

**New findings from conflict archives:
An introduction and methodological framework**

Laia Balcells
Department of Government
Georgetown University

and

Christopher Sullivan
Department of Political Science
Louisiana State University

As they pursue information and deploy violence during conflict, combatants compose, catalog, and preserve a wide variety of records, such as memos, investigative reports, and communiqués. In an increasing number of post-conflict scenarios, these records are being archived and released publicly, quickly becoming a critical new source of data for studies of peace and conflict. The objective of this special issue is to advance a new research agenda focused on the systematic analysis of conflict archives. The contributors each spent significant time collecting original data from often-dusty archives and, in many cases, developed new methodologies for sampling, cataloging, and analyzing historical documents. Their findings reveal how violence simultaneously shapes and is shaped by factors that remain largely unobservable using more conventional sources of conflict data, including clandestine mobilization, bureaucratic accountability, and political identities. By considering these studies in relation to one another, this introduction aims to provide readers with a comprehensive understanding of field research strategies and analytical techniques for studying original data from conflict archives. We conclude that while archival data are subject to their own biases that must be considered, this research agenda addresses significant limitations associated with traditional data sources and, in turn, pushes scholars to rethink many of the mechanisms underlying the causes and dynamics of peace and conflict.

Introduction

For decades, scholarship in the field of peace and conflict studies has relied on data gathered from secondary historical accounts and news media reports. While useful for comparative analysis, these data sources have demonstrable biases towards over-representing easily observable acts of violence (Davenport & Ball, 2002; Davenport, 2009; Sullivan, 2016a; Weidman, 2015a). One important consequence is that theoretical and empirical analyses are often predisposed towards studying the most visible violent actions and processes (e.g. acts of terror, large demonstrations, pitched battles, fatal shootings, etc.). To reveal the hidden mechanisms underpinning political conflict, researchers have begun systematically collecting and analyzing thousands of historical records from conflict archives.

When we talk about ‘conflict archives’ we refer to catalogs of written material systematically produced as part of a political contest between a government force and an armed challenger (either domestic or international) and later collected and preserved for public access. Materials contained within conflict archives commonly include memos, investigative reports, communiqués, and other documents produced by government or rebel forces. Such records are inherent to conflict; they serve as the primary conduits through which combatants acquire, accumulate, and distribute information. Because combatants produce these documents contemporaneously with violence, the records reveal how actors understood the strategic environment in real time. And because the forces producing these documents were on the ground conducting operations in situations of heightened insecurity, their archives reveal information on violent and non-violent behaviors that cannot be easily monitored by third-party observers (e.g. NGOs, the

media). As a result, data from conflict archives typically contain better coverage (across space and time) and more information (on actors, behaviors, beliefs, and context) than nearly any other systematic collection of material.

This special issue highlights empirical findings and theoretical insights emerging from a research program based around conflict archives. In this introduction, we discuss how material found in archived documents compares to more traditional sources of conflict data. By considering the contributing articles in relation to one another, we hope to provide readers with a comprehensive understanding of field research strategies and analytical techniques for studying original data from conflict archives.

The articles contained in this special issue reveal how researchers use data from conflict archives to challenge widely held conventions and/or reveal insights into aspects of violence not often considered due to lack of credible data. In this way, the common elements unifying this research agenda extend beyond methodological concerns. New findings from conflict archives reveal mechanisms and processes previously buried beneath the fog of war. Using data from conflict archives, research in this special issue connects micro-, meso-, and macro- levels of analyses, demonstrates how identities intersect conflict cleavages in varying ways, and depicts the “gray zone of power” between institutions and violence (cf. Auyero, 2007).

The remaining sections of this paper assess the past, present, and future of research on conflict archives. We first review the history of archival research in conflict studies and explain its exponential growth in recent years. We then provide an assessment of the methodological debates emerging from this kind of research and specify a set of best standard practices. Finally, we summarize the contributions to this special issue,

documenting how a new generation of research employs conflict archives to shed light on processes of peace and conflict. The final section concludes.

Research from conflict archives: A short review

The convergence of research from conflict archives across the globe reflects two congruent trends: (1) exponential growth in access to conflict archives combined with (2) conflict studies' increasing attention to data quality and measurement error. With regards to archival access, activists in both the domestic and international arenas continue to push for increased transparency in government. At the same time, practitioners working on peace and reconciliation have progressively come to view the release of conflict archives as key to transitional justice. As access to previously confidential records expands, leading archivists are directing the often years-long processes of diligently cleaning, organizing, and preserving the materials into an expanding list of conflict archives, examples of which can be found in Table I of the Online Appendix.

For both activists and archivists, the systematic preservation and release of records performs truth telling purposes that are independent of their ability to advance scientific research. 'When records of past surveillance and repression are made public,' Wisser & Blanco-Rivera (2016: 126) argue, 'they take on new meaning. Initially used by the state as a means to control citizens and quell opposition, these records can become the evidence used by victims and their relatives for the purposes of accountability, memory making and truth seeking.' Indeed, in the aftermath of a highly violent conflict, the release of archived documents can constitute an important component of truth and

reconciliation processes, revealing to victims' families what happened to their kin and providing important sources of evidence to prosecute perpetrators.

Simultaneous to the expansion of access to conflict archives, conflict scholars have emerged from the methodological debates of the previous two decades with demands for better data along with a renewed interest in combining quantitative and qualitative methods (cf. Davenport, 2010; Gohdes & Price, 2012; Eck, 2012; Morgan, 2013; Arjona, 2016; Balcells, 2017).

In the social sciences, important work pioneered the path to today's research, including Rude (1964) and Tilly et al. (1975), who drew on historic police registers to compile listings of conflict related events, and Weber (1976) and Meyer et al. (1979), who looked into administrative records to examine the expansion of state authority. Like their contemporary counterparts, these studies extracted information methodically from a set of archival materials, so that that information could be compared over time and space. But in many cases, it remains unclear how records were sampled, what the exact coding and measurement procedures were, or what standards were being applied to generate inference.

New research on conflict archives adds rigor to this tradition by systematizing the sampling process, drawing clear and replicable measurement rules, and studying the archival materials with sophisticated analytical techniques. Of particular note is the scholarship of Roger Gould, who serve as an exemplar of how historical archives can be used to solve big questions in social science, and in particular how archives can test theories relevant for contemporary conflict (e.g. Gould 1995, 1999, 2003). Other noteworthy efforts to systematically employ data from conflict archives to assess theories

of peace and conflict include Cunningham (2004), Kalyvas (2006), Lawrence (2013), Davenport (2014), Braun (2016), and Finkel (2017).

To illustrate the transformative effects of this research program, the next section assesses some methodological debates in the study of conflict archives.

Advantages and limitations of data from conflict archives

Archives often provide better coverage across space and time than secondary sources or data produced by third party observers, such as NGOs or the media. This is true partially because the forces responsible for creating the archived documents typically have more resources than outside monitors. But more importantly, violence systematically distorts both the supply and demand of information gathered by third-party observers (e.g. Cohen & Greene, 2012; Sullivan, 2016a; Weidmann, 2015a; Zhukov & Baum, 2016).¹

Reflecting a famous dictum, if a body falls in a conflict and no one is there to observe it, will it impact our results? The short answer is we rarely know; the longer answer is that it will depend on whether, how severely, and in which direction systematic missingness correlates with our explanatory models (King et al., 1994). Several recent examples reveal the implications for peace and conflict. For instance, biases in media data tend to validate reactionary “threat-response” models of repression, whereas Sullivan’s (2016a, 2016b) archival research reveals a hidden layer of “preemptive repression” designed to limit the emergence of future challengers. Similarly, while media sources readily report on observable violence, Zhukov’s (2016) archival analysis reveals that when punishment strategies fail, governments turn to forcible disarmament as a way

¹ The rich literature studying biases in media data is well summarized in Douglass & Harkness (2018).

of denying rebel advances. Further examining situations where violence limited third party monitoring, Kaplan (2013) uses interviews alongside conflict archives to introduce the role of non-violent civilian resistance into a body of conflict research that has long been focused narrowly on rebel-government interaction. In each of these instances, scholars deployed conflict archives to overturn prevailing assumptions based largely upon more conventional data sources.

Beyond increasing the scope of available information, conflict archives also commonly provide greater depth of coverage across event types as well as greater subnational disaggregation. This is particularly the case if local archives are used in conjunction with archives at the national level (e.g. Balcells, 2017). In some cases, recent work has combined historical process tracing with conflict archives to study the mechanisms underlying important political phenomena and resolve persistent theoretical debates (e.g. Lawrence, 2013; Schwartz & Straus, 2018). Archives are particularly useful to undertake process tracing, as they can provide information on the wartime processes that are absent in other data sources.

Finally, archives allow for the study of new topics, including the direct examination of the reporting structures combatants rely on to commit acts of violence. Conflict archives can provide firsthand information on the behavior of participants in a conflict along with new meta-data about how the organization producing the documents understood and reported that behavior. For example, surveillance information is passed across a military to direct offensive strategies towards areas where violence is thought to be most effective. Later, the command staff may catalog those reports to evaluate the impacts of violence. When these documents are preserved in conflict archives, they

convey information both about what was happening on the ground as well as how the combatant(s) understood and interpreted those events. In settings when archive materials are based upon government records, conflict archives allow us to literally see like a state (Scott, 1998).

To be clear—our argument is not that archives should be *the* source of data for studies of peace and conflict, but instead that in many cases conflict archives provide more comprehensive and less biased information on political conflict than any other systematic collection of data. All sources of conflict data are likely to present incomplete representations of political violence. Still, in our assessment raw (i.e. unprocessed) material generated by combatants and collected from conflict archives gets researchers closer to the ground truth of what transpired in a particular event.

Regarding limitations, archival data have political biases, both with regards to their availability (e.g. records can be destroyed or made otherwise unavailable; many are blocked by state secrecy laws), as well as in the underlying materials. Governments do not commonly make a full listing of their records available. And when documents are released, it is commonly a strategic act as the current government seeks to shame their rivals by releasing documents that are particularly damaging to the opposition. Indeed, this is one way in which victors define ‘history.’ The politically motivated release of archival materials may be entirely distinct from their scholarly analytical purpose, and may involve deliberate manipulations of the types of documents that get released and/or the types of information that are collected. State forces can directly affect the flow of materials in an effort to rewrite history, as the Rwandan government did by limiting the collection of information on conflict related violence to acts of genocide (Loyle &

Davenport, 2015). In another example, the *Causa General* in Spain was created specifically to bring only the losers of the civil war to justice (Balcells, 2017).

Archives may also provide information in units of analysis that are not theoretically relevant or useful. For example, governments typically collect their data by administrative districts, and these units may serve as imperfect proxies for testing our theories. Also, units can change over time, and the information on units from period t might not be useful for the understanding of phenomena measured at different units in $t+n$. Further, the types of information gathered in records and reports may shift over time and space as combatants respond to conflict. For example, Douglas (2016) demonstrates how missingness in the attributes collected by US Army Intelligence during the Vietnam War varied systematically across both the type of operation and the demographics of potential targets. There may also be incentives to report acts of government violence differently from rebel violence. Guberek & Hedstrom (2017) show how the terminology used to describe killings in the Guatemalan National Police Archive shifted across time and across political regimes. Thus, even when an archive provides nearly unprecedented access to repressive institutions, important “silences” remain that may impact subsequent analyses (ibid., cf. Sullivan, 2016a; Weld, 2014). Douglass & Harkness (2018) take this as a starting point before demonstrating how machine learning can be used to construct relevant units of analysis based on how relevant information varies as a function of conflict.

Acknowledging the challenges discussed above, we now review several additional best standard practices for scholars sampling, coding, and analyzing data from conflict archives.²

Best practices for generating and analyzing data from conflict archives

Understanding the archive

In order to assess the manner in which sources of potential bias interact, it is crucial that the researchers have an in-depth knowledge of the historical case(s) and, in particular, of the process by which the archives have been generated. Researchers must have been aware of the process by which documents in the archive were composed, preserved, and released. A good method for understanding the materials contained in (and missing from) conflict archives is to spend significant amounts of time speaking with archivists about the structure of the archive, the origins of the documents, how they are preserved, and what findings other researchers have discovered before. Researchers should use this information to develop testable propositions about what is and is not contained in the archives, which can then be used to confirm or invalidate theories of the data generating process (e.g. Douglas, 2016; Guberek & Hedstrom, 2017; Douglass & Harkness, 2018). These tests should be made explicit, and the results should be preserved for subsequent replication by later researchers.

Taking time to understand the contents of an archive and the manner politics influenced the construction and preservation of documents has a second, theory building,

² Conflict scholars have recently developed several excellent guides for quantitative and qualitative data collection, and we refer readers to them as an important source for improving data quality (Salehyan, 2015; Davenport & Moore, 2014; Seybolt et al., 2013). We also wish to echo several of their recommendations with particular reference to conflict archives.

benefit. The construction of the texts contained in conflict archives provides insight into what sorts of information were being collected, what language was used to describe that information, what is conspicuously absent, and how the bureaucratic machine that created and preserved these documents operated (Stoler, 2002; Dimitrov & Sassoon 2014; Art, 2016). Viewing the archive as a system of information collection, transmission, accumulation, reduction, consumption, and preservation can reveal the hidden machinations organizing political violence.

Developing a theory of inference

Researchers must be explicit about the population (i.e. universe of events) they intend their findings to speak to (i.e. generate inferences about). This should be informed by the archival data generating process. For instance, Zhukov & Talibova (2018) focus on a specific subset of political imprisonment records that were contained within the Soviet Archives. Guided by a belief that the Soviets collected these records uniformly and comprehensively, they use their records as a proxy for the universe of deportations that occurred during this period.

When dealing with convenience samples, the data generating process limits population level inferences (Seybolt et al., 2013). And as a general rule, we agree with Kruger et al. (2013: 261)'s assessment that 'When convenience data are used in any capacity, it is best to "stay close to the data," guarding against overly ambitious inferences or quantifying potential biases' (see also Art, 2016). But, given the gravity of the problem of violence, good data are often better than none. We also articulate several methods for analyzing convenience samples below.

Sampling and coding

Given a theory of the data generation process and a theory of inference, researchers must develop sampling and coding procedures to efficiently collect information speaking to the population of interest. Conflict archives typically contain millions of pages of material, and (in the absence of machine coding) important decisions need to be made about how to extract information from a relevant subset of the documents. If the researcher is interested in generalizing only to the universe of events contained in the archive, randomly sampling from the archive will provide a sound and efficient strategy. If, on the other hand, researchers hope to use the archive as a way to speak to the broader universe conflict events, then randomly sampling from the archive will not improve inferential abilities. And in scenarios where the researcher is interested in a specific type of activity (e.g. surveillance records) then purposive sampling may be more efficient for gathering a comprehensive set of data. In all cases, we advise researchers to be as transparent as possible about what sampling strategy was used, what alternatives were available, and what implications their sampling strategy holds for subsequent analyses and inferences.³ The same type of exercise should be done if researchers decide to focus on particular regions of a country to generate inferences about its totality.⁴

Pretesting and comparing several different sampling procedures and coding protocols can provide an understanding of how sampling methods will affect your data, as well as what procedures will be most efficient. Here too we advise researchers document the

³ We do not have any specific guides with regards to sample size as these will vary across studies according to time, funding, analytical methodology, and the objective of the study.

⁴ McLaughlin (2014), for example, studies desertion during the Spanish Civil War using archival data from the province of Santander.

results generated from their comparative pretesting as a way of validating subsequent analyses.

Analytic techniques

Each of the items discussed above should play a pivotal role when it comes time to analyze the data. Acknowledging that different research questions demand different research strategies and that mixed methods often produce the most convincing results, we provide several specific points that pertain to managing the limitations identified above. One way that scholars studying conflict archives can address concerns of bias is to verify the information with other sources. However, the timing of archival releases sometimes makes verification difficult. For example, in Spain the study of archives impossible during Franco's dictatorship, so scholars of the civil war could only verify information through oral sources several decades after the civil war (e.g. Solé i Sabate & Villarroya, 1989). By then, many witnesses had passed away.

In some settings, it can make sense to combine sources, though care must be taken to explain how and why these sources were selected and combined. In other scenarios, conflict archives are combined with variables collected from other datasets, in a way that creates distinct measurement processes for the independent and dependent variables (thus reducing the probability of spurious correlations). In this issue, Hassan & O'Mealia combine conflict archives with satellite data.

Studies of conflict archives need to include an awareness of how the data generating process affects analytical models.⁵ At minimum, we urge scholars in conflict archives to

⁵ Another example of how this can be done can be found in Sullivan (2014) who matches treatment and control units based on observations of the data generating process identified prior to assignment.

be as transparent as possible about their underlying beliefs about how the records were created, preserved, archived, and accessed as well as how each of these facets may affect their results. In the future, the best work in this area will directly incorporate theories and analyses of the data generating process into their research. Fortunately, rigorous new methods are emerging. Most notably, several recent efforts have advocated capture-recapture (also known as multiple systems estimation or MSE) methods for recovering the larger universe of undocumented events (Seybolt et al., 2013; Hendrix & Salehyan, 2015). MSE provides reliable estimates, but also involves some burdensome assumptions and constraints on disaggregation. Another technique for mitigating bias in convenience samples involves extrapolating from conflict archives using population based surveys, though here too there are challenges (Seybolt et al., 2013). Finally, a new estimator derived by Cook et al. (2017: 223) simultaneously models the data generating process and the causal effects of interests, enabling researchers ‘to accurately test theories on both the causes of and reporting on an event of interest.’ We anticipate that future research on conflict archives will formulate theories about how the data generating process might affect empirical results, and then subsequently model that process in a manner that improves inferential validity.

Replication

Each step in the research process beginning with the initial pilot tests of conflict archive and ending with the analysis of complete datasets should be transparent and replicable. For subsequent replication, the coding protocol and sampling procedures need to be made available to the scholarly community, and in the interests of scientific

advancement, we recommend that researchers release the raw disaggregated data, rather than limiting their replication data release to the processed data used in published analyses. We understand that research on conflict archives often requires years of effort, and as a result researchers face professional incentives to retain the results of their efforts as proprietary data. But we suggest that a clear limit be set on this, such that the data are released to the broader scientific community within a reasonably finite period of time. Releasing the information in the widest format possible will expand the research community's access to high quality data and reduce the need for duplicating effort.⁶

Contributions to the special issue

Articles in this issue delve into conflict archives, employing sophisticated methodologies to analyze data and obtain results that speak to big questions in social science. Using data from a wide range of locations, the researchers identify new research questions, challenge conventional wisdom, and present new findings to shape the direction of conflict studies.

A major theme emerging from this research is the organization and operation of state security forces. It should come as no surprise that many of the articles focus on operations and violence perpetrated by governments. Given their superior resources and typically more established organizational structure, governments collect and preserve documents pertaining to their war efforts far more comprehensively than rebels. And while rebel records do get released, those challenging the state face incentives to minimize any paper trails that could allow the state to target them with repression. But, as we discuss further in the conclusion, the focus on government repression extends beyond

⁶ Of course, the research releasing digitalized archival data should be cited in any work using the data.

the origins of conflict archives. Conflict studies has worked hard to match the study of repression to the study of non-state violence. But such efforts have been hampered both by data availability and by the political motivations of (government) agencies funding large research projects. As a result, there remains much to learn about the operation of coercive governments, and conflict archives offer unique opportunities in which to study their operations.

From this work, we learn how political violence informs and reshapes repressive institutions. Kristine Eck uses detailed archival data on British security inputs to colonial territories to explore an interesting hypothesis: The trajectories of conflict observed prior to independence explains the different paths that states take developing security and surveillance capacity. States with conflict experience received more colonial security inputs, enabling them to develop police forces that commit fewer violations of physical integrity rights than states that did not have an experience of conflict. In a way, Eck finds that the experience of counterinsurgency made stronger states, which is consistent with the literature on counterinsurgency and civil war that uses conventional sources of data (e.g. Balcells & Kalyvas, 2014). Eck's (2018) combination of qualitative and quantitative analysis reveals new evidence on previously unobserved mechanisms linking conflict to the expansion of the security apparatus.

Several other contributions in this issue reveal how governments accumulate and deploy intelligence. Hassan & O'Mealia (2018) look at the challenges of managing an effective bureaucracy in a relatively weak state. Given the difficulties of monitoring bureaucratic agents spread out across the state's territory, central governments look for cues regarding who is loyal or disloyal. Somewhat counter-intuitively, the authors argue

that election violence can serve as a signal that influences the career trajectory of individual bureaucrats. Using government records, they show that pro-incumbent election violence signals local bureaucrats' loyalty to the central government, while opposition initiated election violence demonstrates that local agents are either inept or disloyal. Like Hassan & O'Meara (2018), Sullivan & Davenport (2018) look at subjects that are difficult for the state to observe. In this case, the authors examine the government's ability to monitor the clandestine behaviors of political challengers. Employing archival reports from a variety of U.S. intelligence agencies along with documents produced by a Black Nationalist social movement organization, they show how challengers adapt to overt repression in ways that make subsequent state surveillance more difficult. The results demonstrate how overt repression actually drains state intelligence, challenging conventional wisdom on counter-insurgency and counter-terrorism.

Douglass & Harkness's study (2018) can be read either from a purely methodological point of view, or from the view of a theory of state surveillance. As a methods article, the authors provide a new technique for analyzing variance in the information produced within archival sources, which optimizes the spatial units that researchers should use when geo-referencing their data. This minimizes biases produced by the geographic dispersal in document production and the modified area unit problem. As a theoretically relevant piece of conflict scholarship, the article shines light on how difficult it is for governments to identify challenger activities, particularly those occurring in areas with weak government control and a dense landscape. This finding validates predictions from several dominant models of civil war (e.g. Fearon & Laitin, 2003; Kalyvas, 2006), while

simultaneously challenging the assumption that states are capable of precipitating or responding to dissent at will.

Violent conflict has political consequences, and this is often due to the fact that political factors underlie violent dynamics. A second set of articles study dynamics of lethal and non-lethal violence with archival data, and they point to political and ideological factors underlying patterns of violence in different contexts. Scharpf (2018) exploits on archival data from Argentina's Dirty War (1975-1981) to study how ideological leanings of military officers shape sub-national patterns of state violence. He distinguishes between Nationalist officers, who dominated the infantry, artillery, and communications branch, and Liberal-conservative officers, who served in the cavalry branch. Because of their ideological convictions, Scharpf (2018) finds evidence that Nationalist officers were more prone to use large-scale terror against the broader population, while Liberal-conservatives advocated the selective targeting of guerrilla forces and their supporters. Schwartz & Straus (2018) study a slightly different set of ideological tradeoffs in their research focusing on mass atrocities in Guatemala. Following a recent trend in the field (e.g. Kalyvas, 2006; Arjona, 2016; Balcells, 2017), these authors highlight the importance of local-level analysis of patterns of violence. In particular, they emphasize the relevance of local interpretations of military strategy and insurgent threat. They leverage the 359 pages of military plans and communications that comprise Operation Sofía, a one month-long counterinsurgent campaign waged in the northwestern Maya Ixil region. The authors find some evidence that state actors understood the civilian population as fully dedicated to rebel forces, although their results also suggest that troops on the ground perceived Ixil communities as "winnable," and

thus deployed violence for coercive (rather than purely ideological) purposes. Balcells (2018) uses data from the Spanish Civil War (1936-1939) to analyze the relationship between violence and displacement during an ideologically polarized civil war. She finds that refugees tended to cluster by ideology when they settle in hosting localities, and that the arrival of refugees sometimes had the side effect of triggering additional violence against local enemies.

Drawing on data from different historical periods and different continents, a final set of studies in this special issue examine the legacies of violent conflict. Osorio, Schubiger & Weintraub (2018) focus on repression carried out by the Mexican government in the 1960s and 1970s, and its long-term consequences for state consolidation. They focus on a particular type of repression: disappearances. Relying on data gathered from a secret governmental report that was subsequently leaked to the press, they find that disappearances have mixed effects on state consolidation: on the one hand, such abuses are positively correlated with contemporary measures of fiscal and bureaucratic capacity; on the other hand, they are also negatively correlated with measures of state monopoly over the use of force. Zhukov & Talibova (2018) examine archival records of Stalinist political imprisonments and identify their effect on contemporary Russian politics. Using proximity to the railroad as an instrumental variable, they find that deportations have a significant impact on contemporary political behavior: they lead to a decrease in both electoral participation and support for Putin. They also find that deportations has influenced political behavior in Ukraine, where turnout is lower in localities where Stalin-era repression was high.

Conclusion

The unique information contained in conflict archives provides new opportunities for conflict scholars to illuminate previously opaque mechanisms and processes motivating violence. Critically, there is a certain subversive element to the study of conflict archives. To prevent their opponents from accessing sensitive information, governments carefully protect and preserve the data and analysis contained in conflict archives. The fact that conflict archives are typically kept confidential even after the cessation of violence implies that they also contain some residual value for those hoping to challenge the state (Art 2016).⁷ Finding, releasing, and analyzing materials from conflict archives democratizes knowledge that had previously remained proprietary, bringing into the public sphere formerly secretive information on the causes and consequences of violent resistance and/or repression (Nalepa, 2010).

We hope this issue will stimulate additional research by releasing new data collected from archives across the world, which may later be subjected to comparative analysis.⁸ Peace and conflict research has moved away from the dichotomy of micro-level case studies and cross-national statistics and towards analyses that include both within-case and across-case analyses (e.g. Christia, 2012; Balcells, 2017). We hope to further this important analytical endeavor.⁹

Some of the limitations of conventional research are not fully solved in research based on archival data. For example, as discussed above, the selection bias issues related

⁷ Rebels of course produce documentary evidence of their behavior as well, and in some cases these materials are cataloged in conflict archives. But owing in part to their need for secrecy and in part to their limited resources, these records tend to be much more limited. Still, a number of works rely upon recently released documents produced by rebels during armed conflict (e.g. Shapiro, 2013; Sullivan & Davenport, 2018). Also, one must not forget that victorious rebels become the government (e.g. in the case of the Spanish Civil War).

⁸ Data files for this research can be found at <http://www.prio.org/jpr/datasets>.

⁹ We kindly thank a JPR editorial board member for making this point.

to the data generating process are not always well addressed by researchers. Scholars should do a better job describing the data generating process might influence their study, and presenting ways to solve the bias. In addition, we still need to attend to the challenges that all scholars working with observational data face with regards to causal inference.

Overall, with its limitations, the unique information contained in conflict archives provides new opportunities for conflict scholars to identify new findings that illuminate the mechanisms and processes motivating violence. This special issue intends to establish a set of standard procedures by which the study of conflict via archives should be executed, and it presents some novel research using conflict archives in a groundbreaking way.

References

- Arjona, Ana (2016) *Rebelocracy: Social Order in the Colombian Civil War*. New York: Cambridge University Press.
- Art, David (2016) Archivists and Adventurers: Research Strategies for Authoritarian Regimes of the Past and Present. *Social Science Quarterly* 97(4): 974-990.
- Auyero, Javier (2007) *Routine Politics and Violence in Argentina: The Gray Zone of State Power*. New York: Cambridge University Press.
- Balcells, Laia (2012) The Consequences of Victimization on Political Identities: Evidence from Spain. *Politics & Society* 40(3): 311-347.
- (2017) *Rivalry and Revenge: The Politics of Violence during Civil War*. New York: Cambridge University Press.
- (2018). Dynamics of Internal Resettlement During Civil War. Evidence from Catalonia (1936-1939). *Journal of Peace Research* (this issue).
- Baum, Matthew A & Yuri M Zhukov (2015) Filtering Revolution: Reporting Bias in International Newspaper Coverage of the Libyan Civil War. *Journal of Peace Research* 52(3): 384-400.
- Braun, Robert (2016) Religious Minorities and Resistance to Genocide: The Collective Rescue of Jews in the Netherlands during the Holocaust. *American Political Science Review* 110(1): 127-147.
- Cohen, Dara Kay & Amelia Hoover Green (2012) Dueling Incentives: Sexual Violence in Liberia and the Politics of Human Rights Advocacy. *Journal of Peace Research* 49(3): 445-458.
- Cook, Scott J; Betsabe Blas, Raymond J. Carroll & Samiran Sinha (2017) Two Wrongs Make a Right: Addressing Underreporting in Binary Data from Multiple Sources. *Political Analysis* 25(2): 223-240.
- Cunningham, David. (2004) *There's something happening here: The New Left, the Klan, and FBI counterintelligence*. Berkeley, CA: University of California Press.
- Davenport, Christian (2009) *Media Bias, Perspective, and State Repression: The Black Panther Party*. New York: Cambridge University Press.
- (2010) Data from the Dark Side: Notes on Archiving Political Conflict and Violence. *PS: Political Science & Politics* 43(1): 37-41.
- (2014) *How social movements die*. New York: Cambridge University Press.

Davenport, Christian & Patrick Ball (2002) Views to a Kill Exploring the Implications of Source Selection in the Case of Guatemalan State Terror, 1977-1995. *Journal of conflict resolution* 46(3): 427-450.

Davenport, Christian & Will Moore (2014) Creating Conflict Data: Standards and Best Practices (<http://conflictconsortium.weebly.com/standards-best-practices.html>).

Dimitrov, Martin K. & Joseph Sassoon. (2014) State Security, Information, and Repression: A Comparison of Communist Bulgaria and Ba'thist Iraq. *Journal of Cold War Studies* 16(2): 3-31.

Douglas, Rex (2016) Understanding Civil War Violence through Military Intelligence: Mining Civilian Targeting Records from the Vietnam War. In: C.A. Anderton & J. Brauer (ed.) *Economic Aspects of Genocides, Mass Atrocities, and Their Prevention*. New York: Oxford University Press.

Douglass, Rex & Kristen Harkness (2018) Stalin's Terror and the Long-Term Political Effects of Mass Repression, *Journal of Peace Research* (this issue)

Eck, Kristine (2012) In Data we Trust? A Comparison of UCDP GED and ACLED Conflict Events Datasets. *Cooperation and Conflict* 47(1): 124-141.

— (2018) The Origins of Policing Institutions: Legacies of Colonial Insurgency. *Journal of Peace Research* (this issue).

Finkel, Evgeny (2017) *Ordinary Jews: Choice and Survival during the Holocaust*. Princeton, NJ: Princeton University Press.

Gohdes, Anita & Megan Price (2012) First Things First: Assessing Data Quality before Model Quality. *Journal of Conflict Resolution* 57(6): 1090-1108.

Gould, Roger V (1995) *Insurgent Identities. Class, Community, and Protest in Paris from 1848 to the Commune*. New York: Cambridge University Press.

Gould, Roger V (1999) Collective Violence and Group Solidarity: Evidence from a Feuding Society. *American Sociological Review* 64(3): 356-380.

Gould, Roger V (2003) *Collision of Wills*. Chicago, IL: University of Chicago Press.

Guberek, Tamy & Margaret Hedstrom (2017) On or off the Record? Detecting Patterns of Silence about Death in Guatemala's National Police Archive. *Archival Science*. 17(1): 27-54.

Hassan, Mai & Thomas O'Mealia (2018) Uneven Accountability in the Wake of Political Violence: Evidence from Kenya's Ashes and Archives. *Journal of Peace Research* (this issue).

Hendrix, Cullen S & Idean Salehyan (2015) No News is Good News: Mark and Recapture for Event Data when Reporting Probabilities are Less than One. *International Interactions* 4(2): 392-406.

Kalyvas, Stathis N (2006) *The Logic of Violence in Civil War*. New York: Cambridge University Press.

Kaplan, Oliver (2013) Protecting Civilians in Civil War: The Institution of the ATCC in Colombia. *Journal of Peace Research* 50(3): 351-367.

Krüger, Jule; Patrick Ball, Megan Price & Amelia Hoover Green (2013) It doesn't add up: Methodological and Policy Implications of Conflicting Casualty Data. In: Taylor B. Seybolt; Jay D. Aronson & Baruch Fischhoff (ed.) *Counting Civilian Casualties: An Introduction to Recording and Estimating Nonmilitary Deaths in Conflict*. Oxford: Oxford University Press.

Lawrence, Adria (2013) *Imperial Rule and the Politics of Nationalism: Anti-Colonial Protest in the French Empire*. New York: Cambridge University Press.

Loyle, Cyanne & Christian Davenport (2015) Transitional Injustice: Subverting Justice in Transition and Post-Conflict Societies. *Journal of Human Rights* 15(1): 126-149.

Meyer, John W; David Tyack, Joane Nagel & Audri Gordon (1979) Public Education as Nation-Building in America: Enrollments and Bureaucratization in the American States, 1870-1930. *American Journal of Sociology* 85(3): 591-613.

Morgan, T. Clifton (2013) The Secret Ingredient on Iron Chef-Road Kill! *Conflict Management and Peace Science* 30(1): 3-10.

Nalepa, Monika (2010) *Skeletons in the Closet: Transitional Justice in Post-Communist Europe*. New York: Cambridge University Press.

Osorio, Javier; Michael Weintraub & Livia Shubiger (2018) Disappearing Dissent? Repression and State Consolidation in Mexico. *Journal of Peace Research* (this issue)

Rude, George (1964) *The Crowd in History: A Study of Popular Disturbances in France and England, 1730-1848*. New York: John Wiley.

Salehyan, Idean (2015) Best Practices in the Collection of Conflict Data. *Journal of Peace Research* 52(1): 105-109.

Scharpf, Adam (2018) Ideology and State Terror: How Officer Beliefs shaped Repression during Argentina's 'Dirty War'. *Journal of Peace Research* (this issue).

Schwartz, Rachel & Scott Straus (2018) What Drives Violence against Civilians in Civil War? Evidence from Guatemala. *Journal of Peace Research* (this issue).

- Scott, James C (1998) *Seeing like a State. How Certain Schemes to Improve the Human Condition Have Failed*. New Haven, CT: Yale University Press.
- Seybolt, Taylor B; Jay D. Aronson, and Baruch Fischhoff (ed.) (2013) *Counting Civilian Casualties: An Introduction to Recording and Estimating Nonmilitary Deaths in Conflict*. Oxford: Oxford University Press.
- Shapiro, Jacob N (2013) *The Terrorist's Dilemma: Managing Violent Covert Organizations*. Princeton, NJ: Princeton University Press.
- Solé i Sabaté, J M & Joan Villarroya (1989) *La Repressió a la Reraguarda de Catalunya (1936-1939) [Translation]*. Barcelona: Publicacions de l'Abadia de Montserrat.
- Stoler, Ann Laura (2002) Colonial Archives and the Arts of Governance. *Archival science* 2(1): 87-109.
- Sullivan, Christopher M (2014) The (in) effectiveness of torture for combating insurgency. *Journal of Peace Research* 51(3): 388-404.
- Sullivan, Christopher M (2016a) Undermining Resistance Mobilization, Repression, and the Enforcement of Political Order. *Journal of Conflict Resolution* 60(7): 1163-1190.
- Sullivan, Christopher M (2016b) Political Repression and the Destruction of Dissident Organizations: Evidence from the Archives of the Guatemalan National Police. *World Politics* 68(4): 645-676.
- Sullivan, Christopher M and Christian Davenport. 2018. Evidence of Repression, Challenger Adaptation, and Surveillance from US “Red Squad” and Black Nationalist Archives. *Journal of Peace Research*. (this issue)
- Tilly, Charles; Louise Tilly, and Richard H. Tilly (1975) *The Rebellious Century*. Cambridge, MA: Harvard University Press.
- Weber, Eugen (1976) *Peasants into Frenchmen: The Modernization of Rural France, 1870-1914*. Palo Alto, CA: Stanford University Press.
- Weidmann, Nils B (2015a) On the Accuracy of Media-based Conflict Event Data. *Journal of Conflict Resolution* 59(6): 1129-1149.
- Weidmann, Nils B (2015b) Communication Networks and the Transnational Spread of Ethnic Conflict. *Journal of Peace Research* 52(3): 285-296.
- Weld, Kirsten (2014) *Paper Cadavers: The Archives of Dictatorship in Guatemala*. Durham, NC: Duke university Press.

Wisser, Katherine M & Joel A. Blanco-Rivera (2016) Surveillance, Documentation, and Privacy: An International Comparative Analysis of State Intelligence Records. *Archival Science* 16(2): 125-147.

Zhukov, Yuri M (2016) Taking away the Guns: Forcible Disarmament and Rebellion. *Journal of Peace Research* 53(2): 242-258.

Zhukov, Yuri M & Matthew A. Baum (2016) Reporting Bias and Information Warfare. Paper presented at the International Studies Association Annual Convention, Atlanta, GA, 16-19 March 2016.

Zhukov, Yuri M & Roya Talibova (2018) Stalin's Terror and the Long-Term Political Effects of Mass Repression. *Journal of Peace Research* (this issue).

Acknowledgements: We thank Gudrun Østby, Henrik Urdal, Marianne Dahl, Ragnhild Nordås, and the rest of the *JPR* editorial team for helping craft this special issue. We also want to thank the contributors and workshop participants from the International Studies Association and Peace Science conferences for helpful comments.

Biographical Statements

LAIA BALCELLS, b. 1980, PhD in Political Science (Yale University, 2010); Associate Professor of Government at Georgetown University (2017–); research interests: nationalism and ethnic conflict, civil wars, dynamics of political violence; terrorism.

CHRISTOPHER M. SULLIVAN, b. 1983, PhD in Political Science (University of Michigan, 2014); Assistant Professor of Political Science, Louisiana State University (2015–); research interests: order, violence, human rights; recent articles in: *Nature Human Behaviour*, *World Politics*, *Journal of Conflict Resolution*, and *Journal of Politics*.

Filename: Intro-FINAL-resubmission-EdOff-CS_11072017a.docx
Folder: /Users/sullivan/Library/Containers/com.microsoft.Word/Data/Documents
Template: /Users/sullivan/Library/Group Containers/UBF8T346G9.Office/User
Content.localized/Templates.localized/Normal.dotm
Title:
Subject:
Author: christopher sullivan
Keywords:
Comments:
Creation Date: 3/9/18 10:30:00 AM
Change Number: 2
Last Saved On: 3/9/18 10:30:00 AM
Last Saved By: Microsoft Office User
Total Editing Time: 0 Minutes
Last Printed On: 3/9/18 10:30:00 AM
As of Last Complete Printing
Number of Pages: 26
Number of Words: 6,994
Number of Characters: 41,473 (approx.)