

Conflict Contagion – Predicting internal conflicts

By Tom Clarke and Marie Oldfield

Defence Science and Technology Laboratory, UK

Abstract

With an increased emphasis on upstream activity and Defence Engagement, it has become increasingly more important for the UK Ministry of Defence (MOD) and government to understand the relationship between conflict and regional instability. As part of this process, the Historical and Operational Data Analysis Team (HODA) in Defence Science and Technology Laboratory (Dstl) was tasked to look at factors that influenced the regional spread of internal conflicts to help aid the decision making of government.

Conflict contagion is the process by which a conflict in one state (State A) influences the outbreak of conflict in another state (State B). The aim of looking at these factors was to produce a tool to assist in the cost/benefit analysis of intervening in future conflicts.

The task was carried out by quantitatively studying a selection of contagion and non-contagion case studies over a set of 14 numerical variables that covered structural, political, socio-economic and cultural factors of both States A and B. All case studies took place after the end of the Cold War. The dataset used contained both binary and continuous data. A range of statistical techniques were used including correlation, regression and principal component analyses along with graphical analytical methods.

Unfortunately creating a tool to predict contagion has proven difficult and so the scope of the study was changed - as no statistically relevant link was found. However, it was shown that some links do exist between the variables and the existence of contagion, and it also appears it is a combination of variables which has an effect on the likelihood of contagion.

Conflict contagion is a rare event and was only recorded in less than 1% of cases present in the dataset. Of the variables analysed, the most significant factor was the number of refugees present in State B who originated in State A for the year before the case study.

It is recommended that further work be conducted. It is possible that a more detailed investigation, using a broader range of techniques and/or a wider selection of variables might be successful in assessing the likelihood of conflict contagion.

1. Introduction

1.1 Task

Dstl was tasked by and the Ministry of Defence (MOD), to develop a set of easily understood and widely applicable rules of thumb that link measurable variables with the likelihood of regional conflict contagion. These rules were to be used to assist with the cost/benefit analysis of conducting an intervention in future conflicts.

1.2 Technical approach

Conflict contagion is a wide ranging, loosely defined topic and so an approach that narrows down and clearly states what is being looked at was required. It is important to first define what the key terms mean within the scope of the study.

Conflict – ‘An armed conflict between a state and militarized rebels that has caused at least 25 cumulative related deaths’¹. Using the ‘best estimate for battle related deaths figure’ from Uppsala Conflict Data Programme (UCDP) Battle-Related Deaths Dataset².

Conflict contagion – A process by which a conflict in one state (State A) has a causal link to a conflict erupting in another state (State B). This process must take place while the conflict in State A is either still ongoing or within 5 years of the conflict in State A ending. Both states must also be present in the same regional neighbourhood as defined by Table 2 in Annex A.

¹ BLACK, N. 2013. When have violent civil conflicts spread? Introducing a dataset of substrate conflict contagion, *Journal of Peace Research* **50(6)**, 751-759.

² UCDP Battle-Related Deaths Dataset v.5-2014, Uppsala Conflict Data Program, www.ucdp.uu.se, Uppsala University.

State pairing – A pairing of a state in which conflict has/is taking place and a second state in which conflict is erupting that may (or may not) be a case of conflict contagion.

To re-iterate the above, this study is not examining when a given conflict expands to include other states either as participants or locations in which the conflict is fought.

The second aspect of the approach was to find a collection of cases where contagion is known to have happened. To this end the dataset created by Nathan Black³ of Harvard University was used. This contains 21,634 state pairings by year, of which he identified 122 that had confirmed⁴ causal link. Of these only 64 cases took place in the post-cold war period. These 64 cases compare to the 10,166 non-contagion cases that took place post-cold war.

The third aspect was to decide what variables to research and access to see if they had a relationship with the likelihood of conflict contagion. The variables have to be such that they are in, or can be converted into a numerical format and the data on them had to be publically available. The variables selected are in Table 1 below.

For the data collection within the study, the aim was to use as few sources as possible to reduce the disparity. However, due to the range of data topics, numerous different sources had to be used. The variables and main data sources used are listed in Table 1. These sources were picked mainly for their academic reputation but in some cases (regime type, ethnic fractionalization) these were the only places such data exists.

Variable	Notes ('year' refers to the year of possible contagion)	Main Data Source
Population size	Of State B in year.	World Bank Databank.
Conflict intensity	Number of deaths in State A's conflict in year, or closest preceding year if conflict over.	Uppsala Conflict Data Programme (UCDP) Battle-Related Deaths Dataset v.5.
Regime type	Of State B in year.	Polity IV Project using both POLITY and POLITY2 scores.
Regime type normalised	Of State B in year.	
Economic strength	GDP/capita of State B in year.	World Bank Databank.
Refugee flows	In State B from State A in year -1.	The Refugee Project.
Relative refugee flow	As percentage of State B population in year -1.	
Median Age	Of State B population in year.	From 2003 onwards CIA World Factbook before then UNdata.
Male/Female split	Percentage male State B population in year.	World Bank GenderStats.
Ethnic linkage	Do States A and B share a common ethnic group in year?	CIA World Factbook.
Size of shared ethnic group	In State B as percentage of total State B population in year.	
Political status of shared ethnic group	In State B in year.	Ethnic Power Relations dataset v.3.0.
Ethnic fractionalization	Of State B in year.	
Peace Years	Years since last war/violent conflict in State B.	UCDP Battle-Related Deaths Dataset v.5.
Intervention	Did State A's conflict involve international intervention in the preceding 5 years?	UCDP/PRIO Armed Conflict Dataset v.4.
Border effects	The type of border between State A and B.	Correlates of War (COW) Direct Contiguity Data v.3.1.
Trade relationship	Bilateral trade between States A and B as a percentage of total State B trade.	COW Trade Data v.3.0.

TABLE 1. List of data variables and main data sources used.

³BLACK, N. 2013. When have violent civil conflicts spread? Introducing a dataset of substrate conflict contagion, *Journal of Peace Research* 50(6), 751-759.

⁴ Ibid. Black confirmed this using by at least one academic source or area expert, this varied on a case by case basis.

1.3 Caveats

Due to the scope of this study and the datasets used there are caveats to the study's conclusion. The study used an indicative but not exhaustive list of case studies and the conclusions should be viewed with this in mind.

The study looked at only a limited number of variables in relation to the given state pairings. As such it is likely that some factors that may encourage conflict contagion have been missed.

For the cases where either State A or B was in the middle of an ongoing conflict the dataset is likely to contain inaccuracies due to the difficulty of collecting data in a conflict zone. In some cases the data may not exist at all.

The data regarding the ethnicities within a given state was one of the largest areas of difficulty. In some cases (Somalia, Haiti, Swaziland, Lesotho, Papua New Guinea) the ethnicity of the population was simply given as their nationality with no further ethnic breakdown. Within the data on African countries the ethnicity Bantu is often used, this covers 300-600 ethnic groups spread across Central and Southern Africa. This results in a misleadingly large joint ethnicity between some African states. In a few cases the racial mixture of white and black people is used for the ethnicity which is too broad ranging to be included in the analysis.

Within the timeframe of the study it has been impossible to thoroughly investigate every available resource. This means more accurate data might be found in the future that would necessitate an update of the study.

Some data sources used different definitions of 'refugee' and so contain different figures. When different data sources are used to collate the data there is a disparity in some of the results. This then affects the refugee flow and relative refugee flow variables.

Due to the nature of the data collected it is impossible to produce reliable confidence intervals.

2. Method

A structured method was used in the study which is shown in Figure 1 below, followed by a more in depth discussion of each stage.

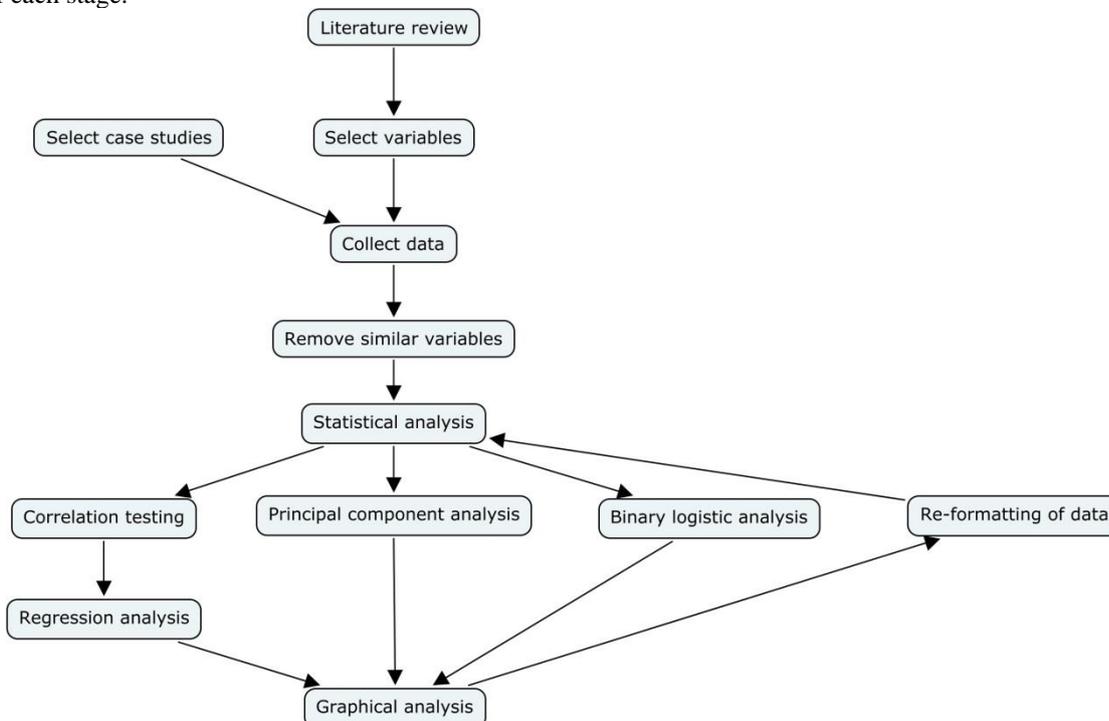


FIGURE 1. Method flow diagram.

A literature review was conducted to select the variables to be studied. Among the academic literature there is general agreement that an integrated environment of permissive conditions⁵ and proximate causes⁶ are important in explaining conflict contagion. Generally, variables fall into four categories as follows: structural, political, socio-economic and

⁵ Permissive conditions are underlying, often structural factors which make a state inherently prone to conflict.

⁶ Proximate causes are catalytic factors which tip a conflict-prone state into violence.

cultural. Variables from each of these categories were selected for analysis based upon results from the literature review, and where there were available datasets that could be analysed.

Sixty-four case studies of non-contagion conflicts were selected, out of a possible 10,166 state pairings.

The data collection was carried out using the main data sources listed in Table 1. During this process some issues were found with the data. The main issue identified was large inconsistencies found within the data gathered for the trade relationship variable. This resulted in this variable being discounted from the statistical analysis.

Once all the data was collated into a database the next stage was to start the statistical analysis. The data used fits into three distinct formats: binary, continuous and continuous within a set range. This meant no one statistical method was a perfect fit and so a range of methods were used.

The first step in the analysis was to decide for the three groupings⁷ of similar variables which one to select. In the first instance, regime type normalised was used as it normalised the outliers and placed all variables within a suitable range for analysis. For the refugees' grouping both variables were considered valid enough to be used in the initial statistical modelling. After the initial statistical model runs the most valid one was to be selected. For ethnicity the ethnic linkage was selected as this was considered the more reliable variable.

A correlation testing was carried out in order to detect any potential relationships. Having found one possible relationship between refugee flow and intervention, a regression analysis was carried out to see how much this variable contributed to the overall model.

Principal component analysis was then used to check how many principal components were present. Binary logistic regression was used on the binary variables, but a better way to interpret them was found in the form of visual output. Scree plots and loading plots were used for the graphical analysis of the data.

As the study progressed consolidation of the data formats was attempted - due to the issues with analysing data in a variety of different formats. The first was to convert all the continuous data into a binary format. The second was to convert all the continuous data into a quaternary⁸ format. Neither of these methods proved useful and were, therefore, not used during the rest of the analysis.

The data in a binary format was investigated separately in more depth, due to the fact contagion is binary. This was done visually with bar charts, main effect plots and interaction plots.

3. Findings

Conflict contagion occurs at a much smaller rate than is generally assumed. From the post-cold war dataset only 0.63% of cases have a causal contagion link. This equated to, within the dataset, only 1 state pairing in every 160 having a causal link.

The study was unable to find a statistically relevant link between the variables and the likelihood of a contagion event taking place. Some of the data suggests that there is a link between a combination of some of the variables and the likelihood of contagion. However, there is not enough statistical evidence to wholly support this finding.

Of the variables studied, refugee flow appears to be the most significant individual factor. However it only explains 15% of the variation within the dataset.

In cases where intervention did not take place and no ethnic linkage exists between states then the likelihood of conflict contagion was low. In cases where there was intervention and ethnic linkage there was a larger percentage of conflict contagion.

One example of a link found between variables is the international interventions took place in nine of the ten case studies with the highest refugee flow⁹. The only case where this did not happen had the lowest refugee flow of the ten cases. The cause and effect linkage is unclear. A large amount of refugees might produce a situation where international actors feel they have to intervene. Or an intervention might take place resulting in a widening of the conflict which then results in a larger number of refugees. A third and more likely option is these two cause and effects links form a feedback loop.

⁷ The three groupings are: regime type and regime type normalised; refugee flow and relative refugee flow; and ethnic linkage, size of shared ethnic group and status of shared ethnic group.

⁸ This is a format where all the data is split into equal quarters and coded with 0, 1, 2, 3 or 4.

⁹ Refugee flow is the number of refugees originating in State A present in State B in the year before the case study.

4. Conclusions

Conflict contagion is a rare event; sixty-four cases out of 10230 in the dataset (or 0.63%) were attributed to conflict contagion. This point should be kept in mind when looking at cases where it might occur in the future.

No statistically relevant link was found, as such creating a tool to predict contagion has proven beyond the scope of this study. It has been shown that some links exist between the variables and the existence of contagion. However, the development of more advanced statistical techniques that can compare binary to continuous data may prove a more effective way of devising a tool that can predict contagion.

The most significant variable was refugee flow which explains about 15% of the variation. This alone is not sufficient to predict conflict contagion. No single variable or simple cluster of variables could be combined to produce a viable method of predicting conflict contagion.

5. Recommendations

While this report does not provide a tool capable of estimating the likelihood of conflict contagion, it does, however, provide an air of optimism that such a tool could be created in the future.

There are a large number of variables that could be affecting the likelihood of conflict contagion, that were not included in the study. A selection of variables that could be analysed in the future, if the information can be sourced, are:

- The media links between two states. Both in terms of the passage of formal news information and that of social media. This would rely on both the interconnectivity within the given state and its links to the outside world. Some work has been done on this already in relation to new media¹⁰ and from a standalone point of view it does not seem to have much of an effect¹¹. This would apply to more recent cases studies only, as these variables were not as prevalent at the end of the Cold War as they are now.
- If there is any cross state linkage between various internal groups of the two states. This is an expansion of the border effects variable, to look more in depth at acts such as training given or weapons supplies traded.
- The corruption of government and political orientation of non-government groups in relation to State A. While the political systems of governance are measured by the Polity IV project these variables were not.
- The religious make-up of the two states looked at separately to ethnicity.

The trade relationship of the two states had to be ignored in this study due to inconsistency issues with the given dataset. If a new dataset could be sourced or created then the inclusion of this variable into the analysis could prove useful.

A major problem with the analysis was not having a suitable statistical technique for the mixed data types. If a technique can be found that allows for the comparison of the three data types used (binary, continuous and continuous with a set range) it should be applied to this dataset, as it looks like a combination of variables (which have different data formats) have an effect on the likelihood of contagion.

A future method could be to partner each case study of known contagion with a case study from the same regional neighbourhood (Table 2) where contagion did not take place. This would allow us to see if there is any regional effect on the likelihood of conflict contagion and allow a more like for like comparison of the countries in question.

This study did not engage international partners or academia directly. In the future these would be a good resource to use, both potentially providing new statistical methods and insights into areas not covered by this study.

REFERENCES

ADAY, S., FARRELL, H., LYNCH, M., SIDES, J., FREELON, D. 2012. Blogs and Bullets II: New Media and Conflict after the Arab Spring, *Peaceworks* 80.

¹⁰ Twitter, Facebook, YouTube, etc.

¹¹ ADAY, S., FARRELL, H., LYNCH, M., SIDES, J., FREELON, D. 2012. Blogs and Bullets II: New Media and Conflict after the Arab Spring, *Peaceworks* 80.

- ALESINA, A., DEVLEESCHAUWER, A., EASTERLY, W., KURLAT, S., WACZUARG, R. 2002. Fractionalization, *NBER Working Paper* **9411**.
- BARBIERI, K., KESHK, O., POLLINS, B. 2009. TRADING DATA: Evaluating our Assumptions and Coding Rules, *Conflict Management and Peace Science* **26(5)**, 471-491.
- BLACK, N. 2013. When have violent civil conflicts spread? Introducing a dataset of substrate conflict contagion, *Journal of Peace Research* **50(6)**, 751-759.
- CEDERMAN, L.E., WIMMER, A., MIN, B. 2010. EPR v. 1.1 'Why Do Ethnic Groups Rebel? New Data and Analysis', *World Politics* **62(1)**, 87-119.
- Central Intelligence Agency. 'The World Factbook' <http://www.cia.gov/library/publications/the-world-factbook/index.html> Last accessed 6/11/2014.
- Correlates of War Project. 'Direct Contiguity Data, 1816-2006. Version 3.0.' <http://correlatesofwar.org> Retrieved 6/8/2014.
- GLEDITSCH, PETTER, N., WALLENSTEEN, P., ERIKSSON, M., SOLLENBERG, M., STRAND, H. 2002. UCDP/PRIO Armed Conflict Dataset v.4-2014a. 'Armed Conflict 1946-2001: A New Dataset', *Journal of Peace Research* **39(5)**.
- GOHDES, A. & PRICE, M. 2012. First Things First: Assessing Data Quality before Model Quality, *Journal of Conflict Resolution*, **57(6)**.
- Google Earth 7.1.1.1888 (2003).
- MARSHALL, M.G., GURR, T.R., JAGGERS, K. 2014. Polity IV Project: Political Regime Characteristics and Transitions, 1800-2013, www.systemicpeace.org/polity4.htm Retrieved 8/7/2014.
- STINNETT, D.M., TIR, J., SCHAFER, P., DIEHL, P.F., GOCHMANN, C. 2002. The Correlates of War Project Direct Contiguity Data, Version 3, *Conflict Management and Peace Science* **19(2)**, 58-66.
- THEMNÉR, LOTTA, WALLENSTEEN, P. 2014. UCDP/PRIO Armed Conflict Dataset v.4-2014a. 'Armed Conflict, 1946-2013', *Journal of Peace Research* **51(4)**.
- The Refugee Project. www.therefugeeproject.org Last accessed 5/11/2014.
- The World Bank. 'GenderStats' <http://datatopics.worldbank.org/gender/> Last accessed 13/11/14.
- The World Bank. World Development Indicators: Population, total <http://data.worldbank.org/data-catalog/world-development-indicators> Retrieved 1/10/2014.
- UN. 'UNdata: Median age' <https://data.un.org/Data.aspx?d=PopDiv&f=variableID%3A41> Last accessed 14/11/2014.
- Uppsala Conflict Data Program. 'UCDP Battle-Related Deaths Dataset v.5-2014', Uppsala University. www.ucdp.uu.se Retrieved 8/7/2014.
- WIMMER, A., CEDERMAN, L.-E., MIN, B. 2009. Ethnic politics and armed conflict. A configurational analysis of a new global dataset, *American Sociological Review* **74(2)**, 316-337.

Annex A Regional neighbourhoods

Neighbourhood	States
North America	Canada, United States, Mexico
Central America	Mexico, Guatemala, Belize, Honduras, El Salvador, Nicaragua, Costa Rica, Panama (plus Cuba as a State A of all states, because of its influence in Latin America)
Caribbean	Cuba, Haiti, Dominican Republic, Jamaica, Trinidad and Tobago
South America	Colombia, Venezuela, Ecuador, Peru, Bolivia, Paraguay, Brazil, Argentina, Uruguay, Chile, Guyana, Suriname (plus Cuba as a State A of all states)
Western Europe	United Kingdom, Ireland, France, Spain, Portugal, Germany, Netherlands, Belgium, Denmark, Switzerland, Italy, Austria, Norway, Sweden, Finland
Eastern Europe	Poland, Czech Republic, Slovak Republic, Bulgaria, Romania, Serbia, Montenegro, Bosnia, Croatia, Macedonia, Slovenia, Greece, Hungary, Albania (plus Russia as a State A of all states)
Former Soviet Union (Eastern Europe)	Moldova (with Romania as a State A), Ukraine, Belarus, Latvia, Lithuania, Estonia (plus Russia as a State A of all states)
Former Soviet Union (Caucasus)	Russia, Georgia, Azerbaijan, Armenia
Former Soviet Union (Central Asia)	Kazakhstan, Uzbekistan, Turkmenistan, Kyrgyzstan, Tajikistan (plus Afghanistan, China, Pakistan, and Russia as State As of all states)
Near East	Turkey, Syria, Lebanon, Israel, Jordan, Egypt, Iraq, Iran, Cyprus
Persian Gulf States	Saudi Arabia, Kuwait, Qatar, Bahrain, United Arab Emirates, Oman, Yemen/North Yemen, South Yemen
Central Asia	Iran, Afghanistan, Pakistan
South Asia	India, Hyderabad, Pakistan, Nepal, Bhutan, Bangladesh, Sri Lanka (plus China as a State A of all states; plus Burma as a State A of India)
Southeast Asia	Burma, Thailand, Laos, Cambodia, Vietnam/North Vietnam, South Vietnam, Malaysia, Singapore, Indonesia, Brunei, Papua New Guinea, Philippines, East Timor (plus China as a State A of all states; plus India as a State A of Burma)
Northeast Asia	Mongolia, China, Taiwan, Japan, North Korea, South Korea (plus Russia as a State A of China, Mongolia and North Korea)
North Africa	Morocco, Algeria, Tunisia, Libya, Egypt, Mali, Niger, Chad, Mauritania
West Africa	Mauritania, Mali, Senegal, Gambia, Guinea-Bissau, Guinea, Sierra Leone, Liberia, Côte D'Ivoire, Burkina Faso, Ghana, Togo, Benin, Niger, Nigeria, Cameroon, Equatorial Guinea, Gabon
Central Africa	Chad, Sudan, Ethiopia, Eritrea, Somalia, Kenya, Uganda, Rwanda, Burundi, Democratic Republic of the Congo, Republic of the Congo, Central African Republic, Tanzania, Djibouti (plus Angola as a State A of both Congo's)
Southern Africa	Angola, Zambia, Zimbabwe, Malawi, Mozambique, Botswana, Namibia, South Africa, Lesotho, Swaziland, Madagascar, Comoros, Mauritius (plus D.R. Congo as a State A of Angola)

TABLE 2. Regional neighbourhoods.¹²

¹² BLACK, N. 2013. When have violent civil conflicts spread? Introducing a dataset of substrate conflict contagion, *Journal of Peace Research* **50(6)**, 751-759.