Localizing conflict spillovers:
introducing regional heterogeneity in conflict studies

Maarten Bosker*and Joppe de Ree†

Word count: 11429

Abstract

Cross-border conflict spillovers are commonly viewed as one of the main instigators of violent civil conflict. We question the alleged robustness of this finding reported by earlier studies that typically ascribe part of the observed spatial clustering of civil war to cross-border conflict spillovers. Spillovers however, are not the only explanation for the spatial clustering of civil war. We show that if we control for two different types of *regional heterogeneity* the evidence on the role of conflict spillovers turns out to be more delicate and not as omnipresent as suggested by previous studies. We allow for regional heterogeneity in unobserved conflict causing variables and in the magnitude of conflict spillovers themselves. We find significant conflict spillovers in Africa, but not in other parts of the world. In Africa, having at least one neighbor at war increases the probability of civil war onset by about 4%.

---

*University of Groningen, The Netherlands. Email: E.M.Bosker@rug.nl (Maarten)
†Corresponding author: DIW Berlin, Mohrenstrasse 58, 10117 Berlin, Germany. Email: jderee@diw.de (Joppe).

1 Introduction

The onsets of civil war are clustered in particular regions around the world. Figure 1 shows that in the second half of the 20th century civil war mostly broke out in (Sub-Saharan) Africa, (South-East) Asia, and Central and South America.

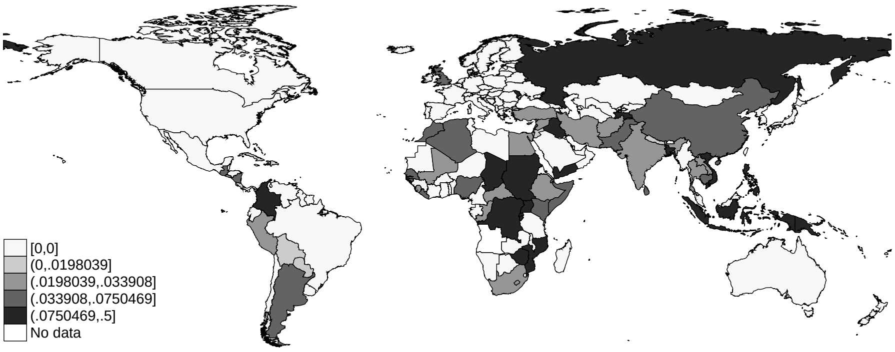

Figure 1: Average onset by country, 1945-2000

Note: Average onset is calculated as the number of onsets divided by the number of years a country had peace at $t-1$.

The clustering of civil war has not gone unnoticed and has attracted the interest of policymakers and political scientists alike. To date, there are two explanations for the observed clustering that have received considerable attention in the literature trying to understand the causes of civil war.

Until recently the clustering of civil war has been explained –mostly without paying explicit attention to the degree of spatial conflict clustering– by the similar spatial clustering of country specific attributes that are thought to cause the outbreak of civil war. Income per capita and democracy are two prominent examples that are clustered in a similar way as civil conflict. Examples of such studies are Hegre, Ellingsen, Gates, and Gleditsch (2001), Elbadawi and Sambanis (2002), Fearon and Laitin (2003), Collier and Hoeffler (2004) and Miguel, Satyanath, and Sergenti (2004).

Recently a second explanation for the observed spatial clustering of civil war has gained considerable popularity in the empirical literature on conflict. This second explanation argues that the spread of conflict from one country to its neighbors (i.e., conflict spillovers) is driving the spatial clustering of civil war [see for example Sambanis (2001), Buhaug and Rød (2006), Salehyan and Gleditsch (2006), Gleditsch (2007) and Buhaug and Gleditsch (2008)]. Indeed, several clear examples of conflicts spilling over from one country to the other do exist. Burundi's civil war spread to Rwanda and subsequently to the Democratic Republic of Congo (also directly involving Uganda and Tanzania among others). Sierra Leone's and Liberia's civil wars were heavily intertwined and the conflict on the Balkan first involved mainly Serbia and Croatia but spread from there to Bosnia and Kosovo.

The second explanation is fundamentally different from the first by adding an international dimension to the causes (and consequences) of civil war. A country may well have its internal institutions and economy in good shape, but when it happens to be located in a "bad neighborhood" its chances of falling into a state of civil war are much higher than were it surrounded by politically stable countries with well functioning economies [see Sambanis (2002), Gleditsch (2007) and Buhaug and Gleditsch (2008)]. In effect, if conflict spillovers are important, the fate of individual countries does not only depend on their own actions to prevent civil war, but also on what is happening in neighboring states around them.

Of course, both explanations for the spatial clustering of civil war are not mutually exclusive. It is not unlikely that both are relevant for explaining the clustering of civil war onset around the globe. Establishing the relevance of both explanations is important however, as policy implications that follow from these two views differ widely. If (observed or unobserved) country specific attributes are driving the outbreak of civil war, the causes of civil war are confined to within each country's own boundaries. As a result, policies to prevent civil war should be country-specific and aimed at lowering the chances of civil war by e.g. alleviating poverty or improving institutions. If instead conflict spillovers are identified as important drivers of conflict it would stress the significance of transnational causes of conflict like refugee movements, rebel groups receiving direct aid from, or safe refuge in neighboring countries [see e.g. Salehyan and Gleditsch (2006); Gleditsch (2007); Buhaug and Gleditsch (2008) and Moore and Shellman (2007)]. This would open up an international dimension to the issue of conflict prevention and emphasizes the importance of policies aimed at preventing the spread of conflict from one state to the other (regional cooperation, border defense, etc.).

Given the important differences in terms of policy implications, we think that the current empirical literature on civil conflict has been too quick in ascribing an important role to conflict spillovers in the process of conflict ignition. Where some authors have already been critical and do not find a big role for conflict spillovers [Hegre, Ellingsen, Gates, and Gleditsch (2001); Fearon and Laitin (2003)], recent work has shown repeatedly that conflict spillovers do seem a reality [see Sambanis (2001), Salehyan and Gleditsch (2006), Gleditsch (2007) and Buhaug and

Gleditsch (2008)]. The latter is exemplified in a recent robustness analysis into the causes of civil war by Hegre and Sambanis (2006) that identifies conflict spillovers as one of the most robust explanations for conflict around the world.

Despite the evidence regarding the relevance of conflict spillovers, we question the alleged robustness of these findings. We argue that the empirical literature insufficiently allows for heterogeneity at the regional level, which potentially results in spurious findings of conflict spillovers. In particular, we focus on two different types of regional heterogeneity that have not received adequate attention in the empirical literature on civil conflict.

The first type of heterogeneity we consider is (regional) heterogeneity in unobserved conflict causing variables. Where most empirical studies into the causes of civil war adequately control for *observables*[1] causing the spatial clustering of civil war –such as per capita income–, no adequate attention has been given to modeling *unobserved* (regional) heterogeneity. If the spatial clustering of civil war is (only partly) caused by unobserved characteristics, standard regression models that include a conflict spillover variable will quite easily –and wrongly so– ascribe part of this variation to the spillover variable and as a result misestimate both its size and significance. This problem is well known in the spatial econometrics and empirical economic geography literature, where it is referred to as the issue of spatial heterogeneity (omitted variables) versus spatial dependence (spillovers) [see e.g. Elhorst (2003), Abreu, Florax, and de Groot (2005) and Breinlich (2006)].

We argue that the extensive temporal and cross-country coverage that typifies civil conflict datasets and the fact that conflict spillovers and (unobserved) regional heterogeneity have inherently different characteristics allow the researcher to model both spillovers and unobserved heterogeneity at the same time. Simple statistical tests are subsequently available to reveal the relevance of both competing explanations for the spatial clustering of conflict. In section 4.1 we derive explicit mathematical expressions for biases in the estimated conflict spillover parameter resulting from not, or inadequately controlling for unobservables. These derivations provide useful intuition about the origins of these biases and facilitate the interpretation of the results we find in the empirical sections of the paper. More specifically, in the empirical section 4 we show the impact of allowing for unobserved heterogeneity at four, ever more detailed, geographical scales.

The second type of heterogeneity we allow for in this study is regional heterogeneity in the relevance of conflict spillovers themselves. Previous studies looking for evidence on conflict

---

[1] Observed spatial heterogeneity.

spillovers assume that the impact of having a neighboring country at war is the same around the world. If this assumption is violated however, we would tend to ascribe too much or too little importance to conflict spillovers in different parts of the worlds. It is not unlikely that conflict spillovers are a reality in some parts of the world, whereas in other regions the relevance of the international dimension of conflict is swamped by individual countries' characteristics. In section 3 we show that one can already gauge the importance of regional differences in the relevance of conflict spillovers by looking at how much individual countries contribute to the estimated effect of conflict spillovers in a very simple econometric onset model. In section 5 we further develop this idea and show that the evidence for conflict spillovers indeed differs a lot across different regions of the world.

Overall, we show that the conclusions regarding the role of conflict spillovers change quite dramatically after allowing for these two simple types of heterogeneity at the regional level. Doing so results in a much more accurate pinpointing into when and where these conflict spillovers seem to be a fact instead of a (statistical) fiction, hereby being of great help in increasing our understanding of when, where and how the international dimension of civil war is important. We find that conflict spillovers are not as omnipresent as suggested by many of the recent empirical papers in the conflict literature. More specifically, conflict spillovers seem to be a reality on the African continent, whereas we find no evidence of conflict spillovers in other parts of the world.

2 Data

Before going into the theoretical and empirical contributions of our paper, we briefly discuss the data that we use throughout the paper. We use the data set from Hegre and Sambanis (2006) as our baseline data set. As a result we define civil war in accordance with their definition that is basically the civil war definition used in Sambanis (2004). A civil war is defined as an armed conflict between an internationally recognized state and (mainly) domestic challengers able to mount an organized military opposition to the state. The war must have caused more than 1,000 deaths in total and in at least a three-year period.

To control for country specific attributes that may be driving the outbreak of civil war (observed regional heterogeneity), we include the following variables as regressors in our most elaborate onset models: gdp per capita, population size, and peace duration (Hegre and Sambanis (2006) take these as their 'core - variables' to be included in any study on civil war onset, see page 512 of their paper), the growth rate of gdp per capita, a measure of ethnic heterogeneity (the Vanhanen (1999) index), an oil exporter dummy and a measure of democracy based on the Polity IV index.

As our conflict spillover variable, $n_{i,t-1}$, we include a dummy variable taking the value 1 if *one or more* of a country's neighbors experienced civil conflict in period $t-1$[2].

Finally, we augment the data set with three measures that we will include to control for unobserved regional heterogeneity: 6 continent dummies, 21 world-region dummies (see Appendix C for their definition), or individual country dummies.

3 Localizing spillovers

Figure 1 shows that the onset of civil war is geographically clustered in different parts of the world. Africa, Asia and Central and South America are the regions where civil conflict most frequently broke out. But also within these broader defined regions we find differences between countries in the number of conflict outbreaks (some countries in Africa and South America have never experienced the onset of civil war, whereas others saw (renewed) outbreaks of civil conflict rather frequently). In effect, figure 1 illustrates the potential relevance of the first type of regional heterogeneity that we are considering in this paper: some regions (or countries) are for some reasons more prone to outbreaks of civil conflict than others. If we fail to allow for this variation in our conflict models (by e.g. including geographically clustered regressors, but also by allowing for country or region fixed effects), we could easily –and mistakenly– ascribe part of the observed spatial clustering to conflict spillovers.

Whereas figure 1 identifies regions where *outbreaks* of civil war occur relatively more frequent than in others, we are also interested in identifying regions where *spillovers* occur more frequently than in others. To graphically illustrate this second type of heterogeneity, we construct a simple country specific measure of spillover intensity. Figure 2 plots this measure and allow us to localize areas where onsets happen more frequently when neighboring countries are experiencing civil war than when neighbors are at peace. If conflict spillovers matter, we would observe that civil war onset happens (significantly) more frequently with one's neighbors at war than with one's neighbors at peace. Studying world wide patterns of this measure offers preliminary insights into whether, and if so, which countries or areas seem more prone to conflict spillovers than others.

This country-specific measure of spillover intensity is calculated as each country's individual contribution to the estimated pooled conflict spillover parameter, $\hat{\rho}^{POLS}$, that results from esti-

---

[2] We will also look at the impact of having more than one neighboring country at war in section 6.

mating the following simple linear probability onset regression (we refer to this estimator as the pooled OLS estimator[3]):

$$c_{it} = \gamma + \rho n_{it-1} + [u_{it}] \quad \text{for } c_{i,t-1} = 0 \tag{1}$$

where c_{it} denotes a dummy variable taking the value 1 if country i experiences civil war at time t, γ denotes a constant, n_{it-1} is a dummy variable taking the value 1 if at least one of country i's neighbors experienced civil war at time $t-1$. u_{it} is an error term.

The pooled OLS parameter estimator $\hat{\rho}^{POLS}$ of model 1 is simply the difference between the empirical probability (or relative frequency) of an onset that *could* be due to a spillover (i.e. the empirical probability of war onset conditional on having at least one neighbor at war at time $t-1$), and the empirical probability of war onset that could never have been due to a spillover (i.e. the empirical probability of war onset when all neighbors are at peace) [see Appendix A for mathematical details].

$$\hat{\rho}^{POLS} = \frac{s_1}{s_3} - \frac{s_2}{s_4} \tag{2}$$

where s_1 is the total number of onsets with at least one neighbor at war, s_2 the total number of onsets without any neighbor at war, s_3 the total number of observations with at least one neighbor at war and s_4 the total number of observations without any neighbor at war.

Next, the right-hand side of equation (2) can be rewritten as the sum of the individual countries' contributions:

$$\hat{\rho}^{POLS} = \frac{s_1}{s_3} - \frac{s_2}{s_4} = \frac{\sum_i s_1^i}{s_3} - \frac{\sum_i s_2^i}{s_4} = \sum_i \left(\frac{s_1^i}{s_3} - \frac{s_2^i}{s_4} \right) = \sum_i \left(\underbrace{\frac{s_1^i}{s_3^i} \cdot \frac{s_3^i}{s_3} - \frac{s_2^i}{s_4^i} \cdot \frac{s_4^i}{s_4}}_{i's \quad contribution} \right) \tag{3}$$

where s_1^i, s_2^i, s_3^i and s_4^i are the country specific counterparts to s_1, s_2, s_3 and s_4 respectively (e.g. s_1^i is the total number of onsets *in country i* with at least one neighbor at war). Each country's contribution to the estimated pooled spillover parameter, $\hat{\rho}^{POLS}$, can thus be interpreted as the difference between the country specific probability of onset that could have been a spillover, $\frac{s_1^i}{s_3^i}$ (multiplied by the country's share in the total number of observations with at least one neighbor at war, $\frac{s_3^i}{s_3}$), and the country specific probability of onset that could never have been due to a conflict spillover, $\frac{s_2^i}{s_4^i}$ (multiplied by the country's share in the total number of observations without any

---

[3] The estimated parameter $\hat{\rho}^{POLS}$ is equivalent to the marginal effects estimator of a probit model with only n_{it-1} and a constant included as regressors.

neighbor at war, $\frac{s_4^i}{s_4}$).

Figure 2 maps the spatial distribution of the country-specific measure of spillover intensity, providing a first glance at where most of the conflict spillover action is coming from.

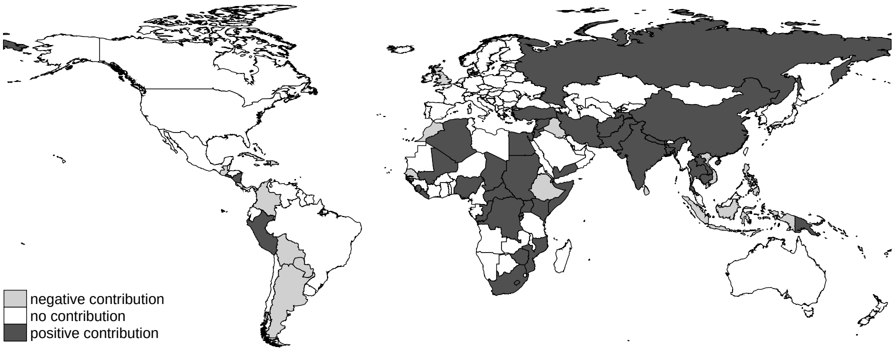

Figure 2: Individual countries' spillover intensity

Notes: Lightly shaded countries contribute ***negatively*** *and the darker colored countries contribute positively to the pooled spillover parameter estimate. White countries do not contribute to the pooled spillover parameter estimate (114 out of the 180 countries in our data do not contribute to the estimated pooled spillover parameter).*

This shows that the countries that contribute most to the pooled OLS estimate $\hat{\rho}^{POLS}$ are mainly located in Africa and Asia. Countries in Europe, North America all contribute zero, mostly because none of these countries experienced civil war. Most interesting, some countries in Latin America and Asia contribute negatively to the simple pooled conflict spillover parameter [see Appendix D for the top ten and bottom ten contributors]. This preliminary exercise strongly suggests that conflict spillovers may be more relevant in some regions than in others.

Figure 1 and 2 each serve to illustrate the relevance of dealing with the two types of regional heterogeneity we consider in this paper. We note however, that it is hard to a priori distinguish between either type on the basis of these two figures. We need to allow for both concepts simultaneously in more elaborate econometric conflict models to estimate their relative importance. Keeping Figure 1 and 2 in mind, we now turn to the econometric estimation of our fully fledged civil war onset models where we allow for both types of regional heterogeneity and assess their respective relevance.

4 Controlling for regional heterogeneity in unobservables

As our baseline model we employ a fairly standard model of civil war onset. We are interested in estimating the conditional probability of conflict in country i in period t, c_{it}, given that it did not experience civil conflict one period before, $c_{i,t-1} = 0$. The onset probability is allowed to depend on a set of country-specific $\mathbf{x}_{it}$ variables (gdp per capita, population size, etc) that we include to control for observed (regional) heterogeneity. Moreover, we model spillovers by including a dummy variable that takes the value 1 if *one or more* of a country's neighbors experienced civil conflict in period $t-1$, $n_{i,t-1}$.[4] Finally, we allow for time invariant country specific effects α_i that capture any unobserved time-invariant heterogeneity at the country level.

$$P\left(c_{it} = 1 \middle| c_{it-1} = 0, n_{i,t-1}, \mathbf{x}_{it}, \alpha_i\right) = F\left(\rho n_{i,t-1} + \mathbf{x}_{it}\beta + \alpha_i\right) \quad \text{for } c_{it-1} = 0 \tag{4}$$

The function F is the normal CDF for a probit model and the identity transformation for a linear probability model. Moreover, we assume (and this is standard in the empirical conflict literature) that the separate arguments $n_{i,t-1}$, $\mathbf{x}_{it}$ and α_i enter additively. ρ is the parameter of interest and captures both the direction and the intensity of conflict spillovers. A positive ρ for example implies that the probability of a civil war onset at period t is increasing if one or more neighbors were experiencing conflict at period $t-1$. β is the parameter vector on the variables capturing observed (regional) heterogeneity. The main interest of this section is on how to obtain consistent estimates of the spillover parameter ρ in the presence of the unobserved effects α_i.

In absence of heterogeneity in the country specific effects there would be no problem in estimating (4) without paying explicit attention to the α_i's. The α_i's are effectively the same for each country (i.e., $\alpha_i = \alpha_j \; \forall i, j$) and sufficiently allowed for by including a constant in the regression. In any other case –where unobserved heterogeneity were present– it is no longer possible to simply assume that the country specific effects are uncorrelated with the explanatory variables of interest. This would normally leave the parameter estimates of the other included variables unchanged. Any heterogeneity in the α_i's produces a correlation with the spatial lag and therefore inconsistent regression outcomes. There is a clear analogy with dynamic panel data models for example, where fixed (or random) effects are *positively* correlated with the lagged dependent variable by construction. However, in models that include a *spatial* lag, the direction of these correlations cannot –in general– be predicted ex ante. In section 4.1 we show that the size and direction of these correlations depend on the relative importance of *within* and *between* region heterogeneity in the α_i's.

Ex post however, we can easily verify the statistical importance of the unobserved effects by

---

[4] We include a lagged neighboring conflict variable to be able to abstract from the need to use complicated spatial econometric techniques, see Beck and Gleditsch (2006), p.40 for more on this. In order for this to work we do have to assume intertemporal independent errors.

simply allowing for them in our conflict regressions in addition to the spillover variable and a vector of controls [see section (4.2) for empirical results]. In the empirical analysis we impose different assumptions on the unobserved effects, by restricting them to be country –, world region –, continent –, or world specific (the latter would effectively be captured by including just a constant in the regression). The use of world-region, continent or country fixed effects is in itself not uncommon in the civil conflict literature (see e.g. Fearon and Laitin (2003)). In fact, (several) region effects are also found to be among the robust explanations of civil war onset (Hegre and Sambanis, 2006). Hegre and Sambanis (2006) conclude that '... region effects have not been adequately explored in the civil war literature.' The significance of included region effects, simply indicates that 'civil war occurs more frequently in some parts of the world than others' [also pointed out by Hegre, Ellingsen, Gates, and Gleditsch (2001) on p.42 of their paper]. Given that they are included to control for unobservables, they do not provide direct insights into what these unobservables are. Yet, they are essential to make sure that one does not wrongly ascribe too much importance to one of the additional regressors. And most importantly for our purpose, they do *not* represent conflict diffusion or contagion and prevent the researcher from making wrong inference regarding the importance of conflict spillovers.

Where Gleditsch and Ward (2008) argue that 'models with region dummies... provide an alternative tot the spatially lagged y model (i.e., a neighboring conflict variable[5])', we argue that country and/or region dummies should be allowed for *in addition* to the spatial lag. Country or regional dummy variables and conflict spillover variables have very different characteristics that allow them to be included simultaneously in empirical models of civil conflict. The fact that country or region dummy variables are fixed within regions or countries ánd over time clearly sets them apart from a neighboring conflict variable, that is both varying within regions and over time. These differences, and the fact that the typical data set used in empirical studies of civil war spans many countries over many years, perfectly allows us to model both conflict spillovers and fixed effects in empirical models of civil war. This empirical strategy subsequently allows for *testing* rather than *hypothesizing about* the relevance of the two *different* concepts. We show in the empirical section that indeed both region and country dummies as well as spatial lags are often significant at the same time.

---

[5] The text in parenthesis is added by us and does not appear in Gleditsch and Ward (2008)

4.1 General and naive specifications of regions

Explicitly allowing for the α_i's at the country level by including country dummy variables in the onset regression typically works well, and produces consistent estimates for the spillover parameter ρ. In this section we mathematically derive the systematic biases that potentially occur in conflict spillover models that *do not* control for the α_i's at the country level. In principle there may be particular geographical patterns in the α_i's that allow us to control for *region* effects rather than for *country* effects, without disrupting the consistency of the spillover estimator.[6] It is well known that allowing for country fixed effects is a rather crude way of using data as it inactivates a lot of potentially interesting variation (the parameters of interest are only identified on the basis of within-country variation over time). In principle therefore one would be interested in economizing on degrees of freedom by controlling for unobserved heterogeneity at a higher geographical scale instead (using e.g. continent or world-region dummies). What are the consequences of using geographically broader definitions of unobserved heterogeneity in the empirical analysis of conflict? More specifically: under what conditions does this introduce a bias in the estimated conflict spillover parameter?

We will show that allowing for unobserved effects at the region level is feasible if countries within these regions are sufficiently homogenous with respect to their individual country specific unobserved effects α_i. If we are able to define regions that are composed of countries with similar α_i's, we may just as well account for this unobserved heterogeneity by including region dummies [as done by e.g. Fearon and Laitin (2003) and Hegre and Sambanis (2006)] or even by including a simple constant (i.e., controlling for unobserved heterogeneity at the world level). A priori defining such regions is inherently difficult as the α_i's are unobserved. If however, there exists substantial heterogeneity with respect to the α_i's within regions (i.e., if there is important *within* region heterogeneity) we are introducing biases by including region dummies. In that case, including e.g. continent or world region dummies may capture some of the unobserved heterogeneity, but the empirical approach is yet too crude to be useful. In fact we show that controlling for regional heterogeneity using an incorrect region specification (so that the countries within the specified regions exhibit substantial heterogeneity in the unobserved effects) may introduce much stronger biases in the conflict spillover parameter than when not controlling for it at all.

To show what happens in this type of models we derive explicit mathematical expressions for

---

[6] Note that including country fixed effects allows for any arbitrary regional classification at a higher geographical scale.

the biases that can arise from estimating a simple stripped down linear probability onset model. The expressions we derive are transparent and highly intuitive. We have first excluded the $\mathbf{x}_{it}$ vector from the analysis. The influence of control variables may alter the size and the direction of the biases we identify in this section. However, the underlying mechanisms we propose persist and are used to understand and interpret the outcomes of the empirical section. Second, we derive formulas for the biases in a linear probability framework. Similar types of biases should arise in nonlinear probit/logit type models, yet deriving explicit formulas for these biases is much harder (and perhaps even impossible).[7] Within this (stylized) linear probability framework we are able to derive formulas for biases that can be easily related to real-world observations.

For the sake of argument we use three different region (r) classifications at an ever more detailed geographical scale: the world, a region (comprising more than one country, e.g. a continent or subcontinent), and a country. Next, suppose that the 'true model' is the model that allows for unobserved heterogeneity at the country level:

$$\begin{aligned} P\left(c_{it} = 1 \middle| c_{i,t-1} = 0, n_{i,t-1}, \mathbf{x}_{it}, \alpha_i\right) &= E\left[c_{it} \middle| c_{i,t-1} = 0, n_{i,t-1}, \mathbf{x}_{it}, \alpha_i\right] \\ &= \rho n_{i,t-1} + \alpha_i \;\text{ for } c_{it-1} = 0 \qquad (5) \end{aligned}$$

This structural model translates into the following regression model:

$$c_{it} = \rho n_{i,t-1} + \alpha_i + [\varepsilon_{it}] \;\text{ for } c_{it-1} = 0 \qquad (6)$$

The parameters of this model may be estimated by regressing c_{it} on the time lagged spillover variable n_{it-1} and a full set of country dummies using OLS (with robust s.e.'s switched on). However, we do need a data set that exhibits a sufficiently large T dimension for consistency [see Appendix B for derivations]. We assume that our data set satisfies this requirement as we are studying civil war data from 1945 up to 2000 [see Nickell (1981) for an analysis about biases in dynamic fixed effects models exhibiting a small T dimension]. The error term ε_{it} is uncorrelated with the regressors by assumption (5), such that estimating equation (6) with OLS applies. We call the OLS estimator for ρ that includes country dummies the fixed effects (FE) estimator and denote it by $\hat{\rho}^{FE}$.

---

[7] In the empirical section we are estimating probit models. All results can be reproduced using linear probability models without obtaining large differences. We take this as evidence that the biases we derive in this section are approximating the biases we get in non-linear models sufficiently well. The reason for showing the probit results is that we ran into problems with off-the-chart predictions in linear probability models for the fixed effects model. Linear probability models tend to produce similar outcomes as probit or logit type models, unless linear probability models are predicting probabilities outside the $0-1$ interval.

When analyzing the consistency of our estimators it is inconvenient to deal with outside conditions (i.e., here the onset condition $c_{it-1} = 0$ that determines which observations should be used in onset regressions). To do so, we first rewrite the simple onset model represented by equation (6). We can simply incorporate the onset selection criterion in the regression model by pre-multiplying all regressors by the following indicator $1\left(c_{it}\middle|c_{it-1} = 0\right)$ [Miguel, Satyanath, and Sergenti (2004) use a similar approach by defining onsets]. Using this transformation equation (6) can be written as follows:

$$\dot{c}_{it} = \rho\dot{n}_{it-1} + \dot{\alpha}_i + [\dot{\varepsilon}_{it}] \tag{7}$$

where $\dot{c}_{it} = c_{it} \cdot 1\left(c_{it}\middle|c_{it-1} = 0\right)$, $\dot{n}_{it-1} = n_{it-1} \cdot 1\left(c_{it}\middle|c_{it-1} = 0\right)$. Performing OLS on the selected *onset* sample [i.e., using regression model (1)] is equivalent to performing OLS on the full sample after transforming the variables used in the estimation in the manner described above.

*Bias in spillover models: including **only a constant** to capture the α_i's:* Equation (7) can be rewritten as follows:

$$\dot{c}_{it} = \rho\dot{n}_{i,t-1} + \gamma + [(\dot{\alpha}_i - \gamma) + \dot{\varepsilon}_{it}] \tag{8}$$

The error term between brackets now consist of the difference between the i specific country effect and the constant (i.e., the world wide effect) in addition to the original error term ε_{it}. The OLS estimate for ρ derived from performing OLS on equation (8) is referred to as the pooled OLS estimator $\hat{\rho}^{POLS}$.

To derive a mathematical expression of the potential biases (or inconsistencies[8]) when estimating (8), we calculate the probability limit ($p\lim$) of the estimated conflict spillover parameter using pooled OLS, $\hat{\rho}^{POLS}$ [see Appendix B for a complete derivation]:

$$p\lim\left[\hat{\rho}^{OLS}\right] = \rho + \overbrace{p\lim\left[\frac{\sum_{i,t}\left[\left(\dot{n}_{it-1} - \bar{\dot{n}}_r\right)\left(\dot{\alpha}_i - \bar{\dot{\alpha}}_r\right) + \left(\bar{\dot{n}}_r - \bar{\dot{n}}\right)\left(\bar{\dot{\alpha}}_r - \bar{\dot{\alpha}}\right)\right]}{\sum_{i,t}\left[\left(\dot{n}_{it-1} - \bar{\dot{n}}_r\right)^2 + \left(\bar{\dot{n}}_r - \bar{\dot{n}}\right)^2\right]}\right]}^{\text{bias of the pooled OLS estimator}} \tag{9}$$

where $\bar{\dot{n}}_r = \frac{1}{N_r^o}\sum_{i\in r}\sum_t \dot{n}_{it-1}$ is the average of the lagged spillover variable $\dot{n}_{it-1}$ over time *within the particular region r that country i belongs to.* $\bar{\dot{n}} = \frac{1}{N^o}\sum_i\sum_t \dot{n}_{it-1}$ is the average of the time lagged spillover parameter over time and *across all countries*. Similarly, $\bar{\dot{\alpha}}_r = \frac{1}{N_r^o}\sum_{i\in r}\sum_t \dot{\alpha}_i$ is the average

---

[8] We are using the term *bias* and *inconsistency* interchangeably. Note that in onset models we never obtain true unbiasedness as these models are inherently dynamic.

country effect $\dot{\alpha}_i$ *within the particular region r that country i belongs to* and $\overline{\dot{\alpha}} = \frac{1}{N^o}\sum_i \sum_t \dot{\alpha}_i$ is the average country effect *across all countries of the world*. N^o is the number of observations in the onset sample and N_r^o the number of observations in the onset sample within region r, or more formally: $N^o = \sum_{it} (1 - c_{it-1})$ and $N_r^o = \sum_{i \in r} \sum_t (1 - c_{it-1})$

The estimator $\hat{\rho}^{OLS}$ is said to be a consistent estimator for ρ if $p\lim\left[\hat{\rho}^{POLS}\right] = \rho$. Equation (9) immediately shows that including only a simple constant will provide a consistent estimator of the conflict spillover parameter if all α_i's are the same across the world (i.e., $\alpha_i = \alpha_j \ \forall i, j$, such that $\dot{\alpha}_i - \overline{\dot{\alpha}}_r = 0 \ \forall i$ and $\overline{\dot{\alpha}}_r - \overline{\dot{\alpha}} = 0 \ \forall i$). Or to put it differently: $\alpha_i = \gamma \ \ \forall i$, i.e. in the absence of any heterogeneity in the α_i's. If this condition is not satisfied, the probability limit of the OLS estimator for $\hat{\rho}^{OLS}$ is equal to the true population spillover parameter ρ *plus* an additional bias term, that is generally non-zero.

Bias in spillover models: including ***region dummies*** *to capture the α_i's:* We now include region specific dummy variables in the regression equation to allow for unobserved heterogeneity at the region level. Next, we again rewrite equation (8):

$$\dot{c}_{it} = \rho \dot{n}_{i,t-1} + \gamma^r + \left[(\dot{\alpha}_i - \gamma^r) + \dot{\varepsilon}_{it}\right] \tag{10}$$

γ^r is a region specific constant that is the same for all countries within a particular region r. The error term now consists of the difference between the country effect $\dot{\alpha}_i$ and the region effect γ^r in addition to the original error term $\dot{\varepsilon}_{it}$. We call the OLS estimator for ρ that is associated with the above defined regression model the REGION OLS estimator, $\hat{\rho}^R$.

Similarly to the case of including only a constant we can derive a formula for the potential bias that is associated with regression model (10) by deriving the probability limit of $\hat{\rho}^R$:

$$p\lim\left[\hat{\rho}^R\right] = \rho + \overbrace{p\lim\left[\frac{\sum_{i,t}\left[\left(\dot{n}_{it-1} - \overline{\dot{n}}_r\right)\left(\dot{\alpha}_i - \overline{\dot{\alpha}}_r\right)\right]}{\sum_{i,t}\left[\left(\dot{n}_{it-1} - \overline{\dot{n}}_r\right)^2\right]}\right]}^{\text{bias of the OLS estimator with region effects}} \tag{11}$$

This expression immediately shows that we get a consistent estimator of the true conflict spillover parameter if $\alpha_i = \alpha_j \ \forall i, j \in r$ for all regions, i.e., if there exists no heterogeneity in the α_i's between countries *within* the same region. Note that this requirement is less strict than for the model that includes only a constant in additional to the spillover variable. However, equation

(11) shows that even if we include region dummies we may find biases if the individual effects α_i differ from their region specific averages in the onset sample $\overline{\alpha}_r$. Note also that –more generally– this bias pops up if we allow for unobserved heterogeneity at the region level using an incorrect definition of regions. An incorrect region specification will inevitably leave some within-region heterogeneity between the countries that are grouped together in incorrectly specified regions.[9]

The formulas for the biases that can emerge from estimating model (8) and model (10) respectively provide us with useful insights into the direction of these biases. The expressions suggest the importance of concepts like *within region* and *between region* heterogeneity in the α_i's. Or more specifically *within region* and *between region* correlation of the spillover variable and the α_i's. Equation (11) shows that the within region covariance between the spillover variable and the individual effects drives the direction of the bias of the REGION OLS estimator [see the numerator of equation (11)]. Equation (9) on the other hand shows that the direction of the bias in the pooled OLS estimator depends on a weighted sum of the within *and* between region covariance of the spillover variable and the individual effect [see the numerator of equation (9)]. These formulas turn out to be useful when interpreting the results of our regression in the next section. In particular they show that the direction of the bias produced by heterogeneity in the α_i's depends on the nature and the degree of spatial clustering of the α_i's.

To illustrate this we consider the following two examples. The first example deals with the case where the α_i's are randomly allocated around the globe. On average, war-prone (relatively high α_i) countries and peaceful (relatively low α_i) countries are neighbors. Obviously, the high α_i countries are experiencing more conflict than their peaceful neighbors on average. We could say therefore that a war-prone country's *neighbor*, is *relatively* peaceful by construction. Random allocation of α_i's therefore produces a negative correlation of the α_i's and n_{it-1}'s within regions. The proposed mechanism produces a downward bias on the estimated spillover parameter, such that we may find negative biases in both the pooled OLS and in the REGIONS OLS estimations.

In the second example we move away from the random allocation of the α_i's by considering the case where the α_i's are clustered in particular regions around the world. Assume that there are particular regions of the world that are struggling with notoriously under-performing regimes causing the repeated outbreak of civil war. If we do not manage to control for the variable 'regime performance' in our onset regressions, we are effectively missing out on an unobserved effect

---

[9] Hereby putting considerable force to the argument of authors who have stressed that the introduction of some set of regional dummies is 'arbitrary' and may result in estimates that can be as wrong as when not controlling for these 'arbitrary' regions.

that is similar for all countries within those regions. One could call this a *neighborhood* effect α_r (i.e., $\alpha_i = \alpha_j = \alpha_r \ \forall i, j \in r$). Because countries within those regions are often neighbors however, neglecting this unobserved neighborhood effect tends to spuriously ascribe a role to spillovers. Within those regions, countries keep ending up in war *not* because conflict is spilling over from neighbors, but simply because countries are subject to similar under-performing regimes. So, if there is a tendency that similar α_i's are clustered in particular regions of the world, the above described mechanism tends to produce an upward bias of the spillover parameter in the pooled OLS estimator. This bias however, may be eliminated by including region dummies.

At this point we note that there are other potential mechanisms producing biases in pooled OLS and REGION OLS estimators. Conflict onset models are inherently dynamic models and hence they require error terms to be uncorrelated over time for consistency, an assumption that is clearly violated if there exists heterogeneity in the α_i's [see e.g., Griliches (1961) and Achen (2000)]. On the other side of the same coin, when estimating an onset model we are basically analyzing a specific selected sample (i.e. those countries with peace at period $t-1$). The selection on peace at period $t-1$ is typically an endogenous selection: where the α_i affects the dependent variable (conflict at period t) it also affects the selection criterion. However, accounting for the heterogeneity in the α_i's solves both the issues specifically related to spatial lags (discussed above) as well as the selection problem. Yet, we think that the issue of selection is an important area for future research.

4.2 The impact of controlling for unobserved regional heterogeneity

Having established the potential impact of inadequately controlling for unobserved heterogeneity at the country level on spillover estimators, we now turn to the estimation of fully-fledged onset models. Here we also control for observed regional heterogeneity, by including several important country-specific variables (*observed* regional heterogeneity) that have been shown to have a significant impact on a country's probability to fall into a state of civil conflict.[10] We specify the following conditional onset model for all our estimations:

$$P\left(c_{it} = 1 \middle| c_{i,t-1} = 0, n_{i,t-1}, \mathbf{x}_{it}, \alpha_i\right) = \Phi\left(\rho n_{i,t-1} + \mathbf{x}_{it}\beta + \alpha_i\right) \quad \text{for } c_{i,t-1} = 0 \tag{12}$$

---

[10] Results without these (observed) controls are available upon request.

Table 1: ONSET Regressions - Full sample

Variables	+ no unobs.het.	+ continents	+ world regions	+countries
	(1)	(2)	(3)	(4)
Neighboring	0.007	0.004	0.002	0.035
conflict $t-1$	(0.003)**	(0.003)	(0.004)	(0.013)***
ln gdp per	-0.008	-0.005	-0.006	-0.025
capita	(0.002)***	(0.002)***	(0.002)***	(0.015)*
ln population	0.003	0.003	0.005	0.025
	(0.001)***	(0.001)***	(0.001)***	(0.017)
Δ ln gdp cap.	-0.051	-0.045	-0.059	-0.124
	(0.019)***	(0.018)**	(0.021)***	(0.053)**
ethnic heterogeneity	0.000	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.001)
oil	0.007	0.003	0.001	-0.006
	(0.007)	(0.005)	(0.005)	(0.023)
polity IV	0.000	0.000	0.000	0.001
	(0.000)	(0.000)	(0.000)	(0.001)
peace years				
≥ 2 years	-0.004	-0.003	-0.005	-0.008
	(0.009)	(0.008)	(0.010)	(0.024)
≥ 3 years	-0.001	-0.002	-0.001	0.009
	(0.007)	(0.007)	(0.008)	(0.015)
≥ 5 years	-0.011	-0.010	-0.011	0.005
nr. obs.	5235	5127	4414	1927
F-test Peace Dur.	p =0.01	p =0.01	p =0.02	p =0.80
log likelihood	-439.54	-430.41	-417.45	-337.72
LR-stat unobs.het.		18.3	25.9	159.5
5% critical value		11.1 $\chi^2(5)$	22.4 $\chi^2(13)$	137.7 $\chi^2(112)$

Notes: Robust standard errors between parentheses. p-values for a joint test on the significance of the variables capturing peace duration are reported. ***, ** and * denote significance at the 1%, 5% and 10% level respectively.

where Φ denotes a normal CDF, so that we use standard probit maximum likelihood techniques to estimate the parameters of the model.[11]

Table 1 shows the results of estimating our baseline model using the full sample[12] both with and without controlling for three differently defined forms of unobserved regional heterogeneity. The first column shows the (standard) results that includes only a constant. The subsequent columns control for unobserved regional heterogeneity at an ever narrower geographical scale going from continents to world-regions [see Appendix C for definitions] to individual countries.

Column (1) reports the standard results from the literature : the neighboring conflict variable is significant and positive (but quite small), indicating the significant presence of conflict spillovers [see e.g., Sambanis (2001), Hegre and Sambanis (2006), Gleditsch (2007) and Buhaug

---

[11] None of the results we present in this paper depend on assuming a normal instead of a logistic distribution (i.e. running logit instead of probit.). Results are available upon request.

[12] Given the available time span of the data we do not worry about the initial conditions, that may influence the results when using a data set that spans only few years.

and Gleditsch (2008)]. Column (2) however shows that once we allow for unobserved time-invariant heterogeneity across continents this result disappears immediately: the parameter on the neighboring conflict variable becomes much smaller and turns insignificant. Moreover the model including continent dummies is statistically preferred over the model that does not take account of unobserved heterogeneity [see the results of a LR-test in the Table[13]]. We obtain a similar result when narrowing down the geographical scale of the region dummies to the level of world-regions. Again the parameter on the neighboring conflict variable is insignificant and is even smaller than in Column (2). The model including world-region dummies is statistically preferred over the model including continent dummies (and thus implicitly also over the model that does not control for unobserved heterogeneity) indicating significant additional heterogeneity within continents [again see the results of a LR-test in the Table[14]].

Under the assumption that the REGION OLS estimators are consistent we would conclude that the net bias of the pooled OLS estimators reported in column (1) of table 1 is positive: using simple pooled OLS overestimates the relevance of the alleged international dimension to civil conflict. Note that REGION OLS estimators are consistent if heterogeneity of the α_i's within regions is absent [i.e., $\alpha_i = \alpha_j \ \forall \ i, j \in r$. see equation (11)]. An upward bias on the pooled OLS estimator suggests that neighbors at war and the unobserved effects are positively correlated, which makes intuitive sense. For example, if in some regions countries are endowed with large α_i's (e.g., Africa and Asia) we would see a lot of war in those regions. As a result we would also see a lot of countries with their neighbors at war. On the contrary, we would hardly see wars in regions where countries are endowed with small α_i's (e.g., Europe and North America) and consequently we would see hardly any countries with their neighbors at war either. So if we were able to specify regions that are sufficiently homogenous with respect to their unobserved effects we would expect to find larger spillover effects in pooled OLS (because of the positive bias) than in OLS that includes region dummies.

Column (4) however shows that when we include country dummies in the regression – that is controlling for the most geographically narrow definition of unobserved heterogeneity– the results are quite different. Conflict spillovers 'reappear' as being significant positive contributors to the probability that civil war breaks out. [And note the LR-test indicates that this model is preferred

---

[13] The LR statistic testing the equality of model (1) and (2) is $2 \times (-430.41 - -439.54) = 18.26$, which is distributed $\chi^2(5)$ as we loose 5 degrees of freedom by including the continent dummies (4 extra parameters are estimated and 1 because North-America predicts the r.h.s. perfectly).

[14] The LR statistic testing the equality of model (2) and (3) is $2\times(-417.45 - -430.41) = 25.92$, which is distributed $\chi^2(13)$ as we loose 5 additional degrees of freedom by including the world region dummies (14 world region dummies are estimated and 4 world regions predict the r.h.s. perfectly, minus the 5 that we had already lost by including continent dummies).

over the models controlling for unobserved regional heterogeneity at a higher geographical scale[15].] Moreover, the spillover effects we find in column (4) are much stronger than the baseline results of column (1). Where column (1) suggests that having at least one neighbor at war increases the onset probability of civil war by less than one percent, the results in column (4) suggest that these effects are around 3.5 percent. The column (4) results emphasize the importance of dealing with unobserved heterogeneity at the country level in conflict onset models. Not doing so –see columns (1), (2), and (3)– results in underestimating the true spillover effect by far.

The combined evidence presented in table 1 shows that allowing for unobserved (fixed) heterogeneity at the region level does not help much for estimating the correct spillover effects. In fact, it seems to make things worse. Both the pooled OLS estimator and the REGION OLS estimator greatly underestimate the spillover effects. Under the assumption that the fixed effects estimator is consistent the net bias terms in both equation (9) and (11) are negative. On top of that, our evidence suggests that the bias of the REGION OLS estimator is even stronger than the bias of the pooled OLS estimator. Using the results of table (1) and the formulas (9) and (11) we suggestively conclude that in the sample of onset observations there exists important *negative* correlation between having a neighbor at war and the unobserved effects *within* regions.

Overall the results clearly show that allowing for various types of unobserved heterogeneity matters a great deal for inference regarding the relevance of conflict spillovers. Allowing for naive definitions of unobserved heterogeneity in column (2) and (3) mitigates the correlations we find in column (1). On the basis of the first three columns we would, mistakenly, draw the conclusion that allowing for a set of region dummies removes the spillover effect. However, when controlling for the time-invariant country effects, conflict spillovers reappear. Our results therefore report the interesting result that controlling for unobserved heterogeneity at the country level identifies a more intense effect of conflict spillovers relative to what we find when not controlling for unobserved heterorgeneity in column (1), or when inadequately controlling for unobserved heterogeneity in column (2) and (3).

---

[15] The LR statistic testing the equality of model (3) and (4) is $2 \times (-337.7 - -417.45) = 159.5$, which is distributed $\chi^2(112)$ as we loose 112 additional degrees of freedom by including the country dummies (61 country dummies are estimated and 91 country dummies predict the r.h.s. perfectly, minus 18 from model (2) and (3) minus an additional 22 countries that we already lost by including world region dummies).

5 Regional heterogeneity in the magnitude of conflict spillovers

So far we have been talking about controlling for heterogeneity in unobservables by allowing for unobserved effects at the continent, world region or country level. We assume however that the spillover effect itself is the same across continents. Figure 2 however, already suggests quite convincingly that there may be differences in spillover effects across continents. More specifically, figure 2 suggests that spillovers are mainly confined to Africa and Asia.

In this section we therefore extend the analysis by allowing spillover effects to differ between continents. It may be that within certain regions, countries' internal affairs are more heavily intertwined than in other regions. Therefore we might see that certain regions are more prone to experiencing conflict spillovers than others. In this section we present estimates that allow us to study the differences in the relevance of conflict spillovers for three of the most conflict ridden continents in the world: Africa, Asia, and Latin America (we did not find spillovers in the other continents Europe, North America or Oceania[16]). To do so, we also allow for heterogeneity in the conflict spillover parameter ρ in addition to allowing for heterogeneity in the α_i's.

Formally, we estimate (4) using data for each of the three continents only, i.e.

$$P\left(c_{it} = 1 \middle| c_{i,t-1} = 0, continent_i = C, n_{i,t-1}, \mathbf{x}_{it}, \alpha_i\right) = \Phi\left(\rho^C n_{i,t-1} + \mathbf{x}_{it}\beta^C + \alpha_i\right) \quad \text{for } c_{i,t-1} = 0 \qquad (13)$$

where C denotes either Africa, Asia or Latin America. Contrary to allowing for the α_i's by including country dummies, estimating separate models for each continent verifies whether the relevance of conflict spillovers *themselves* exhibits regional heterogeneity. That is, the estimated spillover parameter ρ is allowed to differ in size and (possibly) significance across the three different continents. Note that another modeling strategy would be to interact our neighboring conflict variable with each of the three continent dummies and estimate the model on the whole sample. Results of doing this are very similar to the ones we present in section 5.1 (results are available upon request).

5.1 The impact of allowing for regional heterogeneity in the magnitude of conflict spillovers

Table 2 shows the results of estimating our main onset model (13) for Africa, Latin America and Asia separately. First when only controlling for observed regional heterogeneity (including the

---

[16] Regression results for the other continents are available on request

$\mathbf{x}_{it}$-variables), and second by also controlling for unobserved heterogeneity by including world region or country dummies respectively.

Table 2: ONSET BY CONTINENT - Africa, Asia and Latin America

Variables	**Africa**	+ WR	+ FE	**Asia**	+WR	+FE	**Lat. America**	+WR	+FE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Neighboring	0.024	0.019	0.039	0.001	-0.003	0.026	-0.005	-0.007	0.031
conflict $t-1$	(0.008)***	(0.008)**	(0.019)**	(0.009)	(0.008)	(0.017)	(0.010)	(0.009)	(0.026)
ln gdp per	-0.007	-0.003	0.025	-0.014	-0.011	-0.028	-0.004	-0.005	0.003
capita	(0.006)	(0.006)	(0.036)	(0.005)***	(0.005)**	(0.022)	(0.008)	(0.008)	(0.035)
ln population	0.008	0.006	0.059	0.002	0.007	0.034	0.003	0.003	-0.018
	(0.003)**	(0.003)*	(0.025)**	(0.002)	(0.003)**	(0.037)	(0.003)	(0.003)	(0.023)
Δ ln gdp cap.	-0.050	-0.053	-0.131	-0.075	-0.072	-0.116	-0.136	-0.141	-0.325
	(0.053)	(0.046)	(0.104)	(0.043)*	(0.039)*	(0.063)*	(0.064)**	(0.064)**	(0.131)**
ethnic heterogeneity	-0.000	0.000	0.001	0.000	-0.000	-0.001	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.001)**	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
oil	-0.006	-0.005	-0.025	0.020	0.007	-0.009	-	-	-
	(0.009)	(0.009)	(0.021)	(0.019)	(0.015)	(0.036)	-	-	-
polity IV	0.001	0.001	0.004	0.001	0.001	0.003	-0.000	-0.000	-0.000
	(0.001)	(0.001)**	(0.002)***	(0.001)*	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
peace years									
≥ 2 years	0.000	-0.000	-0.009	-0.045	-0.047	-0.058	-0.010	-0.009	0.002
	(0.016)	(0.014)	(0.034)	(0.050)	(0.049)	(0.071)	(0.036)	(0.034)	(0.031)
≥ 3 years	-0.015	-0.013	-0.009	0.013	0.013	0.029	0.001	0.001	-0.002
	(0.024)	(0.022)	(0.031)	(0.013)	(0.010)	(0.015)*	(0.019)	(0.019)	(0.033)
≥ 5 years	-0.008	-0.005	0.023	-0.014	-0.011	-0.005	-0.019	-0.018	-0.006
	(0.013)	(0.011)	(0.014)	(0.018)	(0.015)	(0.020)	(0.023)	(0.021)	(0.021)
nr. obs.	1429	1429	643	1249	1249	731	889	889	501
F-test Peace Dur.	p =0.18	p =0.31	p =0.50	p =0.32	p =0.33	p =0.56	p =0.47	p =0.49	p =0.98
log likelihood	-157.95	-154.78	-117.82	-160.60	-156.30	-134.13	-76.75	-76.46	-69.73

Notes: Robust standard errors between parentheses. p-values for a joint test on the significance of the variables capturing peace duration are reported. ***, ** and * denote significance at the 1%, 5% and 10% level respectively.

Table 2 clearly indicates that allowing for regional heterogeneity in the relevance of conflict spillovers is important. We only find evidence for conflict spillovers on the African continent. In Asia and Latin America conflict spillovers do not seem to be a reality. Even when controlling for country-specific effects the neighboring conflict variable remains insignificant. In Africa on the other hand, the estimated conflict spillover parameter is always positive and significant. Moreover, the estimated spillover parameter is larger as compared with the estimates from table 1. Combining this evidence with the insignificant conflict spillover parameters from Latin America and Asia, strongly suggests that the results using the world wide sample in table 1 are almost entirely driven by Africa.[17]

Within Africa we do not find many differences across specifications. Yet, the estimate of the spillover effect in the FE model is again the largest. From the LR test we conclude that model (2) is not statistically preferred over model (1), but we do find that model (3) is statistically preferred over model (2) (the test statistic is χ^2 (40) = 73.92, which is greater than the 5% critical value 55.73). These results indicate that also in case of Africa world regions are again too broadly defined to filter out all the unobserved heterogeneity in the α_i's. Yet, the biases that this causes seems less of an issue as we still find significant and sizeable spillovers in Africa in column (1) and (2).

These findings most importantly identify that conflict spillovers are not as omnipresent as suggested by previous (world-wide) studies (Sambanis, 2001; Buhaug and Gleditsch, 2008; Gleditsch, 2007; Hegre and Sambanis, 2006). The fact that spillovers only appear to be a reality in Africa is of great interest for future research. Accurately localizing conflict spillovers greatly helps researchers in identifying explanations of why and how civil conflict spills over across international boundaries. Several studies have already tried to open the 'black box' of conflict spillovers, by looking explicitly at conditions (e.g. large refugee flows, ethnic boundaries spanning official boundaries, sharing a long border, or cross-border rebel activity) that make conflict spillovers across countries more likely [see Salehyan and Gleditsch (2006); Gleditsch (2007); Buhaug and Gleditsch (2008)]. These studies all have a global focus. Our evidence points out that we should focus on the African continent if we want to improve our understanding into the (pre)conditions under which conflict is likely to spill over from one country to another. Given that conflict spillovers in the rest of the world are not born out clearly by the data, a focus on Africa, and what sets it apart from the rest of the world[18] will in our view be much more successful in determining the underlying processes

---

[17] Also when only excluding Africa from the world sample, no evidence of conflict spillovers shows up any more. Results available upon request.

[18] We strongly believe that the artificial border creation during the colonial era, is one of the likely candidates.

involved in the international spread of civil conflict.

6 A simple robustness check using the number of neighbors at war

So far we have identified the (ir)relevance of conflict spillovers in different parts of the world by looking at the estimated parameter on a dummy variable indicating whether or not at least one neighboring country experienced civil war a year before. Previous papers have also looked at whether or not the chances on a conflict spillover increase with the number of neighbors in civil war [see e.g. Hegre and Sambanis (2006), Fearon and Laitin (2003) or Gleditsch (2007)]. Generally these papers simply include a variable measuring the total number of neighbors in civil war one year before as opposed to a dummy variable indicating whether at least one neighbor is in civil war one year before, to verify this. They usually find that the results are either quite similar to that when including the simple 'at least one neighbor in civil war' dummy [Fearon and Laitin (2003) and Hegre and Sambanis (2006)] or find unsatisfactory results when including the total number of neighbors in civil war (Gleditsch, 2007).

Simply including the total number of neighbors at war one year before implicitly puts a potentially artificial linear structure on the effect of having additional neighbors in civil war. We therefore use the information on the number of neighbors in civil war differently. We continue to include the standard 'at least one neighbor in civil war' dummy in our onset regressions but we augment the model by also including a 'having at least two'- or 'having at least three'-neighbors in civil war dummies.[19] The estimated coefficients on these dummy variables then tell us whether or not having more than one neighbor at war adds significantly to the chances of a conflict spillover (without pre-imposing the linear structure).

Besides providing potentially interesting information on the increased likelihood of conflict spillovers when multiple neighbors are at war, the coefficients on each of the 'more than one neighbors in civil war' dummies also provide a very simple, but quite appealing robustness check on the original results. If conflict spillovers are a reality, the 'more than one neighbor in civil war' dummies should in our view be either insignificant (which would suggest that having one neighbor in civil war is already bad enough such that an additional neighbor in civil war does

---

[19] There are only very few countries in the sample with four or more or five or more neighbors at war (occurs only 36 times or 10 times respectively). Therefore we only include a dummies for having one-or-more (occurs 849 times), two-ore-more (occurs 314 times), and three-or-more (occurs 126 times) neighbors at war.

Table 3: ONSET - ROBUSTNESS

VARIABLES	Full sample+FE	Africa+FE	Asia+FE
≥ 1 Neighbors	0.036	0.038	0.030
conflict $t-1$	(0.013)***	(0.020)* [p =0.055]	(0.016)* [p =0.051]
≥ 2 Neighbors in	-0.006	0.003	-0.030
conflict $t-1$	(0.014)	(0.021)	(0.015)**
≥ 3 Neighbors in	-0.001	0.001	-0.008
conflict $t-1$	(0.020)	(0.026)	(0.028)
ln gdp per	-0.025	0.025	-0.033
capita	(0.015)*	(0.037)	(0.021)
ln population	0.027	0.058	0.059
	(0.018)	(0.026)**	(0.038)
Δ ln gdp cap.	-0.125	-0.132	-0.122
	(0.052)**	(0.105)	(0.062)**
ethnic heterogeneity	0.000	0.001	-0.000
	(0.001)	(0.001)**	(0.000)
oil	-0.006	-0.025	-0.007
	(0.023)	(0.020)	(0.033)
polity IV	0.001	0.004	0.002
	(0.001)	(0.002)**	(0.002)
peace years			
≥ 2 years	-0.008	-0.010	-0.045
	(0.024)	(0.034)	(0.063)
≥ 3 years	0.009	-0.010	0.027
	(0.015)	(0.031)	(0.015)*
≥ 5 years	0.004	0.024	-0.008
	(0.011)	(0.014)*	(0.020)
nr. Obs.	1927	643	731
F-test Peace Dur.	p =0.82	p =0.45	p =0.63
F-test ≥ 2 neighbors at war	p =0.91	p =0.99	p =0.17
Log likelihood	-337.62	-117.81	-132.38

NOTES: Robust standard errors between parentheses. p-values for a joint test on the significance of the variables capturing peace duration are reported. ***, ** and * denote significance at the 1%, 5% and 10% level respectively.

not add significantly to a country's chances on a conflict spillover) or they should be significant but showing a consistent pattern, i.e. *more* neighbors at war should *increase* the probability of conflict spillovers. One message they should *definitely not* convey is that having more than one neighbor at war decreases the chances on a conflict spillover. Such a finding would be highly unrealistic, shedding considerable doubt on the overall conflict spillover story.

Table 3 shows the results of including the 'more than one neighbors in civil war' dummies to the model. We only show the results controlling for country specific effects.[20] The outcomes can be readily compared to the earlier results shown in column (4) of table 1 and in column (3) and (6) of table 2 respectively.

---

[20] Results including world-region dummies instead are very similar. They are available upon request.

The results show something quite interesting. When considering the whole sample, we find that having more than one neighbor in civil war does not increase a country's chances on a conflict onset compared to having only one neighbor in civil war. The 'two or more neighbors in civil war' dummies and the 'three or more neighbors in civil war' dummies are individually, and also jointly (see the result of the F-test) insignificant. The same insignificance of the 'more than one neighbors in civil war' dummies holds for the African subsample, but differs from the column (1) results by at least having the consistent positive sign (i.e., countries with more than one neighbor at war are more likely to experience an onset). However this effect is very small and far from significant.

The results for Asia are harder to interpret and strongly confirm the unlikeliness of important conflict spillovers in that particular part of the world. When including the 'more than one neighbor in civil war' dummies, we find a significantly positive spillover effect when having only one neighbor at war. However this effect is offset when having at least two neighbors at war: the coefficient on the 'two or more neighbors at war' variable is significantly negative so that when having two or more neighbors at war the evidence for conflict spillover completely vanishes. (We tested the null that a positive effect of spillovers from one or more neighbors at war is exactly offset when having two or more neighbors at war. $H_0 : \rho^{(\geq 1)} = -\rho^{(\geq 2)}$. p-value [0.812]). So if one were to believe these results, this would imply that conflict in Asia *only* spills over across international boundaries when only one of a country's neighbors experiences civil war, but stops spilling over when more than one of a country's neighbors is in civil war. We think that this result is quite hard to defend, and argue that it clearly supports the earlier found absence of conflict spillovers in Asia in table 2.

7 Conclusions

The widely accepted notion that civil war has the potential to spill over from one country to another needs some careful revision in our view. We argue that previous studies have overlooked the issue of unobserved regional heterogeneity. We improve on earlier studies on this topic by eliciting the importance of dealing with unobserved heterogeneity at the regional level when studying cross-border conflict spillovers. Section 4.1 derives and analyzes explicit mathematical expressions of the bias in the estimated conflict spillover parameter if we *do not or inadequately* control for unobserved heterogeneity at the country level. We show that neglecting time-invariant unobserved heterogeneity leads to estimates of the conflict spillover parameter that are typically biased downwards. We argue that the spatial distribution of the country specific effects is at least partly to blame for these findings. We show that important *between* region heterogeneity typically produces an upward bias on the conflict spillover parameter, whereas unobserved *within* region heterogeneity tends to produce a downward bias. We find significantly larger spillover effects when using the (consistent) fixed effects estimator than when using the pooled OLS estimator, or when we control for either continent or world-region effects. The results indicate that unobserved within-region heterogeneity is the dominating force in producing biases, conditional on our definitions of regions.

Also, when allowing for regional heterogeneity in the presence of conflict spillovers themselves, we find that spillovers are highly important in Africa, whereas we do not find evidence for conflict spillovers in other parts of the world. After controlling for time invariant country specific effects (fixed effects) we estimate that having at least one neighbor at war in Africa increases the probability of a civil war onset by about 4%. This is a substantial effect as the average African country has between three to five neighbors. These results seem to indicate that policymakers indeed must worry a great deal about the international dimension of the causes and consequences of civil war on the African continent.

We should be explicit however in saying that our results do not indicate that the international dimension does not matter for understanding the dynamics of civil war across space and time in continents other than Africa. Our findings indicate that *conflicts* do not seem to spill over across international borders outside of the African continent. Neighbors may still have a lot of influence in igniting or preventing civil war in their neighboring countries by mingling in other countries' internal affairs [see for example Gleditsch (2007), Harbom and Wallensteen (2005), Gleditsch and Beardsley (2004) or Moore and Shellman (2007)] through refugee flows, ethnic boundaries that cross official boundaries, providing safe refuge to rebel groups, or actively promoting peace talks between the fighting factions in their neighboring country.

Finally, we would like to note that our findings regarding the importance of dealing with (unobserved) heterogeneity at the region or country level are also of great interest to other discussions in political science that see a big role for spatial interaction between countries [e.g. the spread of democracy (Gleditsch and Ward, 2006), policy diffusion (Simmons and Elkins, 2004), or economic development (Murdoch and Sandler, 2002, 2004)]. It would be very interesting to see whether adequately controlling for (unobserved) heterogeneity also changes the conclusions regarding spillovers in these fields of political science.

A Localizing conflict spillovers

This Appendix how to derive (2) from the pooled OLS estimator for ρ in (1).

First, and as in section 4.1, we incorporate the typical onset-sample selection (i.e., $c_{it-1} = 0$) criterion directly into the regression equation (see section 4.1)

$$\dot{c}_{it} = \gamma + \rho \dot{n}_{it-1} + [\dot{u}_{it}] \tag{14}$$

Performing OLS on the selected *onset* sample (i.e., using regression model (1)) is equivalent to performing OLS on the full sample, but after transforming the data using the transformation defined in section 4.1.

Next, take the average of equation (14):

$$\bar{\dot{c}}_{it} = \gamma + \rho \bar{\dot{n}}_{it-1} + \left[\bar{\dot{u}}_{it}\right] \tag{15}$$

where $\bar{\dot{c}} = \frac{1}{N^o} \sum_{it} \dot{c}_{it}$, $\bar{\dot{n}} = \frac{1}{N^o} \sum_{it} \dot{n}_{it}$, $\bar{\dot{u}} = \frac{1}{N^o} \sum_{it} \dot{u}_{it}$. N^o is the number of observations in the onset sample: $N^o = \sum_{it} (1 - c_{it-1})$.

Subtracting equation (15) from equation (14) yields a new regression equation where γ drops out, that yields the same estimator for ρ as performing OLS on the baseline regression equation (14).

$$\dot{c}_{it} - \bar{\dot{c}} \quad = \quad \rho \left(\dot{n}_{it-1} - \bar{\dot{n}}\right) + \left[\dot{u}_{it} - \bar{\dot{u}}\right] \tag{16}$$

Using (16), the formula for the pooled OLS estimator becomes:

$$\begin{aligned} \hat{\rho}^{POLS} &= \frac{\sum_{it} \left(\dot{n}_{it-1} - \bar{\dot{n}}\right)\left(\dot{c}_{it} - \bar{\dot{c}}\right)}{\sum_{it} \left(\dot{n}_{it-1} - \bar{\dot{n}}\right)\left(\dot{n}_{it-1} - \bar{\dot{n}}\right)} && (17) \\ &= \frac{\sum_{it} \left(\dot{n}_{it-1}\dot{c}_{it} - \bar{\dot{n}}\dot{c}_{it} - \dot{n}_{it-1}\bar{\dot{c}} + \bar{\dot{n}}\bar{\dot{c}}\right)}{\sum_{it} \left((\dot{n}_{it-1})^2 - 2\bar{\dot{n}}\dot{n}_{it-1} + \left(\bar{\dot{n}}\right)^2\right)} && (18) \end{aligned}$$

Now to get at (2), we define:

$$\begin{aligned}
s_1 &= \#\text{onsets with at least one neighbor at war} \\
s_2 &= \#\text{onsets without any neighbor at war} \\
s_3 &= \#\text{observations in the onset sample with at least one neighbor at war} \\
s_4 &= \#\text{observations in the onset sample without any neighbor at war} \\
s_1 + s_2 &= \text{total \# of onsets} \\
s_3 + s_4 &= \text{total \# of observations in the onset sample}
\end{aligned} \tag{19}$$

Using these definitions we can rewrite the separate elements of the pooled OLS estimator (18):

$$\begin{aligned}
\sum_{it} \dot{n}_{it-1}\dot{c}_{it} &= s_1 \\
\sum_{it} \bar{\dot{n}}\dot{c}_{it} &= \frac{s_3}{s_3+s_4}(s_1+s_2) = \frac{s_1+s_2}{s_3+s_4}s_3 \\
\sum_{it} \dot{n}_{it-1}\bar{\dot{c}} &= \frac{s_1+s_2}{s_3+s_4}s_3 \\
\sum_{it} \overline{\dot{n}\dot{c}} &= \frac{s_3}{s_3+s_4}\frac{s_1+s_2}{s_3+s_4}(s_3+s_4) = \frac{s_1+s_2}{s_3+s_4}s_3 \\
\sum_{it} (\dot{n}_{it-1})^2 &= \sum_t \sum_i \dot{n}_{it-1} = s_3 \\
\sum_{it} \bar{\dot{n}}\dot{n}_{it-1} &= \frac{s_3}{s_3+s_4}s_3 \\
\sum_{it} \left(\bar{\dot{n}}\right)^2 &= \left(\frac{s_3}{s_3+s_4}\right)^2 (s_3+s_4) = \frac{s_3}{s_3+s_4}s_3
\end{aligned} \tag{20}$$

Plugging the s definitions into (18), and multiplying both the numerator and the denominator of $\hat{\rho}^{POLS}$ by $(s_3 + s_4)$:

$$\hat{\rho}^{POLS} = \frac{s_1 - \frac{s_3}{s_3+s_4}(s_1+s_2)}{s_3 - \frac{s_3}{s_3+s_4}s_3} = \frac{s_1}{s_3} - \frac{s_2}{s_4} \tag{21}$$

then shows that $\hat{\rho}^{POLS}$ can indeed be rewritten as (2) in the main text.

B Asymptotic behavior of different OLS estimators

In this appendix we are considering the asymptotic behavior (when T goes to infinity) of the three different estimators, discussed in section 4.1, where the true (data generating) model exhibit country specific region effects, establishing under what conditions each of the estimators provides a consistent estimate of the conflict spillover parameter. With consistency of an estimator $\hat{\rho}$ we mean that the estimator converges in probability to the population parameter ρ when T becomes infinitely large, or more formally: $p\lim_{T\to\infty}\hat{\rho}^{POLS} = p\lim\hat{\rho}^{POLS} = \rho$. For expositional purposes we drop the $_{T\to\infty}$ subscript in the remainder of the appendix.

B.1 The pooled OLS estimator

The first estimator we are studying is the pooled OLS estimator for the spillover parameter ρ. The pooled OLS regression model is defined by equation (8) in the main text. Averaging equation (8) over time and countries yields:[21]

$$\bar{\dot{c}} = \gamma + \rho\bar{\dot{n}} + \left[\bar{\dot{\alpha}} - \gamma + \bar{\dot{\varepsilon}}\right] \tag{22}$$

where $\bar{\dot{\alpha}}$, $\bar{\dot{c}}$ and $\bar{\dot{n}}$ are the global averages of α_i, c_{it}, n_{it-1} within the onset sample, i.e.

$$\bar{\dot{\alpha}} = \frac{1}{N^o}\sum_{it}\dot{\alpha}_i, \quad \bar{\dot{n}} = \frac{1}{N^o}\sum_{it}\dot{n}_{it-1}, \quad \bar{\dot{c}} = \frac{1}{N^o}\sum_{it}\dot{c}_{it} \tag{23}$$

where N^o is the number of observations in the onset sample: $N^o = \sum_{it}(1 - c_{it-1})$. Subtracting equation (22) from equation (8) eliminates the constant γ:

$$\dot{c}_{it} - \bar{\dot{c}} = \rho\left(\dot{n}_{it-1} - \bar{\dot{n}}\right) + \left[\dot{\alpha}_i - \bar{\dot{\alpha}} + \dot{\varepsilon}_{it} - \bar{\dot{\varepsilon}}\right] \tag{24}$$

Performing OLS on regression equation (24) yields the same parameter estimate for ρ as performing OLS on the baseline regression equation (8).

The pooled OLS estimator can be written as the sum of ρ and an additional term:

$$\hat{\rho}^{POLS} \quad = \quad \rho + \frac{\sum_{it}\left(\dot{n}_{it-1} - \bar{\dot{n}}\right)\left(\dot{\alpha}_i - \bar{\dot{\alpha}} + \dot{\varepsilon}_{it} - \bar{\dot{\varepsilon}}\right)}{\sum_{it}\left(\dot{n}_{it-1} - \bar{\dot{n}}\right)^2} \tag{25}$$

To verify the consistency of $\hat{\rho}^{POLS}$, we need the conditions under which the additional term convergence to zero. To do this first normalize its numerator and denominator by $\frac{1}{N^o}$ to facilitate

---

[21] We are considering a balanced panel for the derivations.

the analysis.

$$p\lim\left[\hat{\rho}^{POLS}\right]=\rho+\frac{p\lim\frac{1}{N^{o}}\sum_{it}\left(\dot{n}_{it-1}-\overline{\dot{n}}\right)\left(\dot{\alpha}_{i}-\overline{\dot{\alpha}}+\dot{\varepsilon}_{it}-\overline{\dot{\varepsilon}}\right)}{p\lim\frac{1}{N^{o}}\sum_{it}\left(\dot{n}_{it-1}-\overline{\dot{n}}\right)^{2}} \tag{26}$$

The denominator of the fraction in equation (26) is an average. We assume that standard regularity assumptions apply and that it converges to a finite nonzero constant. (This is a standard assumption that basically assumes that the dynamic process is not exploding.) Consistency of $\hat{\rho}^{POLS}$ subsequently depends on the properties of the numerator. As ε_{it} is i.i.d. and n_{it-1} and ε_{it} are uncorrelated it is straightforward to show that equation (26) can be written as follows:

$$p\lim\left[\hat{\rho}^{POLS}\right]=\rho+\frac{p\lim\frac{1}{N^{o}}\sum_{it}\left(\dot{n}_{it-1}-\overline{\dot{n}}\right)\left(\dot{\alpha}_{i}-\overline{\dot{\alpha}}\right)}{p\lim\frac{1}{N^{o}}\sum_{it}\left(\dot{n}_{it-1}-\overline{\dot{n}}\right)^{2}} \tag{27}$$

Further analyzing equation (27). For the purpose of our study we are introducing two new definitions. N_r^o is the number of observations in the onset sample within region r: $N_r^o=\sum_t\sum_{i\in r}(1-c_{it-1})$.

$$\overline{\dot{\alpha}}_r = \frac{1}{N_r^o}\sum_t\sum_{i\in r}\dot{\alpha}_i \tag{28}$$

$$\overline{\dot{n}}_r = \frac{1}{N_r^o}\sum_t\sum_{i\in r}\dot{n}_{it-1} \tag{29}$$

Using these definitions and the simple trick of both adding and substracting $\overline{\dot{\alpha}}_r$ and $\overline{\dot{n}}_r$, we can rewrite equation (26) as:

$$p\lim\left[\hat{\rho}^{POLS}\right] = \rho+\frac{p\lim\frac{1}{N^{o}}\sum_{it}\left[\left(\dot{n}_{it-1}-\overline{\dot{n}}_r\right)+\left(\overline{\dot{n}}_r-\overline{\dot{n}}\right)\right]\left[\left(\dot{\alpha}_{i}-\overline{\dot{\alpha}}_r\right)+\left(\overline{\dot{\alpha}}_r-\overline{\dot{\alpha}}\right)\right]}{p\lim\frac{1}{N^{o}}\sum_{it}\left[\left(\dot{n}_{it-1}-\overline{\dot{n}}_r\right)+\left(\overline{\dot{n}}_r-\overline{\dot{n}}\right)\right]^{2}} \tag{30}$$

Expanding both the numerator and the denominator and using three relationships below

$$\begin{aligned}
\frac{1}{N^o}\sum_{it}\left(\dot{n}_{it-1}-\bar{\dot{n}}_r\right)\left(\bar{\dot{\alpha}}_r-\bar{\dot{\alpha}}\right) &= \frac{1}{N^0}\sum_{r}\left(\bar{\dot{\alpha}}_r-\bar{\dot{\alpha}}\right)\sum_{t}\sum_{i\in r}\left(\dot{n}_{it-1}-\bar{\dot{n}}_r\right) \\
&= \frac{1}{N^o}\sum_{r}\left(\bar{\dot{\alpha}}_r-\bar{\dot{\alpha}}\right)\underbrace{\left[N_r^0\bar{\dot{n}}_r-N_r^0\bar{\dot{n}}_r\right]}_{=0}=0 \qquad (31)
\end{aligned}$$

$$\begin{aligned}
\frac{1}{N^o}\sum_{it}\left(\bar{\dot{n}}_r-\bar{\dot{n}}\right)\left(\dot{\alpha}_i-\bar{\dot{\alpha}}_r\right) &= \frac{1}{N^o}\sum_{r}\left(\bar{\dot{n}}_r-\bar{\dot{n}}\right)\sum_{t}\sum_{i\in r}\left(\dot{\alpha}_i-\bar{\dot{\alpha}}_r\right) \\
&= \frac{1}{N^o}\sum_{r}\left(\bar{\dot{n}}_r-\bar{\dot{n}}\right)\underbrace{\left[N_r^o\bar{\dot{\alpha}}_r-N_r^o\bar{\dot{\alpha}}_r\right]}_{=0}=0 \qquad (32)
\end{aligned}$$

$$\begin{aligned}
\frac{1}{N^o}\sum_{t}\sum_{i}\left(\dot{n}_{it-1}-\bar{\dot{n}}_r\right)\left(\bar{\dot{n}}_r-\bar{\dot{n}}\right) &= \frac{1}{N^o}\sum_{r}\left(\bar{\dot{n}}_r-\bar{\dot{n}}\right)\sum_{t}\sum_{i\in r}\left(\dot{n}_{it-1}-\bar{\dot{n}}_r\right) \\
&= \frac{1}{N^o}\sum_{r}\left(\bar{\dot{n}}_r-\bar{\dot{n}}\right)\underbrace{\left[N_r^o\bar{\dot{n}}_r-N_r^o\bar{\dot{n}}_r\right]}_{=0}=0 \qquad (33)
\end{aligned}$$

we can simplify equation (30) to equation (9) in the main text. i.e.,

$$p\lim\left[\hat{\rho}^{POLS}\right] = \rho+\frac{p\lim\frac{1}{N^o}\sum_{it}\left[\left(\dot{n}_{it-1}-\bar{\dot{n}}_r\right)\left(\dot{\alpha}_i-\bar{\dot{\alpha}}_r\right)+\left(\bar{\dot{n}}_r-\bar{\dot{n}}\right)\left(\bar{\dot{\alpha}}_r-\bar{\dot{\alpha}}\right)\right]}{p\lim\frac{1}{N^o}\sum_{it}\left[\left(\dot{n}_{it-1}-\bar{\dot{n}}_r\right)^2+\left(\bar{\dot{n}}_r-\bar{\dot{n}}\right)^2\right]}$$

B.2 The REGION estimator

The second estimator we study is the REGION OLS estimator for the spillover parameter ρ. This estimator is referred to as $\hat{\rho}^R$. The REGION OLS regression model is defined by equation (10) in the main text. Averaging equation (10) over t and i within region r yields:

$$\bar{\dot{c}}_r=\gamma_r+\rho\bar{\dot{n}}_r+\left[\bar{\dot{\alpha}}_r-\gamma_r+\bar{\dot{\varepsilon}}_r\right] \qquad (34)$$

where $\bar{\dot{c}}_r$ is defined:

$$\bar{\dot{c}}_r=\frac{1}{N_r^0}\sum_{t}\sum_{i\in t}\dot{c}_{it} \qquad (35)$$

Subtracting equation (34) from equation (10) eliminates the region specific constant γ_r:

$$\dot{c}_{it}-\bar{\dot{c}}_r=\rho\left(\dot{n}_{it-1}-\bar{\dot{n}}_r\right)+\left[\dot{\alpha}_i-\bar{\dot{\alpha}}_r+\dot{\varepsilon}_{it}-\bar{\dot{\varepsilon}}\right] \qquad (36)$$

Performing OLS on regression equation (36) yields the same parameter estimate for ρ as performing OLS on the baseline regression equation (10).

The REGION OLS estimator can be written as the 'true' ρ plus an additional term:

$$\hat{\rho}^R = \rho + \frac{\sum_{it}\left(\dot{n}_{it-1} - \overline{\dot{n}}\right)\left(\dot{\alpha}_i - \overline{\dot{\alpha}}_r + \dot{\varepsilon}_{it} - \overline{\dot{\varepsilon}}_r\right)}{\sum_{it}\left(\dot{n}_{it-1} - \overline{\dot{n}}_r\right)^2} \tag{37}$$

To verify the consistency of $\hat{\rho}^R$, we need the conditions under which the additional term convergence to zero. We first normalize its numerator and denominator by $\frac{1}{N^o}$. We again assume that the denominator of the fraction in equation (41) converges to a finite nonzero constant. The consistency of $\hat{\rho}^{FE}$ therefore depends on the asymptotic behavior of the numerator. Next, as ε_{it} is i.i.d. by assumption and n_{it-1} and ε_{it} are uncorrelated, it is straightforward to show that equation (37) can be written as equation (11) in the main text, i.e.

$$p\lim\left[\hat{\rho}^R\right] = \rho + \frac{p\lim \sum_{it}\left[\dot{n}_{it-1} - \overline{\dot{n}}_r\right]\left[\dot{\alpha}_i - \overline{\dot{\alpha}}_r\right]}{p\lim \sum_{it}\left[\dot{n}_{it-1} - \overline{\dot{n}}\right]^2} \tag{38}$$

B.3 The FE estimator

The third estimator we study is the fixed effects (FE) estimator for the spillover parameter ρ. This FE estimator is obtained by performing OLS on the so-called within transformed model, given by equation (6). This subtracts the country specific time mean of all variables that immediately eliminates the country specific effect α_i:

$$\dot{c}_{it} - \overline{\dot{c}}_i = \rho\left(\dot{n}_{it-1} - \overline{\dot{n}}_i\right) + \left[\dot{\varepsilon}_{it} - \overline{\dot{\varepsilon}}\right] \tag{39}$$

where $\overline{\dot{c}}_i = \frac{1}{N_i^o}\sum_t \dot{c}_{it}$ and $\overline{\dot{n}}_i = \frac{1}{N_i^o}\sum_t \dot{n}_{it-1}$ with N_i^o the number of observations in the onset sample on country i: $N_i^o = \sum_t (1 - c_{it-1})$.

The FE estimator can be written as the 'true' ρ plus an additional term:

$$\hat{\rho}^{FE} = \rho + \frac{\sum_{it}\left(\dot{n}_{it-1} - \overline{\dot{n}}_i\right)\left(\dot{\varepsilon}_{it} - \overline{\dot{\varepsilon}}_r\right)}{\sum_{it}\left(\dot{n}_{it-1} - \overline{\dot{n}}_i\right)^2} \tag{40}$$

First we normalize the numerator and the denominator by $\frac{1}{N^o}$.

$$p\lim\left[\hat{\rho}^{FE}\right] = \rho + \frac{p\lim \frac{1}{N^0}\sum_{it}\left(\dot{n}_{it-1} - \overline{\dot{n}}_i\right)\left(\dot{\varepsilon}_{it} - \overline{\dot{\varepsilon}}_i\right)}{p\lim \frac{1}{N^o}\sum_{it}\left(\dot{n}_{it-1} - \overline{\dot{n}}_i\right)^2} \tag{41}$$

Just as in the pooled OLS and REGION OLS cases we assume that the denominator of the fraction in equation (41) converges to a finite nonzero constant. The consistency of $\hat{\rho}^{FE}$ subsequently depends on the asymptotic properties of the numerator. Expanding this numerator and using that $p\lim \frac{1}{N^o}\sum_{it} n_{it-1}\varepsilon_{it} = 0$ (because n_{it-1} and ε_{it} are uncorrelated by assumption yields:

$$p\lim \frac{1}{N^o}\sum_{it}\left(\dot{n}_{it-1} - \bar{\dot{n}}_i\right)\left(\dot{\varepsilon}_{it} - \bar{\dot{\varepsilon}}_i\right) = -p\lim \frac{1}{N^o}\sum_{it}\bar{\dot{n}}_i\dot{\varepsilon}_{it} - p\lim \frac{1}{N^o}\sum_{it}\left(\dot{n}_{it-1} - \bar{\dot{n}}_i\right)\bar{\dot{\varepsilon}}_i$$

Using large T asymptotics we can show that both the above terms equal zero:

$$p\lim \frac{1}{N^o}\sum_{it}\bar{\dot{n}}_i\dot{\varepsilon}_{it} = \sum_i \frac{N_i^o}{N^o}\bar{\dot{n}}_i \times \underbrace{p\lim \frac{1}{N_i^o}\sum_t \dot{\varepsilon}_{it}}_{=0} \tag{42}$$

$$p\lim \frac{1}{N^o}\sum_{it}\left(\dot{n}_{it-1} - \bar{\dot{n}}_i\right)\bar{\dot{\varepsilon}}_i = \sum_i \frac{N_i^o}{N^o}\bar{\dot{\varepsilon}}_i \times \underbrace{p\lim \frac{1}{N_i^o}\sum_t \left(\dot{n}_{it-1} - \bar{\dot{n}}_i\right) = 0}_{=0} \tag{43}$$

So that the FE estimator $\hat{\rho}^{FE}$ is a consistent estimator for ρ, i.e.

$$plim\left[\hat{\rho}^{FE}\right] = \rho \tag{44}$$

C Continents and World Regions

We define Continents and World Regions according to the United Nations classification.
(see http://www.un.org/depts/dhl/maplib/worldregions.htm).

Continents

Africa, Asia, Europe, Latin America, North America and Oceania.

World Regions

Eastern Africa, Middle Africa, Northern Africa, Western Africa, Southern Africa, Eastern Asia, South-central Asia, South-eastern Asia, Western Asia, Eastern Europe, Northern Europe, Southern Europe, Western Europe, Caribbean, Central America, South America, North America, Australia & New Zealand, Melanesia, Micronesia, and Polynesia.

D Top and bottom ten contributors to the pooled conflict spillover parameter

The ten countries that contribute most negatively to the pooled conflict spillover parameter obtained from (1) are (in declining order of magnitude): *Indonesia, Argentina, Colombia, Cyprus, Guatemala, Lebanon, Philippines, Sri Lanka, Yemen AR, and Bolivia.*

The ten countries that contribute most positively to the pooled conflict spillover parameter obtained from (1) are (in declining order of magnitude): *Uganda, Democratic Republic of Congo, Pakistan, China, Sudan, Nigeria, Rwanda, Liberia, Burundi, and Yemen PR.*

References

Abreu, M., R. Florax, and H. de Groot (2005): "Space and growth: a survey of empirical evidence and methods," *Région et Développement*, 21, 13–44.

Achen, C. (2000): "Why lagged dependent variables can suppress the explanatory power of other independent variables," *unpubl. manuscript*.

Beck, N., and K. S. Gleditsch (2006): "Space is more than geography: using spatial econometrics in the study of political economy," *International Studies Quarterly*, 50, 27–44.

Breinlich, H. (2006): "The spatial income structure in the European Union - what role for economic geography," *Journal of Economic Geography*, 6, 593–617.

Buhaug, H., and K. S. Gleditsch (2008): "Contagion or confusion? Why conflicts cluster in space," *International Studies Quarterly*, 52, 215–233.

Buhaug, H., and J. K. Rød (2006): "Local determinants of African civil wars," *Political Geography*, 25, 315–335.

Collier, P., and A. Hoeffler (2004): "Greed and grievance in civil war," *Oxford Economic Papers*, 56, 563–595.

Elbadawi, I., and N. Sambanis (2002): "How much war will we see? Explaining the prevalence of civil war," *Journal of Conflict Resolution*, 46, 307–334.

Elhorst, P. (2003): "Specification and estimation of spatial panel data models," *International Regional Science Review*, 26, 244–268.

Fearon, J., and D. Laitin (2003): "Ethnicity, insurgency and civil war," *American Political Science Review*, 97, 75–90.

Gleditsch, K. S. (2007): "Transnational dimensions of civil war," *Journal of Peace Research*, 44, 293–309.

Gleditsch, K. S., and K. Beardsley (2004): "Nosy neighbors: third-party actors in Central American Conflicts," *Journal of Conflict Resolution*, 48, 379–401.

Gleditsch, K. S., and M. D. Ward (2006): "Diffusion and the international context of democratization," *International Organization*, 60, 911–933.

——— (2008): *Spatial Regression Models*. SAGE Publication, Inc.

Griliches, Z. (1961): "A note on serial correlation bias in estimates of distributed lags," *Econometrica*, 29(1), 65–73.

Harbom, L., and P. Wallensteen (2005): "Armed conflict and its international dimensions," *Journal of Peace Research*, 42, 623–635.

Hegre, H., T. Ellingsen, S. Gates, and N. P. Gleditsch (2001): "Toward a democratic civil peace? Democracy, political change and civil war, 1816-1992," *American Political Science Review*, 95, 33–48.

Hegre, H., and N. Sambanis (2006): "Sensitivity analysis of empirical results on civil war onset," *Journal of Conflict Resolution*, 50, 508–535.

Miguel, E., S. Satyanath, and E. Sergenti (2004): "Economic shocks and civil conflict: an instrumental variable approach," *Journal of Political Economy*, 112, 725–753.

Moore, W., and S. Shellman (2007): "Whither will they go? A global study of refugees' destinations, 1965-1995," *International Studies Quarterly*, 51, 811–834.

Murdoch, J., and T. Sandler (2002): "Economic growth, civil war, and spatial spillovers," *Journal of Conflict Resolution*, 46, 91–110.

——— (2004): "Civil wars and economic growth: spatial dispersion," *American Journal of Political Science*, 48, 138–151.

Nickell, S. (1981): "Biases in Dynamic Models with Fixed Effects," *Econometrica*, 49(6), 1417–1426.

Salehyan, I., and K. S. Gleditsch (2006): "Refugees and the spread of civil war," *International Organization*, 60, 335–366.

Sambanis, N. (2001): "Do ethnic and nonethnic civil wars have the same causes? A theoretical and empirical inquiry (Part 1)," *Journal of Conflict Resolution*, 45, 259–282.

——— (2002): "A review of recent advances and future directions in the quantitative literature on civil war," *Defense and Peace Economics*, 13, 215–243.

——— (2004): "What is a civil war? Conceptual and empirical complexities of an operational definition," *Journal of Conflict Resolution*, 48, 814–858.

Simmons, B., and Z. Elkins (2004): "The globalization of liberalization: policy diffusion in the international political economy," *American Political Science Review*, 98, 171–189.

Vanhanen, T. (1999): "Domestic ethnic conflict and ethnic nepotism: A comparative analysis," *Journal of Peace Research*, 36, 55–73.