Introducing HiSCoD: A New Gateway for the Study of Historical Social Conflict

CÉDRIC CHAMBRO† & PAUL MANEUVRIER-HERVIEU‡

March 2022

Abstract

Social conflict pervades human society and fulfils a number of essential functions in its development and transformation, including the creation of new norms and institutions. The Historical Social Conflict Database (HiSCoD) is an ongoing project designed to provide to scholars and society at large with a set of resources for analysing social conflict from the Middle Ages to the late 19th century. Based on original archival research and existing repositories, the aim is to provide a global database of social conflict in past societies by collecting, aggregating, documenting and harmonising data. As of today, the database contains data more than 20,000 instances of conflict, from fiscal scuffles to urban revolts involving thousands of individuals. For every event, we provide information on the date, location, type of conflict, and, when possible, number of participants, participation of women, and a summary of events. Each individual event is documented through a hierarchical system of forms using XML-EAD technology. This article describes the data collection process and presents some descriptive statistics.

JEL classifications: D74, H56, N00, N90.

Keywords: Event data; Sub-national; Social conflict; Riot; Violence; Protest.

*We are grateful to John Bohstedt, Maïwenn Bourdic, Samuel Cohn, Chantal Coudert, Patrick J. Crawford, Le Désarmement havrais, Renan Donnerh, the Family and Community Historical Research Society, Pierre-Louis Gatier, Julian Langer, Aurélien Lignereux, Jean Nicolas, Michele Rosenberg, Matteo Tiratelli, Rémy Vanier, Maria Waldinger, and Karl-Alexandre Zimmer for kindly sharing data. We also would like to thank Pascal Buléon and the Digital Cluster team (Pierre-Yves Buard, Subha-Sree Pasupathy, Anne Goloubkoff, and Julia Roger) at the Maison de la Recherche en Sciences Sociales at the University of Caen Normandie for its technical support. All errors are our own.

†Cédric Chambru, Department of Economics, University of Zurich, Zürichbergstrasse 14, 8032 Zürich, Switzerland. Email: cedric.chambru@econ.uzh.ch.

‡Paul Maneuvrier-Hervieu, Department of Social and Political Sciences, University of Milan, Via Passione, 13, 20122 Milano, Italy. Email: paul.manevuier@unimi.it.
1 Introduction

Over centuries, social conflict has been at the centre of human relations, fulfilling a number of functions essential to the development and transformation of societies (Cohn 2006; Tilly 1986). As an expression of tensions relating to values, interests and power between or within groups, they shed light on critical issues such as the nature of sovereignty and the power of state, the scope for popular resistance, and the expression of distress. On many occasions, they precipitated major social change by generating new values, new norms, and new institutions or by contributing to the formation of new group identities (Thompson 1964). Analysis of the reactions provoked by social conflict and the nature of its interaction with protest, revolt, and collective violence is fundamental to the study of social systems. In this regard, social conflict is arguably the most relevant indicator for researchers seeking to understand how individuals related to their environment. As March Bloch noted in the early 1930s, “agrarian revolt appears to be as inseparable from the seigneurial regime as strikes from large capitalist factories” [authors’ translation] (Bloch 1931, p. 175).

Since the 19th century, social conflict has been a major feature of social and political research. Historians and social scientists have used it to analyse a wide range of topics such as the construction of modern states (Mousnier 1958; Root 1992), economic crises and the politics of provisioning (Bohstedt 2010; Tilly 1975), the use and the meaning of collective violence (Tilly 2003), and the development of political consciousness (Nicolas 1985; Porchnev 1963; Thompson 1971). Several large-scale projects have indexed information on contemporary social conflict by time and geography and their data is readily available online. For instance, the Armed Conflict Location & Event Data Project (ACLED) gathers information on political violence and protest for a diverse range of countries from the mid-1990s onwards (Raleigh et al. 2010).1 The Social Conflict Analysis Database (SCAD) focuses on non-armed conflict and includes various forms of protest and riots in Central America and Africa between 1990 and 2017 (Salehyan et al. 2012).2 The xSub project takes a different approach, aiming to creating a ‘database of databases’ on armed conflict in order to preserve and consolidate existing large data collections on armed conflict (Zhukov et al. 2019).3 More recently, Kitamura (2021) has released the World Historical Battles Database, which enlarges and expands existing repositories (Brecke 1999; Clodfelter 2017) by providing additional information such as the location, year and outcome of more than 8,000 battles.4 To the best of our knowledge, however, no such large-scale resource has been developed for historical social conflict.

The debate between Boris Porchnev (1954, 1963) and Roland Mousnier (1958, 1967) on the origins of rebellions during the 17th century, and later the groundbreaking work of Edward P. Thompson (1971), generated a large number of monographs, articles and theses on disorder and revolts in Europe and elsewhere. From the 1960s onwards, several research projects began to retrieve data from archival documents (e.g. Aoki 1971; Beik 1997; Bercé 1974; Gailus 1990;

---

4 https://osf.io/mdjzu.
Hobsbawm and Rudé 1968). Despite these efforts, data sets describing these events are sparse and rarely accessible since they were generally put together before the 2000s and the era of digitalisation. The list of thousands of events compiled by Jean Nicolas and his collaborators for France between 1661 and 1789, for example, remained inaccessible until recently (Nicolas 2002).

The Historical Social Conflict Database (HiSCoD) is designed to fill this gap by providing an online tool with a set of resources for analysing social conflict on a global scale from the Middle Ages to the late 19th century. The project brings together three strands of research: the previous consolidation of historical social conflict by scholars—whether in the forms of data sets or within their publications; the collection of new instances of conflict from primary sources; the long-term preservation of data to avoid the loss of knowledge. We hope that this ongoing project will facilitate new research through the provision of reliable qualitative and quantitative data on historical social conflict.

The remainder of the paper is organised as follows. Section 2 provides an overview of the methodology used to assemble the data set, with a particular emphasis on sources, definition, and classification. Section 3 presents a global overview of the database and some descriptive statistics. Section 4 offers a conclusion and discusses limitations.

2 Methodology

As of March 2022, the Historical Social Conflict Database (HiSCoD) contains information on more than 20,000 episodes of social conflict from the Middle Ages to the late 19th century. The events compiled are drawn from a number of large-scale surveys conducted by scholars since the 1950s, dozens of case studies by local historians since the 19th century, and new archival research we have undertaken ourselves. In this section, we briefly summarise the data sources before presenting the methodology and definition used to collect and align all data points.

2.1 Data sources

Events recorded in the HiSCoD database, whether they have been directly retrieved from the archives or collected from previous works, come from a wide range of primary documents traditionally used by historians to study social conflict. Regardless of the country in question, the nature of available sources depends above all on the historical period under study. Information on medieval and early modern social conflict is found in chronicles, annals, journals, judicial records and diaries, whereas information from the 18th and 19th centuries can be found in social movements, protests, and strikes.

Footnotes:

5 For a recent overview of this literature, see Ruff (2020). For an overview of the literature on uprisings and revolts during the Middle Age, see Firnhaber-Baker and Schoenaers (2017).
6 The data were retrieved, cleaned, aligned and geo-coded by Chambru (2019b), before we released them online in November 2020 in the beta version of the HiSCoD project.
7 https://www.unicaen.fr/hiscod.
8 We use the XML-EAD standard, a machine-readable format and an international archival standard, which allows information to be standardised within and across repositories, to ensure the interoperability and sustainability of researchers’ work while making data easily accessible.
be retrieved from administrative correspondence, judicial records, municipal council records, private papers, and/or local newspapers.\footnote{For an example of archival materials used in this project, see online Appendix Figure A.1. For an extensive description of the sources available and their limitations for the Middle Ages, see Cohn (2006, pp. 13–21). For the 17th and 18th centuries, see Fletcher and Steveson (1985, pp. 26–31) and Nicolas (2002).}

Given the wide range of sources, the level of information available on individual instances of social conflict is heterogeneous and disparate. It is sometimes possible to follow the course of events hour by hour through police reports and trial records, while in other cases evidence identifies only the location of the event. To address this imbalance we use a simple form to standardise existing data sets and collect new data (see Section 2.5). This means that we choose not to include information such as participants’ occupations, to allow us to pull together different data sets and offer a unified database.\footnote{For instant, Lignereux (2008) reports quite systematically detailed information on the presence of police force, an information rarely included in other data set. It should be noted that we maintain a unique identifier allowing to backward link each event in HiSCoD with its corresponding form in the original databases.}

### 2.2 Definition

The issue of what comprises a social conflict is one of the most highly debated topics among historians. It is difficult to reach a consensus on the characteristics and criteria required to define an event as a social conflict because they differ vastly in form and scale (Bercé 1974; Charlesworth 1983; Foisi 1970; Fourquin 1972; Hilton 2003; Mollat and Wolff 1970; Neveux 1988; Nicolas 2002; Randall 2006, chap. 12; Bohstedt 1983, 2010). In his study on food riots in England, Bohstedt (2010, p. 17) doubted that “anyone will ever compile a complete record of all the riots in a period” because it is “impossible ‘by definition’ for two lists to agree if they did not agree on a definition of the events they were collecting in the first place”. This was precisely the task we set ourselves: to merge and combine these disparate data sets on riots, revolts, popular disturbances, seditions and the like into a unified, analysis-friendly format within a database.\footnote{In this regard, the HiSCoD project closely follows the objectives put forward by the xSub project (Zhukov et al. 2019).} We use the expression ‘social conflict’ as a generic category to encompass all the expressions used by scholars and/or mentioned in the sources. We were also guided by Jean Nicolas’ seminal work, relying on a broad definition when categorising an event as social conflict (Nicolas 2002, p. 75). This allows us to grasp infinitesimal instances of discontent and protest over time and geography and to incorporate as many archives as possible created by scholars since the 19th century.

We consider social conflict to be any event involving a group of at least three individuals belonging to different families and which either perpetuates violence or threatens violence against one or more members of a different group or against representatives of political, religious, and economic power; or any event involving an attack on property, buildings, furniture, papers or other signs symbolising such powers.\footnote{We deliberately used a low number of participants, as an inclusion criterion, not exclude many studies with stricter criteria. For instance, Bohstedt (2010) only considers events with more than 50 participants, whereas Bercé (1974) listed events with more than 20 participants. Aoki (1971) used no such threshold (see White (1995, Appendix 1) for a discussion of Aoki’s data in English). For a discussion of what constituted a popular revolt as discussed by today’s historians, see Zhukov et al. (2019).}
Table 1: Overview of the main existing datasets included in HiSCoD

<table>
<thead>
<tr>
<th>Geographic area</th>
<th>Date range</th>
<th>Type of conflict</th>
<th># of obs.</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>France</td>
<td>1661–1789</td>
<td>All</td>
<td>8,528</td>
<td>Nicolas (2002)</td>
</tr>
<tr>
<td>England</td>
<td>1830–32</td>
<td>All</td>
<td>2,205</td>
<td>Holland (2005)</td>
</tr>
<tr>
<td>Italy, France, Flanders</td>
<td>1090–1435</td>
<td>All</td>
<td>1,112</td>
<td>Cohn (2006)</td>
</tr>
<tr>
<td>England</td>
<td>1347–1819</td>
<td>Food riot</td>
<td>1,008</td>
<td>Bohstedt (2010)</td>
</tr>
<tr>
<td>Normandy</td>
<td>1709–1817</td>
<td>Food riot</td>
<td>916</td>
<td>Maneuvrier-Hervieu (2020)</td>
</tr>
</tbody>
</table>

Notes: Bohstedt (2010) provides lists of food riots for various years between 1347 and 1819.

2.3 Data collection

The first step of the project was to collect and align several large-scale, unpublished data sets (Table 1). It is worth noting that we did not deliberately limit the geographical coverage of the data sets to Western Europe; this, with the exception of China and Japan, reflects the current literature on historical social conflict.

To improve the geographical coverage, we retrieved data from several dozen regional studies carried out by historians and local historians since the 19th century. In some cases, we assembled data directly from lists and tables in the paper/book in question, whereas in other cases, we carefully transcribed information on the events mentioned in such documents. Most of the data on social conflict during the French Revolution, for example, comes from the comprehensive descriptions provided by Ado (1996) and Lefebvre (1973) in their work on French peasantry. In addition, we also occasionally undertook detailed searches for social conflict in archival sources. Considering the quantity and scope of archival materials available across countries, this effort is necessarily limited without the enlistment of international scholars. We therefore opened up HiSCoD to volunteer contributions from external scholars who had collected information on social conflict in the course of their own archival research (e.g. Tiratelli 2019). We carefully reviewed and curated all these data, which ranged from one event to few hundred episodes of social conflict, before integrating them in the database.

2.4 Typology

Pre-industrial popular disturbances included a broad spectrum of events from grain and bread riots to crafts revolts, tax riots, religious conflicts and many more. Tilly (1976, p. 375) insists opposed to actions to be taken isolated individuals or families against outsiders, see Cohn (2006, pp. 5–13).

The HiSCoD project is an ongoing effort and does not include a comprehensive list of existing data set. There are many avenues to improve the scope of the database (e.g. Aoki 1971; Gailus 1990; Griffin 2001).

For instance, on the lack of survey for on pre-industrial Germany, see Gailus (1990, p. 160). For an overview of the Chinese historiography on peasant rebellions, see Wakeman and Grant (1975). Yang (1975) provides a useful discussion on the scope and coverage of the data for 19th-century China.

These two monographs have significantly helped to deepen the knowledge of popular disturbances and their geographical distribution across France during the early years of the 1790s. Needless to say, they are not exhaustive.
Table 2: Classification used in the HiSCoD database

<table>
<thead>
<tr>
<th>Category</th>
<th>Numeric values</th>
<th>Description</th>
<th># of obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food riot</td>
<td>1</td>
<td>Includes riots protesting against high prices, popular taxation, interception of food transport, and food looting</td>
<td>4,226</td>
</tr>
<tr>
<td>Tax riot</td>
<td>2</td>
<td>Includes smuggling-related events, and riots protesting against taxes (local, feudal or national)</td>
<td>4,297</td>
</tr>
<tr>
<td>Religious conflict</td>
<td>3</td>
<td>Includes conflict relating to the practising of religion</td>
<td>649</td>
</tr>
<tr>
<td>Conflict with local or national authorities</td>
<td>4</td>
<td>Includes conflict with the military (including deserters and draft dodgers), police, marshals, parliaments, municipal authorities, militias or any conflict protesting against a decision taken by authorities</td>
<td>3,550</td>
</tr>
<tr>
<td>Feudal conflict</td>
<td>5</td>
<td>Includes attacks on castles and conflicts concerning the abolition of feudal rights</td>
<td>940</td>
</tr>
<tr>
<td>Slave revolt</td>
<td>6</td>
<td>Includes slave revolts either on boats, at gatherings, and/or maroons</td>
<td>3</td>
</tr>
<tr>
<td>Political conflict</td>
<td>7</td>
<td>Includes all conflicts protesting against a political regime during a revolutionary or transitional period, and any opposition to a political regime aiming to change it</td>
<td>1,089</td>
</tr>
<tr>
<td>Labour conflict</td>
<td>8</td>
<td>Includes strikes, collective bargaining, conflict against guilds or between different guilds, wage riots, etc.</td>
<td>1,403</td>
</tr>
<tr>
<td>Banditry</td>
<td>9</td>
<td>Includes robberies by a group of individuals or events related to organised crime</td>
<td>0</td>
</tr>
<tr>
<td>Other forms of conflict</td>
<td>10</td>
<td>Includes all conflicts which do not fit in previous categories</td>
<td>4,458</td>
</tr>
</tbody>
</table>

Notes: We categorised events related to seigneurial taxes as tax riots rather than feudal conflict. In 334 instances, the lack of information in the sources pre-empted the attribution to an exact category.

that “to categorize is a first step on the way to identifying what there is to explain, and therefore on the way to explaining it”. Every historian who has written a monograph or conducted a large-scale survey of historical social conflict has thus been confronted with the challenge of classifying the diverse events she has identified. This means that the literature includes dozens of typologies, sometimes involving very different categories depending on the social and historical context. Both Cohn (2006) and Nicolas (2002), for instance, identify no less than 70 different types of conflict in their data.\footnote{For a discussion on the creation of typology and the changing patterns of social conflict, see e.g. Cohn (2006, pp. 76–107), Pillorget (1973, pp. 141–154), Shoemaker (1987, pp. 77–80) and Tilly (1976).}

The main objective of the HiSCoD project is to act as gateway for the study of historical social conflict by gathering existing information from a wide range of studies and sources. Therefore, we chose to use a relatively simple classification to encompass a wide range of geographic areas and a long time span (Table 2). We used the description of the course of events and, when relevant, the researchers’ original typology to assign each social conflict to one of the 10 types. It should be reminded that historical social conflict involved a variety of participants, often with a range of different motives, with riots protesting against high crop prices occasionally leading to further protests about wages or taxation, for instance, and/or
conflict with local authorities about setting arrested rioters free. We would highlight here that our classification is first and foremost intended as a tool to guide users through the myriad of events reported.

2.5 Structure of the database

Whenever possible, we created one entry for any social conflict on a specific day in a given location. To situate each event in space and time, we recorded time units (year, month, date), historical and contemporary location, and geographic coordinates (longitude, latitude), along with an indicator for geographic precision. In addition to the summary of the course of events, we added two variables: the number of participants and an indicator for the participation of women. Table 3 provides an overview of information available for each entry.

If several sources and/or researchers reported the same event, we merged all information to create a unique entry.

In few instances, we were not able to identify the location of the event. We kept the observations in the database, but left the geographic coordinates empty. For more details on the procedure to add coordinates to any event, see online Appendix Section A.1.

Online Appendix Table A.1 provides a detailed list of variables contained in the database.
3 The HiSCoD database

3.1 Data preservation

We designed the HiSCoD project with the aim of creating a platform to ensure that data on historical social conflict was preserved and to make sure that such data can easily located, accessed and reused. To meet our requirements for interoperability and sustainability, we chose to encode the database using the EAD standard. The *Encoded Archival Description* standard is a public domain XML schema created and used by archivists, and museums and national libraries, for encoding archival finding aids. Its markup structure makes it a machine-readable format, enabling data to be shared easily across platforms.

This solution guarantees interoperability and that data will be preserved and accessible. Leveraging the encoding system, we generated a standardised form for each social conflict that can be downloaded in either English or French. Each form contains information on the date of the event, its location (at the municipal level) and the name of the historical political entity in which the event occurred. Whenever possible, we also provided details on the number of participants, on the participation of women, and a brief summary, which narrates the course of the action. We also included bibliographic references and archival signatures for users looking for more detailed information about the event.

3.2 Data accessibility

Users can access data in one of three ways. First, they can navigate the interactive map on the project website to identify events of interest and access individual forms.\(^{20}\) Each form is available in both English and French and can be manually downloaded as pdf.\(^{21}\) Second, they can access a user-friendly semantic repository, based on the EAD standard, to either search for keywords or explore data by interacting with a data access system. Finally, the most up-to-date version of the data set can be download as csv file from the project’s Github repository.\(^{22}\) A detailed description of the variables included in the data set is available in online Appendix Table A.1.

3.3 Data overview

This section provides an overview of the database as of March 2022 and presents a few descriptive statistics on the 20,000 or so events to introduce users to the current scope of the project. Figure 1 shows the distribution of historical social conflict across countries from the Middle Ages to the late 19th century. The vast majority of events are located in Europe, and in particular France and England (about 92 per cent). As noted in Section 2.3, this bias to some extent reflects the availability and/or accessibility of data in the literature. We are continuing to

\(^{20}\) [https://www.unicaen.fr/hiscod](https://www.unicaen.fr/hiscod).

\(^{21}\) It should be noted that to avoid the overcrowding of information on pdf forms, we chose to report only the most important variables. For accessing the full data set, we advise using the csv files available on the project’s Github repository.

\(^{22}\) [https://github.com/hiscod](https://github.com/hiscod).
work to remedy this issue and expand the geographical coverage of the database. The absence of recordings of social conflict for a country should not, however, necessarily be interpreted as resulting from missing data. It might also be the result of the historical context particular to a country. For instance, whereas food riots were endemic in France and England during the 18th century, they were relatively uncommon in Spain before 1766 (Rodríguez 1973, p. 145).

Figure 2 shows the temporal distribution of the events recorded in the database. The unequal distribution of social conflict originates both from survival bias in relation to sources and available data, and from the fact that the 17th and 18th centuries witnessed evolution and transformation in the forms of popular disturbance (Tilly 1976). The large-scale peasant or urban revolts inherited from the medieval and early modern period, such as the events of the Peasants’ War in Germany or the revolts in France in the first decades of the 17th century, were gradually succeeded by smaller and more localised episodes of social conflict (Bercé 1974; Cohn 2006; Nicolas 2002). We used a 10-category classification system to categorise social conflict based on primary motives, forms, targets and actions (Section 2.4). We believe this variable is an effective way of entering the database and navigating the thousands of events it includes. The lack of information in the sources also pre-empted attribution to an exact category in a small number of entries. In the future, we hope to be able to refine the entries for these events further by adding information from new sources.

Figure 3 provides an overview of the number of entries for which we were able to document the main variables included in the database: number of participants, participation of women, summary of the course of events, archival signatures, and bibliographic references. 12,201 entries (58 per cent) contain a description of more than five words in length in either English
or another language. Unfortunately, in most cases, it is impossible to document the number of participants and/or the participation of women with any certitude due to the paucity of information in the primary and secondary sources. We believe these entries are still useful for users wishing to identify particular subsets of events. Finally, we also report the number of entries in respect of which we have at least one archival signature and two bibliographic references. While we do our best to include as many references as possible, it is difficult to identify systematically all sources of information on a specific instance of social conflict. Interested users can use these references as an entry point to the historiography of specific events and consult the bibliographies of these monographs and articles.

3.4 Data use cases

A number of scholars have already used historical data on social conflict to study a wide range of questions in economics, history, and/or political science. Although replication is now common practice in the social sciences, such studies have rarely included raw data, which makes it difficult to re-use them beyond the scope of the publication, e.g. in order to analyse social conflict at city level rather than county level.

Aidt and Franck (2015), for instance, investigate the effect of exposure to violent social conflict on voting behaviour during the 1830s in England. They use data on the Swing riots (Holland 2005) to show that voters’ and patrons’ fear of rioters in the immediate neighbourhood of their constituency shifted their allegiance towards the reform-friendly Whigs during the 1831 election. In another paper, Aidt, Leon-Ablan et al. (2021) use the same

---

Figure 2: Number of social conflicts per quarter century

---

To facilitate research, we included a direct link to online archive catalogues for more than 2,650 events.
The database includes 9,1616 entries for which we can ascertain the involvement of women. The database includes 12,201 entries with a description of more than five words in length. The database includes 2,935 entries with at least two bibliographic references.

Figure 3: Available information for individual social conflicts

data set to analyse information flows and the impact of news about repression on participation in riotous events. They also document the role of local organisers and personal and trade networks in the geographic distribution of these riots.

In another strand of work, Chambru (2019a) and De Juan and Wegenast (2020) use data from Nicolas (2002) and Bohstedt (2010) respectively to analyse the effect of temperature variability on social conflict in early modern France and England. Chambru further documents the key role of state institutions and their agents in the alleviation of the adverse effects of weather shocks. Touzery (2022) studies the rise of fiscal capacity in France between the 14th and the late 18th century and the increasing opposition to the tax burden on the part of people and communities.

As of March 2022, information on most of the events listed in the HiSCoD was not publicly and easily accessible. We hope that this ongoing project will facilitate new research on social conflict during the pre-industrial era. Scholars could use it to investigate topics such as the changing forms of conflict over time, the geographical distribution of social conflict, the resilience of societies before climate variability and natural disasters, the relationship between the rise of the modern state and social conflict, and many other topics.

4 Conclusion

This article presents a new gateway for the study of worldwide historical social conflict from the Middle Ages to the late 19th century. To the best of our knowledge, the HiSCOD project is
the first attempt to produce such a database and make it available to the academic community. One of our main objectives is to create a tool for ensuring that knowledge is not lost and that data is preserved in the long term. To meet this goal, we created standardised forms containing the essential information on more than 20,000 events and encoded them using the EAD standard, a public domain XML schema created by archivists.

It should be noted that HiSCoD is an ongoing project with limitations and avenues for improvement, which implies that caution should be exercised when using it. Despite the efforts devoted to the collection of data, the heterogeneity of the surviving archives and previous large-scale surveys on historical social conflict means that there are significant disparities in the database’s geographic coverage. For instance, as of today, it is difficult to undertake cross-country comparisons going beyond the food riots in England and France or of the medieval revolts collected by Cohn (2006). This caveat arose from the aggregation of the works of different scholars in different times and places. To circumvent this issue, we continuously update existing entries with additional information and add new episodes of social conflict to expand the temporal and geographical coverage of the database. The existence of several repertoires of historical social conflict in countries, such as China (Yang 1975), England (Griffin 2001), France (Sottocasa 2004), Germany (Gailus 1990), Japan (Aoki 1971), and Nordic countries (see Mikkelsen et al. 2018) provide many avenues for further development. One of the main challenges will nonetheless be to include geographic areas for which the historiography and archives on historical social conflict are limited. We would like to call on users, researchers and citizen scientists who are interested in expanding knowledge of social conflict in past societies to contribute to this ongoing project by sharing any relevant information.

24 Cohn (2019). Cohn (2006, p. 22) reminds that one should not forget the existence of bias because “it comes as little surprise that the cities with the strongest chronicle traditions are often those with the greatest number of revolts”. For a recent attempt at comparative history, see Cohn (2019). For earlier attempts, see e.g. Mousnier (1967), Bin Wong (1983) and Davies (1973).
References


Online appendix

Introducing HiSCoD: A New Gateway for the Study of Historical Social Conflict

CÉDRIC CHAMBRU
University of Zurich

PAUL MANEUVRIER-HERVIEU
University of Milan
A  Additional information on HiSCoD

In this section, we provide additional information on the construction of the database. Section A.1 describes the procedure used to attribute geographic coordinates to events. In Section A.2, we detail how we coded information on the participation of women. In Table A.1, we summarise the variables contained in the database.

A.1 Geographic coordinates and administrative units

Attributing geographic coordinates to thousands of historical locations is obviously a challenging exercise. Over centuries, many locations merged with neighbouring ones, changed their spelling, or simply disappeared because of war and plagues. This makes automatic geo-coding with services such as Geonames very inaccurate.\(^1\) Douglass and Harkness (2018) and Hammond and Weidmann (2014) both discuss the spatial accuracy of machine-coded datasets at the sub-national level and how it could create additional sources of bias in empirical analyses. For instance, Füglister (2020) uses data from Clodfelter (2017) on historical battles and shows that 20 per cent of the machine-coded battles in Europe have a spatial error of more than 22.5 km— a number rising to 88 km after the 85th percentile (p. 48).

We therefore relied on historical and contemporary records of places to identify locations where events occurred as precisely as possible.\(^2\) We then manually matched it with its current municipality and retrieve data on coordinates through either Wikipedia, Google Maps, or data sets from national statistical offices, such as Admin Express from the French National Institute of Geographic and Forest Information (IGN) and GeoBasis-DE from the German Federal Agency for Cartography and Geodesy (BKG).\(^3\)

As a rule of thumb, we tried to attribute to locations coordinates that corresponded to the centre of the present-day municipality, usually near the city hall, market place, historical church or similar.\(^4\) This means that there could be few hundred meters or, in the case of merger between two municipalities, a few kilometres difference between the coordinates and the exact historical location. In most cases, we have been able to identify a precise location.\(^5\)

However, in a few instances, the sources did not record a specific location, but rather the area or the administrative unit (county, district, etc.). In such cases, we placed the coordinates at the capital city or the most of city of the administrative unit. We flagged these points by attributing the value “2” to the variable geo_precision.

---

\(^1\)  http://www.geonames.org

\(^2\)  For instance, the Gazetteer of British Place Names (https://gazetteer.org.uk) and the French Topographic Dictionary (https://dicotopo.cths.fr).

\(^3\)  https://geoservices.ign.fr/adminexpress,


\(^4\)  The only exception is Great Britain for which the Index of Place Names, produced by Office for National Statistics, sometimes allow for geo-coding at a finer level. https://www.ons.gov.uk/methodology/geography/geographicalproducts/otherproducts/indexofplacenamesipn.

\(^5\)  When needed, we noted additional information about the location and geocoding procure in the variable comments.
In a smaller number of cases, the sources only provide broad geographic information, such proximity to the coast, a river or provinces.\(^6\) Whenever possible, we used the coordinates of the capital/main city of the administrative unit and flagged it by attributing the value “3” to the variable geo_precision. These entries typically do not have a contemporary record of location. We use these coordinates only for mapping purposes and extreme caution should be exercised before undertaking any spatial analysis with them.

Finally, we could not retrieve the location of a small proportion of events because the place name corresponded to nothing or the description was too vague. We included these events in the database but left the coordinates fields empty.

To document contemporary administrative units, we matched our data points with vector data for administrative boundaries provided by the relevant national institute for geographic information (German BGK, French IGN, Italian ISTAT, British ONS, Swisstopo, etc.). Where necessary, we retrieved additional information from the Wikipedia page for the municipality. We also made use of the historical administrative boundaries provided by Chambru (2019) and Satchell et al. (2018) to add information on the historical administrative divisions of France and England.

A.2 Participation of women

Identifying the exact composition of a crowd is often a difficult exercise, even if the information would be particularly useful for researchers. For instance, how should one interpret the omission of women from the summary of the course of events. Did women actually not take part into the conflict or did the witness omit them? To circumvent this issue, we used a three-level variable to indicate the participation of women: “Yes” means that the sources or the researcher clearly state that women took part in the event; “No” means that they clearly did not; “Unknown” means we cannot claim that they did or they did not on the basis of the available information.

\(^6\) This is particularly the case for many social conflicts during Medieval times, which were large-scale events encompassing entire regions. It should be noted that we did not include event that were reported as nationwide.
<table>
<thead>
<tr>
<th>Variable name</th>
<th>Type</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>id</td>
<td>Numeric</td>
<td>Numerical identifier</td>
<td>1 to 99999</td>
</tr>
<tr>
<td>id riot hiscod</td>
<td>Character</td>
<td>Unique identifier</td>
<td></td>
</tr>
<tr>
<td>id riot original database</td>
<td>Character</td>
<td>Unique identifier from previously existing data set</td>
<td></td>
</tr>
<tr>
<td>title</td>
<td>Character</td>
<td>Title of the event based on the location where it occurred and its date</td>
<td></td>
</tr>
<tr>
<td>riot date</td>
<td>Date</td>
<td>Date of the event</td>
<td>1734-05-22, etc.</td>
</tr>
<tr>
<td>year</td>
<td>Numeric</td>
<td>Year</td>
<td>1000 to 1900</td>
</tr>
<tr>
<td>month num</td>
<td>Numeric</td>
<td>Month</td>
<td>1 to 12</td>
</tr>
<tr>
<td>month</td>
<td>Character</td>
<td>Month</td>
<td>January, February, etc.</td>
</tr>
<tr>
<td>day</td>
<td>Numeric</td>
<td>Day</td>
<td>1 to 31</td>
</tr>
<tr>
<td>day week</td>
<td>Character</td>
<td>Day of the week</td>
<td>Monday, Tuesday, etc.</td>
</tr>
<tr>
<td>riot type hiscod num</td>
<td>Numeric</td>
<td>Numeric ID for classification</td>
<td>1 to 10</td>
</tr>
</tbody>
</table>

| riot type original database 1 | Character           | Original classification used by researchers       | 5. Feudal conflict            |
| riot type original database 2 | Character           | Original classification used by researchers       | 6. Slave revolt              |
| riot type original database 3 | Character           | Original classification used by researchers       | 7. Political conflict         |
|                               |                     |                                                  | 8. Labour conflict           |
|                               |                     |                                                  | 9. Banditry                  |
|                               |                     |                                                  | 10. Other forms of conflict  |
Table A.1: List of variables included in the HiSCoD database (cont’d)

<table>
<thead>
<tr>
<th>Variable name</th>
<th>Type</th>
<th>Description</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>nb participants</td>
<td>Character</td>
<td>Number of participants involved in the event</td>
<td></td>
</tr>
<tr>
<td>women participation</td>
<td>Character</td>
<td>Indicate whether women were mentioned among the participants</td>
<td>Yes, No, Unknown</td>
</tr>
<tr>
<td>city source</td>
<td>Character</td>
<td>Name of the location mentioned in the sources</td>
<td></td>
</tr>
<tr>
<td>city name</td>
<td>Character</td>
<td>Name of the current municipality</td>
<td></td>
</tr>
<tr>
<td>city code</td>
<td>Character</td>
<td>Codes used by national statistical office to identify municipality</td>
<td></td>
</tr>
<tr>
<td>city latitude</td>
<td>Numeric</td>
<td>Latitude</td>
<td></td>
</tr>
<tr>
<td>city longitude</td>
<td>Numeric</td>
<td>Longitude</td>
<td></td>
</tr>
<tr>
<td>geo precision</td>
<td>Numeric</td>
<td>Code to indicate the precision of coordinates</td>
<td>1: exact, 2: nearest location, 3: imprecise</td>
</tr>
<tr>
<td>country name</td>
<td>Character</td>
<td>Name of the present-day country where the event occurred</td>
<td>France, Italy, etc.</td>
</tr>
<tr>
<td>admin level 1 type</td>
<td>Character</td>
<td>Type of the highest-level administrative division</td>
<td>Region, Land, State, Province, etc.</td>
</tr>
<tr>
<td>admin level 1</td>
<td>Character</td>
<td>Name of the highest-level administrative division</td>
<td></td>
</tr>
<tr>
<td>admin level 2 type</td>
<td>Character</td>
<td>Type of the secondary administrative division</td>
<td>Department, County, Powiat, etc.</td>
</tr>
<tr>
<td>admin level 2</td>
<td>Character</td>
<td>Name of the secondary administrative division</td>
<td></td>
</tr>
<tr>
<td>historical political entity</td>
<td>Character</td>
<td>Name of the political entity in existence when the event occurred</td>
<td>Holy Roman Empire, etc.</td>
</tr>
<tr>
<td>historical admin level 1 type</td>
<td>Character</td>
<td>Type of the historical highest-level administrative division</td>
<td></td>
</tr>
<tr>
<td>historical admin level 1</td>
<td>Character</td>
<td>Name of the historical highest-level administrative division</td>
<td></td>
</tr>
<tr>
<td>historical admin level 2 type</td>
<td>Character</td>
<td>Type of the historical secondary administrative division</td>
<td></td>
</tr>
<tr>
<td>historical admin level 2</td>
<td>Character</td>
<td>Name of the historical secondary administrative division</td>
<td></td>
</tr>
<tr>
<td>description event english</td>
<td>Character</td>
<td>Summary of the course of events in English</td>
<td></td>
</tr>
<tr>
<td>description event original</td>
<td>Character</td>
<td>Summary of the course of events in language used by researchers or the archival sources</td>
<td></td>
</tr>
<tr>
<td>Variable name</td>
<td>Type</td>
<td>Description</td>
<td>Values</td>
</tr>
<tr>
<td>---------------</td>
<td>----------</td>
<td>-------------------------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>primary sources</td>
<td>Character</td>
<td>Archive signatures</td>
<td></td>
</tr>
<tr>
<td>url</td>
<td>Character</td>
<td>Link to the archival repository, catalogues, and search aids</td>
<td></td>
</tr>
<tr>
<td>bibliography</td>
<td>Character</td>
<td>Bibliographical references</td>
<td></td>
</tr>
<tr>
<td>pages</td>
<td>Character</td>
<td>Pages where the event is mentioned</td>
<td></td>
</tr>
<tr>
<td>comments</td>
<td>Character</td>
<td></td>
<td></td>
</tr>
<tr>
<td>author</td>
<td>Character</td>
<td>Name of the author(s) of the form</td>
<td></td>
</tr>
<tr>
<td>contributor</td>
<td>Character</td>
<td>Name of the contributor(s)</td>
<td></td>
</tr>
<tr>
<td>date creation</td>
<td>Date</td>
<td>Date to indicate the creation of the form</td>
<td></td>
</tr>
<tr>
<td>date edition</td>
<td>Date</td>
<td>Date to indicate the most recent modification</td>
<td></td>
</tr>
</tbody>
</table>

*Notes:* For all variables related to municipality and administrative unit, we used 2021 as year of reference. Information on city code are derived from official classification by national statistical offices such as INSEE code for France, ISTAT code for Italy, OFS number for Switzerland.
In this section, we provide an overview of the geographical and temporal coverage for which there are relatively consistent data on historical social conflict. For each set of data, we list the type of conflicts included; the definition, if any, used by the researchers to identify the event; and the main references on the data set. We will periodically update this list as we continue working to expand the scope of the HiSCoD project.

1. France

(a) Area: France
   Year(s): 1661–1789
   Typology: All
   Definition: “[...] the threshold of collective violence is met when a group of at least four individuals, not belonging to the same family, directly perpetuate violence against one or more representatives of a political, religious, economic power, etc., or attacks property, buildings, furniture, papers, various signs symbolising these powers.” (Nicolas 2002, p. 75)\(^7\)

(b) Area: France
   Year(s): 1789–94
   Typology: All
   Definition: No definition given by the authors
   Main references: Ado (1996) and Lefebvre (1973)

(c) Area: France
   Year(s): 1800–59
   Typology: All
   Definition: “Only collective and violent confrontations against gendarmes were retrieved. The threshold was set at least three individuals involved in the assault, beyond the possible spectators. Although not very high, it nevertheless represents the number that can be ascertained by excluding turbulent arrests. It is also in line with the legislation, whether this concerns unauthorised gatherings or the levels of incrimination under the penal code. Violence is another required condition. Finally, assaults against police officers, soldiers,

---

\(^7\) Original definition: « [...] le seuil de la violence collective est franchi dès lors qu’un groupe d’au moins quatre individus n’appartenant pas à la même famille s’en prend directement (violences diverses, gestes, mots) à un ou plusieurs représentants d’un pouvoir politique, religieux, économique, etc., ou encore s’attaque aux biens, aux bâtiments, aux meubles, aux papiers, aux signes divers symbolisant ces pouvoirs.»
national guards or customs officers are excluded from the data set, unless they were accompanied by gendarmes.” (Lignereux 2008, p. 19)\(^8\)

**Main references:** Lignereux (2008)

(d) **Area:** Normandy, France  
**Year(s):** 1709–1817  
**Typology:** Food riot  
**Definition:** “[…] the threshold of collective violence is met when a group of at least four individuals, not belonging to the same family, directly perpetuate violence against one or more representatives of a political, religious, economic power, etc., or attacks property, buildings, furniture, papers, various signs symbolising these powers.” (Nicolas 2002, p. 75)  
**Main references:** Maneuvrier-Hervieu (2020)

(e) **Area:** Provence, France  
**Year(s):** 1579–1660\(^9\)  
**Typology:** All  
**Definition:** “It is precisely these insurrectionary events, defined as ‘breakaway’ conflicts, which pit parts of the population either against a legitimate authority, or against its agent, that constitute the topic of our study.” (Pillorget 1975, p. 146)\(^10\)  
**Main references:** Pillorget (1975)

(f) **Area:** South-West, France  
**Year(s):** 1600–60\(^11\)  
**Typology:** All  
**Definition:** “[…] I propose that the term popular revolt should be given to the formation of an armed group of the populace, involving participants from separate households and lasting for more than one day.” (Bercé 1974, p. 674)\(^12\)

---

\(^{8}\) Original definition: « Seuls les affrontements collectifs et violents contre des gendarmes ont été retenus. Le seuil a été fixé à trois individus au moins impliqués dans les voies de fait, au-delà de l’éventuel cercle de spectateurs. Peu élevé, il délimite néanmoins un corpus qui reste appréhensible en excluant les simples arrestations mouvementées. Il respecte la législation en vigueur, qu’il s’agisse des rassemblements non autorisés ou des paliers d’incrimination du code pénal. La violence est une seconde condition impérative. Enfin, les agressions contre des policiers, des soldats, des gardes nationaux ou des douaniers sont exclus du décompte, à moins qu’ils aient été accompagnés par des gendarmes.»

\(^{9}\) For instances of social conflict occurring 1660, see dataset 1(a).

\(^{10}\) Original definition: « Ce sont précisément tous ces mouvements insurrectionnels, définis comme les conflits ‘de rupture’ qui opposent une partie d’une population soit à une autorité légitime, soit à son représentant ou à son agent, qui constituent l’objet de notre étude.»

\(^{11}\) For instances of social conflict occurring 1660, see dataset 1(a).

\(^{12}\) Original definition: « je propose d’appeler révolte populaire la formation d’une troupe populaire armée, qui réunisse dans son sein des participants venus de plusieurs distinctes communautés d’habitats et qui se maintiennent sur pied pendant plus d’un jour.»
Main references: Bercé (1974)

(g) **Area**: Provincial cities, France  
**Year(s)**: 1590s–1660\(^{13}\)  
**Typology**: All  
**Definition**: No definition given by the author\(^{14}\)  
**Main references**: Beik (1997)

2. **United Kingdom**

(a) **Area**: England, Scotland and Wales  
**Year(s)**: 1740–1, 1756–17, 1766–7, 1791–6, 1810–9\(^{15}\)  
**Typology**: Food riot  
**Definition**: “By riot I mean an incident in which a crowd of fifty or more people acted in hostile fashion to damage or seize property, to attack persons physically, or to coerce individuals to perform or desist from some immediate action.” (Bohstedt 1983, p. 4)  
For this census, my definition of a riot is: an episode of crowd force or violence, meaning a collective assault on persons or property, illegal seizure of property, and/or coercion of a person to do something he/she would not otherwise do. By a crowd I mean at least several dozen rioters, on the assumption that such an assemblage does social politics, acting on widely-held values, interests, and consequences, that transcend the more ‘private’ interests that might animate a brawl or family feud. [...] For each riot I have recorded where possible the following parameters: date; place; characteristics and actions of the crowd; identities and actions of the forces of order; and both criminal trials and relief measures. Duration and location also define riots in my censuses. I consider a riot to be bounded by rough ‘dramatic unity’, meaning a set of actions that involved more or less the same group of actors. Hence I treat as one event all riotous violence within contiguous territory (parishes, etc.), within a week’s time, because the actors in a riot are not so much individuals as communities.” (Bohstedt 2010, p. 16).  
**Main references**: Bohstedt (2010)

(b) **Area**: England, Scotland and Wales (Swing riots)

---

\(^{13}\) For instances of social conflict occurring 1660, see dataset 1(a).

\(^{14}\) “[... a wide range of well-documented cases without following any rigorous principle of selection. It was more important to find well-documented cases which provided narratives from several perspectives than it was to achieve an ideal chronological or geographical balance. [...] a collection of carefully read instances of popular protest, ranging from the most spontaneous incident to the most complex political struggle” (Beik 1997, pp. 9, 11).

\(^{15}\) Bohstedt’s data also include events for various years between 1347 and 1739.
Year(s): 1830–2
Typology: All
Definition: No definition given by the authors
Main references: Hobsbawm and Rudé (1968) and Holland (2005)

3. Italy
   (a) Area: Italy, France
   Year(s): 1090–1435
   Typology: All
   Definition: “[... this study defines popular protest as collective action (violent or non-violent) against those of higher social status, whether they were wool bosses, landlords, or representatives of the state.” (Cohn 2006, p. 8)
   Main references: Cohn (2006).

4. Germany
   (a) Area: Germany, Poland
   Year(s): 1847
   Typology: Food riot
   Definition: No definition given by the author
   Main references: Gailus (1994)

5. Netherlands
   (a) Area: Netherlands
   Year(s): 1600–1776
   Typology: Labour conflict
   Definition: No definition given by the author
   Main references: Dekker (1987, 1990)

6. Spain
   (a) Area: Spain
   Year(s): 1766
   Typology: Food riot

---

16 It should be noted that we excluded several cases from the original data set as they do not fit our definition of a social conflict. These were events such as ‘creating an atmosphere of fear’, ‘petition’, ‘robbery’, ‘sending anonymous threat’ or ‘seditious notice’.

17 Cohn also includes events for Flanders and Liégeois but without any information on their exact locations.
Definition: No definition given by the author
Main references: Rodríguez (1973)
Example of archival evidence used to retrieve social conflict

Source: Arch. dép. Seine-Maritime, C 109, n° 27, 4 May 1775.

Figure A.1: Archival records reporting a food riot at Étrépagny on 2 May 1775
“Lettre reçue le 4 may 1775,

Monseigneur,

La halle d’Etrepagny a été pillée aujourd’hui comme celle de Gisors hier, la populace ameutée a été ensuite chés les laboureurs du lieu et s’est fait livrer le bled au prix qu’elle a taxé elle même, les uns ont païé à raison de trois livres le boësseau les autres ont emporté sans rien paier : cette populace dit que si les halles ne sont pas fournies comme à l’ordinaire, ils iront chés les laboureurs dans les campagnes.

J’ay l’honneur d’être avec respect,

Monseigneur,

Votre très humble et très obeissant serviteur,

Courtoin, le 2 may 1775.”

“Letter received on 4 May 1775,

My Lord,

The market at Etrepagny was looted today as the one in Gisors was yesterday. The crowd assembled then went to the local farmers and had the wheat delivered at the price they had set themselves. Some paid three livres per bushel, others took it away without paying anything. The populace says that if markets are not supplied as usual, they will go to ploughmen in the countryside.

I remain, my Lord, your most humble and obedient servant.

Courtoin, 2 May 1775.”
Online Appendix References


