Impact of the Fragile States on Stability and Development of the Arab Region: The Channels of Conflict Contagion [1]

Fernando Cantú

Emerging and Conflict Related Issues Division (ECRI)

United Nations Economic and Social Commission for Western Asia (UN-ESCWA)

[1] The opinions expressed herein are those of the author and do not necessarily reflect those of the United Nations or the UN Economic and Social Commission for Western Asia.

Impact of the Fragile States on Stability and Development of the Arab Region:

The Channels of Conflict Contagion[2]

Fernando Cantú

cantu@un.org

United National Economic and Social Commission for Western Asia (UN ESCWA)

This version: 9 December 2016

Abstract

Recent cases of armed conflict in Iraq, Syria, Libya, Nigeria and other places show that armed conflict has the potential of crossing national borders and spread to other states, both close and distant. Motivated by these observations and based on the previous literature on the subject, this paper studies the channels of conflict contagion and the channels through which it can move from one territory to another. It does so by using a flexible framework based on the notions of the gravity model to determine spatial weights, and which can be extended to study diverse instances of the neighbourhood effect. The empirical estimations provide strong evidence of conflict contagion. Geographic proximity does play a role, but conflict can extend well beyond only adjacent states. In addition, cultural affinity is an important channel of contagion and, when it acts as a vector, conflict ignores distances and can travel around the planet. Given a country's domestic determinants and the security situation of its neighbourhood, the model can provide an estimate of the likelihood of conflict onset. This shows, for example, how the Arab region is particularly prone to the eruption of conflict due to, among other determinants, a contagion effect. These results show the importance of considering a wider sphere of influence for neighbouring conflict, particularly when cultural factors could be at play, when designing early warning systems of other tools to monitor armed conflict.

[2] This document is a work in progress. Please do not cite or copy without written consent from the author. Comments are welcome, please send them to cantu@un.org

I. Introduction: violent neighbours

People, goods and ideas move on a global scale, unremittingly crossing borders and moving from one country to another. This is evidenced by the vast volume of world trade, the increasing flow of financial assets, an ever more mobile labour force and the wave-like progress in worldwide democratic transitions. However, the same global character also applies to negative arenas such as the illegal flow of weapons and illicit products, human trafficking, external influences threatening national sovereignty and instances of violent conflict. These positive and negative externalities have been grouped under the term neighbourhood effect.

UN ESCWA (2014) proposes a framework to study this phenomenon. It is defined around four dimensions: the source in the neighbouring country, the affected domestic variables, the specific channels through which these two are linked, and the nature of the neighbourhood. This structure can be used to organize the study of all instances of such indirect influences originating in the neighbourhood. For example, positive economic developments in one member country of the European Union may have a positive influence on the economy of other members through trade, intra-industry linkages, increased investment and other financial flows, tourism, and so on. This occurs because of geographic proximity, but also because of economic complementarity and an extensive network of economic and political agreements and alliances among the member countries. As another example, as evidenced by Danneman and Ritter (2014), the outbreak of a conflict in a country may lead to a certain degree of democratic regression in nearby states as they increase civil and political repression in an attempt to thwart the start of similar conflicts at home. This occurs when neighbouring conflict is close-by, but also when its motivation (resistance to cultural oppression, separatist claims, reaction to unequal economic or political opportunities) can be readily emulated domestically.

Although the neighbourhood effect, by definition, only has an indirect influence on domestic determinants, its impact can be substantial. This is particularly true of episodes of armed conflict. Not only may they lead to negative political consequences, as exemplified above, but they may also damage bordering infrastructure, disrupt production and trade links among economic partners, increase risk perception in the entire region, divert productive investment to defence expenditure and lead to population displacement. Murdoch and Sandler (1992, 1994), Collier et al. (2003) and Gates et al. (2010) document the impact of neighbouring conflict on socioeconomic development. In addition to this, there is also the potential of conflict contagion. There are many contemporary cases of this, from the Boko Haram insurgency in Nigeria spreading to Cameroon and also threatening neighbouring Chad and Niger to the internal conflict in Syria spilling over its borders into Lebanon, Jordan, Turkey and Iraq. In fact, the Arab countries provide multiple examples of conflict contagion and it could be argued that the neighbourhood effect in this sphere is of particular consequence for this region.

In case an episode of internal conflict breaks out in one location, how likely is it that it will spread to neighbouring countries? How far can its influence be felt? And is it only a matter of geographic proximity or other channels of contagion could also be at play? Could conflict also spread, for example, through economic, cultural or political links? What is the actual risk of conflict contagion in the Arab region at present?

The objective of this report is to revise these questions in light of the evidence presented by historical instances of internal armed conflict across the world. After controlling for the main domestic determinants that may contribute to the onset of violent episodes, it will assess the significance of contagion from neighbouring conflict through a variety of channels that could be at play. For this, it will propose and make use of a model that defines the extension and nature of the neighbourhood.

The rest of the report is organized as follows. The next chapter briefly describes the current state of affairs in the Arab world and it argues that the probability of conflict contagion could be especially elevated for this region. Then, a short review of the existing studies that document this effect will be presented, with an emphasis on the definition of the neighbourhood and the channels of transmission considered by the authors. The fourth chapter will present a model that could guide the specification of the neighbourhood, both in terms of proximity and strength of influence, for this and other studies of this phenomenon. This model will be applied in an empirical analysis of the effect of neighbouring conflict on the likelihood of domestic conflict onset; this will be estimated with data from all recent episodes of conflict across the world. A final chapter will summarise the main conclusions.

II. Recent significance for the Arab region

The Arab region is arguably one of the locations in the world where the spectre of conflict contagion is constantly present. In the five years between 2009 and 2013, 41% of the Arab countries experienced at least one episode of conflict.[3] Some of these are long-lasting confrontations that have expanded to larger and larger territories and crossed state borders. For reasons described below, all of them pose threats of contagion to their neighbours.

The most immediate reference is the current war in Syria who has already caused palpable effects in bordering countries and over the entire region. What started as a series of peaceful protests against the regime of Bashar Al-Assad, part of the political movements that swept the Arab region in 2011, soon transformed into a complex conflict between the government, backed by different foreign actors, and a multitude of groups and militias with equally multifaceted sources of support, funding and personnel. Against this backdrop rose the Islamic State (IS), a radical group initially affiliated to Al-Qaeda that proclaimed itself as the start of a caliphate and that holds violent, exclusivist principles. This group swiftly extended its influence to Iraq, a country struggling with political divisions, confessional tensions and a legacy of violence and instability since the United States intervention. There, it rapidly overcame military positions and claimed a large area in North-Western Iraq. The IS has also extended its operations across the borders to Lebanon, Jordan and Turkey. Distant locations such as the Sinai Peninsula in Egypt, Libya and Nigeria have also witnessed violent episodes attributed to IS-affiliated actors.

In addition to the cross-border expansion of the IS and the involvement of other groups, the conflict in Syria has had other repercussion that also translate into a higher risk of conflict in neighbouring countries. Millions of Syrian refugees have migrated to Turkey, Lebanon and Jordan and this has a

[3] UN ESCWA calculations based on data from UCDP/PRIO Armed Conflict Dataset.

direct impact in the likelihood of conflict in the host communities.[4] The conflict has also disrupted economic activity, mainly in bordering areas, and the resulting increase in unemployment and poverty could reduce the opportunity cost of criminal activities and participation in violent groups, particularly among the youth. In addition, there is an upsurge in the flow of illicit weapons and an increase in militant recruitment from close and distant origins.

The crisis in Syria is not the only conflict in an Arab country that is causing negative externalities and threatens to export its violence to neighbouring regions. Libya is now in the middle of a protracted battle between the internationally recognised government, now exiled from the capital and based in the eastern city of Tobruk, a rival Islamist coalition that has elected a new National Congress (still largely unrecognised by the international community), and numerous militias and armed groups, some of them in declared alliance with the IS. Under this turmoil, the country has become a magnet of foreign fighters and an epicentre of arm smuggling. Direct attacks to Egyptian citizens residing in Libya, as well as this government's involvement in favour of the Tobruk-based Libyan government, have threatened to spark episodes of conflict inside Egypt. This comes in addition of the already volatile situation in the Sinai. Tunisia, a fragile democracy that is challenged by its own radical groups and their militant activities, is also jeopardised by the bordering conflict.

The military insurgency of the Houthis against the central government in Yemen has also involved regional players: while Iran and groups such as Lebanon-based Hezbollah have overtly supported the Houthi groups, Saudi Arabia led a coalition of countries in an intervention aiming to debilitate Houthi positions and in support of the president Abed Rabbuh Mansur Hadi, in exile in the Saudi capital. This crisis adds to the unrest and instability that were already afflicting the country, the ongoing radicalisation of many tribal groups and the violence imposed by Al-Qaeda and affiliated groups. The hostilities in Yemen are threatening to spill over not only to bordering countries, but to become a proxy war involving the entire region.

As revealed by the examples reviewed above, most of the recent Arab conflicts involve non-state groups that operate transnationally in terms of their hubs of influence, supply of funding and weaponry, and source of recruitment, and whose movement across countries is facilitated by long, porous borders. Importantly, some of them echo an exclusivist discourse along confessional lines that readily finds adepts from one extreme of the region to the other. This is a factor of concern that could facilitate the movement of violence from one country to another.

Moreover, some common traits of the Arab region could make it prone to conflict contagion. There is a generalised governance deficit that, to one degree or another, limits accountability of those in power and restricts the political participation of most citizens. In connection to this, there is a widespread income inequality and limited economic opportunities for large segments of the population. These factors have already sparked protests and violent episodes that swiftly found common ground and spread to several countries. In addition to the Arab-Israeli conflict that led to a shared history of conflict and the legacy of the displaced Palestinian population, current conflicts persist and the number of affected persons that are forced into displacement keeps growing. As mentioned above, this is a significant feature that could export violence to neighbouring countries. Finally, as witnessed by the current crises in Syria, Yemen and Libya, conflicts in this region act, to an important extent, as proxy wars between regional and external actors seeking to maintain or extend

[4] See Salehyan and Gleditsch (2006) and UN ESCWA (2015).

their influence. And this confrontation can lead to new or renewed clashes emerging in other locations.[5]

Borrowing the term from Collier et al. (2003), the Arab countries could be at the verge of a regional equivalent of the conflict trap, in which a conflict in one state could lead to the contagion of violence to neighbouring countries, returning through a negative feedback that further increases the likelihood of remaining in conflict and infecting other, even more distant territories. Extrapolating from the harmful consequences of a conflict trap and the difficulty of escaping such a vicious circle at the country level, the possibility of a regional version of this phenomenon remains a daunting prospect for the Arab region.

III. Documented evidence of the movement of conflict across borders

Evidence of the systematic occurrence of conflict contagion has been well documented in the literature. Sambanis (2001), Gleditsch (2002, 2007), Hegre and Sambanis (2006) and Buhaug and Gleditsch (2008) all conclude that a conflict in the neighbourhood significantly increases the likelihood of conflict onset. This result is robust to different specifications of the model, particularly in relation to the domestic and regional determinants considered; this holds even when the possible clustering of attributes in neighbouring countries is taken into account.

Other studies have focused on specific elements of conflict contagion among neighbours. Salehyan and Gleditsch (2006) and Forsberg (2009) find evidence that the arrival of refugees increases the probability of a conflict onset in the host country. Therefore, conflict-driven population displacement could be a trigger of conflict transmission among countries.

The likelihood that violence in the neighbourhood will turn into violence at home differs from country to country depending on many factors. One of them is the capacity of the state: Braithwaite (2010) argues that more capable (i.e., stronger) states are better able to resist the spread of violence from the neighbours; the empirical evidence robustly supports this hypothesis. Another factor is the political configuration: Maves and Braithwaite (2013) conclude that autocracies with elected legislatures are more likely to be affected by the contagion from a neighbouring conflict. This arises from the fact that elected legislatures may give voice and formal power to the opposition relative to the autocrats. This fact would support the commonly-found evidence that democracy has a quadratic effect on the onset of violence: conflicts habitually occur in hybrid regimes (“flawed democracies” or “imperfect autocracies”) while more complete forms of autocracy or democracy experience internal conflict less frequently.

There is also a differentiated influence of the neighbouring conflict that depends on its type. Fox (2004), Buhaug and Gleditsch (2008), Cederman et al. (2009), de Groot (2011), Beiser (2012), Bosker and de Ree (2014) and Forsberg (2014) have studied the particular case of ethnic (or ethno-religious or ethno-linguistic) conflicts. They all find evidence that this type of conflicts move most frequently across borders and affect neighbouring states. Indeed, ethnic conflicts in a nearby territory may trigger a demonstration effect in ethnic groups at home, particularly when they have reasons

[5] See UN ESCWA (2014) for more information on these common factors of the Arab region that could elevate its risk of conflict contagion.

(repression or discrimination) and resources (group concentration and relative size) to mobilise and start fighting. This is enhanced if the ethnic group involved in the neighbouring conflict, or one with a high degree of affinity, is also represented at home. The transmission of information between territories also seems to play a significant role.

All studies mentioned in the sample above include an explicit or inferred definition of what is meant by neighbouring states. Some go as far as considering alternative demarcations (through a different distance threshold, for example), as part of their robustness checks. These definitions contain the assumptions about the channels of conflict contagion from the neighbourhood that the author is considering.

Regarding the size of the neighbourhood, most studies consider only bordering states or those within a specific distance. The latter is defined as the minimum distance between the countries[6] or as inter-capital distances. A few papers apply a global definition of neighbourhood and include all countries in the world, but weighted by the inverse of the distance between them. In this sense, all of the studies define the neighbourhood through a distance measure (contiguity, minimum distance or inter-capital distance) and therefore assume that geographic proximity is one (and frequently the only) channel of conflict contagion.

In addition, some studies incorporate additional channels of contagion, depending on the hypotheses they aim at testing. For example, some of them assume that cultural affinity (ethnic, linguistic and/or religious) has a role to play. Others consider that conflict-driven population displacement may be source of conflict transmission from one country to another. Other channels, such as economic linkages, flow of information and length of borders (as a way to measure the potential points of contact between countries) have also been proposed. Table 1 summarises this information.

IV. How does conflict spread? A gravity model

As illustrated above, there seems to be a consensus in the literature that geographical proximity plays an important role in conflict contagion. And this clearly makes sense: it is expected that nearby violence could have a stronger effect than distant fighting. However, many studies assume that only adjacent conflicts could lead to contagion. When a wider radius is considered (through a minimum distance threshold, for example), this is frequently set arbitrarily.

Table 1. Definition of the neighbourhood considered in the sample of studies

Article	Span of neighbourhood	Channel of contagion
Sambanis (2001)	Bordering states	Geographic proximity (contiguity)
Gleditsch (2002)	Bordering states; states within 50, 475 or 950 km*	Geographic proximity (contiguity, minimum distance), economic interdependence
Fox (2004)	Bordering states	Geographic proximity (contiguity), cultural affinity (religious)

[6] These can be calculated directly with the popular CShapes dataset (Weidmann et al., 2010).

Salehyan and Gleditsch (2006)	States within 100 or 950 km*	Geographic proximity (minimum distance), population displacement
Gleditsch (2007)	States within 950 km**	Geographic proximity (minimum distance)
Buhaug and Gleditsch (2008)	Bordering states; all states*	Geographic proximity (contiguity, minimum distance), border length, size (population), cultural affinity (ethnic), population displacement
Cederman et al. (2009)	States within 500 km**	Geographic proximity, cultural affinity (ethnic)
Forsberg (2009)	Bordering states	Geographic proximity (contiguity), population displacement
Braithwaite (2010)	Bordering states	Geographic proximity (contiguity)
De Groot (2011)	Bordering states	Geographic proximity (contiguity), border length, cultural affinity (ethnolinguistic)
Beiser (2012)	All states (except bordering)	Geographic proximity (inter-capital distance), information flow
Maves and Braithwaite (2013)	All states	Geographic proximity (contiguity, minimum distance)
Bosker and De Ree (2014)	Bordering states	Geographic proximity (contiguity)
Forsberg (2014)	Bordering states	Geographic proximity (contiguity), cultural affinity (ethnic)

* This study includes several estimations spanning different definitions of the neighbourhood.

** Alternative thresholds were considered as part of robustness checks.

In addition to proximity, there could be other factors that influence the likelihood of conflict diffusion. Following the hypotheses of the articles listed in Table 1, these could be cultural affinity, border length, economic interdependence, population displacement or information flow. However, this influence has not been studied in a systematic way. In some occasions, it is incorporated into the calculation of the neighbours' weights (such as in De Groot, 2011); sometimes it is reflected through a different coding of the neighbouring conflict variable (such as in Fox, 2004); in yet other cases, it is included as additional explanatory variables that characterise the neighbourhood (such as in Buhaug and Gleditsch, 2008). Moreover, these channels are rarely evaluated consistently and most robustness checks only investigate the differences within these choices (by including or excluding the single hypothesised channel, or by varying the minimum distance threshold to select neighbours, for example) rather than between them. As noted by Zhukov and Stewart (2013), the statistical inferences regarding the importance of the neighbourhood are sensitive to these choices and they must therefore be carefully considered.

This paper proposes a framework to study the diffusion of conflict across borders based on two notions. Borrowing the terms from Zhukov and Stewart (2013, p. 271), the first one refers to the connectivity choices: who should be considered a neighbour. The second specifies the spatial weights, the channel and the extent through which the selected neighbours could affect the risk of conflict contagion.

The decision of which countries that should belong to the neighbourhood (the connectivity choice) follows, in this context, a geographic criterion. Instead of selecting a fixed number of neighbours or other arbitrary decision rules, a minimum distance threshold will be used. Therefore, the neighbourhood is defined by all those states that are located within a certain distance from the country being studied. This is calculated through the minimum distance option of the aforementioned CShapes dataset. The choice of threshold is inevitably arbitrary, so different options were considered: bordering states,[7] 100 km, 500 km, 1000 km, and all countries in world. For example, to obtain the neighbourhood with a minimum distance threshold of 500 km, an area is drawn extending 500 km from all points along the border of a country; all other countries with at least part of their territories within this radius are then considered as neighbours. This is summarised in C_t, a matrix of zeros and ones that indicates, for each year t, which countries are included in the neighbourhood of each state. This is a symmetric matrix of dimension N_t, where N_t is the number of countries in the world at time t.

Regarding the spatial weights, I follow the basic scheme of a gravity model, according to which the influence of neighbouring units depends directly on their mass and inversely on their distance. In other words, larger states should have a greater influence than smaller states, and nearer countries should have a larger impact than more distant countries. This is a common model that has been used extensively to study, among other, issues of political influence and international trade.[8] The proximity and mass measures are encapsulated in the quantities Π_t and Λ_t, both symmetric matrices of dimension N_t.

All these components are collected into a spatial connectivity matrix, denoted by W_t, which summarises the weight of each country in the neighbourhood. It combines the three elements described above, the connectivity choices, the proximity component and the mass measure, in the following manner.

$$W_t = C_t \cdot \Pi_t \cdot \Lambda_t$$

The elements of this matrix depend on the units of the proximity and mass matrices, as well as on the number of countries that fall inside the neighbourhood. This means that it may not be comparable between countries, between different time periods or between different choices of connectivity, proximity or mass. In order to ensure comparability, the matrix W_t is row-standardised so that the sum of each of its rows is equal to one. This will be denoted as $\widetilde{W}_t$.

It quickly becomes evident that the operationalisation of the gravity model in this context is a data-intensive process, as it requires dyadic information for all countries of the world for each year

[7] This was calculated through a 10 km threshold in the CShapes R package with a tolerance parameter equal to 0.05. This is because of the approximation error given this tolerance level and also to consider cases of countries that, although separated by small geographical features, are adjacent for all practical matters, such as Malaysia and Singapore.
[8] The idea of a gravity model to study the determinants of conflict contagion comes from Buhaug and Gleditsch (2008). However, these authors do not study the full implications of the model and do not account systematically for the proximity and mass components of the neighbourhood influence. The present study attempts to bridge these gaps.

considered. The total number of observations for each set of weights (Π_t or Λ_t,) and each period *t* is then[9]

$$\frac{N_t(N_t - 1)}{2}$$

Taking into account that the average number of independent countries in the world from 1946 to 2012 is around 150, this translates into an average information requirement of 11 175 distinct observations per year. Some of the variables that could be proposed to estimate the proximity and mass effects in the gravity model are simply not available at this level of disaggregation and would therefore have to be excluded from the analysis.

Several options could be followed for the proximity measure of the gravity model. A natural idea would be to consider geographical distance, with the assumption that conflicts in nearer states have a higher probability of having an impact. This study does not using geo-referenced data that would allow the exact placement of the conflict, so this leaves the three choices available in the CShapes dataset: minimum distance, inter-centroid distance and inter-capital distance. Since minimum distance was already used to determine the connectivity choices, and in order to incorporate a measure of geographical (centroid) and geopolitical (capital) proximity, the latter two were considered.

However, proximity goes well beyond pure geographical notions. For example, two countries can be close together if their economies are interlinked (measured through the amount of trade between them), or if there is a high cultural affinity between their population (ethnic, religious and/or linguistic), or if there is a mutual history of political alliances and military pacts. These alternative concepts of proximity lead to the following alternatives for the proximity measure Π_t: inter-centroid distance, inter-capital distance, trade, cultural affinity and alliances.

Similarly, several alternatives could be proposed for the mass matrix Λ_t. The hypothesis is that bigger countries have more opportunities of contact and exert a larger emulation or other type of influence. The immediate way to measure it is through population. Other alternatives could be the size of the economy (measured through total GDP) and border length (although this implies that the connectivity choice is only adjacent states). For both the proximity and mass measures, an additional alternative considered is that of the identity matrix. This implies that one or both factors have no impact on the likelihood of conflict contagion.

These options do not cover the whole spectrum of possibilities that could be proposed. For example, a proximity matrix could be based on the amount of information exchange between countries, under the hypothesis that countries that are more connected would tend to have a larger impact on each other. Or it could be argued that migration flows between countries (including refugees) is a relevant way to measure proximity, with the assumption that what matters for conflict contagion is the size of the population that is forced out of the country and into its neighbours. Mass weights could also be measured through conventional measures of strength, either through traditional (hard) military

[9] This result is obtained since the unstandardised connectivity matrix W_t is symmetric (i.e., the unstandardised weight of A on B is equal to the unstandardised weight of B on A) and its diagonal is zero (i.e., the weight of A on itself is zero).

power or soft (cultural attraction) power. Complete dyadic information at this level for this and other alternatives is not available so that their evaluation is left for future research.[10]

All the combinations of the spatial connectivity matrix given the alternatives mentioned above are summarised in Table 2. The rows indicate the proximity weight Π_t, the columns indicate the mass weight Λ_t while the different connectivity choices are listed inside each cell. For instance, the spatial weight *w*(1000, centroid, GDP) indicates that all states within 1000 km are considered as part of the neighbourhood, that they are weighted (inversely) by their inter-centroid distance to account for proximity, and by their GDP to represent mass.[11] This gives a total of 62 possible spatial connectivity matrices. As described in the following section, they will be used to weight the incidence of conflict in the neighbourhood and then test its significance for the onset of internal conflict (i.e., conflict contagion from the neighbourhood). The sources for the data required to construct these matrices is detailed in Appendix B.

It must be noted that the connectivity choices and the gravity model described here are designed to study the impact that a conflict could bring about to neighbouring countries in terms of an increased probability of conflict contagion. They are not intended to test the influence of other neighbouring variables (such as regional development or democratic status) on the onset of domestic conflict. However, as noted in ESCWA (2014), all these spheres are potentially interlinked and it could be argued that events such as a large economic crisis or a democratic breakdown in a state could increase the risk of conflict for the affected country as well as for its neighbours. To test such hypothesis, a similar procedure to the one followed in this study could be replicated: select connectivity choices, proximity

[10] Note that there exists dyadic information for refugees available from the United Nations High Commissioner for Human Rights (UNHCR) and the United Nations Relief and Works Agency for Palestinian Refugees in the Near East (UNRWA). However, this database is complete from the year 2000. Although some data is available from 1960 (obtained through direct correspondence with UNHCR), it tends to be incomplete, up to the point where data points far in the past cover only a few cases of population displacement. For this reason this option for the proximity matrix was not pursued.

[11] Note that some combinations are not considered. For example, border length as mass weight can also be applied when considering adjacent states (a threshold of zero for the connectivity choice). Additionally, when only bordering countries are considered, no proximity weight is applied since that would add little information. Finally, trade and alliances as proximity weights are only constructed when the connectivity choice includes all states in the world, since these are not geographic criteria (however, preliminary estimations were calculated with other connectivity choices and they are available from the author).

Table 2. Combinations of connectivity choices, proximity weight and mass weight used to construct the spatial connectivity matrix W_t

		Λ_t			
		None	**Border length**	**Population**	**GDP**
Π_t	**None**	*w(*0, none, none) *w(*100, none, none) *w(*500, none, none) *w(*1000, none, none) *w(*world, none, none)	*w(*0, none, border)	*w(*0, none, popn) *w(*100, none, popn) *w(*500, none, popn) *w(*1000, none, popn) *w(*world, none, popn)	*w(*0, none, GDP) *w(*100, none, GDP) *w(*500, none, GDP) *w(*1000, none, GDP) *w(*world, none, GDP)
	Inter-centroid distance	*w(*100, centroid, none) *w(*500, centroid, none) *w(*1000, centroid, none) *w(*world, centroid, none)		*w(*100, centroid, popn) *w(*500, centroid, popn) *w(*1000, centroid, popn) *w(*world, centroid, popn)	*w(*100, centroid, GDP) *w(*500, centroid, GDP) *w(*1000, centroid, GDP) *w(*world, centroid, GDP)
	Inter-capital distance	*w(*100, capital, none) *w(*500, capital, none) *w(*1000, capital, none) *w(*world, capital, none)		*w(*100, capital, popn) *w(*500, capital, popn) *w(*1000, capital, popn) *w(*world, capital, popn)	*w(*100, capital, GDP) *w(*500, capital, GDP) *w(*1000, capital, GDP) *w(*world, capital, GDP)
	Cultural affinity	*w(*0, culture, none) *w(*100, culture, none) *w(*500, culture, none) *w(*1000, culture, none) *w(*world, culture, none)	*w(*0, culture, border)	*w(*0, culture, popn) *w(*100, culture, popn) *w(*500, culture, popn) *w(*1000, culture, popn) *w(*world, culture, popn)	*w(*0, culture, GDP) *w(*100, culture, GDP) *w(*500, culture, GDP) *w(*1000, culture, GDP) *w(*world, culture, GDP)
	Trade	*w(*world, trade, none)		*w(*world, trade, popn)	*w(*world, trade, GDP)
	Alliances	*w(*world, alliances, none)		*w(*world, alliances, popn)	*w(*world, alliances, GDP)

Note: the three quantities that define the spatial weights are denoted as *w*(connectivity choice, proximity weight, mass weight).

and mass weights that are relevant for the type of contagion under analysis and use them to construct spatial connectivity matrices; these could then serve to test the hypothesised neighbourhood effect on conflict onset given the appropriate channels of transmission.[12]

V. The main channels of conflict contagion

This study uses the evidence from the history of internal armed conflicts in the world since the end of the Second World War to analyse the main channels of conflict contagion. This is done through the estimation of a quantitative model based on a large dataset of economic, social, political and conflict variables. The neighbourhood effect on the probability of internal conflict onset was documented through the extensive definition of the neighbourhood outlined in the previous section.

First, a statistical model was selected in order to study the main research questions related to the impact of neighbouring conflict. It also includes a set of controls for conflict onset that has been identified in the literature; this has the purpose of isolating the impact of the main research variable and reducing the possibility of misattributing to it the significance of other factors. Then the gravity model outlined previously will be applied to obtain different characterisations of the neighbourhood and, through them, elucidate the main channels of contagion.

V.1. Statistical model and data sources

The basic unit of analysis is the country-year; that is, one observation per country per year. All independent countries in the world are included from 1946 to 2012. However, questions of data availability constrained the time coverage to a shorter period. The main description of data will be presented below, but the full details of the construction of the dataset and the data sources are left for Appendix B.

The dependent variable is the onset of internal conflict in country *i* at time *t*, where conflict is defined as "a contested incompatibility that concerns government and/or territory where the use of armed force between two parties, of which at least one is the government of a state, results in at least 25 battle-related deaths."[13] This is the standard definition of the Uppsala Conflict Data Project (UCDP) and Peace Research Institute Olso (PRIO) dataset on armed conflict.[14] Only internal conflicts are considered since inter-state wars generally follow different dynamics and processes of potential contagion. The analysis concentrates on transitions from a phase of peace to a period of conflict (i.e., a conflict onset); the determinants for remaining in conflict and for transitioning from conflict to peace are dissimilar and not the focus of study here.

---

[12] More generally, the same idea could also be used to study other types of spatial diffusion processes, such as the diffusion of democracy or the spillovers from neighbouring economic events.

[13] Gledistch et al. (2002, p. 618-619).

[14] This study uses version v.4-2014a, see Themnér and Wallensteen (2014).

The model is designed to study how neighbouring conflict may affect the likelihood of an internal conflict onset. As described previously, the gravity model applied here gives 62 different ways to assign weights to the surrounding countries, resulting in as many spatially weighted measures of conflict in the neighbourhood. However, in order to study the specific implication of each of these combinations on the channels of conflict contagion, the model should also include a set of domestic control variables that may affect the risk of conflict onset. These are the standard control variables identified in the literature as potentially motivating violent incidents and they will only be succinctly described here. The theoretical and empirical rationale for their inclusion has already been extensively discussed in Hegre and Sambanis (2006) and the other studies previously mentioned (and the references therein).

Population has been identified as an important determinant of conflict onset. This because the definition of conflict used employs an absolute threshold that is simply easier to meet in more populated countries. But also because a larger population leads to increased competition for resources, a broader pool to recruit militants and possibly a higher heterogeneity in terms of culture or income. In order to mitigate the high disparity in total population among the countries in the world, this variable enters the model after a logarithmic transformation.

It has also been argued that a large proportion of youth relative to the general population, the so-called youth bulges, could be linked to conflict since they may not find sufficient opportunities for economic, social and political participation and these grievances may drawn them into violent groups and conflict. The recent observations in some Arab countries seem to suggest this could be the case. As suggested by Urdal (2004), youth bulges are calculated as the proportion of the population aged 15 to 24 relative to the adult population.

In more developed countries, states are strong and the economic opportunity cost of violence is high; therefore, they are less likely to experience internal conflict. This link between development and conflict has been robustly confirmed in many studies. The level of development is commonly approximated through GDP per capita or a vital health indicator such as infant mortality rate. However these only account partially for the state of development of a country and it is not difficult to find cases where a high level of aggregate income has not translated to better standards of living, or vice versa. The Human Development Index (HDI) proposed by the United Nations Development Programme (UNDP)[15] measures development through a geometric mean of indicators for health (life expectancy), education (average years of schooling and expected years of schooling) and income (gross national income per capita). Even if this composite index is an interesting alternative, it cannot be used in this estimation due to data availability issues. Instead, I construct a development indicator inspired by the HDI, but through an unobserved components model (to avoid the attribution of exogenous weights to the variables) for the three components: GDP per capita, a sub-composite health indicator (also constructed through an unobserved components model for life expectancy, infant mortality rate and under-five mortality rate) and average years of schooling. To avoid potential endogeneity issues with conflict, the first lag of this variable is considered in the model.

[15] This index has been published by UNDP since 1990. It shows information for most countries in the world for the period 1980-2005 at five-year intervals and then annually until 2013. For the latest version, see UNDP (2014).

One disadvantage of using a composite index for development is that it only partly incorporates the information contained in GDP per capita. In the context of conflict studies, this variable has been used as a proxy for development or welfare (both of them better studied through the more comprehensive indicator proposed above), but also as a proxy of state capacity. To bridge this gap, the relative political capacity measure proposed by Arbetman-Rabinowitz and Johnson (2008) is included. This is a measure of efficiency that attempts to assess "the ability of a government to extract resources from a population given their level of economic development."[16]

Another concept that has consistently been linked to internal conflict is the type of regime. Most studies have argued that both strong autocracies and strong democracies successfully avoid conflicts (this is in part linked to their political capacity considered in the previous variable) and that most violent episodes occur in the middle: hybrid regimes or flawed democracies. To account for this, the standard polity 2 score and its square are included. As with development, these variables enter the equation as their first lag in order to mitigate potential endogeneity issues.[17]

As documented in Fearon and Laitin (2003), Montalvo and Reynal-Querol (2005), Hegre and Sambanis (2006) and Wimmer et al. (2009), the effect of ethnicity heterogeneity on conflict has been subject to debate. Montalvo and Reynal-Querol (2005) find evidence that it is ethnic polarisation, more than mere fractionalisation, which is correlated to violent episodes. Following this finding, the model estimated here includes a measure of ethnic polarisation as an additional determinant of the start of conflict. The data concerning the ethnic groups per country is obtained from the variable "people cluster" of the Joshua Project database.[18] This is a variable with only one observation per country (i.e., no time series are available) and including it is equivalent to assuming that the ethnic composition of the country remains invariable over time. Although this is a clear simplification and one could think of some recent cases that refute this, it arguably remains a credible hypothesis. Therefore, and given the lack of more complete databases, this variable will be included in the analysis.

Two additional variables that have been included as regressors in armed conflict models are oil production per capita and the percentage of the territory covered by mountains. For the former, it has been argued (Ross, 2004; Humphreys, 2005; Wimmer and Min, 2006) that rents from oil give financial incentives for conflict initiation, in particular for secessionist conflicts when the resources are concentrated geographically. Additionally, rents from resource extraction may give rise to clientelism,

---

[16] Arbetman-Rabinowitz and Johnson (2008, p. 2)

[17] The polity 2 varies from -10 (strong autocracy) to 10 (strong democracy). Additionally, it codes three special cases: anarchy, transition and foreign interruption. The model initially included an indicator variable that took the value of one for these special situations. Predictably, it was significant with the expected positive sign (i.e., the probability of conflict onset increases when one of these situations is present). However, I felt that this amounted to trying to predict a variable with itself and, since the rest of the estimated coefficients and their significance remained unchanged, it was not included in the final model. The results for this preliminary model are available upon request.

[18] In addition to the distribution of ethnic groups per country, this database also includes information on languages and religion. Polarisation measures for these concepts were also constructed, but they turned out to be not significant. Parallel measures of ethnic, religious and linguistic fractionalization were also attempted, but the estimated coefficients were mostly statistically insignificant. These results are available from the author.

exclusion and potential mobilisation. Finally, economies whose resources depend on oil exports could be severely weakened during a price shock, potentially leading to mobilisations and conflict. On the other hand, and in relation to the concept of state capacity described before, states with high rents obtained from the extraction of natural resources may have additional resources to prevent conflict (by either appeasing or repressing the groups of concern). The logarithm of oil production per capita is included to account for this potential link between oil rents and armed conflict. The source of data is the United States Energy Information Agency, Wimmer and Min (2006) and other minor sources. The percentage of a country's land area covered by mountains serves as a proxy of rough terrain that could give fighters hidden and/or difficult to access shelter and centres of operation. The data is mainly obtained from Fearon and Laitin (2003).

In order to account for temporal dependence, the model also includes the number of peace years since the last conflict, as suggested by Beck et al. (1998). Since the effect of the duration of peace on conflict onset is unlikely to be linear, the authors suggest including it through a smoothing spline. However, it is difficult to decide the degree of smoothness, as the spline can lead to overfitting by approximating the observed frequency between conflict and peace duration as perfectly as desired, therefore leaving no explanatory power to any other variables. In addition, using the same set of data to estimate the smoothing spline and then to estimate the coefficients may generate some statistical concerns. For this reason, time dependence is included through an exponential density as in Gleditsch (2007).

V.2. Estimation methodology

The dependent variable, onset of internal conflict, is an indicator variable: it takes the value of one if an armed conflict, according to the definition adopted in this study, breaks out in country *i* at year *t*, and zero in all other cases. This will be fitted through a logistic regression model using the variables described above as explanatory variables, including the measures of weighted incidence of conflict in the neighbourhood.

However, the data does not originate from a random sample. It has a panel structure since a number of countries are observed through time. The combination of limited dependent variable and panel data can be challenging. If the time-series-cross-section structure is ignored through a pooled regression, the results could suffer from omitted variable bias. The common solution, the fixed effect estimator, raises several issues. First, all time-invariant variables (mountainous terrain and ethnic polarisation in our case) cannot be estimated since they are confounded with the country-specific intercept; in the same vein, this estimator is very inefficient when there is little within-country variability (such as in the cases of population or the indicator of regime type). Equally worrisome is the fact that conflict data is sparse: most of the countries in the sample do not experience any internal conflict and, as a consequence, the fixed intercept would perfectly predict the outcome. This amounts to discarding the majority of the sample and keeping only those that transitioned at least once from a state of peace (dependent variable equal to zero) to a state of internal armed conflict (dependent variable equal to one). As noted by Beck and Katz (2001), this is a high price to pay in terms of efficiency.

The alternative solution, the random effects estimator, has been mostly rejected since it imposes the mostly unrealistic assumption that all regressors are independent of the country error component. However, recent studies (Hegre and Nordkvelle, 2014; Bell and Jones, 2015) have upheld this estimator by noting that it overcomes some of the concerns with the fixed effects solution (particularly in relation to limited within-country variability) and that the trade-off between bias and efficiency may be not be as pronounced in moderate samples. That is, even if the exogeneity assumption does not hold introducing some bias in the estimates, this could be a minor price to pay in relation to the gains in efficiency. In addition, an explicit modelling of the dependence between the country random component and the regressors by including the within-mean as additional variables in the regression (a solution first proposed by Mundlak, 1978) could control for the endogeneity bias of the random effects estimator. Based on these results, a logistic regression with country-specific random effects was selected as the estimation method.[19]

It must be noted that a spatial model such as this one could introduce endogeneity in the model through another channel: if it is true that neighbouring conflict has an effect on the onset of internal armed conflict, then the beginning of an episode of conflict at home would also affect the probability of violence outbreaks in the neighbours. In other words the dependent variable and this regressor follow an endogenous relationship and this would introduce a bias in the regression. This could be directly corrected through the use of an autologistic regression model, as done in Gleditsch (2007). However, this complicates the interpretation of the results as the estimated coefficients would only amount to the "first wave" of the effect, without considering the back and forth feedback effect on the likelihood of conflict between the country and the neighbourhood. In addition, it could be argued that neighbouring conflict does not have an instantaneous effect and that the transmission mechanisms take some time to become active. For this reason, I include the lagged measure of neighbouring conflict, thus making the assumption that it is a predetermined variable.

V.3. Estimation results

Even if the conflict dataset extends from 1946 to 2013 for all countries of the world, other variables are only available for shorter periods of time. The common denominator for all of them is the period 1960 to 2012. This, in addition to countries appearing at a later stage or disappearing before the end of the period, plus a proportion of cases with missing data, results in a final sample of 6 276 country-year observations. Over this period, a total of 278 onsets of armed conflict were counted. The results of the estimation for this sample are presented in Table 3.[20, 21]

[19] A Monte Carlo simulation was developed in order to compare different estimators (pooled regression, fixed effects and random effects) with a setting comparable to the one relevant for this study (in terms of sample size, sparseness, presence of time-invariant variables, etc.) The results overwhelmingly supported the use of random effects, especially when accompanied by Mundlak's correction. The full report of this simulation is available from the author.

[20] As described before, to guard against the risk of endogeneity bias originating from any eventual dependence between the explicative variables and the random effect, the within-country means were included as additional regressors. Wald tests consistently rejected their significance, so these variables were dropped. The results of the tests are available from the author.

Table 3. Determinants of internal conflict onset

	(1)	(2)	(3)	(4)
	Only domestic determinants	With neighbouring conflict		
		w(w,culture,GDP)	*w(w,centroid,none)*	*w(w,culture,none)*
Constant	-9.7399***	-9.3553***	-9.2640***	-9.4723***
	(1.2497)	(1.2387)	(1.2332)	(1.2409)
Population	0.5065***	0.4875***	0.4705***	0.4859***
	(0.0825)	(0.0819)	(0.0812)	(0.0819)
Youth population	2.4400	1.2331	0.4343	1.1425
	(1.8624)	(1.9054)	(1.9682)	(1.9215)
Development	-0.2775**	-0.4153***	-0.3928***	-0.3728***
	(0.1397)	(0.1483)	(0.1455)	(0.1446)
State capacity	-0.2260	-0.1791	-0.2266	-0.2424
	(0.2208)	(0.2218)	(0.2195)	(0.2211)
Regime type	-0.2768*	-0.2316	-0.2162	-0.2072
	(0.1677)	(0.1686)	(0.1673)	(0.1694)
$(\text{Regime type})^2$	-0.5658***	-0.5426***	-0.5531***	-0.5747***
	(0.1646)	(0.1637)	(0.1623)	(0.1649)
Ethnic polarisation	1.2442**	1.1290**	1.2350**	1.2088**
	(0.5745)	(0.5732)	(0.5635)	(0.5713)
Oil production per capita	0.0344	0.0401	0.0493	0.0387
	(0.0708)	(0.0710)	(0.0698)	(0.0707)
Mountainous terrain	0.7078	0.6800	0.5819	0.7266
	(0.5669)	(0.5648)	(0.5561)	(0.5613)
Peace duration	1.2458***	1.1707***	1.2193***	1.2014***
	(0.3198)	(0.3193)	(0.3227)	(0.3213)
Neighbouring conflict		1.8112***	4.6043***	3.2974***
		(0.6138)	(1.5742)	(1.1554)
Observations	6276	6276	6276	6276
Log-likelihood	-817.13	-812.82	-812.94	-813.06
AIC	1658.25	1651.65	1651.88	1652.11

Notes: The dependent variable is the onset of internal armed conflict. The three variations of neighbouring conflict are calculated with the three best performing weights, as shown in Table 4. The method of estimation is logistic regression with country-specific random effects. Standard errors indicated in parentheses. The symbols ***, ** and * indicate significance at 1%, 5% and 10%, respectively. AIC stands for Akaike Information Criteria. For details on the sources of data as well as the transformations applied to the variables, see Appendix B.

[21] Table 3 uses the regular standard errors of the estimated coefficients to test their significance. Contrary to other authors, I do not report "robust", sandwich-type standard errors as they lack any sense in the context of a logistic regression. In a nonlinear model, heteroskedasticity does not lead to inefficient but unbiased estimators, as in a linear model, but it will also cause the estimators to be biased. Because of this, the use of traditional correction methods for the standard errors could be misleading. I then implicitly make the assumption that the inclusion of all the domestic determinants, the time dependence variable and the random intercept all contribute to a sufficiently correct specification of the model.

The first column refers to a model with all the domestic control determinants described above but without considering the influence of the neighbourhood. All the variables have the expected sign but not all of them have a significant impact on internal conflict onset. As expected, a larger population is one of the major predictors of conflict; moreover, a population that is ethnically polarised is particularly linked to violent episodes. However, the estimation shows that the “youth bulge” factor does not play a consistent role according to the evidence.

Our measure of development, which summarises progress along social and economic dimensions, also features prominently: higher developed states are generally less prone to experience conflict. State capacity, although exhibiting the expected sign, has a high p-value, so that, at least as it is measured here and given the other determinants of included in the model, a weak state is not a significant determinant of conflict. Nonetheless, the hypothesis that both autocracies and democracies experience fewer conflicts holds, so that “hybrid regimes” or “flawed democracies” are the regimes with a higher probability of suffering from an armed conflict. A rough terrain and the importance of oil production are not significantly linked to conflict when the rest of the controls are included. Finally, the exponential function of the number of peace years is highly significant, indicating that conflict onset shows time-dependence: the longer the duration of peace, the less likely that a new conflict will break out.

The measure of neighbouring conflict was now considered to study its impact on the probability that contagion. There are 62 possible combinations of spatial weights and these were introduced in the model one at a time. Of these, 16 proved to be significant and they are listed in Table 4. Several conclusions can be drawn from these results.

i. Regarding the connectivity choice, all but two of the weights define the neighbourhood as countries within 500 km or more (some consider all countries in the world). This shows that conflict does not only cross direct borders, but that it can also have an impact in more distance places. It would then be important to extend the network beyond adjacent countries when studying conflict contagion.
ii. Geography does play an important role in the diffusion of conflict, as proximity weights related to this criterion (inter-centroid distance and inter-capital distance) are frequently represented. However, cultural affinity is also a major determinant as more than 30% of the significant weights are defined by it. Moreover, when culture appears in the proximity weight, it is usually tied to broader connectivity choices (1000 km or all countries in the world), indicating that cultural ties as channels of conflict contagion ignore distances and can have an influence even in far-away locations. Trade and alliances are, on the contrary, irrelevant to study this phenomenon.
iii. The mass of the neighbours in conflict hardly matters. With one exception each featuring border length, population or the size of the economy, the selected weights choose the identity matrix as mass weights. This means that, when considering conflict contagion, neighbours large or small have a comparable influence.

The three best performing weights (as chosen by the AIC) are included in columns (2)-(4) of Table 3. The estimates of the other determinants hardly change compared to model (1), except for a higher significance for the development indicator, and the variable for neighbouring country is highly significant

Table 4. Neighbourhood weights that lead to significant coefficients

Significance at the 1% level	Significance at the 5% level	Significance at the 10% level
w(world, centroid, none) 1651.88	*w*(500, centroid, none) 1655.81	*w*(0, none, none) 1657.25
w(world, culture, none) 1652.11	*w*(1000, centroid, none) 1654.94	*w*(0, none, border) 1657.49
w(world, culture, GDP) 1651.65	*w*(1000, capital, none) 1656.31	*w*(500, none, none) 1656.60
	w(1000, culture, none) 1655.83	*w*(500, capital, none) 1656.89
	w(world, capital, none) 1656.17	*w*(500, culture, none) 1657.04
	w(world, culture, popn) 1653.93	*w*(1000, none, none) 1657.14
		w(world, none, none) 1657.09

Notes: The significance level of the neighbouring variable is obtained by keeping the rest of the model constant and changing only the spatial weights applied to neighbouring conflict. For the definitions of the weights, see Section IV. The figure below each of the weights is the corresponding AIC (Akaike Information Criteria) of each of the models.

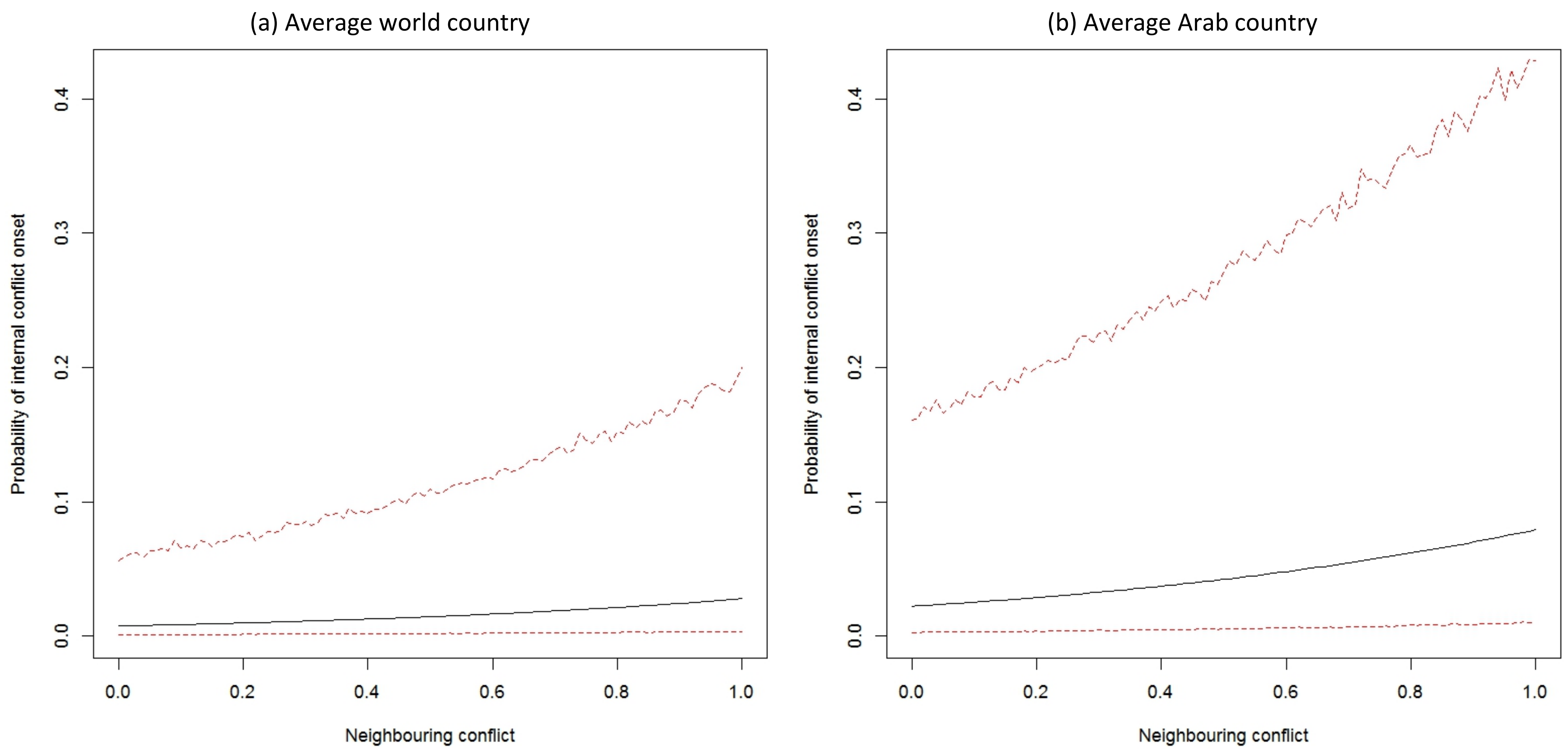
and with a positive sign. It can therefore be concluded that neighbouring conflict, according to the definition of the spatial weights, does play a crucial role in increasing the probability of conflict onset.[22]

Even if the coefficients of the model included in Table 3 can be used to evaluate the direction and significance of each variable, as well as aiding in model selection, they cannot be interpreted directly because they are based in a (nonlinear) logistic model. Appendix C lists the marginal effects at the mean for each variable for the four models included in Table 3.

As an alternative way to interpret the impact of neighbouring conflict obtained in the regression model, the first panel of Figure 1 presents the impact of changes in this variable of the best model (column 2) for the average country in the world in the most recent year (i.e., by fixing the domestic determinants at the world mean during 2012). The solid line represents the point estimate for the probability of the onset of internal conflict obtained by varying the value of neighbouring conflict. The effect of conflict in the neighbourhood is small but significant. However, note the dashed lines around it, which represent the (10%, 90%) confidence region.[23] When there is no neighbouring conflict, the model assigns a probability of conflict onset for the average country of between 0.08% and 5.84% with 80% confidence. However, if 20% of the neighbourhood of this hypothetical average country is undergoing conflict, these probabilities raise to 0.12% and 7.53% respectively. If 50% of the neighbourhood is experiencing conflict,

[22] The results of the estimation using the other significant weights are comparable to those presented in Table 3. They are available from the author.
[23] This is not a smooth line because it is obtained through simulation. In fact, this confidence region is obtained by sampling 10 000 values of the coefficients given their estimated distribution and then calculating the 10% and 90% quantiles by fixing the rest of the variables at their mean in the most recent year (2012).

Figure 1. Effect of changes in the variable "neighbouring conflict" on the probability of internal conflict onset, 2012



Notes: these figures are based on the model listed in column 2 of Table 3. All other regressors are fixed at their mean in 2012. The dashed lines indicate the (0.10, 0.90) confidence region obtained by sampling 10 000 values of the coefficients and calculating the corresponding quantiles.

the probability range increases to (0.18%, 10.62%). While the lower bound remains close to zero, the upper bound increases exponentially. As more of the surrounding countries get involved in conflict, the risk of contagion clearly increases.

As it has been argued in this document and given the recent events, it could be hypothesised that the Arab region follows a distinctive pattern of conflict contagion. To study this question, it would be ideal to replicate the same analysis over a sample restricted to the Arab countries (or through a variable coefficient model). However, even if this region has accounted for almost 18% of all internal armed conflicts in the world, the absolute number of onsets is only 49. This is a very limited amount of information that restricts any region-specific analysis.

However, the underlying characteristics of this region differ from the rest of the world and this model allows studying the likelihood of conflict given such domestic and neighbourhood determinants. The second panel of Figure 1 illustrates the effect of the neighbouring conflict on the probability of conflict onset for the average Arab country (again, the rest of the variables are fixed at their mean for the Arab region during 2012).[24] It could be noted that, even if there is no regional conflict, the probability of an armed conflict in this region ranges from 0.26% to 15.90% with 80% confidence. The impact of a neighbouring conflict also rises accordingly. If 20% of the neighbouring states are experiencing violence, this probabilities range reaches (0.34%, 20.38%). If half of the neighbourhood is in conflict (a not-so-unlikely eventually for this region), the probabilities of a conflict outbreak at home increase to 0.52% and 27.55%, respectively. Therefore, the possibility of conflict contagion is noticeably higher in the Arab region, compared to the rest of the world, given its underlying characteristics.

V. Conclusions

The theory and empirical evidence presented in this study support the hypothesis of a significant neighbourhood effect of conflict that manifests itself, among other consequences, as a higher probability of contagion to neighbouring states. The evidence shows that this influence extends beyond just the bordering states and can have an impact even in distant territories. This seems to be particularly true when cultural links (ethnic, religious or linguistic) act as the channel of contagion; in such situations, conflict can travel great distances. On the contrary, other factors such as economic interdependence, political alliances and the size of the conflict-affected state, seem to have a negligible impact in the transmission of violence.

In addition to the domestic determinants of conflict, such as low level of development, an ethnically polarised society or a recent history of violence, the incidence of conflict in the neighbourhood can have a small but significant impact on the probability of experiencing an episode of armed conflict at home. Given the internal characteristics and the security situation of the Arab countries, this region seems particularly vulnerable to conflict contagion and it could be well underway of falling in a regional conflict

[24] The Arab region is defined as the 22 current member countries of the League of Arab States.

trap with extremely detrimental prospects for the future development of this part of the world and, through the many channels of action of the neighbourhood effect, the entire planet.

These findings are important when designing tools to monitor resilience or early warning systems of conflict. It is not sufficient to consider only the most immediate states as risk factors, but also take into account that the influence of a conflict can travel great distances, particularly if there is a cultural vector involved.

Finally, this paper proposed a framework based on a gravity model that could guide the study of the neighbourhood effect in general and its channels of transmission. This is a highly flexible structure that was used to study several proximity and mass determinants of conflict contagion, but that could easily be adapted to study other factors, such as the inflow of refugees or military expenditure, subject to data availability. It could also be used in the analysis of other processes of inter-state diffusion, such as political changes, economic spillovers, emulation of ideas, and others.

VI. References

Arbetman-Rabinowitz, Marina and Kristin Johnson (2008). “Relative political capacity: empirical and theoretical underpinnings.” Paper presented at the 49th Annual Convention of the International Studies Association, San Francisco, CA, 26-29 March.

Barro, Robert and Jong-Wha Lee (2010). “A new data set of educational attainment in the world, 1950-2010.” *Journal of Development Economics* 104: 184-198.

Beck, Nathaniel, Jonathan Katz and Richard Tucker (1998). “Taking time seriously: time-series-cross-section analysis with a binary dependent variable.” *American Journal of Political Science* 42 (4): 1260-1288.

Beck, Nathaniel and Jonathan Katz (2001). “Throwing out the baby with the bat water: a comment on Green, Kim, and Yoon.” *International Organization* 55 (2): 487-495.

Beiser, Janina (2012). “Looking beyond borders: identification, information, and ethnic conflict contagion.” *University College London working paper*. Retrieved from http://privatewww.essex.ac.uk/~ksg/esrcjsps/Beiser.pdf (accessed 8 May 2015).

Bell, Andrew and Kelvyn Jones (2015). “Explaining fixed effects: random effects modeling of time-series cross-sectional and panel data.” *Political Science Research and Methods* 3 (1): 133-153.

Bosker, Maarten and Joppe de Ree (2014). “Ethnicity and the spread of civil war.” *Journal of Development Economics* 108 (C): 206-221.

Braithwaite, Alex (2010). “Resisting infection: how state capacity conditions conflict contagion.” *Journal of Peace Research* 47 (3): 311-319.

Buhaug, Halvard and Kristian S. Gleditsch (2008). “Contagion or confusion? Why conflicts cluster in space.” *International Studies Quarterly* 52 (2): 215-233.

Cederman, Lars-Erik, Luc Girardin and Kristian S. Gleditsch (2009). “Ethnonationalist triads: assessing the influence of kin groups on civil wars.” *World Politics* 61 (3): 403-37.

Collier, Paul et al. (2003). *Breaking the conflict trap: civil war and development policy*. Washington DC: World Bank.

Danneman, Nathan and Emily H. Ritter (2014). “Contagious rebellion and preemptive repression.” *Journal of Conflict Resolution* 58 (2): 254-279.

De Groot, Olaf (2011). “Culture, contiguity and conflict: on the measurement of ethnolinguistic effects in spatial spillovers.” *Journal of Development Studies* 47 (3): 436-454.

Fearon, James and David Laitin (2003). “Ethnicity, insurgency and civil war.” *American Political Science Review* 97 (1): 75-90.

Feenstra, Robert, Robert Inklaar and Marcel Timmer (2013). "The next generation of the Penn World Table." *National Bureau of Economic Research Working Papers* 19255. Available at www.ggdc.net/pwt.

Forsberg, Erika (2009). “Refugees and intrastate armed conflict: a contagion process approach.” Paper presented at the 50th Annual Conference of the International Studies Association, New York, 15-18 February.

Forsberg, Erika (2014). “Transnational transmitters: ethnic kinship ties and conflict contagion 1946-2009.” *International Interactions* (forthcoming).

Fox, Jonathan (2004). “Is ethnoreligious conflict a contagious disease?” *Studies in Conflict & Terrorism* 27 (2): 89-106.

Gates, Scott et al. (2012). “Development consequences of armed conflict.” *World Development* 40 (9): 1713-1722.

Gibler, Douglas (2009). *International military alliances, 1648-2008*. CQ Press.

Gleditsch, Kristian S. (2002). *All politics is local: the diffusion of conflict, integration, and democratization*. Ann Arbor, MI: Univer sity of Michigan Press.

Gleditsch, Kristian S. (2007). “Transitional dimensions of civil war.” *Journal of Peace Research* 44 (3): 293-309.

Gleditsch, Nils Petter et al. (2002). “Armed conflict 1946-2001: a new dataset.” *Journal of Peace Research* 39 (5): 615–637.

Hegre, Håvard and Nicholas Sambanis (2006). "Sensitivity analysis of empirical results on civil war onset." *The Journal of Conflict Resolution* 50 (4): 508-535.

Hegre, Håvard and Jonas Nordkvelle (2014). "How much goes out with the bath water? Omitted-variable bias, inefficiency, and over-confidence in country-year studies with dichotomous dependent variable. Paper presented to the International Studies Association Annual Conference, Toronto, March.

Humphreys, Macartan (2005). "Natural resources, conflict and conflict resolution: uncovering the mechanisms." *Journal of Conflict Resolution* 49 (4): 508-537.

Joshua Project. *Datasets*. Available at http://joshuaproject.net/resources/datasets.

Lutz, Wolfgang et al. (2007). "Reconstruction of populations by age, sex and level of educational attainment for 120 countries for 1970-2000." *Vienna Yearbook of Population Research* 2007 (5): 193-235.

Marshall, Monty, Ted Robert Gurr and Keith Richards (2014). Polity IV project: political regime changes characteristics and transitions, 1800-2013." *Center for Systemic Peace*. Available at http://www.systemicpeace.org/inscrdata.html

Maves, Jessica and Alex Braithwaite (2013). "Autocratic institutions and civil conflict contagion." *The Journal of Politics* 75 (2): 478-490.

Montalvo, José and Marta Reynal-Querol (2005). "Ethnic polarization, potential conflict, and civil wars." *American Economic Review* 95 (3): 796-816.

Mundlak, Yair (1978). "On the pooling of time-series and cross-section data." *Econometrica* 46 (1): 69-85.

Murdoch, James C. and Todd Sandler (2002). "Economic growth, civil wars, and spatial spillovers." *Journal of Conflict Resolution* 46 (1): 91-110.

Murdoch, James C. and Todd Sandler (2004). "Civil wars and economic growth: spatial dispersion." *American Journal of Political Science* 48 (1): 138-151.

Ross, Michael (2004). "What do we know about natural resources and civil war?" *Journal of Peace Research* 41 (3): 337-356.

Sambanis, Nicholas (2001). "Do ethnic and non-ethnic civil wars have the same causes? A theoretical and empirical inquiry (part 1)." *Journal of Conflict Resolution* 45 (3): 259-282

Salehyan, Idean and Kristian S. Gleditsch (2006). "Refugees and the spread of civil war." *International Organization* 60 (Spring): 336-366.

Themnér, Lotta and Peter Wallensteen (2014). "Armed conflict, 1946-2013." *Journal of Peace Research* 51 (4): 541-554.

UN DESA (2013). *World population prospects: the 2012 revision*. New York: United Nations. Available at http://esa.un.org/wpp.

UN Inter-Agency Group for Child Mortality Estimation (2014). *Levels and trends in child mortality: report 2014*. New York: UNICEF. Available at http://www.childmortality.org.

UNDP (2014). *Human development report 2014. Sustaining human progress: reducing vulnerabilities and building resilience*. New York: United Nations Development Programme. Available at http://hdr.undp.org/en/2014-report.

UN ESCWA (2014). "Beyond governance and conflict: measuring the impact of the neighbourhood effect in the Arab region." *UN ESCWA Working Paper*, available at http://www.escwa.un.org/divisions/ecri_editor/Download.asp?table_name=ecri_documents&field_name=id&FileID=272.

UN ESCWA (2015). "Refugees and economic growth: challenges and opportunities for the Arab region." *UN ESCWA Working Paper*, available at

Urdal, Henrik (2004). "The devil in the demographics: the effect of youth bulges on domestic armed conflict, 1950-2000." *World Bank, Social Development Papers* 14.

Weidmann, Nils, Doreen Kuse and Kristian S. Gleditsch (2010). "The geography of the international system: the CShapes dataset." *International Interactions* 36 (1): 86-106.

Wimmer, Andreas and Brian Min (2006). "From empire to nation-state: explaining wars in the modern world, 1816-2001." *American Sociological Review* 71 (6): 867-897.

Wimmer, Andreas, Lars-Erik Cederman and Brian Min (2009). "Ethnic politics and armed conflict: configurational analysis of a new global data set." *American Sociological Review* 74 (2): 316-337.

Zhukov, Yuri and Brandon Stewart (2012). "Choosing your neighbors: networks of diffusion in international relations." *International Studies Quarterly* 57: 271-287.

Appendix A. The calculation of spatial weights: the case of Egypt

As an example of the construction of spatial weights according to the methodology described in Section IV and the effect of the different alternative of connectivity choices, proximity and mass weights, this section presents some of the calculations for the case of one country: Egypt.

By taking the different minimum distance thresholds proposed for the connectivity choice, the neighbourhood of Egypt is composed of

- Bordering states: Israel (IL), Libya (LY), Palestine (PS), Sudan (SD).[25]

[25] For the purpose of this paper, Israel and Palestine are included as two states with coinciding borders.

- States within 100 km: Israel, Libya, Palestine, Sudan, Jordan (JO), Saudi Arabia (SA).
- States within 500 km: Israel, Libya, Palestine, Sudan, Jordan, Saudi Arabia, Chad (TD), Cyprus (CY), Eritrea (ER), Greece (GR), Iraq (IQ), Lebanon (LB), Syria (SY).
- States within 1000 km: Israel, Libya, Palestine, Sudan, Jordan, Saudi Arabia, Chad, Cyprus, Eritrea, Greece, Iraq, Lebanon, Syria, Albania, Ethiopia, Niger, Turkey, Yemen.
- All states

Suppose the threshold of 500 km is selected. This translates into a neighbourhood for Egypt comprising 13 states; the rest of the countries of the world receive a weight of zero. If the proximity and mass matrices are set to identities, implying that these factors do not matter on the influence of the neighbouring countries, the row of the spatial connectivity matrix for Egypt in 2012 would be equal to

$$W_{2012}^{\mathrm{EG}}(500,\mathrm{none},\mathrm{none}) = \begin{bmatrix} \mathrm{IL} & \mathrm{LY} & \mathrm{PS} & \mathrm{SD} & \mathrm{JO} & \mathrm{SA} & \mathrm{TD} & \mathrm{CY} & \mathrm{ER} & \mathrm{GR} & \mathrm{IQ} & \mathrm{LB} & \mathrm{SY} & \mathrm{Other} \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

This is not comparable between countries (or even through time if the set of nearby countries around Egypt suffers any change). In order to correct this, the row is standardised so that its sum is equal to one.

$$\widetilde{W}_{2012}^{\mathrm{EG}}(500,\mathrm{none},\mathrm{none}) = \begin{bmatrix} \mathrm{IL} & \mathrm{LY} & \mathrm{PS} & \mathrm{SD} & \mathrm{JO} & \mathrm{SA} & \mathrm{TD} & \cdots \\ 0.0769 & 0.0769 & 0.0769 & 0.0769 & 0.0769 & 0.0769 & 0.0769 & \\ \cdots & \mathrm{CY} & ER & \mathrm{GR} & \mathrm{IQ} & \mathrm{LB} & \mathrm{SY} & \mathrm{Other} \\ & 0.0769 & 0.0769 & 0.0769 & 0.0769 & 0.0769 & 0.0769 & 0.0000 \end{bmatrix}$$

This means that each of the thirteen neighbours of Egypt would receive an equal weight of 7.69% so that the sum of all weights is equal to 100%.

Depending on the context, an equal weight to all neighbours may not be realistic. For this reason, we introduce the proximity and mass weights of the gravity model. If we assume that geographical proximity as measured through centroid distance is the relevant measure (and still assuming that the mass of the neighbouring countries has no relevance), the spatial weights matrix becomes

$$\widetilde{W}_{2012}^{\mathrm{EG}}(500,\mathrm{centroid},\mathrm{none}) = \begin{bmatrix} \mathrm{IL} & \mathrm{LY} & \mathrm{PS} & \mathrm{SD} & \mathrm{JO} & \mathrm{SA} & \mathrm{TD} & \cdots \\ 0.1154 & 0.0744 & 0.1154 & 0.0748 & 0.1016 & 0.0580 & 0.0496 & \\ \cdots & \mathrm{CY} & ER & \mathrm{GR} & \mathrm{IQ} & \mathrm{LB} & \mathrm{SY} & \mathrm{Other} \\ & 0.0870 & 0.0563 & 0.0547 & 0.0570 & 0.0865 & 0.0693 & 0.0000 \end{bmatrix}$$

Among the 13 states, Israel, Palestine, Jordan, Cyprus and Lebanon are the closest geographically and consequently they receive a higher weight; on the other hand, Chad, Greece and Eritrea are located farther away and their weight is reduced. With the introduction of the proximity factor, the weights of the surrounding states are no longer equal and they now reflect the element of geography. If we now decide that the size of the neighbours is also an important factor and we choose GDP as the way to measure it, the spatial weights end up as

$$\widetilde{W}_{2012}^{\mathrm{EG}}(500,\mathrm{centroid},\mathrm{GDP}) = \begin{bmatrix} \mathrm{IL} & \mathrm{LY} & \mathrm{PS} & \mathrm{SD} & \mathrm{JO} & \mathrm{SA} & \mathrm{TD} & \cdots \\ 0.1983 & 0.0125 & 0.0174 & 0.0644 & 0.0316 & 0.3477 & 0.0104 & \\ \cdots & \mathrm{CY} & ER & \mathrm{GR} & \mathrm{IQ} & \mathrm{LB} & \mathrm{SY} & \mathrm{Other} \\ & 0.0213 & 0.0031 & 0.1205 & 0.0796 & 0.0511 & 0.0421 & 0.0000 \end{bmatrix}$$

This causes the weights to change substantially. For instance, Israel and Palestine, states that previously had the same weight since they occupy the same territory, now have a very dissimilar effect, because the Israeli economy is considerably larger than the Palestinian. Saudi Arabia, the largest economy of the region, receives the highest weight (34.77%) despite being relatively more distant from Egypt; the same phenomenon can be observed for other large economies, such as Greece and Iraq. The weight of countries with a low GDP, like Chad and Eritrea, almost disappears.

If we now argue that the relevant channels are cultural affinity as proximity measure and population as size, the spatial weights for Egypt's neighbours are transformed in the following way.

$$\widetilde{W}_{2012}^{\mathrm{EG}}(500, \mathrm{culture}, \mathrm{popn}) = \begin{bmatrix} \mathrm{IL} & \mathrm{LY} & \mathrm{PS} & \mathrm{SD} & \mathrm{JO} & \mathrm{SA} & \mathrm{TD} & \cdots \\ 0.0293 & 0.0402 & 0.0258 & 0.2414 & 0.0465 & 0.1748 & 0.0504 & \\ \cdots & \mathrm{CY} & ER & \mathrm{GR} & \mathrm{IQ} & \mathrm{LB} & \mathrm{SY} & \mathrm{Other} \\ & 0.0016 & 0.0087 & 0.0065 & 0.2163 & 0.0402 & 0.1432 & 0.0000 \end{bmatrix}$$

In this situation, Arab countries with a large population (Sudan, Iraq and Saudi Arabia) receive the highest weights because of their cultural proximity and the impact of their size. The weight of culturally dissimilar countries with relatively small populations (Cyprus, Greece and Eritrea) approaches zero despite their geographical nearness.

These simple examples demonstrate the flexibility of the gravity model in determining the differentiated weight of each neighbouring state. It is possible to adjust proximity and mass weights (in addition to the initial connectivity choice used to select the neighbours) to match the hypothesised channels of transmission that could be in action. But they also show the importance of carefully selecting the weights according to the specific phenomenon under study. If we assume that cultural similarity and a larger population are essential determinants of conflict contagion, Cyprus would have an influence on Egypt of only 0.16%; if we neglect these weights, its influence would unjustifiably raise to 7.7% (in the unweighted matrix). If we are studying the spillover of neighbouring economic policy and we judge that geographical proximity and the size of the foreign economy are the relevant factors, the influence of Saudi Arabia on Egypt would be 34.77%, compared to the 7.7% of the unweighted matrix.

In 2012, five out of the 13 neighbours of Egypt was undergoing an armed conflict: Iraq, Israel, Palestine, Sudan and Syria. By keeping 500 km as the connectivity choice, there are 12 combinations of proximity and mass weights proposed in this study. The resulting values for "neighbouring conflict" for Egypt in 2012 are included in Table 5.[26] They range from a minimum of 0.3206 (with culture and GDP as proximity and mass weights, respectively) to a maximum of 0.6409 (with inter-capital distance and population). If we relax the definition of a neighbourhood restricted to 500 km and we explore the complete 62 combinations available, the values of "neighbouring conflict" vary from just 0.1256[27] with

[26] Because of the construction of this variable, using row-standardised weights, it can vary from zero (a completely peaceful neighbourhood) to one (all neighbours are involved in conflict).
[27] This because the main trading partner of Egypt (China, United States, Italy, France, Turkey, Saudi Arabia and others) are mostly not in conflict, particularly when weighted by GDP.

w(world, trade, GDP) to 1.0000 for any of the weights that use alliances as proximity weights.[28] Clearly, the selection of appropriate weights, and robustness checks under alternative options, becomes a critical step of the analysis.

Table 5. Value of the variable "neighbouring conflict" under a connectivity choice of 500 km and different combinations of proximity and mass weights

Weights	Neighbouring conflict	Weights	Neighbouring conflict
w(500, none, none)	0.3846	*w*(500, capital, none)	0.4789
w(500, none, popn)	0.5742	*w*(500, capital, popn)	0.6409
w(500, none, GDP)	0.3261	*w*(500, capital, GDP)	0.4708
w(500, centroid, none)	0.4319	*w*(500, culture, none)	0.4471
w(500, centroid, popn)	0.6068	*w*(500, culture, popn)	0.6370
w(500, centroid, GDP)	0.4017	*w*(500, culture, GDP)	0.3206

Source: ESCWA calculations based on data as described in Appendix B.

Appendix B. Details on the sources of data and the construction of the variables

The empirical study described in Section V required a large database of domestic determinants of conflict. The sources of these variables as well as the transformations through which they entered the logistic regression model are summarised in Table 6. The rest of this appendix presents in more detail the definition of the variables, the main sources of the data and any pre-processing action that was applied to them.

Onset of internal armed conflict. This is obtained from the "Armed Conflict Dataset" jointly released by the Uppsala Conflict Data Program (UCDP) of the Uppsala University and Peace Research Institute Oslo (PRIO), version v.4-2014a. Only onsets of conflict, minor and major, are considered; i.e., cases where a country was at peace at period *t-1* and experienced the start of a conflict at year *t*. The second and subsequent years of a conflict were dropped from the database as they belong to a different situation in which a country that is at war either remains at war or transitions into peace (both of them outside of the area of interest of this study). Finally, only intra-state conflicts are considered since inter-state wars arguably have different determinants and follow different processes and should therefore be modelled separately. A long list of criteria was applied to assign conflicts to specific locations; this is available from the author on request.

[28] According to the data source used (see Appendix B), Egypt had two active alliances in 2012: a defence pact signed in 1990 with Yemen and a nonaggression pact signed in 1979 with Israel (the Egypt-Israel Peace Treaty signed after the Camp Davis Accords). Both of these countries were involved in armed conflict in 2012.

Table 6. Summary of variables included in the model and their sources

Variable	Construction and transformations	Source of original data
Internal conflict onset	None	UCDP/PRIO Armed Conflict Dataset
Population	Logarithm	UN DESA
Youth population	Proportion of population aged 15-24 relative to adult population	UN DESA
Development	Scores obtained from an unobserved components model on GDP per capita, a sub-composite index of health indicators (itself an unobserved components model on infant mortality rate, under five mortality rate and life expectancy) and an education indicator (average years of schooling). Logarithm of the first lag	GDP per capita: Penn World Tables, World Development Indicators, others. Infant mortality rates and under-five mortality rates: United Nations Inter-Agency Group for Child Mortality Estimation. Life expectancy: UN DESA. Average years of schooling: International Institute for Applied Systems Analysis, Barro and Lee.
State capacity	None	Arbetman-Rabinowitz and Johnson
Regime type	First lag, linear and squared	Polity IV
Ethnic polarisation	$4\sum_{i=1}^{P} p_i^2(1-p_i)$ where p_i is the proportion of each of the P people clusters relative to total population	Joshua project
Oil production per capita	Logarithm of (oil production per capita + 1)	United States Energy Information Administration, Wimmer and Min (2006), others
Mountainous terrain	None	Fearon and Laitin (2003), Wikipedia
Peace duration	$e^{-\text{peace years}/4}$	UCDP/PRIO Armed Conflict Dataset.

Population. This variable is obtained from “World Population Prospects: The 2012 Revision”, prepared by the United Nations Department of Economic and Social Affairs (UN DESA).

Youth population. Similarly, this variable is obtained from the UN DESA publication “World Population Prospects: The 2012 Revision.” It is defined as the ratio of the population aged between 15 and 24 years of age and to the total adult population.

Under-five mortality rate. This variable measures the probability of dying between birth and exactly five years of age per 1000 live births. The main source of data is the UN Inter-agency Group for Child Mortality.

Life expectancy. The data on life expectancy at birth measured in number of years is based on the "World Population Prospects: the 2012 Revision" published by UN DESA. A few countries are not included in this database so their data was imputed from the World Bank's World Development Indicators (WDI) with some limited interpolation.

Mean years of schooling for adults. This variable measures the number of years of formal schooling received on average by adults aged 25 and more. Most of the data comes from International Institute for Applied Systems Analysis (IIASA) and Vienna Institute of Demography (VID)'s "Reconstruction of population by age, sex and level of education for 120 countries for 1970-2000 using demographic back-projection methods". An additional 37 countries were sourced from Barro and Lee's "Educational Attainment Database". Even contemplating both sources, several countries were still lacking data. For 31 of these countries, the UNDP's Human Development Report Office (HDRO) has data based on the UNESCO Institute of Statistics, but for a limited number of recent years. These data points were used for each of the 31 countries, and the remaining trajectory was estimated by applying the growth rate observed in a country with similar results in this indicator from the same region. Since all sources are published at five-year intervals, the data was interpolated linearly to obtain a consistent annual dataset.

GDP per capita. The GDP per capita data is primarily based on the most recent version of the Penn World Table (PWT, version 8.0 published 2013). This source provides the purchasing power parity (PPP) adjusted GDP per capita at constant 2005 US dollars. Some countries are completely missing from the PWT 8.0 and were thus imputed from the K. S. Gleditsch's "Expanded Trade and GDP Data" version 6.0 (published in November 2013), which is based mainly on several version of the PWT and the Maddison Project Database. Both sources have information until 2011 and, in addition, the data for the last year was only provisional for some countries. For this reason, the rate of growth of GDP per capita in constant PPP USD obtained from the World Bank's WDI was applied to all countries for 2011 and 2012; for a few countries with no information in this source (Argentina, Cuba, Myanmar, Syria and Taiwan), the rate of growth for 2011-2012 was extracted from the Economist Intelligence Unit's CountryData.

State capacity. The source of this variable is the replication database of the paper by Arbetman-Rabinowitz and Johnson (2008), available from http://hdl.handle.net/1902.1/16845. Several alternatives for state capacity were available; the variable *rpe_gdp* was chosen since it applicability extend to a more diverse set of countries than the other options.

Regime type. This variable refers to the score in the variable polity2 of the Polity IV database for the year 2013, obtained directly from the Polity IV Project. This score varies from -10 (absolute autocracy) to 10 (absolute democracy). A country is normally considered to be a democracy if they have a score of six or higher, or an autocracy if it receives a score of -6 or lower.

Ethnic polarisation. The source of information to calculate this indicator is the Joshua Project, which intends to map the different peoples of the world along with some of their characteristics (main language, religion, ethnic group and inhabitants per country). This is an updated database that unfortunately does not provide a time series. Of the two variables referring to ethnicity, "people cluster" was preferred over “affinity bloc”, since the latter includes only very aggregated population groups that could mask the real diversity of the population of each country.

Oil production per capita. The starting source of data was the dataset “Production of crude oil including lease condensate” from the International Energy Statistics published by the United States Energy Information Administration (EIA), which provides information for the period 1980-2013. The only oil-producing country missing was South Sudan, which was imputed with data from the CIA World Factbook. The time series was completed by applying backwards the rate of change of the variable available in the replication dataset of Wimmer and Min (2006) and Wimmer et al. (2009). Some missing data was imputed with information from BP Statistical Review as well as diverse sources. The population variable required to calculate the oil production in per capita terms was obtained from UN DESA as described above.

Mountainous terrain. The replication dataset of Fearon and Laitin (2003) was used as the main source of information. For missing countries, the variable was imputed with by using the elevation span as predictor. This requires deleting data for which log(elevation span) < log(600) (and assign them a value of 0, as these observations were truncated by the original author, according to whom a mountain must be 600 meters above sea level, see http://www.stevepickering.net/mountains.asp) and performing a linear regression on the rest.

Peace duration. This refers to the number of years since the last internal conflict ended (or since1946, start of the database, in case there is no recent history of conflict). For the source of data, see the first variable “onset of internal armed conflict” above.

Minimum distance, inter-centroid distance and inter-capital distance. These measures were used to calculate the connectivity choice and the proximity weights. All are obtained from the CShapes R package version 0.4-2.

Bilateral trade. The main source of data for the period 2000-2012 was UN Comtrade Database published by the United Nations Conference on Trade and Development (UNCTAD). Some information missing from this database was imputed using the IMF’s Direction of Trade Statistics. The data was then filled backwards by applying the rates of change available from K. S. Gleditsch’s “Expanded Trade and GDP Data” version 6.0 (published in November 2013).

Political alliances. This variable refers to the existence of any type of alliance (defence, neutrality, nonaggression or entente) between two states. This comes from the Correlates of War Formal Alliance Dataset version 4.1, developed by Gibler (2009).

Total GDP. This variable was obtained by multiplying the GDP per capita described above by the total population from the UN DESA database.

Appendix C. Marginal effects

As described in Section V, the coefficients obtained in the logistic regression model can be used to study the direction and significance of the regressors, but they cannot be interpreted directly since this is a nonlinear model in which the effect of one variable depends on all the variables included in the regression. In order the facilitate the interpretation of the determinants of conflicts, Table 7 presents the marginal effects for each of the variables obtained by setting all other variables at their mean for each of the models included in Table 3.

Table 7. Marginal effects at the mean of the determinants of internal conflict onset

	With neighbouring conflict		
	w(w,culture,GDP)	*w(w,centroid,none)*	*w(w,culture,none)*
Population	0.0062	0.0061	0.0062
	(0.0006, 0.0491)	(0.0007, 0.0460)	(0.0006, 0.0477)
Youth population	0.0157	0.0057	0.0145
	(-0.0057, 0.2360)	(-0.0170, 0.1683)	(-0.0070, 0.2230)
Development	-0.0053	-0.0051	-0.0047
	(-0.0369, -0.0005)	(-0.0333, -0.0005)	(-0.0324, -0.0005)
State capacity	-0.0023	-0.0030	-0.0031
	(-0.0156, 0.0018)	(-0.0182, 0.0006)	(-0.0194, 0.0004)
Regime type	-0.0029	-0.0028	-0.0026
	(-0.0182, -0.0001)	(-0.0166, 0.0000)	(-0.0164, 0.0001)
$(\text{Regime type})^2$	-0.0069	-0.0072	-0.0073
	(-0.0443, -0.0008)	(-0.0446, -0.0009)	(-0.0458, -0.0009)
Ethnic polarisation	0.0143	0.0161	0.0153
	(0.0009, 0.1155)	(0.0013, 0.1219)	(0.0011, 0.1197)
Oil production per capita	0.0005	0.0006	0.0005
	(-0.0009, 0.0046)	(-0.0006, 0.0050)	(-0.0008, 0.0043)
Mountainous terrain	0.0086	0.0076	0.0092
	(-0.0002, 0.0669)	(-0.0009, 0.0595)	(0.0000, 0.0696)
Peace duration	0.0149	0.0159	0.0152
	(0.0017, 0.1007)	(0.0019, 0.1045)	(0.0018, 0.1028)
Neighbouring conflict	0.0230	0.0601	0.0418
	(0.0024, 0.1606)	(0.0067, 0.3937)	(0.0042, 0.2852)

Notes: The marginal effects were obtained by transforming the estimates of the coefficients according to the logistic function and fixed all other regressor at their mean. In parentheses, the 90% confidence interval of the marginal effects obtains by sampling 10 000 values from the estimated coefficients. The explanation of the determinants of conflicts is presented in Section V and Appendix B.