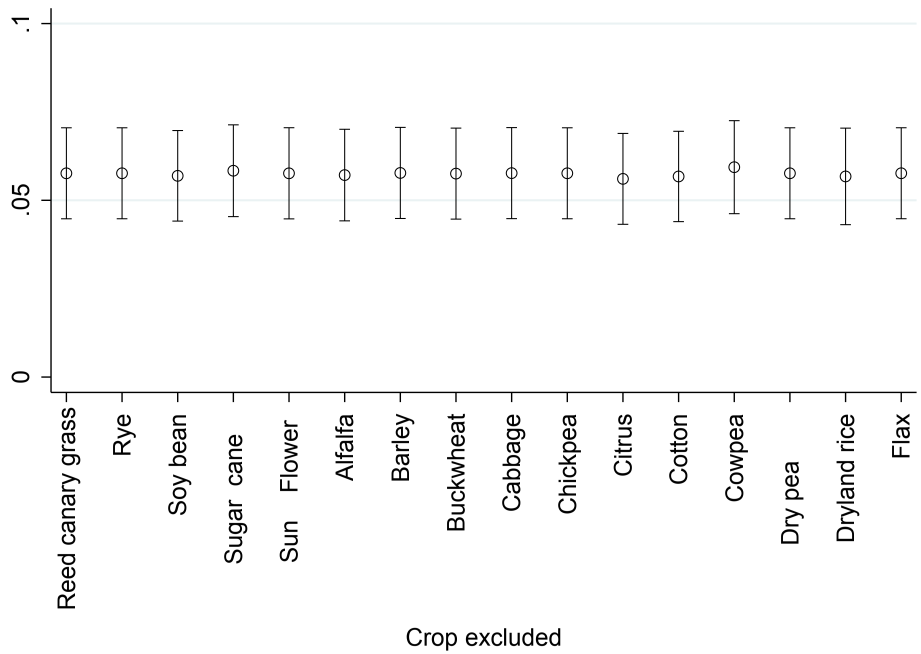
Note: Figure A and B report estimates of Table 2, column (2), when each year is excluded alternatively. Estimated coefficients and confidence interval at 90% are reported. Each estimated coefficients and confidence interval come from a single estimation.

Figure A.15: Exclusion of crop one by one

A- Estimate of *ln fertilizer price*



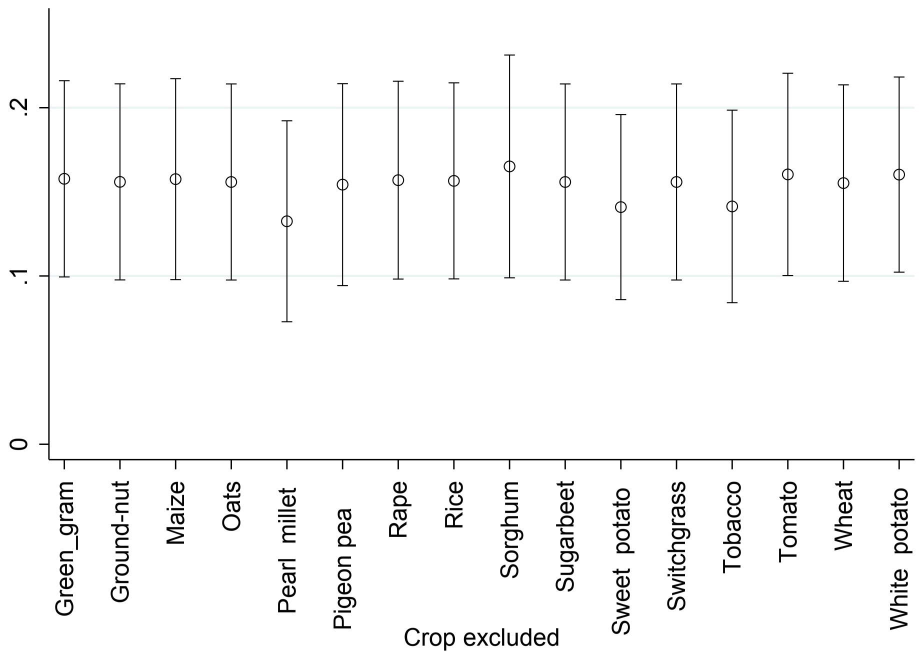
B- Estimate of *ln fertilizer price*$\times$ $\mathbb{V}$(Fertility)



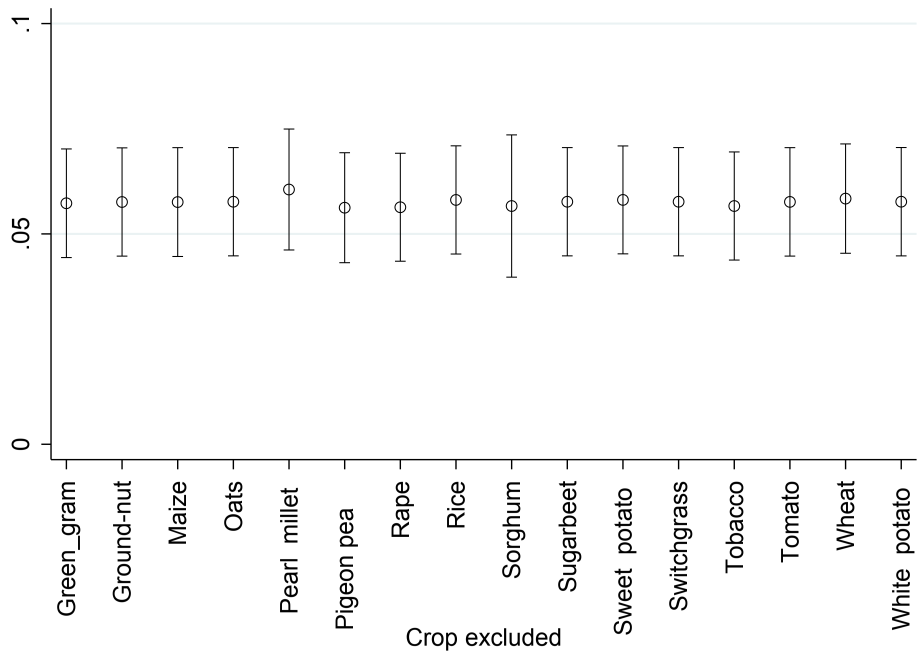
Note: Figure A and B report estimates of Table 2, column (2), when each crop is excluded alternatively. Estimated coefficients and confidence interval at 90% are reported. Each estimated coefficients and confidence interval come from a single estimation.

Figure A.16: Exclusion of crop one by one

A- Estimate of *ln fertilizer price*



B- Estimate of *ln fertilizer price* $\times$ $\mathbb{V}$(Fertility)



Note: Figure A and B report estimates of Table 2, column (2), when each crop is excluded alternatively. Estimated coefficients and confidence interval at 90% are reported. Each estimated coefficients and confidence interval come from a single estimation.

F.4 Robustness: Omitted variables

Table A.18: Alternative specifications

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var.			Conflict incidence			
ln fertilizer price	0.086^a	0.072^b	0.114^a	0.098^a	0.063^a	0.072^a
	(0.029)	(0.030)	(0.031)	(0.031)	(0.018)	(0.019)
$\times$ $\mathbb{V}$(Fertility)		0.025^a		0.024^a		0.003^a
		(0.006)		(0.007)		(0.001)
$\times$ $\overline{\text{Fertility}}$		-0.004		-0.005		0.003^a
		(0.003)		(0.003)		(0.000)
Cell FE	Yes	Yes	Yes	Yes	Yes	Yes
Country$\times$Year FE	Yes	Yes	No	No	No	No
Country-specific time trends	No	No	Yes	Yes	No	No
Country$\times$Region$\times$Year FE	No	No	No	No	Yes	Yes
Observations	111605	111588	111605	111588	111265	111248

^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. Conley (1999) standard errors in parentheses, allowing for spatial correlation within a 500km radius and for infinite serial correlation. A region is defined at as the country-specific second administrative units.

F.5 Robustness: Land inequality across and within ethnic groups

Table A.19: Land inequality within and between ethnic groups: robustness

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. var.	Conflict incidence					
Ethnic groups data	Murdock			GREG		
ln fertilizer price	0.147^a	0.063^b	0.053^c	0.151^a	0.059^b	0.051^c
	(0.035)	(0.030)	(0.030)	(0.035)	(0.030)	(0.030)
$\times$ $\mathbb{V}_W$(Fertility)	0.046^a	0.016^a	0.009	0.063^a	0.024^a	0.022^b
	(0.007)	(0.006)	(0.006)	(0.009)	(0.007)	(0.009)
$\times$ $\mathbb{V}_B$(Fertility)	0.072^b	0.058^b	0.006	0.020	0.048	0.020
	(0.031)	(0.027)	(0.029)	(0.039)	(0.037)	(0.049)
$\times$ Ethn. Frac.	0.016	0.025^a	0.021^b	0.020^c	0.048^a	0.040^a
	(0.010)	(0.009)	(0.009)	(0.012)	(0.010)	(0.011)
$\times$ Density pop.			-0.001			-0.000
			(0.001)			(0.001)
$\times$ $\mathbb{V}_B$(Fertility) $\times$ Density pop.			0.023^b			0.016
			(0.011)			(0.017)
$\times$ $\mathbb{V}_W$(Fertility) $\times$ Density pop.			0.009^a			0.010^a
			(0.002)			(0.004)
Cell FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	No	No	Yes	No	No
Country $\times$ Year FE	No	Yes	Yes	No	Yes	Yes
Observations	111316	111316	97580	110755	110755	97410

LPM estimations. ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%. Conley (1999) standard errors in parentheses, allowing for spatial correlation within a 500km radius and for infinite serial correlation. ln fertilizer price $\times$ $\overline{\text{Fertility}}$ is included as a control variable. $\mathbb{V}_W$(Fertility) is the variance of soil fertility (nutrient availability) within ethnic groups, computed by cell. $\mathbb{V}_B$(Fertility) is the variance of soil fertility between ethnic groups, computed by cell. Density pop. is the log of population density in 1990. Ethn. frac. is an index of ethnic fractionalization in the cell. Data on ethnic groups come from Murdock (1959) in cols 1-3, and from the Geo-referencing of Ethnic Groups (Weidmann and Cederman (2010)) in cols. 4-6.

G Additional references

- Meiyappan, P. and Jain, A. K. (2012), "Three distinct global estimates of historical land-cover change and land-use conversions for over 200 years", *Frontiers of Earth Science*, vol. 6 no 2: pp. 122-139.

- Nunn, N. and Qian, N. (2011), "The Potato's Contribution to Population and Urbanization: Evidence From A Historical Experiment", *Quarterly Journal of Economics*, vol. 126 no 2: pp. 593-650.

- Portmann, F. T., Siebert, S. and Doll, P. (2010), "MIRCA2000-Global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling", *Global Biogeochemical Cycles*, vol. 24 no 1.

- Reynal-Querol, M. (2014), "Data on ethnic and religious fractionalization and polarization", dataset, (http://www.econ.upf.edu/en/people/onefaculty.php?id=p1330).

- Vicente-Serrano, S. M., Begueria, S. and Lopez-Moreno, J. I. (2010), "A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index", *Journal of Climate*, vol. 23 no 7: pp. 1696-1718.