Supplementary Information (SI)

"Views to a war: systematic differences in media and military reporting of the war in Iraq"

Karsten Donnay* and Vladimir Filimonov, ETH Zürich, Switzerland

Email: Karsten Donnay*- kdonnay@ethz.ch; Vladimir Filimonov - vfilimonov@ethz.ch;

*Corresponding author

1 Data

In our analysis we rely on detailed event data from two Iraq-specific datasets: Iraq Body Count (IBC), a web-based data collection initiative administered by Conflict Casualties Monitor Limited (London) [1], whose data can accessed at http://www.iraqbodycount.org, and U.S. military (SIGACT) data downloaded from *The Guardian* website [2]. In this section we outline details on data format and preparation.

The IBC database records violent events resulting in civilian deaths from January 1, 2003 onward with records updated continuously until the present day. Our analysis relies on the publicly available version of the IBC records that does not disaggregate by perpetrator group. Data used in this study was downloaded on November 15, 2011 and provides the following information on each incident: (i) a unique "IBC code", (ii) "Start date" and "End date" of the incident, (iii) "Time" information (if known), given either as time of day with resolution of half an hour (e.g. 9:30 AM), or as time interval (9:00–10:00 AM) or as approximate time of the day (AM or PM). Each data entry also contains (iv) a verbal description of the "Location" (e.g. "al-Thaqafiyah, north of Mosul"), (v) information on the "Target" (e.g. "civilian car driven by mobile phone store owner") and (vi) which "Weapon" (e.g. "magnetic bomb attached to car") was used. The (vii) number of casualties is given as a range between "Reported minimum" and "Reported maximum". Finally, IBC provides (viii) a "Source" field with the name of the news source(s) used to code the incident.

The IBC dataset contains a number of events with a one month interval between "Start Date" and "End Date". Generally, the "End Date" of these entries falls on the last day of the month and the entries are usually recognizable as aggregate monthly casualty counts because the event location is coded, for example, as "19 Baghdad hospitals". Though the number of civilian fatalities reported in such aggregated counts can be quite large (up to several hundreds in early 2006–2008), we excluded them from our analysis because they do not code individual, recognizable conflict events. For the same reason we excluded all events in the IBC dataset where "Start Date" and "End Date" fields differ by more than a day. Note that this amounts to excluding less than 1.5% of all entries in our period of analysis.

In order to reliably extract the location information in the IBC dataset we used a comprehensive dictionary of locations in Iraq that codes hamlets, villages, city quarters etc. to the city or settlement in the direct geographic proximity. This, of course, also allows for an efficient extraction of the Baghdad subset that our analysis rests on. The automated dictionary-based routine recognizes over 99% of IBC locations—we then additionally ensured that none of the entries that could not be automatically location-coded corresponds to locations in Baghdad. As outlined in the article we further restricted our analysis to the period June 1, 2004 to February 28, 2009—a period covered by both datasets without any gaps. We provide this data in a .csv file that contains the "IBC code", "Start Date", "End Date", "Time", "Reported Minimum" and "Reported Maximum" of civilian casualties for each incident. In our analysis we did not use the "Time" information as it is only available for a small subset of events. All events therefore carry a "00:00" timestamp. Note further that, as detailed in the article, we used the "Reported Minimum" of casualties for our analysis since it is the more conservative estimate. Also, where "Start Date" and "End Date" of events differ we use "Start Date" to mark the timestamp of events. In section 3 of this supplementary information we demonstrate that none of these coding choices affect our substantive findings.

The data made available through *The Guardian* contains information on all "significant actions" (SIGACTs) reported by units of the U.S. military in Iraq that resulted in at least one casualty. The dataset covers the period January 1, 2004 until December 31, 2009 but is missing 2 intervals of 1 month length each (from April 30, 2004 to June 1, 2004 and from February 28, 2009 to April 1, 2009) [2], which restricts our period of analysis to the period June 1, 2004 to February 28, 2009 (see also above). Data used in this study was downloaded on September 3, 2013 and provides the following information for each incident: (i) the "Report Key", (ii) its "Date and time" with a resolution up to minutes, (iii) the "Type" of incident (e.g. "Explosive Hazard" or "Enemy Action"), (iv) a "Category" of events the incident is coded to (e.g. "Attack" or "Raid"), (v) the "Title" of the incident with detailed information on its occurrence, (vi) the military regional command or "Region" the incident was reported in, (vii) information on the target of the attack coded as "Attack on" either "NEUTRAL", "ENEMY" or "FRIEND", (viii) casualty counts—both *killed-in-action (KIA)* and *wounded-in-action (WIA)*—disaggregated by "Coalition forces", "Iraq forces", “Civilians” and “Enemy”, (ix) the total number of casualties and (x) the longitude and latitude of where the incident was reported. These geo-coordinates are truncated at a tenth of a degree (about 10 km) for Iraq outside of Baghdad and at a hundredth of a degree (about 1 km) for the military zone of Baghdad. In order to be able to compare it to IBC we restricted the SIGACT data to entries pertaining to deadly violence directed at civilians. As outlined in the article, focusing only on civilian casualties rather than also including incidents that wounded civilians may lead to a biased view of the violence dynamics. To control for this, we performed robustness checks in which we additionally included the number of wounded civilians reported in SIGACT. These results are provided in section 3 of this supplementary information demonstrating that this does not affect our substantive conclusions.

In selecting for events in the Baghdad area we rely on two different criteria outlined in the article. On the one hand we use the U.S. military’s definition of the greater Baghdad area and the corresponding regional command “MND-BAGHDAD”. We also performed each of our analysis for subdatasets generated by selecting all events that fall within a radius of 20 km, 30 km and 40 km from the city center (LON 44.422, LAT 33.325). These four dataset are provided in separate .csv files that contain the “Report key”, “Date”, “Latitude”, “Longitude”, “Region”, “Coalition forces wia”, “Coalition forces kia”, “Iraq forces wia”,“Iraq forces kia”, “Civilian wia”,“Civilian kia”, “Enemy wia” and “Enemy kia” for each incident. Notice that any detailed information on the type of event, target and details on the incident have been intentionally removed from these data.

Note that SIGACT data on Iraq was already published at the time we downloaded the corresponding IBC records. In principal, IBC records may thus have been updated and/or added based on these new informations. In fact, IBC did analyze the correspondence of the casualty records with SIGACT data in detail in 2010 (see http://www.iraqbodycount.org/analysis/numbers/warlogs/). If SIGACT information did indeed enter the IBC database it at best led to a better correspondence of the two datasets and at most our comparative analysis may thus provide a more conservative estimate of the original reporting differences.

2 Event matching algorithm

In section 3.3 of the article we compare the day-by-day match of SIGACT to IBC events using an automated event matching algorithm. Note that we group events with a given casualty count (s) in broad categories and then match each category independently. Specifically, we consider the following categories: $S_1 = \{1\}$, $S_2 = \{2, 3\}$, $S_3 = \{4, 5, 6\}$, $S_4 = \{7, 8, 9, 10\}$, $S_5 = \{11, 12, \ldots, 19, 20\}$ and $S_6 = \{21, 22, \ldots\}$.

Given that the resolution of IBC is days, i.e, events all carry the timestamp "00:00", we also round SIGACT to daily resolution for this comparison. The matching algorithm then proceeds as follows. For each SIGACT event at date t_{SIGACT} in given category S, we select all IBC events within the same size category and with dates in the range $t_{SIGACT} - w + 1 \leq t_{IBC} \leq t_{SIGACT} + w$, where w is the allowed tolerance in days. $w = 1$ then selects only IBC entries that are recorded on the same calendar day as the SIGACT event. For $w = 2$ we consider all events on the same day and on the previous and subsequent day, i.e., ± 1 days timestamp uncertainty. Similarly, $w = 3$ allows ± 2 days of uncertainty, etc. Among these possible matches, we then randomly select one IBC event (without replacement) and mark the original SIGACT event as "matched" in our records. This procedure is repeated for the next unmatched event in the SIGACT database wherein only previously unmatched IBC events are considered (because we selected without replacement).

Once all SIGACT events are processed, we count the number of events per month that could be successfully matched. In order to avoid possible suboptimal solutions through our random "matching" algorithm, we use a Monte-Carlo approach: we simply repeat the random matching procedure 100 times and then select the best match achieved. The method is significantly faster than considering all possible combinations, and at the same time provides similar results. For larger windows w we, of course, expect to obtain a better match. For the article we considered $w = 2$, which most closely corresponds to the manual matching prescription used in a study performed at Columbia University [3] where IBC events were matched to SIGACT entries within 24h prior and 48h following the IBC event. Note that we also alternatively centered our search for matches on SIGACT instead of IBC entries using the full SIGACT timestamp. We find that this has no systematic effect on the quantitative results.

The results for $w = 2$ are discussed in the article. Table S1 summarizes the results of matching SIGACT events to IBC using $w = 1$, i.e., only considering events reported on the same date. Table S2 presents results for $w = 4$, which allows ± 3 days of uncertainty in timestamps. Decreasing the timestamp tolerance significantly decreases the number of events that can be matched, while increasing it improves the quantitative match, as expected. Interestingly, for extreme events ($s > 20$) in 2004–2005 and 2008–2009 and for very

Table S1: Number of SIGACT reports matched to IBC entries, $w = 1$

Casualties	2004-05 & 2008-09			2006-07		
	matched	total	%	matched	total	%
s = 1	1166	1473	79.15	2890	11871	24.34
s = 2, 3	278	417	66.66	1479	3054	48.42
s = 4–6	75	133	56.39	420	693	60.60
s = 7–10	17	45	37.77	125	202	61.88
s = 11–20	16	36	44.44	69	143	48.25
s > 20	15	23	65.21	47	67	70.14

Table S2: Number of SIGACT reports matched to IBC entries, $w = 4$

Casualties	2004-05 & 2008-09			2006-07		
	matched	total	%	matched	total	%
s = 1	1316	1473	89.34	2942	11871	24.78
s = 2, 3	375	417	89.92	1579	3054	51.70
s = 4–6	97	133	72.93	518	693	74.74
s = 7–10	28	45	62.22	161	202	79.70
s = 11–20	18	36	50.00	97	143	67.83
s > 20	15	23	65.21	61	67	91.04

small events ($s = 1$) during the escalation of the conflict in 2006–2007, the quality of matching remains almost unchanged for different timestamp uncertainties.

Note that the matching results reported thus far are always expressed as the fraction of SIGACT reports. The analysis in the article, however, suggests that especially for large events IBC reports significantly more events than SIGACT. We have thus also considered the matches for $w = 2$ expressed as fraction of IBC entries (Table S3). Note that we here correspondingly centered our search on IBC rather than SIGACT events. The high match of IBC entries with few casualties and the low match of IBC entries with many casualties in the period 2006–2007, simply reflects the fact that IBC reports substantially less small events and more large events than SIGACT respectively. The generally lower match in the other periods simply reflects the fact that there IBC overall reports more events than SIGACT.

Table S3: Number of IBC entries matched to SIGACT reports, $w = 2$

Casualties	2004-05 & 2008-09			2006-07		
	matched	total	%	matched	total	%
s = 1	1225	1757	69.72	2921	2974	98.21
s = 2, 3	314	630	49.84	1544	2019	76.47
s = 4–6	83	202	41.08	456	680	67.05
s = 7–10	18	74	24.32	134	257	52.14
s = 11–20	18	59	30.50	76	167	45.50
s > 20	15	34	44.11	54	151	35.76

3 Sensitivity Checks

We performed extensive sensitivity checks in order to guarantee that the substantial findings reported in the article do not depend on particular coding choices. Wherever applicable we report the results for each of the following variations of our data (see section 1 of this supplementary information for details):

(a) instead of the start date of an event in IBC we use its end date as timestamp (if these are different)

(b) instead of the lower IBC casualty estimate we use the upper casualty estimate

(c) instead of civilian KIA we consider civilian KIA + WIA in the SIGACT dataset

(d) instead of "SIGACT Baghdad" we use "SIGACT 20km", "SIGACT 30km" or "SIGACT 40km", i.e. the datasets that cover all events in a 20, 30 or 40 km radius around Baghdad.

The sensitivity checks are grouped according to the corresponding figures and tables in the article. Note that we only report tables or figures for results that differ noticeably from those presented in the article.

Table 2

In Table 2 of the article we show a detailed comparison of the total number of events in IBC and SIGACT and used a two-sample Anderson-Darling test to evaluate their quantitative agreement. The results in Table S4 and S5 confirm that for data variations (b) and (c) the pairwise comparison of the distribution of casualties in SIGACT and IBC does not differ substantially from those reported in the article. For large events (threshold of 40 and more casualties) we find a slightly improved distributional agreement for (c), simply because SIGACT KIA + WIA contains more events with many casualties than SIGACT KIA. Data variation (a) does not affect the aggregate statistics and (d) is already accounted for in the table.

Table S4: Results of the pairwise comparison of the distributions of casualties. The datasets are (i) "IBC Baghdad", (ii) "SIGACT Baghdad", (iii) "SIGACT 20km", (iv) "SIGACT 30km" and (v) "SIGACT 40km". We used a two-sample Anderson-Darling tests (adjusted for ties) for comparison (see the caption for Table 2 of the article for details), data variation (b)

Threshold	Number of events					A^2 statistic			
	(i)	(ii)	(iii)	(iv)	(v)	(i)-(ii)	(i)-(iii)	(i)-(iv)	(i)-(v)
1	9068	18157	17533	18548	19369	1275.05	1279.05	1273.51	1268.11
2	4442	4813	4611	4940	5201	126.03	122.69	130.00	128.16
5	1284	876	851	901	952	8.25	8.87	9.73	9.81
10	548	323	310	325	340	7.20	6.71	6.60	6.69
15	335	159	154	161	169	**1.10**	**1.04**	**0.86**	**1.10**
20	227	105	100	105	108	**1.61**	**1.20**	**1.03**	**0.98**
25	173	77	75	79	82	**2.30**	**2.05**	**1.80**	**1.87**
30	135	47	47	51	52	**1.54**	**1.54**	**1.37**	**1.39**
40	79	29	29	31	32	**2.41**	**2.41**	**2.60**	**2.46**

Table S5: Results of the pairwise comparison of the distributions of casualties. The datasets are (i) "IBC Baghdad", (ii) "SIGACT Baghdad", (iii) "SIGACT 20km", (iv) "SIGACT 30km" and (v) "SIGACT 40km". We used a two-sample Anderson-Darling tests (adjusted for ties) for comparison (see the caption for Table 2 of the article for details), data variation (c)

Threshold	Number of events					A^2 statistic			
	(i)	(ii)	(iii)	(iv)	(v)	(i)-(ii)	(i)-(iii)	(i)-(iv)	(i)-(v)
1	9004	18504	17854	18919	19782	359.92	381.87	355.70	328.97
2	4273	6313	6013	6477	6877	9.49	9.80	8.97	9.49
5	1163	1880	1795	1922	2052	27.87	29.26	25.90	24.07
10	484	992	957	1010	1067	8.70	9.18	8.27	7.03
15	296	675	653	682	715	3.81	4.08	3.91	3.16
20	206	503	490	509	526	**1.44**	**1.35**	**1.47**	**1.32**
25	159	392	382	394	406	**1.34**	**1.25**	**1.42**	**1.33**
30	123	294	287	299	307	**2.59**	**2.34**	**2.45**	**2.37**
40	69	175	168	176	180	3.82	3.89	4.14	4.04

Table 4

The results in Table S6 to S11 confirm that the day-by-day correspondence of IBC and SIGACT (Table 4 of the article) does not critically depend on data variations (a), (b) and (d). However considering both KIA and WIA events in SIGACT (variation (c)), results in a slight improvement in the day-by-day match of small events ($s = 1$) and at the same time significantly decreases the match for large events ($s > 7$) compared to the analysis reported in Table 4 of the article. Considering KIA+WIA thus does not make IBC and SIGACT more consistent.

Figure 3

Data variation (a) has by definition no influence on the aggregate casualty statistics, and (b) and (d) do not result in significant changes to Figure 3 of the article. We would expect variation (c) to affect the overall

Table S6: Number of SIGACT reports matched to IBC entries, data variation (a), $w = 2$

Casualties	2004-05 & 2008-09			2006-07		
	matched	total	%	matched	total	%
s = 1	1263	1473	85.74	2921	11871	24.60
s = 2, 3	337	417	80.81	1558	3054	51.01
s = 4–6	83	133	62.40	486	693	70.12
s = 7–10	22	45	48.88	148	202	73.26
s = 11–20	18	36	50.00	82	143	57.34
s > 20	15	23	65.21	55	67	82.08

Table S7: Number of SIGACT reports matched to IBC entries, data variation (b), $w = 2$

Casualties	2004-05 & 2008-09			2006-07		
	matched	total	%	matched	total	%
s = 1	1248	1473	84.72	2857	11871	24.06
s = 2, 3	341	417	81.77	1578	3054	51.66
s = 4–6	87	133	65.41	487	693	70.27
s = 7–10	25	45	55.55	150	202	74.25
s = 11–20	21	36	58.33	91	143	63.63
s > 20	16	23	69.56	57	67	85.07

Table S8: Number of SIGACT reports matched to IBC entries, data variation (c), $w = 2$

Casualties	2004-05 & 2008-09			2006-07		
	matched	total	%	matched	total	%
s = 1	1039	1135	91.54	2883	11056	26.07
s = 2, 3	396	546	72.52	1721	3298	52.18
s = 4–6	113	251	45.01	522	848	61.55
s = 7–10	30	115	26.08	168	343	48.97
s = 11–20	29	118	24.57	125	317	39.43
s > 20	28	127	22.04	126	350	36.00

Table S9: Number of SIGACT reports matched to IBC entries, data variation (d), 20km, $w = 2$

Casualties	2004-05 & 2008-09			2006-07		
	matched	total	%	matched	total	%
s = 1	1185	1320	89.77	2912	11602	25.09
s = 2, 3	314	377	83.28	1547	2944	52.54
s = 4–6	80	120	66.66	472	671	70.34
s = 7–10	21	42	50.00	144	200	72.00
s = 11–20	19	35	54.28	80	137	58.39
s > 20	15	21	71.42	52	64	81.25

Table S10: Number of SIGACT reports matched to IBC entries, data variation (d), 30km, $w = 2$

Casualties	2004-05 & 2008-09			2006-07		
	matched	total	%	matched	total	%
s = 1	1274	1488	85.61	2924	12120	24.12
s = 2, 3	348	427	81.49	1576	3139	50.20
s = 4–6	86	130	66.15	487	719	67.73
s = 7–10	21	44	47.72	150	210	71.42
s = 11–20	19	37	51.35	85	144	59.02
s > 20	16	24	66.66	54	66	81.81

Table S11: Number of SIGACT reports matched to IBC entries, data variation (d), 40km, $w = 2$

Casualties	2004-05 & 2008-09			2006-07		
	matched	total	%	matched	total	%
s = 1	1345	1626	82.71	2932	12542	23.37
s = 2, 3	376	470	80.00	1612	3275	49.22
s = 4–6	93	148	62.83	495	749	66.08
s = 7–10	23	51	45.09	157	223	70.40
s = 11–20	20	40	50.00	88	152	57.89
s > 20	16	25	64.00	56	68	82.35

casualty statistics in SIGACT though, most notably because it significantly increases casualty counts for many events. Figure S1 confirms that KIA + WIA casualty counts do not feature the same robust power law scaling as reported in Figure 3 of the article and, qualitatively, the shape of the ccdf is more similar to that of IBC. However, the visual similarity is somewhat misleading: the Anderson-Darling tests robustly rejects the null hypothesis of agreement for all thresholds between 20 and 40 casualties per event (see also Table S5). Note further that the tail behavior is also considerably different: the dashed lines correspond to power law fits to the tail of the data with exponents of 3.5 for IBC and 2.79 for SIGACT.

Figure 4

Variations (a) and (b) do not result in significant changes to Figure 4 of the article and variation (d) is already accounted for in the figure. Considering civilian KIA + WIA events in SIGACT, we find that the dynamics of the number of casualties per month more significantly differs from the IBC datasets for all thresholds (see Figure S2) compared to the dynamics reported in the article. In fact, other than in Figure 4(b) where the number of casualties per month agreed for a threshold of 2 and IBC reported more casualties per month than SIGACT for all larger thresholds, we here find that SIGACT always reports more casualties than IBC. Using KIA + WIA counts thus certainly does not render IBC and SIGACT more consistent.

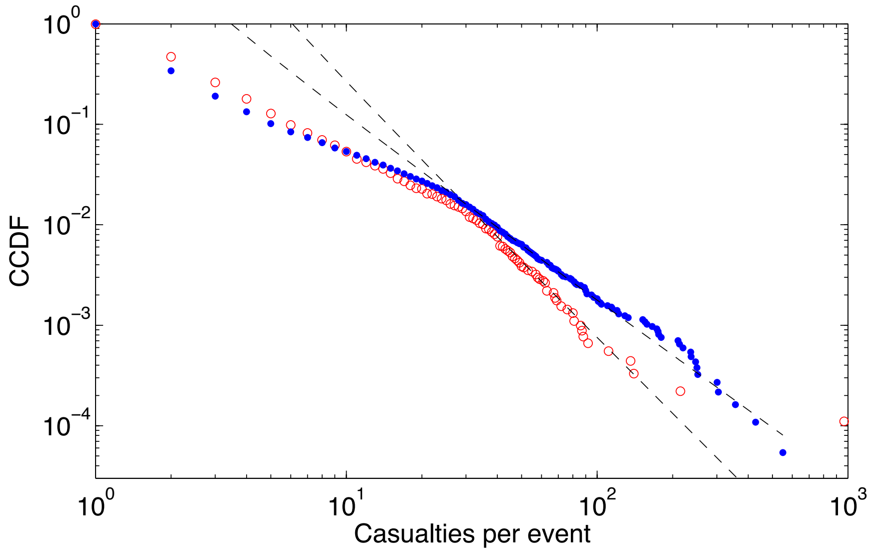

Figure S1: Complementary cumulative distribution function (ranking plot) of number of casualties in the "IBC Baghdad" (red circles) and "SIGACT Baghdad" (blue dots) datasets. Dashed lines correspond to power law fits using maximum likelihood estimation (also see the text of the article), data variation (c).

Figure 5

Data variations (a) and (d) do not result in significant changes to Figure 5 of the article. However, relying on the upper casualty estimates in the IBC dataset (data variation (b)) or KIA + WIA casualty counts in the SIGACT dataset (data variation (c))—or also both data variations taken together—generally decreases the agreement between the dynamics of the number of events per day in IBC and SIGACT. This is visible both in the RMS difference and the results of the Anderson-Darling tests, especially for large thresholds (see Figures S3 and S4).

Figure 7

Data variations (a), (b) and (d) do not result in significant changes to Figure 7 of the main article. Data variation (c), i.e. considering KIA + WIA casualties in the SIGACT dataset, almost insignificantly increase the number of small events ($s = 1$) in the SIGACT dataset that can be matched to events with the same number of casualties within ± 1 day in the IBC dataset. At the same time, however, it significantly decreases the fraction of large events matched (Figure S5).

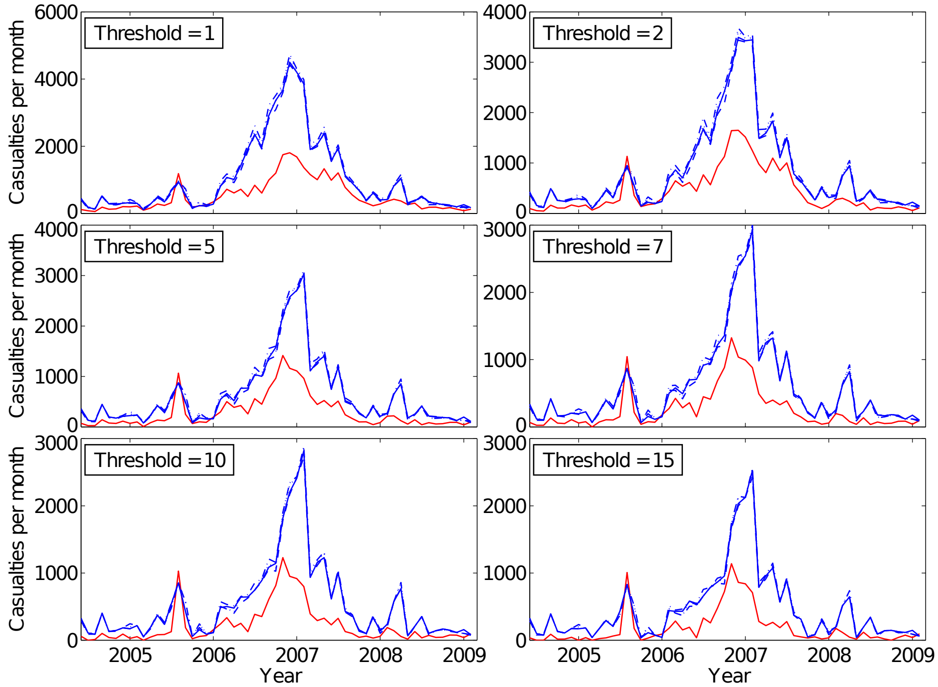

Figure S2: Dynamics of the number of casualties per months in “IBC Baghdad” (red line), “SIGACT Baghdad” (solid blue line), “SIGACT 20km” (dashed blue line), “SIGACT 30km” (dotted blue line) and “SIGACT 40km” (dash-dotted blue line). The panels correspond to subsets of events for thresholds of 1, 2, 5, 7, 10 and 15 casualties respectively. Note that the plots for the different SIGACT datasets (blue lines) are almost indistinguishable. Data variation (c).

Figure 8

The results reported in Figure 8 of the article are not significantly affected by data variations (a), (b) and (d). However, considering KIA + WIA casualty counts results in an increase of the non-trivial timing structure in the SIGACT dataset. In Figure S6 this is reflected in the fact that the null hypothesis of the Poisson (i.e. trivial random) dynamics can be rejected over much broader period of analysis, in particularly for large thresholds.

Figure 9

We find that neither of the data variations has a significant impact on the results reported in Figure 9 of the article.

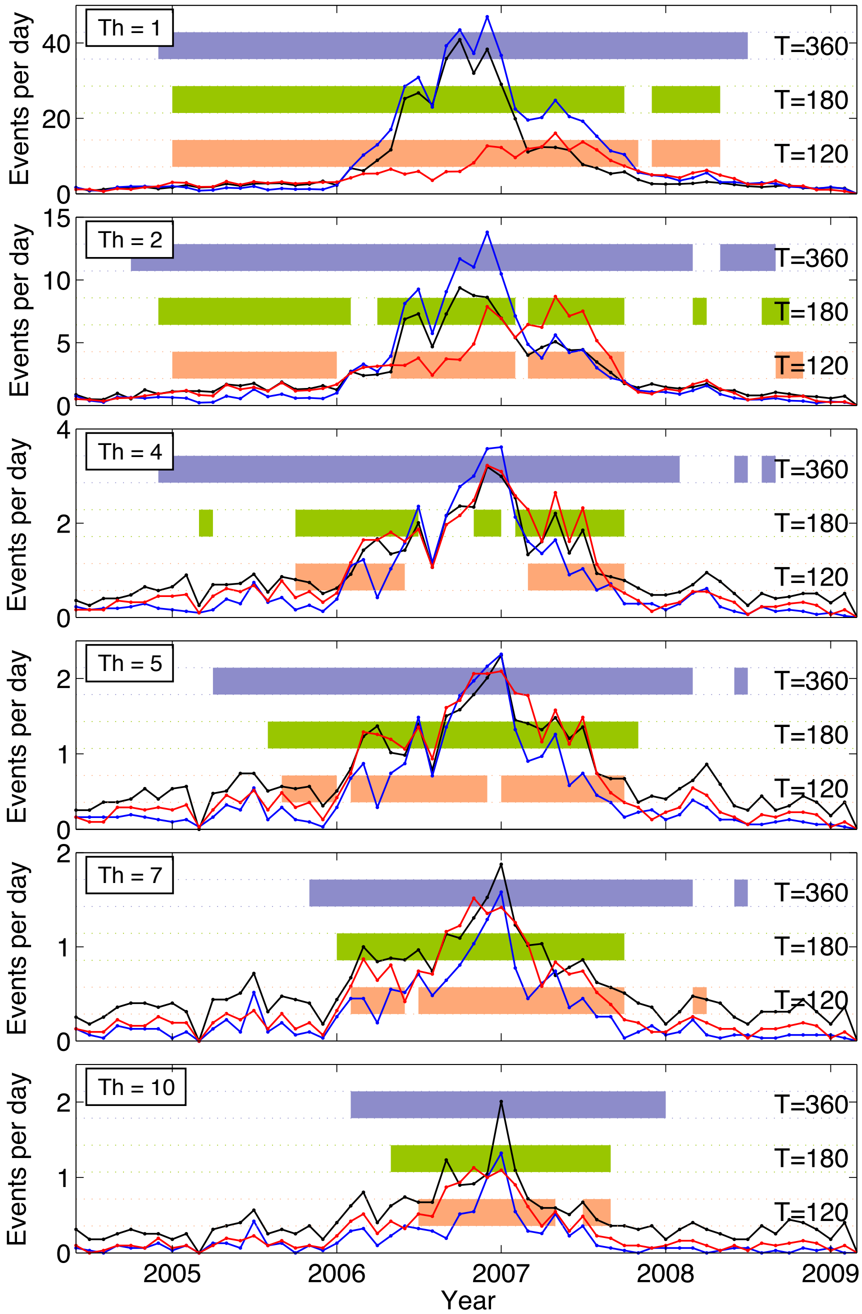

Figure S3: Distributional agreement of "IBC Baghdad" and "SIGACT Baghdad". Color bars illustrate the results of a 2-sample Anderson-Darling test for the distribution of number of events for time windows of $T = 120$ days (orange bars), $T = 180$ days (green bars) and $T = 360$ days (violet bars) for thresholds equal to 1, 2, 4, 5, 7 and 10 casualties. The bars indicate the center of those time windows for which the hypothesis of agreement of the distribution of events per day can be rejected at a 5% significance level. The black line represents the RMS difference between "IBC Baghdad" and "SIGACT Baghdad", red and blue lines are the monthly averages of the number of events per day for the two datasets respectively. Data variation (b).

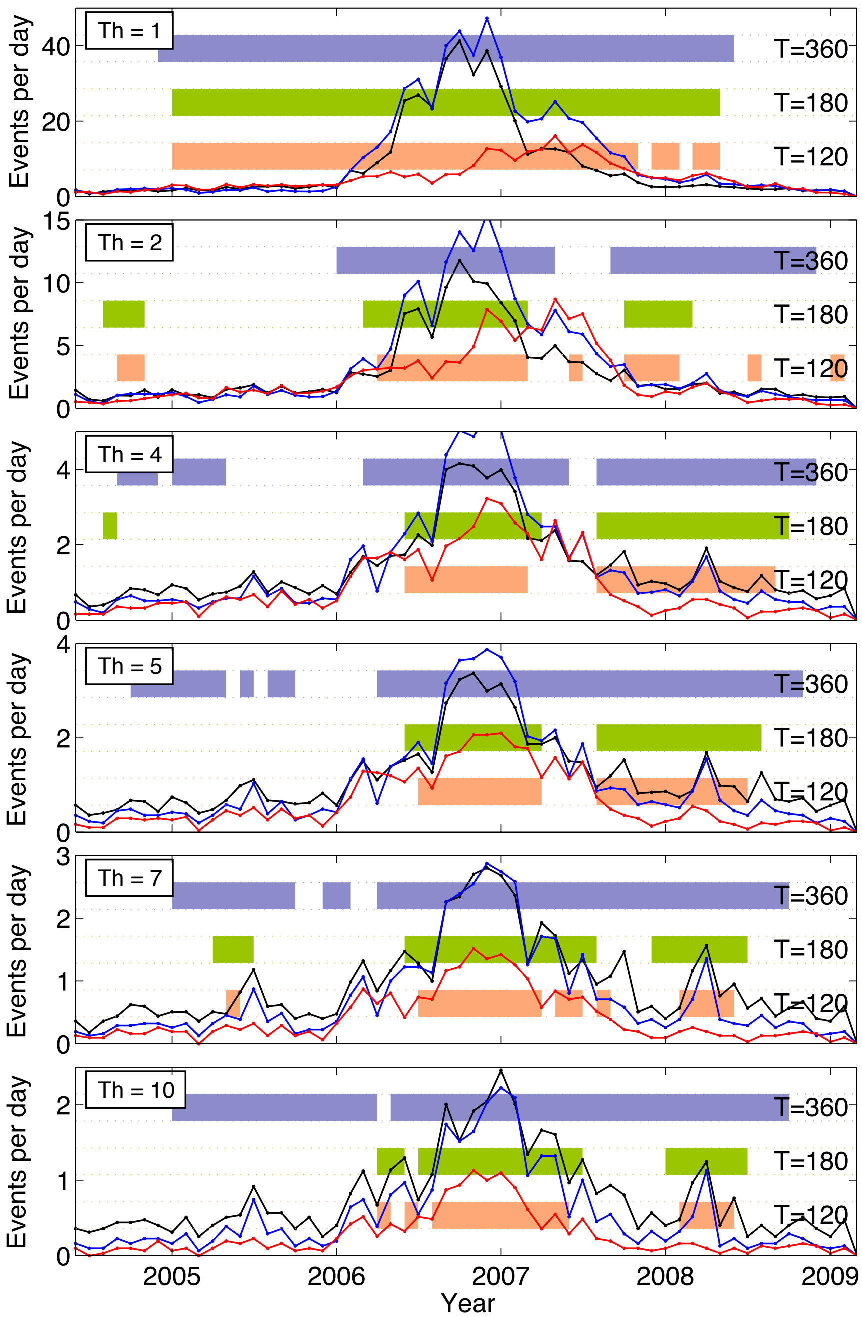

Figure S4: Distributional agreement of "IBC Baghdad" and "SIGACT Baghdad". Color bars illustrate the results of a 2-sample Anderson-Darling test for the distribution of number of events for time windows of $T = 120$ days (orange bars), $T = 180$ days (green bars) and $T = 360$ days (violet bars) for thresholds equal to 1, 2, 4, 5, 7 and 10 casualties. The bars indicate the center of those time windows for which the hypothesis of agreement of the distribution of events per day can be rejected at a 5% significance level. The black line represents the RMS difference between "IBC Baghdad" and "SIGACT Baghdad", red and blue lines are the monthly averages of the number of events per day for the two datasets respectively. Data variation (c).

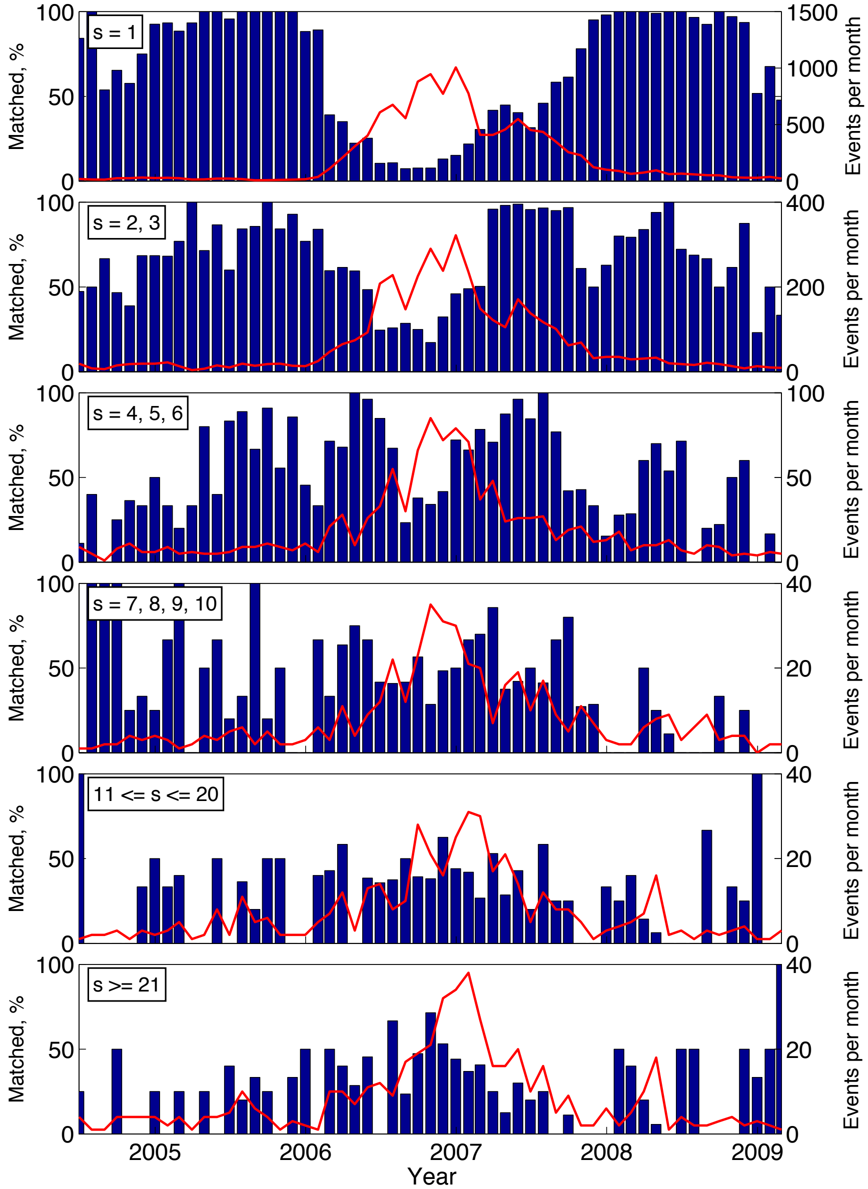

Figure S5: Day-by-day match of events of a given size s in "SIGACT Baghdad" to entries in "IBC Baghdad". Blue bars indicate the number of matched events as a fraction of the total number of events in SIGACT for every months in the dataset (left axis), the red line illustrates the overall number events per months for the given casualty sizes (right axis). When matching events we allow for a timestamp uncertainty of ± 1 day. Data variation (c).

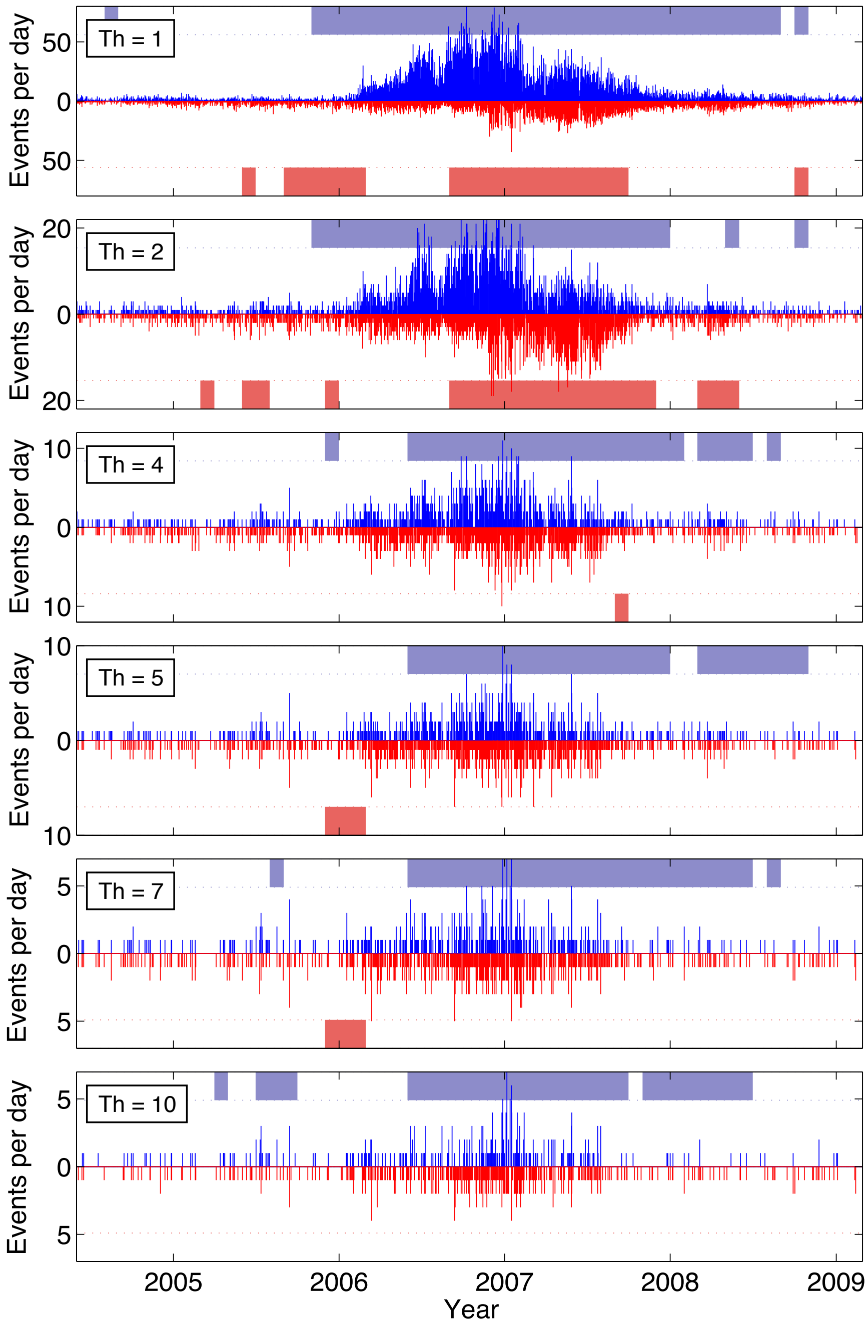

Figure S6: Inter-event timing signatures. Color bars illustrate the results of a KS-test for exponential distribution of the inter-event times in time windows of $T = 180$ days for thresholds equal to 1, 2, 4, 5, 7 and 10 casualties (see text for details). The bars indicate the center of those time windows for which the hypothesis of agreement of the distribution of inter-event times with an exponential distribution can be rejected at a 5% significance level. (i.e., the datasets exhibits a non-trivial timing structure). The graph also shows the dynamics of the number of events per day in "IBC Baghdad" (red) and "SIGACT Baghdad" (blue). The vertical axis for the IBC dataset was mirrored for clarity purposes. Data variation (c).

4 Distribution of events per day

In the daily time series comparison (section 3.3 of the article) we emphasize that the distributions of events per day do not have fat-tails and typically decay almost exponentially. Figure S7 demonstrates this for both "IBC Baghdad" and "SIGACT Baghdad" at various thresholds.

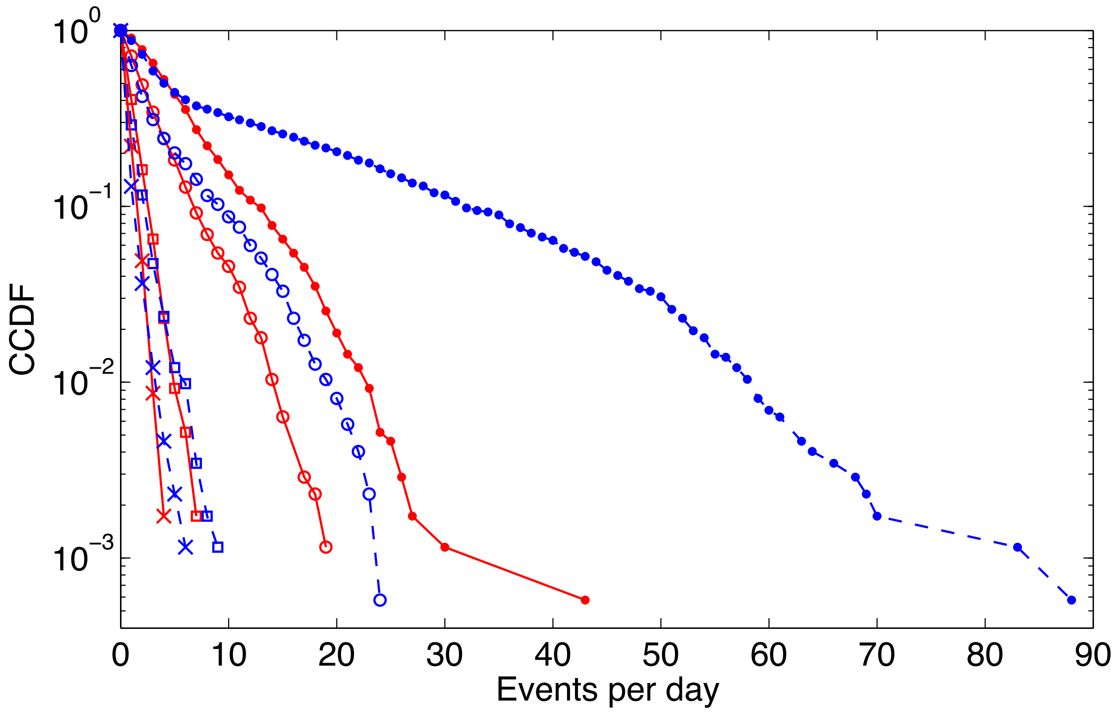

Figure S7: Complementary cumulative distribution function (ranking plot) of number of events per day in the datasets "IBC Baghdad" (red solid line) and "SIGACT Baghdad" (blue dashed line) for thresholds equal to 1 (solid circles), 2 (open circles), 5 (squares) and 10 (crosses) casualties per event.

5 Sensitivity analysis for distributional comparisons

In our analysis of distributional signatures in IBC and SIGACT (section 3.4 of the article) we test the distribution of inter event times against the null hypothesis of exponential distribution, which indicates Poisson dynamics for the process. In order to verify that results of Figure 8 of the article for larger thresholds (more than 2 casualties per event) are not an artifact of small sample size, we applied the same method for much larger moving time windows of 360 days.

Figure S8 shows the results of this analysis. One can clearly see that due to the non-stationarity of the data within the larger time window we can now reject the hypothesis of feature-less dynamics in much wider time intervals, as one should expect. This is clearly visible for both IBC and SIGACT at thresholds of 1 and 2 casualties. However, for the IBC dataset and large thresholds (larger than 2 casualties per event) we can – despite the non-stationarity – for most of the time period analyzed not reject the null hypothesis of exponential distribution. Notice in particular that this is true for the period in which the conflict escalated (second half of 2006 and first half of 2007). The results thus confirm the featureless dynamics of IBC for larger thresholds.

Additionally, in section 3.4 we have also emphasized that testing the null hypothesis of the Poisson distribution of events per day leads to substantially equivalent results. Figure S9 and Figure 8 of the article indeed yield very consistent estimates of where both datasets exhibit non-trivial timing structures. Notable exceptions are short time windows in 2005 and 2006 where the event per day statistics suggest more non-trivial timing structure in IBC (for low thresholds) and more trivial timing structure in SIGACT (for high thresholds) compared to the inter-event statistics.

Notice that both tests effectively complement each other with respect to statistical power. In case of large number of observed events per window the test for exponential distribution of inter-event times provides much more robust results. However, if the samples are small (such as in 2005 or 2008–2009 and in case of large thresholds) the test for Poisson distribution of events per day is more powerful and can reject the null hypothesis of Poisson dynamics even when the clustering is moderate. This gives us additional confidence in the results of Figure 8, in particular for the periods with lower intensity of violence.

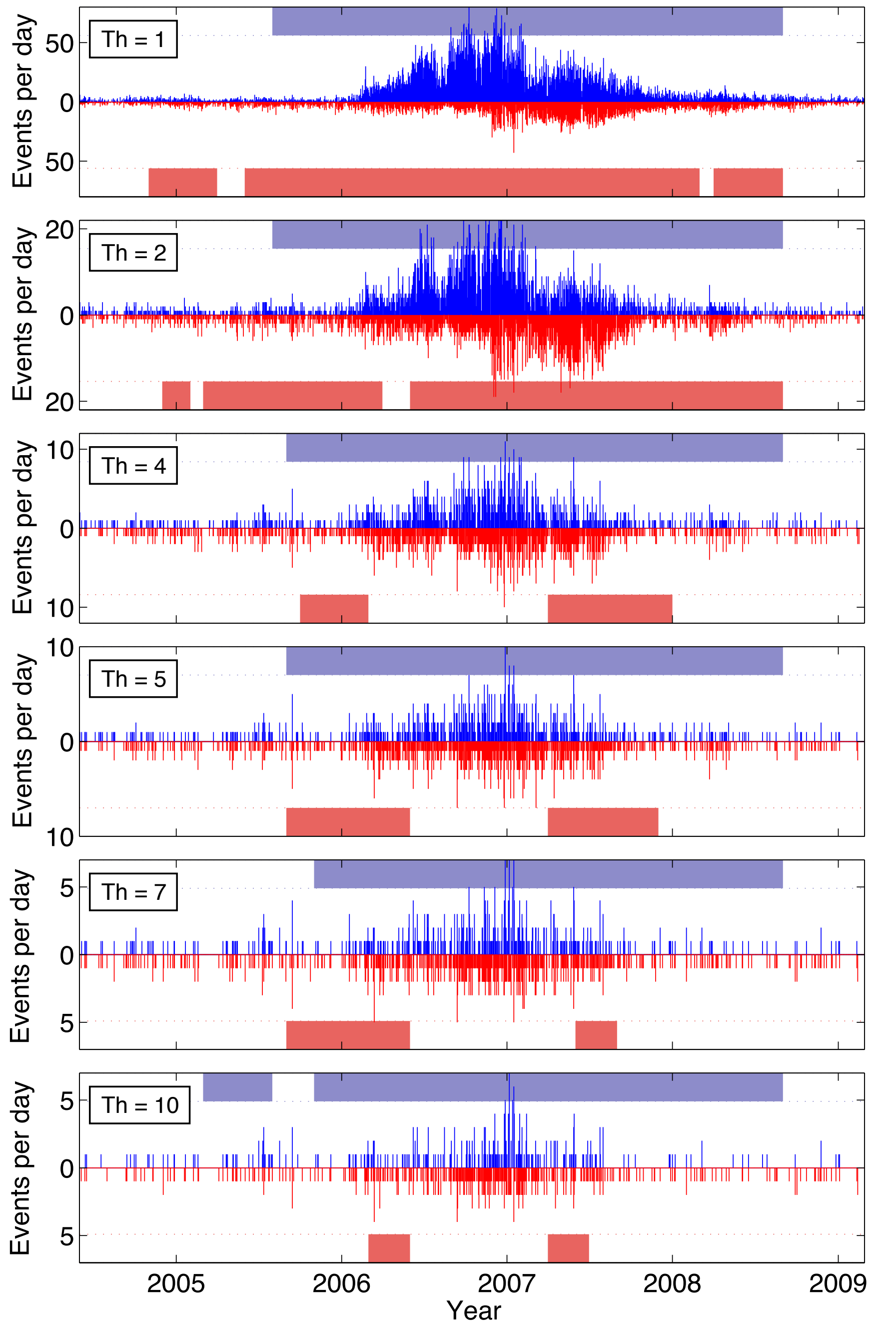

Figure S8: Inter-event timing signatures. Color bars illustrate the results of a KS-test for exponential distribution of the inter-event times in time windows of $T = 360$ days for thresholds equal to 1, 2, 4, 5, 7 and 10 casualties. The bars indicate the center of those time windows for which the hypothesis of agreement of the distribution of inter-event times with an exponential distribution can be rejected at a 5% significance level. (i.e., the datasets exhibits a non-trivial timing structure). The graph also shows the dynamics of the number of events per day in "IBC Baghdad" (red) and "SIGACT Baghdad" (blue). The vertical axis for the IBC dataset was mirrored for clarity purposes.

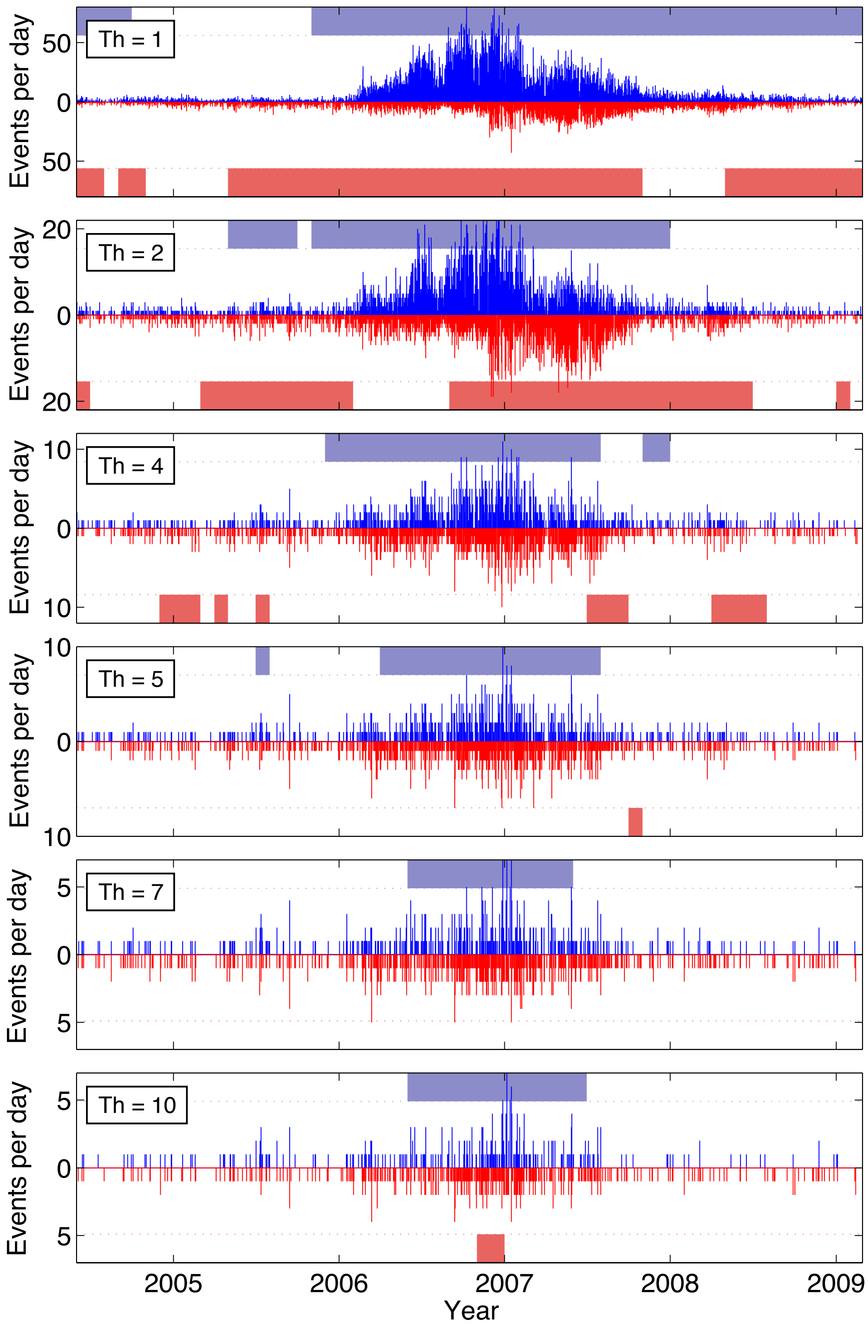

Figure S9: Number of events per day signatures. Color bars represents results of the chi-square test for the Poisson distribution for both datasets and time window of $T = 180$ days for thresholds equal to 1, 2, 4, 5, 7 and 10 casualties (see text for details). The bars indicate the center of those time windows for which the null hypothesis of Poisson distribution for the numbers of events per day can be rejected at a 5% significance level. (i.e., the datasets exhibits a non-trivial timing structure). The graph also shows the dynamics of the number of events per day in "IBC Baghdad" (red) and "SIGACT Baghdad" (blue). The vertical axis for the IBC dataset was mirrored for clarity purposes.

References

1. IBC: **Iraq Body Count**. *http://www.iraqbodycount.org/* 2014.
2. Rogers S: **Wikileaks Iraq: Data Journalism Maps Every Death**. *http://www.theguardian.com/news/datablog/2010/oct/23/wikileaks-iraq-data-journalism (accessed: 09/03/2013)* 2010.
3. Carpenter D, Fuller T, Roberts L: **WikiLeaks and Iraq Body Count: the Sum of Parts May Not Add Up to the Whole-A Comparison of Two Tallies of Iraqi Civilian Deaths**. *Prehospital and Disaster Medicine* 2013, **28**(3):223–229.