

The comparative approach to the study of lynching uses inventories with limited information about a large number of victims to examine the distribution of lynchings over time or across space, attempting to link temporal and spatial variation in lynching to corresponding variation in social, economic, political, or demographic factors. Most comparative analyses of lynching have tested competing theories about the antecedents of lynching (see e.g., Beck and Tolnay 1990; Beck, Massey, and Tolnay 1989; Brundage 1993; Corzine, Creech, and Corzine 1983; Corzine, Huff-Corzine, and Creech 1988; Olzak 1990; Soule 1992; Tolnay and Beck 1992a, 1995; Tolnay, Beck, and Massey 1989; Tolnay, Deane, and Beck 1996). We have learned a great deal about southern lynchings from these different methodological approaches. While there is not full consensus, prior work indicates that southern lynchings (primarily of African Americans) were more likely with (1) a proportionately larger black population, (2) an economically disadvantaged white population, (3) a strong Democratic party, (4) the perception among whites that blacks threatened their economic or social position, and (5) a low level of black out-migration.

One distinct limitation of previous comparative analyses of lynching is their heavy reliance on cross-sectional spatial variation (usually over counties) or temporal fluctuations (usually across years) in order to draw conclusions about the conditions that created a climate favorable to mob violence. The dependent variables in these studies have been either the number of lynchings (or the rate of lynching) that occurred within a given county over a specified period of time, or the number of lynchings that occurred within a given period of time for a relatively large geographic area – for example, an entire state. The nature of existing lynch victim inventories has precluded work in which individuals are the units of analysis. A second limitation to previous analyses is the minimal amount of information that is available about the victims, generally restricted to the victim’s race and gender, the date and location of the incident, and the purported reason for the lynching. For example, Table 1 presents the demographic information currently available for victims listed in the Beck-Tolnay inventory. We are working to link existing inventories of lynch victims with Census data on individuals and the households in which they lived, creating the potential for quantitative analyses that use lynch victims as the primary unit of analysis, and allowing for comparisons between those known to have been lynched and other members of their surrounding communities. Our goal is to bring the individual victim into research on lynchings in a more prominent way.

Table 1. Demographic Characteristics of Victims in Lynching Inventory

	Male	Female	Unknown	Total
Black	2092	67	22	2181 (88%)
White	249	5	0	254 (10%)
Other	5	0	0	5 (<1%)
Unknown	42	0	1	43 (2%)
Total	2388	72	23	2483

In this project, we link information about lynch victims from an existing inventory (Beck and Tolnay 2004) with information about those same individuals from the manuscript records of the United States Census. Beginning with the Beck-Tolnay inventory of confirmed lynching victims from ten southern states between 1882 and 1930 (N=2,483),¹ we are attempting to find

¹ Because the original 1890 census manuscripts were destroyed by fire, we are somewhat limited in our ability to link individuals who were lynched during the 1890s with their census records. We have opted, therefore, to search

Census records for those individuals using Ancestry.com's online, scanned and searchable digital images of the original census enumerators' manuscripts for the Census immediately preceding the lynching incident (e.g., the 1880 Census for an 1885 lynching). For successfully linked cases, the Census information will create a new, richer database of victims and their individual and household characteristics, as well as geographic identifiers. Based on an estimated 40% success rate from our pilot testing, we expect to eventually yield detailed individual and household information on 990 lynch victims. This database will make a new approach to lynching research possible: quantitative analyses of lynching using individuals as the primary unit of analysis, linking information on these individuals with the characteristics of the local context in which the victim resided or was lynched. The number of victims represented in our new database will be sufficiently large to support more qualitative studies of southern lynching.

Our process works as follows: Using information about each lynching victim that is contained in the Beck-Tolnay inventory, we attempt to locate the same individual in the appropriate decennial census. This search begins by using the victim's first and last name, searching in the state and county in which s/he was lynched. Our initial criteria for a successful linkage also include race, sex, and age, when it is available. Of course, it is possible that the victim was not lynched in the same county in which s/he was enumerated in the previous census. For that reason, we extend the search beyond the county in which the lynching occurred. To restrict the geographic focus for our search and limit the potential for "false positives," we employ an "adjacency" criterion to expand the number of counties to be searched. For example, if a victim was lynched in Jones County, Georgia, we would search the census records for Jones County and all adjacent counties (Monroe, Jasper, Putnam, Baldwin, Wilkinson, Twiggs, and Bibb). For counties on a state line, the adjacency criterion would include counties in both states. We also utilize additional information for more difficult cases, including death certificates, newspaper accounts, and draft registration information. We conduct a "forward search" into the census immediately following the lynching for all potential matches we locate, in order to eliminate false positives. Clearly, if an individual was lynched in 1903, we should not be able to locate him/her in the 1910 census manuscripts.

Based on such factors as age, location of residence, and spelling of first and last names, for matches in which only one individual has been identified, we assign a rating of "high" or "medium" to describe our confidence in the linkage. In cases where there is more than one potential match in the census manuscripts, and it is impossible to determine with acceptable certainty which enumerated individual is the same as the lynch victim, using information such as race, age, location, and name spelling, we assign a score that represents our best, but ultimately subjective, estimate that each potential match is the same individual as the lynch victim. These subjective probabilities will sum to 1.0 for any given lynch victim. For cases with multiple matches, we record information for all possible matches, up to a maximum of five. If there is no reason to believe that one individual is more likely than another to be the "true match" then all potential matches will receive the same subjective, estimated probability. This will allow researchers to select only matches with a certain level of confidence for their analyses. It will also permit sensitivity tests in which analyses are conducted with one threshold of confidence, and then repeated with matches of lower certainty to observe any changes in the results.

When a successful linkage is made between a victim included in the lynching inventory and his or her census records, all of the available information in both sources is entered into the

in the 1880 census manuscripts lynch victims from the 1890s and to limit our searching to those individuals lynched between 1890 and 1895. This N reflects that methodological limitation.

database. From the lynching inventory this includes the sex, date, state, county, and purported reason for the lynching. From the census manuscripts it includes all of the information contained on: (1) the household record, and (2) the person records, for all individuals residing in the household. Examples of key household information include the following: county of residence, group quarters type, and for the 1900 census, ownership of dwelling and farm status. Important information for individuals residing in the household includes: name, relationship to head of household, age, race, sex, marital status, birthplace, mother's birthplace, father's birthplace, literacy, and occupation. Appropriate geocoding (i.e., FIPS codes and ICPSR-compatible codes) is included so that the records in the database can be linked to existing state-level or county-level data that are already available in machine readable form, or to such aggregate data that might be prepared subsequently.

As with any attempt to link records across two unique sources, our effort to link southern lynch victims with their census records face a variety of challenges. These include systematic under enumeration of the black male population (Rosenwaike et al. 1998; Coale and Rives 1973), the possibility that lynch victims had migrated across county lines between the year of census enumeration and the year in which they were lynched (Tolnay 1999) or that they were lynched outside their county of residence, and irregularities in the recorded spelling of victims' names and/or multiple matches for common names.

At present, we have searched for 660 of the 2,483 individuals in our lynch inventory. Of these, we have completed the case selection and match identification processes for 405. Data entry has just begun, but we anticipate having searched for, identified, and entered data for roughly 1,000 of the victims in our inventory by the time of the PAA meetings. Again, given our expected yield rate of 40%, this should mean confirmed data on roughly 400 individual victims.² The paper we plan to prepare for PAA will present a preliminary assessment of the demographic characteristics of lynch victims. This primarily descriptive data will be compared to the characteristics of a randomly-selected control group of individuals – stratified by race and sex – from the communities in which the lynch victims were located. This analysis will allow us to determine whether individuals who are “known” to be lynch victims vary systematically along important demographic, socioeconomic, or other dimensions from “typical” members of their communities. Additionally, we will present a preliminary assessment of the risk of being lynched, based on a logistic regression analysis in which the dichotomous outcome variable indicates whether the individual was lynched or not, and the predictor variables include age, marital status, literacy, occupation, distance between location of lynching and location of enumeration, birthplace, and for some decades, home ownership.

Our first batch of completed data is based on efforts to search for 52 victims from a nine-year period in Georgia. We successfully identified matches for 20 of these individuals. Of these, 17 are “single high” matches, for a yield rate of 33%. This means that for one in three cases, we have narrowed the field of potential matches down to a single individual identified in the census records, and that we have a high degree of confidence that this individual is, indeed, the lynch victim. For an additional three cases (or 6% of the total we attempted to link), we identified multiple matches and were unable to restrict the number of potential matches to one. Eighteen of the 20 victims we identified were black, and two were white. Two of the black victims were lynched by black mobs, but all other victims were lynched by mobs composed primarily of whites. Descriptive statistics for these victims appears in Table 2.

² In addition to the authors listed here, we have a half-dozen undergraduate research assistants working on this project, and will add two more in January.

Table 2. Selected Characteristics of Georgia Lynch Victims³ Identified in Census Records

Head of Household	13	Enrolled in School	1
Mean Age	30	Able to read and write	5
Married	12	Born out of state	2
Non-farming occupation	1	Father born out of state	5
Disability	0	Mother born out of state	4

Clearly, having this level of information on lynch victims provides a much more detailed view of their lives and circumstances. Embedding these individuals within their households also allows us to identify unusual – or mundane – characteristics of their wives, children, and coresident parents. This project has the capacity to shed light on a previously unknown aspect of lynching culture in the United States. Namely, we are seeking to uncover why, given that structural conditions were ripe for a hate crime of this type to occur, a particular individual was selected as the lynch victim. If particular demographic characteristics appear to distinguish victims from other members of their communities, we may gain greater insight into the phenomenon of lynching and its role in maintaining the south’s racial caste system.

³ Calculated for single-high matches only.