

The Library of Babel Problem: Hypothesis Testing With Archival Sources

Alexander Lee*

February 29, 2016

Abstract

The use of archival sources is widely understood to be an important tool research tool, but attempts to use archival material for hypothesis testing, as opposed to description or theory development, are troubled by the large size and often opaque structure of archives; factors which lead to misinterpretations of evidence and a tendency to confirm the author's expectations. This paper discusses common features of archival materials and shows how they can compound traditional research design problems. It then proposes a set of best practices for avoiding these problems which differ from the types of archival analysis common within the historical profession, most notably the use of strong and explicit sampling procedures. These practices are illustrated using a brief discussion of material from the National Archives of India on the 1975 Emergency, which demonstrates how these methods produce different conclusions from a more traditional archival analysis.

Key words: Archives, Emergency, Sample Selection

Word Count: 9575

*Department of Political Science, University of Rochester

1 Introduction

Thousands of cities contain archives, which hold millions of pages of unpublished records of a wide variety of institutions and individuals. These records represent a unique resource for inquiry into the functioning of political institutions, since they contain internal material unlikely to be furnished by contemporary political actors, and because they provide insight into historical periods so remote that little other material is available.

Archival sources, like other types of qualitative and textual evidence, present substantial challenges in their interpretation and the making of causal inferences. Sources often are unreliable, disagree with one another, reflect the biases of their creators. All of these problems are familiar to historians, who place solving them at the center of their professional practice. Advice on archival methods targeted to political scientists tends to recommend that they follow historical best practices ([Trachtenberg, 2009](#)). Similarly, more general work on process tracing and case study research emphasizes finding solutions to interpretive problems and the difficulties in making reliable causal statements ([Bennett and Checkel, 2014](#); [Kreuzer, 2010](#); [George and Bennett, 2005](#); [Beach and Pedersen, 2013](#)).

However, archival material presents several additional unique and poorly understood challenges. Archives often feature a very large amount of poorly organized material, making it difficult for scholars to access the full extent of the material produced by the institution. In addition, the preserved information is often unrepresentative of that originally produced. and the patterns of missingness are not always obvious. These shortcomings present temptations for scholars to make inferences based on a biased subsample of the archive.

For many of the purposes for which archival material is conventionally used, these problems are relatively benign. Scholars who aim to explain specific outcomes, or describe events, may avoid these problems if they confine themselves to small and well-

preserved portions of the archival record. However, in many cases, political scientists seek to use archival material to develop or test causal hypotheses. Historical techniques, which emphasize an inductive and descriptive approach to archival material, can produce biases when used to test these sorts of hypotheses, most commonly bias towards confirmation.

To address these problems, it is necessary to analyze archival material in a way that integrates its unique features into standard methods of hypothesis testing. This paper highlights three methods that can be helpful in the analysis and presentation of archival evidence, and are widely used in non-archival studies. Firstly, and most importantly, scholars should develop an explicit *a priori* rule for what types of material will and will not be analyzed. This reduces the danger (encouraged by both traditional archival filing systems and modern content management software) that they only analyze material that supports their existing ideas, and enables them to be more aware of underlying patterns of missingness within the archive. Secondly, scholars, following standard historical practice, should become familiar with the internal procedures and personalities of the institution that produced a particular set of sources, so as to be familiar with both the biases inherent in the written material, and what parts of the archive are appropriate for study. Finally, scholars should acquaint their readers with the exact procedures that they used to gather material, enabling readers to critique these procedures and allowing the readers to understand how specific contentions are backed by evidence.

These methods are all designed to ensure that the scholar does not base his or her judgements on an unrepresentative subsection of the archive, and as such leave aside several important questions in research design, notably how to interpret historical documents with the analyzed subset, and how to “scale” interpretations of individual documents into larger causal arguments. These problems, and in particular the relative merits of quantitative and qualitative methods for inference making, are already the subject of voluminous and at times contentious literature (King, Keohane

and Verba, 1994; Mahoney and Goertz, 2006; Schram and Caterino, 2006). Rather than arguing that one particular interpretive approach is always preferable, this paper presents techniques useful for scholars with a wide range of approaches to source analysis and hypothesis testing, including both the gathering of quantitative data, detailed case studies, and causal process observations (Collier, Brady and Seawright, 2004).

This paper will explain both how archives might be used better in both qualitative and quantitative work. Section Two will define what archival material is, describes several problems with contemporary usage and relates them to well known problems in classic works of research design, and how these problems apply across both quantitative and qualitative scholarship. Section Three discusses a set of practices that address these problems, while Section Four uses these suggestions to lay out a detailed blueprint for how to conduct archival research. Section Five provides an illustration of these methods, analyzing a set of archival sources relating to the 1975 Indian Emergency, and showing that the methods advocated here produce sharply different theoretical conclusions than a more inductive approach. Section Six concludes with a discussion of the place of archival work within the discipline.

2 Common Problems in The Use of Archives

2.1 Definitions

Archives are accumulations of unpublished historical records, usually of an institution or individual. The term also refers to the specialized facilities in which these accumulations are held, though some agencies may continue to maintain their archives directly, and many archival facilities contain large amounts of published material. By their nature, archives are separate from the files actually in use by the institution at the present time. After a term of years or some change in the institutional environment, these working files are then transferred to the archive and (ideally) opened to scholars. Archival material

is thus usually somewhat removed from current events. It compensates for this by being publicly available in a way that contemporary government documents, which are often held to be politically sensitive, will never be.

Some social scientists casually refer to all source material on historical topics as archival, even published secondary work. This usage is misleading, since as part of the process of publication much factual material is inevitably excluded. While published primary source material is often much easier to access, and more clearly presented, than archival material, it is usually less rich, and more biased, than unpublished papers. Published secondary source material adds to this selection effect the interpretive biases of its authors, making any attempt to interpret them as historical source material more difficult (Lustick, 1996).¹

2.2 Confirmation Bias and the Library of Babel

One of the most common features of archives is their vast size relative to the interpretive capacities of scholars. The American National Archives contains over 10 billion pages of material, while even the humble Guatemalan Police archives total some 10 million pages. Even if the topic is narrowed, the capacity of institutions to produce paper runs well ahead of the capacity of scholars to understand it. A scholar interested in the foreign policy of India between 1947 and 1955 has 42,654 files to choose from, the majority handwritten. The problem of size is compounded by problems of institutional organization. In most archives, files can only be requested in finite amounts through a time-consuming procedure, browsing of material is forbidden, and indexes and other finding aids are often fragmentary and cryptic. Scholars must thus invest considerable time in obtaining relevant material before they are able to analyze it.

¹Lustick's complaint about the abuse of subsections of the historical record as evidence thus mirrors an some of the major arguments below, though in his formulation the bias comes from the interpretive decisions of historians rather than from the inherent features of archives themselves.

Some selection of what material to view is thus imperative in any type of archival work. This selection is, however, very dangerous from a research design perspective, since any selection of cases may potentially lead to biased findings. While bias resulting from case selection is a well-known danger in all forms of social scientific inquiry, it is particularly insidious in the archival context, due the opacity of archives to both scholars and those accessing their work. A scholar selecting countries for detailed case studies or congressmen for interviews is aware in a general way of the characteristics of the universe of countries and congressmen, and can thus select a representative and theoretically relevant sample, or at least be aware of the potential biases of the sample they do obtain. The selection criteria are ideally explained in the work itself, where they can be assessed by readers. An archival scholar, by contrast, may have little sense of what the “universe” of files may look like, and his reader even less. A scholar may thus fasten upon a particular subsection of material that supports her hypothesis, without having any sense of its wider validity. Given the vast size of archives, most contain material that can be used to support a wide variety of plausible hypotheses.

While scholars doing tightly focused or descriptive work, also face the problem of choosing what archival material to read, they are less troubled with bias from file selection. To the extent that scholars choose topics on which the archival material is definite and finite, they may be able to read all the available material on the topic, reducing the probability that they will base their conclusions on an unrepresentative portion of the source base.

However, many political scientists begin their archival research with a strong, clearly articulated hypotheses about the findings they will encounter. While such hypotheses are a key part of social scientific inquiry, their existence means that scholars able to process only a small amount of information may consciously or unconsciously choose information that supports their existing ideas.

To see what such unconscious selection bias might look like, consider again the case

of a hypothetical scholar who wishes to understand the effect of colonialism on religious identity. Let us imagine that our scholar has a strong belief that colonial-era policies are a key cause of religious tensions in this area. After searching an index for entries on “religion,” she encounters a wide variety of materials showing colonial involvement in religious affairs—missionary education, subsidized Muslim ulema, job quotas etc. From this she may conclude that the colonial state was heavily involved in the promotion of religious difference, and return home. While such a perspective may be correct, it may ignore archival material in which the colonial government may promote alternate types of cultural differences, or material in which religions considerations are simply not present. Such archival selection is normally difficult to detect, since readers, especially in the social sciences, have little idea of the universe of files from which scholars are selecting their evidence.

The easier availability of electronic indexes, digitized archives and content management software all promise to remove much of the drudgery from archival work, but may at the same time compound the problem of selective reading, by making it easier for scholars to isolate files, or even individuals sentences, that support their thesis, and making it less likely that they will encounter alternate perspectives. [Neitzel and Welzer \(2013\)](#), for instance, use text analysis of transcribed POW conversations to assess the involvement of German soldiers in atrocities during the Second World War. While their technique finds a shocking number of incidents in which atrocities were discussed, it is difficult to access their status as evidence, since the authors have not analyzed the much larger number of “ordinary” conversations. ²

²To the extent that scholars use archives for hypothesis development rather than testing, confirmation bias can be reduced, but never wholly eliminated. Some files must be selected for viewing, and these files will usually reflect some existing ideas about the data. And the use of archives for hypothesis development only postpones the problem of finding evidence, either within the archive or outside of it to confirm or reject the hypothesis.

2.3 Missing Data

Not only are scholars unable to read or code every file in an archive, but the files in an archive are inevitably a subset of the files produced by the institution, and represent an even smaller subset of the information available to the institution. In archival contexts, as in other types of social scientific inquiry, missing data can produce bias in both qualitative and quantitative inference. This bias takes three primary forms:

1. “Survival Bias” is a product of the long periods of time, and the occasionally uncertain environments, that intervene between the producers of material and the scholar’s desk. During this time, there are abundant opportunities for material to be lost, stolen or destroyed. Sometimes these factors are environmental, like the termite infestation at the West Bengal State Archives that has severely eroded our ability to understand early colonial administration in India. Sometimes they are human, as in the USAAF’s 1945 air raid on Potsdam, which destroyed the archives of the German Army and severely limited study of the German role in the First World War. On a lower level, archival institutions, especially in the developing world, suffer low level leakage of material: Files are checked out and not returned (and occasionally found a few decades later in a professor’s attic), files are returned but misplaced, files stored in the front of boxes become flaky and illegible, and maps are taken for sale to dealers.

While these sources of bias may seem random and ignorable, this is not always the case. At every stage of the preservation process, chances favor the survival of material considered important. The Potsdam Air raid, for instance, did not affect certain classes of papers that had been moved away for safekeeping, while the WBSA staff appears to have kept the termites away from files with obvious relevance to left-wing political activism. Similarly, institutions and regions which remain stable over time tend to produce more complete archival records, since wars, revolutions and reorganizations are obvious opportunities for archives to be destroyed or neglected.

2. “Transfer Bias” is a product of the institutional gap between the individuals or

institutions that produce material and the archival institutions that store and disseminate them. Producing institutions may be reluctant to transfer material that they deem embarrassing or sensitive, which they retain in their “active” files.” Britain’s Security Service, for instance, has released no material since the 1950s (Andrew, 2009), while many US presidential libraries embargo material that they consider sensitive. Even assuming perfect goodwill, the period of transfer often becomes a focal point for the discarding and dispersion of material.

3. “Source Bias” is the most common and problematic of the three types of bias. It stems from the fact that government agencies collect and record certain types of knowledge in preference to others. For obvious reasons, governments tend record information about their own activities rather than presenting a record of society as whole: We know far more, for instance, about taxation in the Middle Ages than we do about gender relations. Controversial matters, for instance, tend to produce multiple submissions, appeals to higher authority, and thus a much larger paper record than “routine matters.” Divorces, classically, are better recorded than happy marriages.

A more worrisome tendency is for archives to record instances in which government power was exercised successfully, rather than areas in which it was ineffective. Not only does nobody wish to be reminded of their failures, but successful power relationships require more communication and documentation than weak ones. A colonial government, for instance, may record in some detail its attempts to tax and control local chiefs, but will have little to say about chiefs who pay it no taxes, or about whom it has little knowledge. Casual readers of archives thus tend to gain a somewhat exaggerated impression of the power and efficacy of the agencies that produce them.

What makes these sources of bias particularly damaging for hypothesis testing is that the exact extent and direction of the bias is difficult to determine easily. While a scholar analyzing a survey dataset, for instance, may know exactly which respondents failed to answer a particular question, and an interviewer of elites may know which

individuals avoided his requests, archives sometimes fail to make it obvious what types of information have been lost. Faced with such a situation, social scientists will often cite whatever information is available as the “archival” truth without considering its place within the larger historical record.

2.4 Different Goals in the Use of Archives

As with other types of qualitative and quantitative sources, scholars may have many different goals in the interpretation of archival material. These may include description of particular events, the generation of broadly valid causal hypotheses, and the testing of those hypotheses (Beach and Pedersen, 2013). It is the last of these goals, however, that is most severely impacted by the two biases discussed above. If the goal of a project is to describe a discrete event, a scholar may well be able to read a significant proportion of about the event, just as a historian would. Similarly, scholars can usually avoid writing about events, or aspects of events, that are poorly covered by the archival record—another trend that is very marked in historical writing. Theory builders are also less hurt by these types of bias, since even a small sample of archival material can often be very useful in generating ideas.

Those seeking to test theories using archival material face these problems in a severe form. To the extent that they seek to make broadly relevant causal claims, they must take cognizance of the whole universe of potential data, even when it is not available. While a scholar interested in the origins of a specific war has little to fear from missingness even if only the material on his war happens to be preserved, but a scholar interested in the causes of wars in general would face problems challenges attempting to reason from a single case. More subtly, the very existence of a strong prior opinion about a topic may shape the type of data collected. Hypothesis testing efforts are thus particularly likely to be hampered by the availability of too much or too little material.

Obviously, even if a scholar had every possible document available and was granted

infinite time to read all of them, her task would not be done. She would have to decide the correct interpretation of individual documents that are often ambiguous, one sided, or cryptic. She would then have to decide how to “scale” these interpretations into a larger argument about the material and the truth of the hypotheses. Neither of these tasks is trivial, and a voluminous literature has grown up to advise scholars on how to accomplish them (King, Keohane and Verba, 1994; Mahoney and Goertz, 2006; Schram and Caterino, 2006; Bennett and Checkel, 2014; Kreuzer, 2010; George and Bennett, 2005; Beach and Pedersen, 2013). This is particularly true since neither problem is unique to archives: Many kinds of social scientific information require interpretation before they can become useful for making causal arguments.

Several approaches have emerged to solve these problems. Quantitative scholars for instance, may reduce interpretive facts to observations and analyze them statistically, while more qualitative scholars may attempt to build process tracing arguments based on the connections between particular facts. *Both* strategies, however, are vulnerable to the types of bias outlined in this section. Quantitative analyses based on a biased subsample may lead misleading results, and process tracing with incomplete information is similarly problematic. The advice presented below is thus potentially relevant for both qualitative and quantitative scholarship.

3 Best Practices for Archival Social Science

3.1 Understanding How the Data Was Created

A scholar seeking to use archival material to test a hypothesis should first emulate existing historical best practice by making herself familiar with how the material was presented and preserved. Such familiarization is essential in order to collect the most relevant information from the archive, understanding the biases inherent in the availability or non-availability of material, and the internal conflicts within the organization. An

understanding of these biases, and the general structure of the archive, is also essential in selecting a sampling frame, and in assessing the scope of missing data issues.

At its most basic, this background research involves identifying, through the finding aids in the archives and secondary sources, what institution or individual created the records and how they were organized. In some cases this is simple: The US Supreme Court, for instance, has maintained a fairly constant organizational profile for over two centuries. The responsibility for enforcing federal alcohol tax laws, by contrast, has been shuffled through a bewildering number of agencies in the same period, and these agencies have been reorganized and expanded in accordance with political convenience and administrative fashion. Generalization such as “Federal Authorities” may conceal considerable variation in institutional cultural and incentive structure.

The organization of the institution is closely related to the background of its members. These backgrounds are informative in estimating the nature of source bias and the ways in which conflicts are resolved within organizations. A reader of 19th century Indian official documents, for instance, might be tempted to lump together the documents’ European authors as “colonial officials.” This would minimize differences between numerous families of bureaucrats exercising similar jobs: Elite, college-educated Indian Civil Service Officers, high school educated Indian Police Officers, Army officers on civil duties, survivors of the East India Company’s service, and non-covenanted local recruits, to say nothing of differences in family background, province of employment, and facility in local languages. Such differences can be important in understanding the structure of archival holdings, and political conflicts. In the Indian case, for instance, frequent complaints within the police sources about the force’s poor performance, which might be interpreted as low colonial state capacity, but might also be a result of resentment within the IPS towards the privileges of other sectors of the civil service and an attempt to win a larger share of resources.

After understanding the structure of the institution, the scholar should try to un-

derstand the internal flow of paper. This means stepping back from viewing archives as masses of sources, and seeing them as their creators saw them—as individual cases flowing from desk to desk, or a pile of letters to be responded to. Some of the important questions in such an analysis are who initiates matters, who receives information, and who makes key decisions. In some organizations, peripheral agents may be focused on implementing central policy dictates, while in others the center contents itself with providing resources for field agents. Understanding this flow is important in choosing which documents to focus on. While it may be tempting to quote an impassioned memorandum on one side of a question (particularly if it agrees with the hypothesis) this memorandum may be misleading if it was produced by someone with little effect on decision-making.

Finally, scholars should consider the history of the archive itself: What areas of the archives have suffered the most from Survival and Transfer bias. The easiest way to do this is to examine the handlist or index of the archive itself, and examine the relationship between the amounts of material: Why the Education Department has 57 boxes of files from the 1900-1920 period but only 4 between 1920 and 1940? Such disproportions can potentially mask differences in preservation that may affect attempt to use archival sources for causal inference. They may also be important in determining what material to focus on in the main analysis.

3.2 Creating a Sampling Frame

An important step towards unbiased archival hypothesis testing, is to make explicit at the beginning of the main analysis what material will be reviewed and what will not. All data within this frame should be reviewed, and become part of the conclusion, either qualitatively or quantitatively. Such planning serves a number of important purposes. While it does not increase the amount of material that can be read, it guarantees that the scholar make explicit, at least to himself, the criteria by which the material to be

read is selected. This reduces the potential for the author to shape the findings by choosing what files to read. Similarly, while an explicit sampling frame cannot recreate the data destroyed by the various types of archival bias, it can make more obvious to the scholar what is missing and why.

In development of sampling frame, scholars should be guided by the structure of the archival material, with which they will be familiar after the exploratory study recommended above. A scholar might choose to read all files from a specific department for a specific period, or all of a specific class of document, or all cases under consideration at a specific time. To take some examples, `citetblaydes2013compliance` collects all instances of a specific type of document, the school report on political reliability, for a given year, [Lee \(2011\)](#) uses as observations all Indian Nationalists listed as political suspects in a given year, and [Hainmueller and Hangartner \(2013\)](#) use all immigration cases decided by a Swiss Canton in a given time period. While in all these cases some elements of the sampling frame were determined by data availability (the specific year, the specific Canton) the inclusion of individual files in the data frame was determined by the structure of the archive itself.

Such an approach should be contrasted with the more formless search-based approach common in descriptive political science and history today, in which scholars review a set of files that appear relevant to the topic or contain relevant search terms. As we have seen, such selection can be vulnerable to biases both on the part of the researcher and of the structure of the archive itself.

The sampling frame may or may not be a one-to-one mapping to the theoretical population about which the scholar is trying to draw inferences. In some cases, as in Hainmueller's Swiss immigration cases, one file is equivalent to one case. In other studies, such the the Indian Emergency examples below, a single file may contain details on many individuals, reflecting aggregation of information by the institution, in this case the grouping of the discussion of cases from within the same state. In still other

situations, many files contain information about a single significant case or actor.

Such pre-selection of the material to be examined appears alien to scholars with a descriptive focus, who have a more inductive relationship to archival material, and may often change their reading strategy during research. It is made necessary, however, by the need to test pre-set hypotheses, since the testing of specific theories in an unspecified sample magnifies the danger of the sorts of selection bias discussed in Section Two. The use of a pre-set sampling frame reduces the possibility of such sampling bias. While it cannot reduce missing data bias, it at least forces the scholar to gain a sense of which material within the frame he is unable to read.

After the main analysis, a scholar may still have many questions about the data that might not be answered within a narrow sampling frame. This might include questions about the applicability of the finding to other contexts, or the situation in areas or time periods that are missing in the main data. These questions may become the basis for the construction of additional sampling frames and further analysis. Even if the additional material is not useful for a sustained analysis, a scholar may wish to consider a brief examination of the additional material to see if it is consistent with the findings from the main analysis.

3.3 Providing Context for the Reader

The understanding of bias and the construction of sampling frames are of little use to scholars if they are not communicated to readers. While scholars are habituated to sharing with readers the details of the construction of quantitative datasets, or the organization of anthropological research, they are unwilling to share similar levels of information about their archival research, making their claims sometimes less than fully convincing. Compare two different ways of presenting the same set of hypothetical facts.

(1) To examine the motivations for this unprecedented expansion of primary education, I examined the files of the educational department. The archival evidence shows that

government officials saw the expansion primary education as a method for reinforcing national unity, rather than as a way of enhancing economic development.

(2) To examine the motivations for this unprecedented expansion of primary education, I examined all files produced by the construction subsection of the education department between 1949 and 1965, a total of 6,000 pages of material. Overwhelmingly, they mention as the motivation for construction the need to tie remote villages into a national community, rather than the economic gains from education.

While both these statements present the same facts, the second is far more convincing, since it gives the reader a basis for judging the solid evidentiary basis on which the statement rests.

The gains to openness can be made even larger when scholars are willing to acknowledge that their sources are not always unanimous and discuss the extent of the disagreement:

(3) To examine the motivations for this unprecedented expansion of primary education, I examined all files produced by the construction subsection of the education department between 1949 and 1965, a total of 6,000 pages of material. Overwhelmingly, they mention as the motivation for construction the need to tie remote villages into a national community, rather than the economic gains from education. While national identity value of school construction is mentioned in dozens of instances, on only two occasions are the economic benefits of education mentioned. Both these mentions are toward the end of the period, and concern small villages near the capital.

A necessary part of such openness is discussing the biases present in the sources: Why certain files are present and others aren't, and how the backgrounds of the creators of the archives have influenced the availability of data. Such acknowledgement should ideally be tied to a discussion of how the biases do not overly influence the main results. In certain circumstances, statistical methods like multiple imputation may be used to accomplish this, and counter the "swiss cheese" nature of many archival sources (King et al., 2001). In addition, authors may wish to present their findings from archival material outside the main analysis frame which may provide some insight on the missing material.

4 An Illustration: The Indian Emergency, 1975-1977

4.1 Historical Context

In the summer of 1975, Indian Prime Minister Indira Gandhi was facing a severe political crisis, the product of a souring economy and an energetic non-violent protest movement against her government. These crises came to a head when a district court ruled Gandhi was ineligible to serve in parliament due to technical election code violations. Her response was to temporarily suspend Indian democracy by declaring a state of emergency. The leaders and activists of rival political parties were detained, the newspapers censored, and detention without trial introduced. For the next eighteen months, Gandhi and the group of civil servants around her ruled India by decree. The Emergency was initially popular, particularly among an urban middle class who saw authoritarianism as the antidote to the instability and policy drift of the previous decade. However, the highhandedness and brutality with which the government implemented its policies soon alienated most people, especially in areas where their enforcement was more energetic. The united opposition received a golden opportunity in 1977 when Gandhi, convinced of her own popularity, ended the emergency and called a general election, which she lost decisively.

In terms of the population affected, the Emergency is the most significant democratic-to-autocratic regime transition ever to occur. It also occurred in a country widely noted in the social science literature (both before and after 1975) as having strong democratic norms and institutions relative to other developing countries. The Emergency is thus an important case in understanding how democratic states adjust to authoritarianism. Archival sources, which take us inside this process of adaptation, are crucial for answering this question, particularly given the post hoc-reticence that has led few participants to

talk about their experiences and the censorship that meant there are relatively few contemporary published sources.

In understanding how the institutions of a democracy adopted so easily to authoritarianism and abuses of human rights, two perspectives have been especially influential. The first, which I will call the *personalistic*, focuses on the role of Indira Gandhi, her son Sanjay, and the small group of courtiers around them. In this account, these individuals used the over-centralization of the administrative system to pursue their own political (and, in Sanjay's case personal and eccentric) ends (Henderson, 1977; Kalhan, 1977; Dayal and Bose, 1977) The emergency, and its "excesses" in this view, cannot be understood without reference to Mrs. Gandhi's own authoritarian personality.³

The second perspective, the *embedded* shifts the focus from the Gandhis to the bureaucracy as a whole and to state-level politicians. In this view, the Indian political class, and especially the elite Indian Administrative Service, was a product of colonial rule and the subsequent period of single party dominance, and had a strong cultural predisposition to centralized, non-democratic procedures. The Emergency, in this view, was less an imposition on the bureaucracy and on local political bosses than a liberation, allowing them to pursue long-desired priorities.⁴ Unlike the personalistic view, which emphasizes the contingent nature of authoritarian transitions, this view emphasizes that these transitions, and the policy changes which follow them have structural causes that should be common to a wide variety of post-colonial countries.

³Pro-Gandhi accounts have also tended to view the Emergency as political in nature. See Dhar (2000).

⁴This perspective dovetails with Emma Tarlo's excellent account of how rehousing policy was implemented by the bureaucracy at the municipal level (Tarlo, 2003).

4.2 The Archives

Any attempt to adjudicate these views must confront the flawed and incomplete nature of the archival record available in the National Archives of India today.⁵ The main source in terms of bulk is the records of the Shah Commission, the investigative panel set up by the Janata government to investigate emergency abuses. These can be supplemented with the small number of files that have been directly released by government departments, especially the Home Ministry (concerned with internal order and relations with state governments) and the Prime Minister's office. While this material is modest in size relative to that available on other topics, it still comprises some 30,000 pages, presenting a formidable challenge to the analyst.⁶

Overall, survival bias appears to have been negligible, in part due to the recent nature of the events, and the lack of interest in these files among historians: Of the files at the NAI, the staff are able to locate over 90%, a very high amount relative to other types of files in South Asian archives. Transfer bias, on the other hand, is extremely high. The files released by the PMO and Home Ministry appear to have been carefully chosen, and their file numbers reveal them to be a tiny subset of a much larger whole. In particular, the PMO files for these years appear to have been shorn of anything remotely controversial, and primarily concern routine economic matters such as the permitting of factories.

Similarly, source bias in these files is very marked. Under normal circumstances, bureaucrats tend to write what their superiors want to hear. These incentives were reinforced under the emergency, when the power of higher authority to damage individual careers had been enhanced—several files mention official campaigns to dismiss politically unreliable civil servants.⁷ Civil servants were thus unwilling to transmit information that

⁵While state archives would also provide much useful material on this topic, they have even less relevant information available than the national archives.

⁶The size problem is compounded by the NAI's severe limits on the number of files that can be requested at any one time.

⁷National Archives of India. Shah Commission Papers. "Communications received by the Govern-

reflected poorly on the regime—the figures on arrests for circulating illegal newspapers, for instance, are implausibly low given the size of the country and the wide variety of material recorded as seized ⁸. Civil servants also knew that requests for things like detention orders or police firings were far more likely to be believed if they articulated themselves using the regime’s language, and framed their work as advancing the regime’s priorities.

A wholly opposite set of incentives affected the IAS officers put in charge of compiling material for the Shah Commission. The commission was not intended as a way of gathering comprehensive information about the emergency, but rather of gathering information for the prosecution of its leaders. The commission’s report and its files thus display an almost maniacal focus on actions with which Indira and Sanjay Gandhi could be directly associated. Even within the broader categories of dictatorial or embarrassing behavior by the authorities, emphasis was placed on behavior that was technically illegal or “irregular” under the law of the period.

The pressure to focus on the irregular was enhanced by the structure of commission’s inquiry, in which IAS officers were dispatched to state capitals in the monsoon season of 1977 for brief visits to “scrutinize” the files of the state government and select cases for further action. On his visit to Madhya Pradesh, for instance, PS Mehta viewed the files on “about 450” cases of administrative detention out of the 5,620 that had occurred, and submitted notes on 92 irregular ones to his superiors.⁹ It is obvious that these cases do not reflect anything resembling a representative sample of the incidence of detention during the period.

An additional source of bias was the concern of the Shah commission staff to avoid embarrassment to other IAS officers, their colleagues and professional peers, to whom they were linked by a host of horizontal and vertical ties (Potter, 1996). This concern

ment of Rajasthan Regarding Detentions During the Emergency.”

⁸National Archives of India. Ministry of Home Affairs. General IV. 19949/76SS(IS)

⁹National Archives of India. Shah Commission Papers. “Use of MISA in Madhya Pradesh during emergency (Note on general pattern).”

is reflected at several points in the Shah Commission papers. Mehta, for instance, recommended that the Commission refrain from interviewing the district magistrates and police superintendents who had actually written illegal detention orders, but rather focus on the state government that had approved them.

4.3 A Personalistic Interpretation

Let us imagine that a scholar, perhaps one familiar with Henderson or Dhar's arguments, entered the NAI intending to test the hypothesis that the Delhi government's political agenda was central to the human rights violations of the Emergency period. This scholar, following standard inductive procedures, then searches for documents related to the Gandhis, the Congress, and the rights violations incidents most prominent in the existing secondary literature, and perhaps also peruses a few other files that look especially interesting or relevant.

The sources of bias within the NAI material mean that this there is considerable support for this personalistic interpretation. Notably, the Shah Commission spent nearly a thousand pages (in six parts) gathering information on the brutal suppression of local protest against one of Sanjay's prestige projects (at the Turkoman gate in Delhi) and much additional time running down leads to show that Sanjay had received special favors from administrators on a trip to Lucknow. A scholar seeking support for a thesis stressing Sanjay's importance would thus have a great deal of supporting material ready to hand.

The central role of the Gandhi, is also supported by an superficial reading of the communications between state governments and the Home Ministry. The central Home Ministry communicated frequently with the state governments fairly frequently (thirteen times in the first 35 days of the emergency), directing details of detention and censorship. While only one of these documents is signed by Mrs. Gandhi herself, the others are signed by S.L. Khurana, a civil servant closely associated with Mrs. Gandhi, who would

subsequently receive a series of high promotions under the Congress governments of the 1980s, culminating in the governorship of Tamil Nadu. These documents show a central government that, while it liked to hide behind convoluted conditional phrasing, expressly ordered the detention of specific classes of individuals, notably Members of Parliament (6/26/75), members of the Hindu Nationalist RSS (6/26/75), journalists (6/28/75) and students (7/5/75).¹⁰

Finally, many of the available documents heavily stress the use of detention and repression against political opponents of Mrs. Gandhi, notably the RSS and the socialist and communist opposition. The Shah Commission scrutiny of detentions in Uttar Pradesh, for instance, began its discussion with a detailed listing of 72 cases of abusive detention for political beliefs. Such unfortunates as Subedar Singh, a socialist party worker were detained for shouting “Burn the Delhi government” and “Stop buses,” and would spend months in prison without access to the court system.¹¹ The political bias inherent in this process is underlined by several cases in which individuals had their detention revoked after they were able to produce “certificates from Congressmen and responsible persons about their good conduct and support for Congress policies.”¹² Detention was thus clearly a stick used to repress opposition to the Congress. The perception that the emergency had a strong partisan political orientation is supported in the files of the Home Ministry, which track in exhaustive detail the seizures of specific types of banned literature and their contents.

¹⁰National Archives of India. Shah Commission Papers. “Communications received by the Government of Madhya Pradesh Regarding Detentions During the Emergency.”

¹¹National Archives of India. Shah Commission Papers. “A Report on MISA detentions in Uttar Pradesh.” P.16.

¹²National Archives of India. Shah Commission Papers. “A Report on MISA detentions in Uttar Pradesh.” P.25.

4.4 An Alternative Interpretation

Choosing a Sampling Frame

While a political or partisan interpretation of the emergency has strong textual support, it obviously accords with the agendas of those who produced and censored the available material—with civil servants during the emergency wishing to stress their cooperation with central policy, and civil servants afterwards seeking to blame the Congress and the Gandhis. To reduce this bias, an analysis could focus on a source which appears to contain few omissions. The most interesting of these are the records of detentions under the Maintenance of Internal Security Act (MISA). Detention under the act was indefinite, subject to confirmation by the state government. Detention under MISA was the centerpiece policy of the emergency, and was used for a wide variety of purposes other than the detention of political opponents, notably the enforcement of various types of criminal laws for which the official legal process was considered too slow and procedurally bound to be functional. Detention was also of central importance to the Shah Commission investigators, who preserved a great deal of information on its incidence.

For each state, the commission collected a “scrutiny” file of the MISA detentions, which give basic statistics on MISA, a synopsis of major trends, and descriptions of cases the investigators found troubling. As we have seen, the Shah Commission tended to focus on cases which were politically controversial or legally irregular. Fortunately however, they preserved figures on the total number and breakdown of detentions. Their detailed cases studies also include many cases which violated the technical terms of the MISA statute, meaning that cases involving a wide variety of motivations of detention are described in some detail. Focusing the analysis on the detention files of the Shah commission thus allows us to concentrate on a type of authoritarian behavior where documentation is voluminous, the terms of reference are constant across India, and where transfer and source bias, while not absent, are at least known and quantifiable.

Analysis

The main analysis examined the Shah commission's "scrutiny" files for every state. The raw figures show that political detention tended to vary with the population of the state and the strength of the opposition to the Congress. However, not all detentions were political. In the states where detention was most common, political offenders were only a small minority of those detained under MISA. In Uttar Pradesh, only 19.5% of detentions were for political reasons (1405/7185), and 25.4% in Bihar (593/2333).¹³ In other states, especially in the south, political detentions tended to dominate—321 of 477 in Karnataka, 709 of 1017 in Tamil Nadu ([Government of India, 1978](#)).

The non-political detentions cited a wide variety of crimes, most of which enjoyed popular opprobrium well before 1975. A few were detained for agitating against family planning, a "non-political" act with obvious political implications in the circumstances of the time. Given the scarcity of examples of this, however, this seems to have been rare relative to more general types of criminal and anti-social behavior, or any behavior that would bring individuals into conflict with the district magistrate. In many cases, these were ordinary crimes for which prosecution was difficult or impossible within the dysfunctional Indian judicial system, including armed robbery, burglary, murder and corruption. Typically, in these cases the District Police cited old criminal complaints as evidence, even if the accused had not been convicted.¹⁴ In other cases, the offenses were moral in nature, as in the substantial number of detentions of bootleggers and keepers of gambling dens. Still others had conducted actions that had a negative effect on the government's economic plans, such as workplace indiscipline, the misspending of official loans, or the obstruction of the government wheat purchasing program.¹⁵

¹³National Archives of India. Shah Commission Papers. "A Report on MISA detentions in Uttar Pradesh." and "Report of the Scrutiny of the MISA files of Biha."

¹⁴National Archives of India. Shah Commission Papers. "Report of the Scrutiny of the MISA files of Bihar."

¹⁵National Archives of India. Shah Commission Papers. "A Report on MISA detentions in Uttar Pradesh."

In other cases, however, it is impossible to escape the conclusion that the district officials were detaining people simply for making their job more difficult. Shri Mohan Lal, of Pratapgarh District, for instance, was a government contractor detained for providing low quality bricks rather than the high quality ones specified in his contract.¹⁶ Sajjad Khan of Farrukhabad was detained for occupation of land intended the construction of a government office building and contesting their rights in the courts. Johri Lal Are of Farrukhabad was the editor of a local paper who had criticized local officials. In other cases, detention was used to encourage adherence to regulations, as with Narendra Kumar Goel, whose bus had crashed into a river and was detained “that bus owners may keep their vehicles in order in the future.”¹⁷

In a few cases, there is clear evidence that the district government detained individuals for purely self-interested reasons. Ashok Kumar “at the instance of some persons” was falsely accused by the police of associating with criminals. In one extraordinary case in Bihar, Radhe Shyam Pandey was detained by his own corrupt colleagues in the Tubewell division, who resented his attempts to expose their corruption and persuaded the DM to sign the order.¹⁸

The casual grounds for many detentions are not surprising when we consider the process by which detentions were issued. The District Superintendent of Police, in cooperation with the District Magistrate, issued a list of proposed detainees, though there are indications that some DMs, like that at Lucknow, were “very casual in perusing the material placed before him by the police authority.”¹⁹ These grounds were in turn reviewed by officials the State Government, usually the Chief Minister, with advice from the Home Secretary and Inspector General of Police. District officials thus had almost

¹⁶National Archives of India. Shah Commission Papers. “A Report on MISA detentions in Uttar Pradesh.” P.44.

¹⁷Ibid.

¹⁸National Archives of India. Shah Commission Papers. ‘Report of the Scrutiny of the MISA files of Bihar.’ P.33

¹⁹National Archives of India. Shah Commission Papers. “A Report on MISA detentions in Uttar Pradesh.” P.53.

total discretion over detentions, subject to state review. This echoed procedures during the colonial period, especially those used against nationalists during the Second World War.

Further Context

Other files in the NAI provide context on the high level of non-political or idiosyncratic detentions found in the main analysis. The central government was well aware that many officials were using MISA for reasons that had nothing to do with the goals of the Emergency. Khurana sternly lectured the Chief Secretary of Rajasthan:

There are other states where detentions are still being made at the behest of subordinate police officers...MISA is being used to remedy all kinds of situations although they may be totally unrelated to the emergency. One could understand the detention of a few top anti-socials elements to create an impact but the detention of bootleggers, gamblers and goondas [thugs] in hundreds under MISA is certainly not in consonance with the objections of the act.²⁰

The available circulars and memoranda issued by the state and central governments also show that the focus on non-political offenses came from the initiative of local officials, rather than some central policy. While the Shah Commission emphasized a state circular in Uttar Pradesh that called for the detention of criminals, this appears both exceptional in terms of official communication and to have had a limited impact, since many “criminals” were detained before the circular’s appearance.²¹ The record of circulars to from the Inspector general to DSPs in Madhya Pradesh, which appears fairly complete, shows that 87 directives were sent out during the emergency, the vast majority concerning detention. Of these 87, only 5 concern non-political activity, while 50 concern the campaign against the RSS, the most powerful opposition organization

²⁰National Archives of India. Shah Commission Papers. “Communications received by the Government of Rajasthan Regarding Detentions During the Emergency.” P.30.

²¹National Archives of India. Shah Commission Papers. “A Report on MISA detentions in Uttar Pradesh.”

in MP. ²² Similarly, the available central directives to the government of MP strongly emphasize political repression, though they are less valuable for having been edited by the Shah commission staff.

Discussion

An analysis of the archival material supports the supposition that during the Emergency, the political agendas of the central government were often secondary to the personal and institutional agendas of state politicians local policemen, who used the opportunities provided by the removal of democratic checks and balances enforce their own vision of an ordered society. Given that both the vision and the legal tools they employed have deep roots within Indian government, it is not unfair to say that the emergency allowed preexisting authoritarian tendencies within the Indian bureaucracy to express themselves.

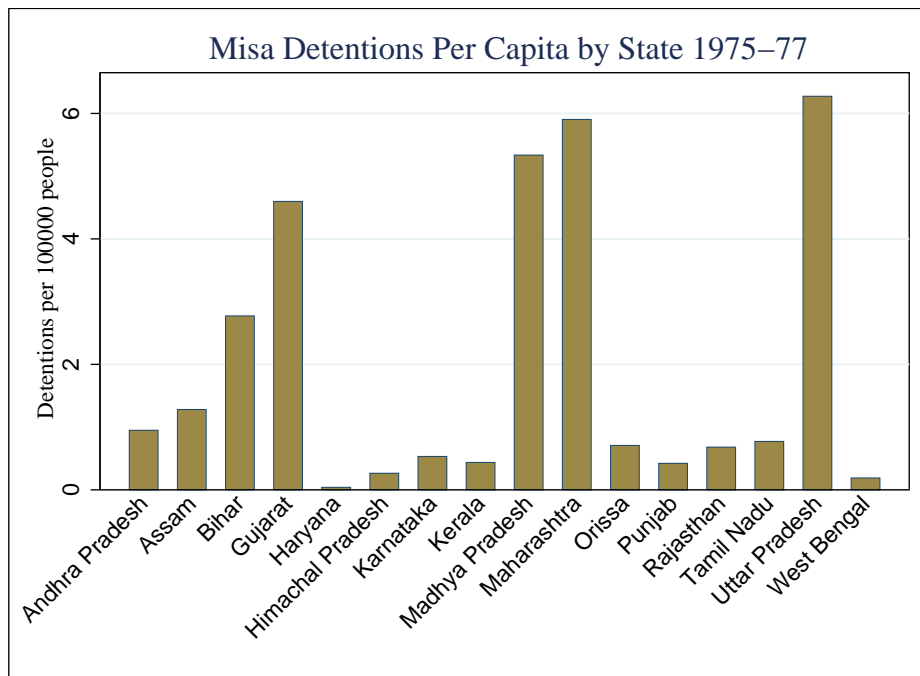
One of the interesting features of the findings is that the extent of this lower-bureaucratic autonomy varied considerably from state to state. However, this question is difficult to answer with the archival material at our disposal. The differences do not appear to come from the bureaucratic personnel, since the Indian Police and Administrative Services are recruited nationally. Similarly, they do not seem to stem from differences in the bureaucratic procedure for reviewing detentions, which seem to have been uniformly cursory, with few or no cases being overturned by the state government. Even in Karnataka, where state government review was the most aggressive in India, orders were usually confirmed “to uphold the authority of the District Magistrate.” (Government of India, 1978, p. 75)

This raises the obvious question of why the bureaucracy was given wide autonomy to use detention in some states, while in some states they confined themselves to interning a small number of political prisoners. While a full discussion of this question is outside the

²²National Archives of India. Shah Commission Papers. “List of Instructions Issued by I.G. police, M.P. during Emergency.”

scope of a methodological paper, relating archival and extra-archival data can show that differences in the ability of lower officials to exploit the emergency are associated with the origins and incentives of state Chief Ministers, the only real political check in the MISA system. Within the Congress system, certain chief ministers possessed strong electoral base independent of the national party, and had more autonomy from the national party, while others were less secure. (Manor, 1978). Jaganath Mishra, the Chief Minister of Bihar, was widely reputed to have been chosen for his loyalty to the Gandhis, while his counterpart in Karnataka, Devraj Urs, was a scion of the local royal family with a somewhat antagonistic relationship to Delhi government (which would later lead to his leaving the party) (Kohli, 1990). The more dependent CMs had both less incentive to avoid local unpopularity, less experience in dealing with the bureaucracy, a weaker set of connections to the local political elite, and more incentive to demonstrate loyalty to the center.

Figure 1: Non-Political MISA Detentions by State, 1975-77



Each bar shows the average number of non-political MISA detentions in each state during the Emergency. Taken from [Government of India \(1978\)](#)

Figure 1 shows the levels of non-political MISA detention during the emergency, five states stand out as having very high levels of detention: Bihar, Gujarat, Madhya Pradesh, Maharashtra, and Uttar Pradesh. These are also the five states in which the chief minister during the bulk of the emergency had only recently been appointed to office in 1975, either immediately before the Emergency or in its first few months.²³ In the other states, by contrast, the chief minister had been in office for a substantial period of time, usually since the 1971 elections. While these results can only be suggestive, they indicate a counterintuitive finding. The appointment of non-autonomous regional governments thus appears to be associated with a reduction in the degree to which detention was used to further the regime's political objectives. Even under the Emergency, Democracy, and the information flow and legitimacy that it brings, appears to be an important tool in controlling the bureaucracy.

5 Conclusion

Archival material can potentially be a valuable source for political scientists, allowing them to test hypotheses about institutional behavior and information gathering that are inscrutable in a contemporary setting. However, the size, disorganization, and non-comprehensiveness of most archives create problems for causal inference, especially selection and missing data. This paper proposes a set of simple suggestions for dealing with these problems: An understanding of the biases inherent in the sources, the development of an explicit frame for reading and analyzing the material, and sharing these methods and findings with readers. As the analysis of the NAI Emergency files shows, such principled methods can produce dramatically different results than an undirected dig through the material.

²³The chief minister of Uttar Pradesh, Hemvati Nandan Bahuguna, had been in office since 1973, but was suspected of disloyalty and was deposed in November 1975 in favor of the more pliable N.D. Tiwari.

While certain of the approaches discussed here, such as an understanding of the producing institution, are useful in a wide variety of contexts, these techniques are not intended to be a universal prescription for archival research. Scholars interested in process tracing and description may find the construction of a hard sampling frame to be unduly burdensome and narrowing. However, these techniques are very appropriate for those who seek to use archives to test existing theories. By enabling causal claims to be made more credibly, broadening the influence of archival evidence within the discipline, and improving our understanding of the inner workings of political institutions.

Bibliography

Andrew, Christopher. 2009. *Defend the realm: The authorized history of MI5*. Random House LLC.

Beach, Derek and Rasmus Brun Pedersen. 2013. *Process-tracing methods: Foundations and guidelines*. University of Michigan Press.

Bennett, Andrew and Jeffrey T Checkel. 2014. *Process tracing*. Cambridge University Press.

Blaydes, Lisa. 2013. “Compliance and Resistance in Iraq under Saddam Hussein: Evidence from the Files of the Ba ‘th Party.”

Collier, David, Henry E Brady and Jason Seawright. 2004. “Sources of leverage in causal inference: Toward an alternative view of methodology.” *Rethinking social inquiry: Diverse tools, shared standards* pp. 229–66.

Dayal, John and Ajoy Bose. 1977. *For Reasons of State: Delhi Under Emergency*. Orient Paperbacks.

- Dhar, Prithvi Nath. 2000. *Indira Gandhi, the "emergency", and Indian democracy*. Oxford University Press New Delhi.
- Firmin-Sellers, Kathryn. 1995. "The politics of property rights." *American Political Science Review* 89(04):867–881.
- George, Alexander L and Andrew Bennett. 2005. *Case studies and theory development in the social sciences*. Mit Press.
- Government of India. 1978. "Shah Commission of Inquiry: Third and Final Report."
- Hainmueller, Jens and Dominik Hangartner. 2013. "Who gets a swiss passport? a natural experiment in immigrant discrimination." *American political science review* 107(01):159–187.
- Henderson, Michael. 1977. *Experiment with untruth: India under emergency*. South Asia Books.
- Jha, Saumitra and Steven Wilkinson. 2012. "Veterans, Organizational Skill and Ethnic Cleansing: Evidence from the Partition of South Asia."
- Johnson, Timothy R, Paul J Wahlbeck and James F Spriggs. 2006. "The influence of oral arguments on the US Supreme Court." *American Political Science Review* 100(1):99–113.
- Kalhan, Promilla. 1977. *Black Wednesday: Power Politics, Emergency, and Elections*. Sterling Publishers.
- King, Gary, Anne Joseph O'Connell, James Honaker and Kenneth Scheve. 2001. "Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation." *American Political Science Review* 95.
- King, Gary, Robert O Keohane and Sidney Verba. 1994. *Designing social inquiry: Scientific inference in qualitative research*. Princeton University Press.

- Kohli, Atul. 1990. *Democracy and discontent: India's growing crisis of governability*. Cambridge University Press.
- Kreuzer, Marcus. 2010. "Historical knowledge and quantitative analysis: The case of the origins of proportional representation." *American Political Science Review* 104(02):369–392.
- Lee, Alexander. 2011. "Who Becomes a Terrorist?: Poverty, Education, and the Origins of Political Violence." *World Politics* 63(02):203–245.
- Lustick, Ian S. 1996. "History, historiography, and political science: Multiple historical records and the problem of selection bias." *American Political Science Review* pp. 605–618.
- Mahoney, James and Gary Goertz. 2006. "A tale of two cultures: Contrasting quantitative and qualitative research." *Political Analysis* 14(3):227–249.
- Manor, James. 1978. "Where Congress Survived: Five states in the Indian general election of 1977." *Asian Survey* pp. 785–803.
- Neitzel, Sönke and Harald Welzer. 2013. *Soldaten: On Fighting, Killing, and Dying: The Secret World War II Transcripts of German POWs*. Knopf.
- Potter, David C. 1996. *India's Political Administrators, from ICS to IAS*. Oxford University Press Delhi.
- Schram, Sanford and Brian Caterino. 2006. *Making political science matter: Debating knowledge, research, and method*. NYU Press.
- Tarlo, Emma. 2003. *Unsettling memories: Narratives of the emergency in Delhi*. Univ of California Press.
- Trachtenberg, Marc. 2009. *The craft of international history: A guide to method*. Princeton University Press.