A Database on Third Party Intervention: Estimates of a Conflicts Salience, the Ascendant Power's Strength and The Expected Gains from Continued Fighting.

Francis MacDonnell
M.A. Norman Patterson School of International Affairs
March 8, 2004

The following document provides a description of each of the three workbooks contained in the file *Indicators.xls*. Each of these workbooks contains data that pertains to the construction of a database for testing the propositions of the 3rd party intervention model proposed in Carment and Rowlands, 1998. The document is divided into 3 sections.

Section 1: Provides a description of the quantitative data used, the source data sets, and the concepts behind them. All these indicators for the entire sample of conflicts can be found in the workbook entitled *quantitative proxies* in the file *Indicators.xls*.

Section 2: Provides a brief description and analysis of the survey data collected in order to assess subjective perceptions of the salience, strength and expected gains for each conflict in the sample. This data can be found in the workbook entitled *survey data* in the file *Indicators.xls*.

Section 3: Provides a comparison between the data used in sections 1 and 2, as well as concrete recommendations on how to best use this data in order to assign, for each of the three variable pertinent to a specific conflict (Salience, Strength, and Gains) a binary score of either high or low. The distilled data, based on these recommendations, can be found in the workbook entitled *Indicators* in the file *Indicators.xls*.

Table of Contents

Section 1: Quantitative Proxies

Pat Regan's database on 3rd party intervention (see Regan, 2002) provides the foundation upon which all these quantitative proxies have been generated, as it is from this database that the sample of conflicts has been drawn. This database is important in testing the model developed in Carment and Rowlands 1998 because it provides rich chronological information with regards to any intervention (type, origin, whether or not it was resisted, etc.) into all intrastate conflicts from 1946-1999. Thus, using the numerical code corresponding to each conflict, this information can be accessed through electronic queries, saving both times and effort.

This database however does not contain much information that would allow us to proxy the three independent variables in the Carment and Rowlands model. The goal here is to estimate, for each conflict in Regan's database, the strength of the ascendant military, the salience of the conflict, and the expected gains of the ascendant military from continued fighting. This has been accomplished by cross referencing conflicts in the sample to information from additional data sets include SIPRI's 'Armed Conflict Dataset' (Gleditsh et al., 2002), the ICB system level data (Brecher and Wilkenfeld, 2000), and the COW national material capabilities version 3 (Singer, 1987).

For a better understanding of the data that is available in Regan's data set, see his codebook. The following describes the additional information from other datasets, how it has been coded, and how it can be used to measure the Salience of a conflict, Expected Gains, and the Strength of an ascendant power.

1. i). Salience

The geo-strategic salience of a conflict is based on its repercussions to other countries. In this sense, an intrastate conflict only becomes salient if there is "spill over". This can take the form of a concurrent interstate crises involving the main actors of the conflict, or in more extreme case it can involve the internationalization of a domestic conflict.

Variable number 49 in the ICB system level dataset (described below) gives each crisis a score from 1 to 5 depending on how many systems it affects. The following explanation has been copied from the ICB code book.

Variable: 49
Name: GEOSTR
Location: Record 2, Column 55
GEOSTRATEGIC SALIENCE

This variable refers to the significance of the location of an international crisis in terms of its natural resources, distance from power centers, etc., measured by the level and number of international systems which are affected by a crisis.

Values and Illustrations

*(1) **One subsystem** (The Chaco II Crisis from 1932 to 1935 was salient to the South American subsystem only).*
*(2) **More than one subsystem** (The Indonesian Independence crises of 1945-47, 1947-48, and 1948-49 were salient to the Southeast Asian and West European subsystems).*
*(3) **Dominant system and one subsystem** (The Marshall Plan Crisis of 1947 was directly relevant to the East European subsystem, along with the dominant East/West system).*
*(4) **Dominant system and more then one subsystem** (The Munich Crisis of 1938 was highly salient to the dominant international system, as well as Eastern and Western Europe).*
*(5) **Global system** (The German attack on the Soviet Union in 1941, the Barbarossa Crisis, was salient to the global system).*

Based on this data, the following information was coded in worksheet entitled *quantitative proxies* in the file *Indicators.xls*.

ICB Crisis Codes: Give the numerical code of any ICB system level crisis corresponding to the conflict in question. For coding purposes a crisis was defined as concurrent with a given conflict if it occurred at the same time and was directly related to the conflict. Crises that had occurred after the conflict had begun (according to at least one source) but before it had reached Regan's start threshold of 200 fatalities, are also recorded as concurrent.

Salience (GEOSTR): Gives the GEOSTR value (from 1 to 5) of any ICB crisis related to the conflict, where there is no concurrent ICB crisis conflicts are given a score of 0.

1. ii). Strength

Data on the strength and capabilities of rebel groups is hard to find and often inaccurate. Thankfully, the model we are trying to test requires that we get information on strength for the ascendant military. We can assume that in most cases this would be the government (an assumption that can be tested by looking at "sizeopp" in Regan's data and comparing it with the "MilPer" variable in COW capabilities). In the cases where this assumption appears to be reasonable, we can draw from a wealth of data contained in the COW National Material Capabilities data-set which includes variables such as:

"MilPer" Military Personnel (thousands)
"MilEx" Military Expenditures, thousands of current year US Dollars.

Based on this data, the following information was coded in worksheet entitled *quantitative proxies* in the file *Indicators.xls*.

NMC Primkey: A numerical code corresponding to the row of the NMC spreadsheet in which the data for the country in question has been coded for the first year of a given conflict.

Strength (Milper): Number of military personnel (in thousands) the country's government had on the first year of the conflict.

Strength (Milthreat): Composite of a countries government's military personnel (in thousands) and expenditures (in thousands of U.S. dollars) for the first year of a conflict. It is the square root of *Milper*MilEx*.

1. iii). Gains

This is by far the trickiest variable to capture. One framework that does provide a way of determining whether the expected gains form continued fighting (as opposed to a negotiated settlement) are high or low is the Win Payoff Matrix (See Appendix B).
In order to calculate gains according to this model, the following variables were coded in the worksheet entitled *quantitative proxies* in the file *Indicators.xls*

sizeopp: Taken from Regan's data base it reflects the estimated size of the opposition forces. It should be noted that for this variable, data is incomplete for 44 of 100 observations.

Incompatibility: Taken for the SIPRI data-set (Gleditsh et al, 2002), it is a general coding of the conflict issue. The incompatibility of a conflict can either be over government (2) or over territory (1).

Territory: Taken for the SIPRI data-set (Gleditsh et al, 2002), this variable indicates the name of the contested territory if the value for incompatibility is 1. Because all conflicts in this sample are intrastate, this variable thus always refers to a territory that is attempting to secede.

a: Used in the Win Payoff Matrix, this value is the size of the territory (if any) listed in the previous variable, relative to the entire area of the state. The values are generally taken from the table Admin98.

p: Used in the Win Payoff Matrix, this value is equal to Milper/(Milper + sizopp)

C: A binary valuable representing the ascendant force. It is coded as 1 if the government is the ascendant for, and 0 if the opposition is (i.e. only 0 is sizopp > Milper)

Gains: The value for gains calculated using a win-payoff matrix, excluding the discount factor for secessionist conflicts (**a**).

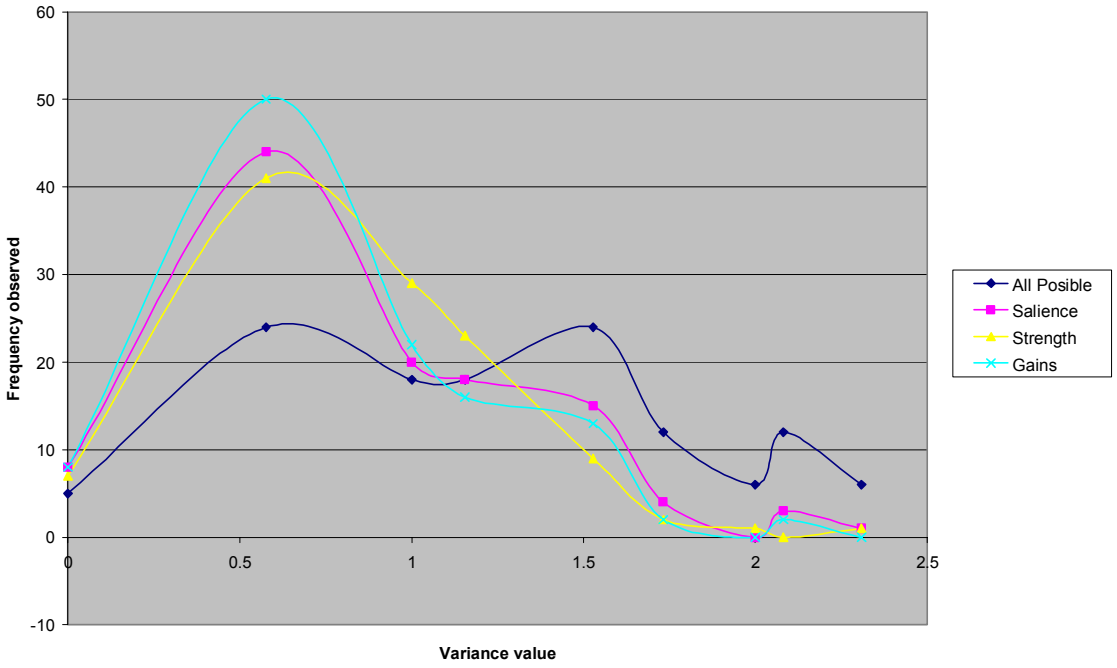
Gains (with area): The value for gains calculated using a win-payoff matrix as described in Appendix B.

Section 2: Survey Data

The Survey data presented in the workbook of the same name was collected based on the expert opinions of Drs. Carment, Rowlands and Choi. For each conflict, the respondent ranked the Salience, Strength of the ascendant power, and expected Gains from continued fighting on a scale of 1-5. The case of Kosovo conflict was used as a benchmark and given a score of 3-3-3.

The purpose of collecting this data was to provide a basis for some kind of comparison with the numeric proxies attempting to quantify the same variables. An analysis of the survey results lends credibility to the finding that there seems to be at least some general consensus among our three experts as to values for these three indicators. To illustrate this, we turn to Figure 1 (below)

Figure 1: Frequency Distributions of Variance



For each variable, in each conflict, the surveys recorded 3 values, one from each of our respondents. Thus for variable x in conflict y it is possible to estimate the variance (standard deviation) of the responses. Figure 1 plots the frequency distribution of these estimates of variance for each of the three variables over the sample of 112 conflicts. Compared to these three plots, the line marked "All possible" represent the plot of the variance for 125 observations representing all possible combinations of 3 numbers ranging from 1 to 5.

It is clear looking at Figure 1 that the estimates of Salience, Strength and Gains, produced by our respondents tended to have significantly smaller variance that one would expect

three randomly selected numbers between 1 and 5 to vary. From this we can conclude that there is at least some consensus among respondents as to which conflicts are of high or low salience/strength/gains.

Another test that was performed on the survey data to ensure that there were no significant problems with the assessments was to check for correlation between any of the variables in each respondent's survey. From a theoretical point of view, one would not expect any strong correlation between any of these three variables. Consequently any observed correlation could indicate that the assessment of a given variable affected the perceived value of another.

Table 1 below gives the R-squared values when a simple linear regression is performed between any two of the three variables assessed. The results show that, with the exception of 1 correlation between salience and gains, the 3 variables seem to have been coded with a relative degree of independence.

Table 1: R-squared Values of Simple Linear Regression Between any 2 of the Coded Variables.

Dr. Carment			
	Salience	Strength	Gains
Salience	1		
Strength	0.437352	1	
Gains	0.451181	0.427012	1
Dr. Choi			
	Salience	Strength	Gains
Salience	1		
Strength	0.485696	1	
Gains	**0.848291**	0.450728	1
Dr. Rowlands			
	Salience	Strength	Gains
Salience	1		
Strength	0.398639	1	
Gains	0.011363	0.009026	1

Section 3: Analysis and Recommendations

By comparing the proxies from section 1 to the corresponding survey data from section 2, I aim to provide concrete recommendations on how to best use all of this data to assign, for each conflict, a simple binary value of high or low for each of the 3 variables (Salience, Strength and Gains).

Summary of Recommendations

1. Salience should be coded using the GEOSTR score as a starting point, and editing the resulting list of highly salient conflicts to remove those that are judged to be, although salient to their "subsystem", located in a subsystem that is in and of itself simply not salient.

2. Strength is the one variable that can be objectively quantified. For this I would advocate relying almost exclusively on the data from the COW national material capabilities database, and in particularly the variable Milthreat.

3. Because the win-payoff matrix does not work well, the best solution is to code gains as high or low, according to the simple set of rules as established below.

3. i). Salience

The geo-strategic salience of a conflict is a variable that is inherently subjective. Consequently it is necessary in this case to attach importance to the survey data. While the proxy chosen (GEOSTR) does correspond, at least in some degree, to the salience of a conflict as subjectively assessed in the surveys (see table 3), it can be seen from Table 4 that other factors such as the region of the world in which the conflict is taking place have an equally important bearing on the subjectively assessed salience of a conflict.

Consequently I would recommend using this ICB GEOSTR data only as a beginning point in the assessment of the geo-strategic salience of a conflict. In other words I would recommend using the GEOSTR score unless a subjective judgment is made of a particular crisis, that although salient to its "subsystem" it is located in a subsystem that is in and of itself simply not salient. Thus, conflicts with a GEOSTR score ≥ 1 should be coded as of high salience, unless:

> *They have a GEOSTR score of 1 and the subsystem they affect is considered peripheral to any prospective intervener (including regional organizations).*

Conflicts with a GEOSTR score of 0 (i.e. no corresponding ICB crisis) be classified as of low salience.

Table 2: Descriptive Statistics for Various Salience Indicators

INDICATOR	Salience (Survey)	Salience (GEOSTR)	Salience (Binary)
MIN	1.000	0.000	0
MAX	4.667	5.000	1
AVERAGE	2.440	0.714	0.429
STDEV	0.811	1.094	na

Table 3: Distribution of GEOSTR and Corresponding Average Survey Scores

ICB GEOSTR SCORE	0	1,2	3,4,5
Number of Conflicts	64	35	13
Average Survey Salience	2.20	2.58	3.23

Table 4: Average Survey Scores for Conflicts Located in Different Geographic Locations

	Americas	Africa	Rest of the World
Nuber of Conflicts	13	44	55
Average Survey Salience	2.74	2.04	2.69

3. ii). Strength

Unlike salience, the strength of the ascendant military is relatively easily quantified through objective means. Consequently I would advocate relying almost exclusively on the data from the COW national material capabilities database. Furthermore, because military might is a combination of manpower and resources, I would advocate using the variable Milthreat, which is an aggregate of military personnel and expenditure. Using only Milper would lead to a situation where the strength of countries like Ethiopia, which had a huge but poorly equipped army, would be overestimated.

If Milthreat is used to proxy the strength of the ascendant belligerent, one only needs to pick a cut-off value below which a country will be said to be of low strength. This is involves some sort of subjective decision. The simplest way of doing it is to decide how may observations we want to classify as high strength, and use the value of the n-th largest observation as our cut-off. For example, if we use the mean value of Milthreat (12688) as the cut-off, we have 23 conflicts where the ascendant belligerent is of high strength and 89 where it is of low strength. Using the median value of 1920 would give us both 56 high strength and 56 low strength belligerents. For the data I have prepared, I used Milthreat = 4000 as the cut-off for high strength. This gives us 40 conflicts where player C had high strength.

Table 5: Descriptive Statistics for Various Strength Indicators

INDICATOR	Strength (Survey)	Strength (Milper)	Strength (Milthreat)
MIN	1.00	1	45
MAX	4.33	4158	367927
AVERAGE	2.07	205	12688
STDEV	0.61	582	38995

Table 6: Distribution of Milper and Corresponding Average Survey Scores

Milper Value	≤ 10	50 ≥ x > 10	200 ≥ x > 50	≥ 200
Number of Conflicts	30	31	26	25
Average Survey Strength	1.81	1.91	2.06	2.57

Table 7: Distribution of Milper and Corresponding Average Survey Scores

Milthreat Value	< 100	1000 > x ≥ 100	10,000 > x ≥ 1,000	100,000 > x ≥ 10,000	≥ 100,000
Number of Conflicts	4	42	39	24	3
Average Survey Strength	2.17	1.83	1.97	2.42	3.67

3. iii). Gains

Tables 8 shows that there is a huge disconnect between the survey scores for expected gains and the two quantitative proxies developed. Furthermore because of gaps in Regan's data for the variable "sizeopp", these indicators could only be generated for 68 of 112 conflicts. Clearly another approach is warranted.

Table 8: Distribution of Gains and Corresponding Average Survey Scores

Gains Value:	< 0.1	0.1 ≤ x < 0.2	0.2 ≤ x < 0.3	≥ 0.3
Number of Conflicts:	26	32	4	6
Average Survey Gains	3.06	3.09	3.33	2.94

At the same time, as the analysis of the survey results has shown, subjectively quantifying gains can also be problematic. The solution that I envisage is that gains be coded merely as high or low, according to the simple set of rules below.

If people are willing to go to war over a subject, it is clearly a matter of great importance to them. Saying that an internal conflict is an important issue to a party involved does not establish high expected gains, because it is a truism. The question becomes what is the

magnitude of the threat posed to the ascendant force's goals/power/order by the opposition? And therefore, to what extent would the ascendant power gain from a total victory of this opposition as opposed to seeking a negotiated solution (and by so doing acknowledging the legitimacy of this opposition)? Based on this I would propose the following set of rules for determining which conflict have particularly high gains.

1) Assume government is player C unless there is clear evidence based on Regan's variable "sizeopp" that they are not.
2) **If** government is player C **and** the conflict is secessionist, code the conflict as low gains **unless** the state involved has a history of or is likely to incur significant fragmentation beyond this particular territorial dispute (Balkanize), in which case it should be coded as high gains.
3) **If** the government is player C **and** the conflict is over government, code it high gains **unless** there is a strong case to be made that the insurgents in question pose little more than a nuisance level threat to the government's current power (e.g. the LRA in Uganda).
4) **If** the insurgents are player C, we code the gains as high because regardless of their objective, they stand a realistic chance of achieving it, and would not want to settle for negotiations based on the *status quo ante*.

Appendix A: List of Quantitative Proxies

ICB Crisis Codes: Give the numerical code of any ICB system level crisis corresponding to the conflict in question. For coding purposes a crisis was defined as concurrent with a given conflict if it occurred at the same time and was directly related to the conflict. Crises that had occurred after the conflict had begun (according to at least one source) but before it had reached Regan's start threshold of 200 fatalities, are also recorded as concurrent.

Salience (GEOSTR): Gives the GEOSTR value (from 1 to 5) of any ICB crisis related to the conflict, where there is no concurrent ICB crisis conflicts are given a score of 0.

NMC Primkey: A numerical code corresponding to the row of the NMC spreadsheet in which the data for the country in question has been coded for the first year of a given conflict.

Strength (Milper): Number of military personnel (in thousands) the country's government had on the first year of the conflict.

Strength (Milthreat): Composite of a countries government's military personnel (in thousands) and expenditures (in thousands of U.S. dollars) for the first year of a conflict. It is the square root of *Milper*MilEx*.

sizeopp: Taken from Regan's data base it reflects the estimated size of the opposition forces. It should be noted that for this variable, data is incomplete for 44 of 100 observations.

Incompatibility: Taken for the SIPRI data-set (Gleditsh et al, 2002), it is a general coding of the conflict issue. The incompatibility of a conflict can either be over government (2) or over territory (1).

Territory: Taken for the SIPRI data-set (Gleditsh et al, 2002), this variable indicates the name of the contested territory if the value for incompatibility is 1. Because all conflicts in this sample are intrastate, this variable thus always refers to a territory that is attempting to secede.

a: Used in the Win Payoff Matrix, this value is the size of the territory (if any) listed in the previous variable, relative to the entire area of the state. The values are generally taken from the table Admin98.

p: Used in the Win Payoff Matrix, this value is equal to Milper/(Milper + sizopp)

C: A binary valuable representing the ascendant force. It is coded as 1 if the government is the ascendant for, and 0 if the opposition is (i.e. only 0 is sizopp > Milper).

Gains: The value for gains calculated using a win-payoff matrix, excluding the discount factor for secessionist conflicts (**a**).

Gains (with area): The value for gains calculated using a win-payoff matrix as described in Appendix B.

Appendix B: Win Payoff Matrix

Assumption 1:
The value **p** proxies the relative size of the government powerbase be it through popular support or economic clout relative to that of the entire country. This assumption is reasonable if we assume that in a conflict where vital security interests are threatened, rational actors marshal most of the resources at their disposal.
Let: **p** equal

$$\frac{\text{Government Forces (“MilPer”)}}{\text{Government Forces (“MilPer”) + Insurgent Forces (“sizeopp”)}}$$

where $0 \leq \mathbf{p} \leq 1$

Assumption 2:
The probability of the government forces winning in any given time period (**PV**) is a function of **p**. It is equal to 1 when the government has all the force (**p** = 1) and 0 when the government has none of the force (**p** = 0). Between these intervals, **PV** < **p** because even with superior force your odds of an outright victory in any specific time period should be lower than your numerical force odds. Hence a reasonable approximation of this function is $\mathbf{PV} = \mathbf{p}^2$.

From the point of view of the combatant in a civil war, we can conceive of 3 possible outcomes, a military loss, a military victory, and a negotiated settlement. An outright win by either side would be unacceptable to any even partially neutral intervener, who would seek to facilitate or impose a negotiated settlement of some kind.

Thus, from the point of view of player C in the game modeled in the JCR paper, its choices are between a win (continued fighting) and a negotiated settlement (accommodate I). A loss is inconceivable because player C is the ascendant force and it would always choose to accommodate player I before contemplating a loss. Therefore we can conceive of C’s gains as the difference between a win and a negotiated settlement. Which, depending on the type of conflict and the nature of player C can be expressed in the payoff matrix below.

	Revolutionary War	Secessionist War
Government is C	***(1 – p) PV***	***a (1 – p) PV***
Insurgents Group is C	***p (1 – PV)***	***p (1 – PV)***

Note: this payoff matrix is consistent with arguments made by Mason, Weingarten, and Fett that (from a governments'perspective) revolutionary wars are for higher stakes than secessionist wars; because in the latter the state can be neatly partitioned and still remain viable[1].

Revolutionary Wars

In a revolutionary war, the very existence of the government is brought into question. A win by C, if C is the government, implies that they retain 100% of the power of the state. On the other hand if C is the insurgent group, a win means that they gain 100% control of the state. A negotiated settlement however would imply some form of power sharing between the two groups in a new government.

Share of Power

Government		Insurgents	
Win	Negotiated Settlement	Win	Negotiated Settlement
1	**p**	1	(1 - **p**)

We assume that a negotiated settlement will involve some sort of democratic model, or at least a division of power within the government based on proportional representation. So if the government is C it stands to retain a share of power of **p** in the event of a negotiated settlement and of **1** in the event of a win. Conversely the rebel group stands gain a share of (1 – **p)** in the event of a settlement and **1** in the event of a win.

Each side's respective payoff being the difference between a win and a negotiated settlement (discounted by the probability of victory PV) we have the government's payoff a being (1 – **p**) PV and the rebels as being **p** (1 – PV).

Secessionist Wars

In a secessionist war, the government's existence is not in question, simply its control over a portion of its territory. Thus, the state can be neatly partitioned and still remain viable[2]. Let ***a*** denote the area of the area in dispute divided by the total area of the state. A negotiated settlement in this case would involve some form of regional autonomy or federalism, but not the creation of a sovereign state.

We can therefore apply the same logic to this conflict as to that of a revolutionary conflict. Simply discounting the expected gains for a government C by **a** the proportion of the state represented by the secessionist territory

[1] Madson et al. 1999.
[2] Ibid

Works Cited

Brecher, Michael and Jonathan Wilkenfeld. A Study of Crisis. Ann Arbor: The University of Michigan Press. 2000.

Carment, David and Dane Rowlands. "Three's Company: Evaluating Third-Party Intervention in Intrastate Conflict". *Journal of Conflict Resolution,* Volume 42, Issue 5 (Oct, 1998), 572-599.

Gleditsch, N.P., Havard Strand, Michael Eriksson, Margareta Sollenberg, and Peter Wallensteen. "Armed Conflict 1946-2001: A New Data Set" Journal of Peace Research. Vol. 39, No. 5, 2002 pp. 615-637.

Madson, David T., Joseph P. Weingarten, and Patrick J. Fett. "Win, Lose, or Draw: Predicting the Outcome of Civil Wars" Political Research Quarterly, v.52, No.2, 1999 pp. 239-268.

Regan, Patrick M. "Third Party Interventions and the Duration of Intrastate Conflict", *Journal of Conflict Resolution*, Volume 46, Issue 1 (Feb, 2002), 55-73.

Singer, J. David. (1987). "Reconstructing the Correlates of War Dataset on Material Capabilities of States, 1816-1985" *International Interactions*, 14: 115-32.