The Settlement Abandonment Database

Authors: Mathew E. Hauer[1,2]*, Valerie Mueller[3,4], Casey Breen[5], Yang Zhong[1,2]
Affiliations:
[1] Department of Sociology, Florida State University, Tallahassee, FL, USA 32306.
[2] Center for Demography and Population Health, Florida State University, Tallahassee, FL, USA 32306.
[3] School of Politics and Global Studies, Arizona State University.
[4] International Food Policy Research Institute.
[5] Leverhulme Centre for Demographic Science, University of Oxford

*Correspondence to: Mathew E. Hauer, mehauer@fsu.edu

Abstract:

Up to 3 billion people will reside outside of the human climate niche due to climate change by the end of the century and the wholesale abandonment of communities is now a reality. However, our understanding of settlement abandonment relies on small-scale case studies because no database detailing abandoned settlements presently exists. In this paper, we digitize historical US Censuses to to (1) develop a comprehensive database of the population counts for all 34,000 places enumerated in at least one decennial Census between 1890-2020, (2) identify the places missing in at least one Census (a "candidate abandoned settlement"), and (3) crosscheck each candidate to ensure it was unaffected by other processes (e.g., city annexation). Building such a comprehensive database will enable researchers to better study the drivers of settlement abandonment and deploy new methodological approaches to identify the *causal effects* of abandonment.

Keywords: climate change, carbon emissions, demography, IPAT, Kaya identity, mitigation, life course

Acknowledgements: This research was supported by the National Institutes of Health's National Institute on Aging, grant R21AG088598.

Main Text

Up to 3 billion people will reside outside of the human climate niche due to climate change by the end of the century (McLeman 2011; Xu et al. 2020). The wholesale abandonment of communities and emphatic declarations of 'climate refugees' are now a reality (Farbotko and Lazrus 2012; Hino, Field, and Mach 2017) . As the Fourth National Climate Assessment Report emphatically stated, retreat will be "unavoidable" in parts of the United States (Reidmiller et al. 2017). These abandonments are already underway in the United States. Since 1989, the Federal Emergency Management Agency (FEMA) has undertaken managed retreat in over 1,100 US counties (Siders 2019). Despite abandonment's growing importance, we know very little about the processes or consequences of settlement abandonment.

We are embarking on purposeful settlement abandonment without knowing long-term implications. The US federal government is presently facilitating the wholesale abandonment of Isle de Jean Charles in Louisiana, amongst other well-known managed retreats (Siders 2019; Simms et al. 2021). Identification of long-term impacts are identified by either well-known, large-scale environmental events such as Hurricane Katrina (Hori and Schafer 2010) or through historical case studies (Addyman 1989; Atherwood 2022). Furthermore, the drivers of abandonment are multifaceted (Black et al. 2011; McLeman 2011), driving outcomes across multiple dimensions, including economic (Xie, Wang, and Yao 2014) , stress-related (González-Tennant 2019), environmental (Gill et al. 2007; Romm 2011), and conflict-related issues(Arkush 2008), amongst others. A single study on long-term health consequences found that those who stayed in agricultural regions during the dustbowl of the 1930s had longer longevity than those who migrated (Atherwood 2022). This could suggest that settlement abandonment, which is by definition migration, could yield worse health outcomes. Relying on a singular case study understates the accumulation of disadvantages for persons who are forced to abandon a settlement. Current state, local, and federal policies facilitating abandonment are presently underway without crucial knowledge to inform them.

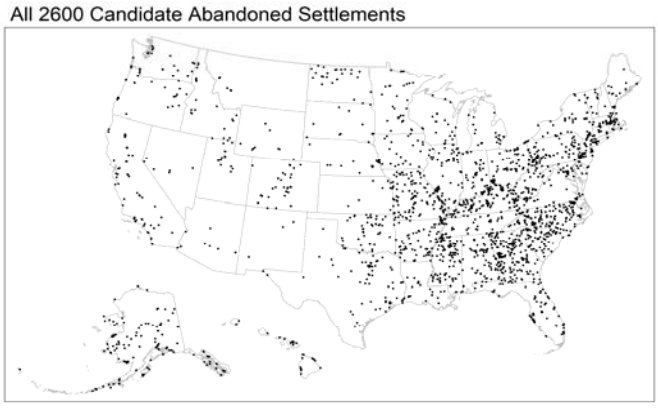

Figure 1. All 2600 Candidate Abandoned Settlements. Here are all identified candidate abandoned settlements missing in at least one decennial census since 1890. Note that many of these candidate abandoned settlements are located in presently occupied areas. These 2600 candidate abandoned settlements will be whittled down to only those that are verified as abandoned.

No comprehensive database of US places exists that can identify abandoned settlements. Building such a comprehensive database will enable researchers to better study the drivers of settlement abandonment and deploy new methodological approaches to identify the *causal effects* of abandonment on socioeconomic outcomes. We propose to build a comprehensive database of the population counts for all 34,000 places enumerated in at least one decennial Census between 1890 and 2020. Crucially, such a comprehensive database will allow for the creation of appropriate counterfactuals for causal identification. Preliminary analysis identified 2600 candidate abandoned settlements between 1890 and 2020, underscoring the feasibility and potential significance of systematically documenting settlement abandonment and the life course of places. Thus, the creation and dissemination of such a comprehensive database will serve as

significant ‘public good’ for other research on forced migration, managed retreat, and other forms of settlement abandonment.

Data: This research here draws on two primary data sources: historical census data of population counts and historical record of places in the decennial Census from NHGIS.

Historical record of Places. The National Historic Geographic Information System (NHGIS) maintained and published by International Public Use Microdata System (IPUMS) contains GIS point files for each decennial Census from 1900 to 2020 . The Place Points File contains all places, incorporated, unincorporated, and census-designed places identified in the Census Bureau’s published datasets and volumes. While this file contains the point location of each place at the time of the decennial census, it does not contain the population enumerated in places in decennial censuses. Each place is given a unique NHGIS place identifier which will form the basis of our universe of places.

CITY, TOWN, OR VILLAGE.	County.	1910	1900	1890
Abbeville town	Henry	1,141	889	405
Adamsville town	Jefferson	649	.....	.....
Alabama City town	Etowah	4,313	2,276	.....
Albertville town	Marshall	1,544	.....	.....
Alexander city	Tallapoosa	1,710	1,061	679

CITY OR TOWN.	County.	1920	1910	1900
Abbeville town	Henry	1,207	1,141	889
Akron town	Hale	675	.....	.....
Alabama City	Etowah	5,432	4,313	2,276
Albany city	Morgan	7,652	6,118	4,437
Albertville town	Marshall	1,666	1,544	.....

Figure 2. Example of an "Abandoned Settlement." Adamsville Town, AL has a population recorded in the 1910 Census of 649 persons but is missing from the 1920 Census.

Historical Census Data. The US Census Bureau maintains all historic publications related to census data collected from 1790 to 2020 freely available online in pdf format. The historic counts of places are yet to be digitized in machine readable format. We will convert these online pdfs into machine readable, tabular format to produce a complete time series of place populations – both abandoned and continuously occupied -- for the period 1890-2020. **Figure** shows an example of a candidate settlement “abandoned” between Census 1910 and 1920. Our preliminary analysis identified more than 34,000 unique places, appearing in at least one Decennial Census between 1890 and 2010, and more than 30,000 places that appear in all Decennial Censuses between 1890 and 2010, suggesting approximately 4,000 candidate abandoned settlements.

Research Plan

We will develop a comprehensive dataset of abandoned settlements in the United States between 1890-2020. We will use historical US Censuses to (1) develop a comprehensive database of the population counts for all 34,000 places enumerated in at least one decennial Census between 1890-2020. This dataset will be a time series of population counts for each place totaling a

dataset with approximately 34,000 rows and 14 columns – one column for each decennial census – digitized from online, historical census data. Once this database is constructed, we will (2) identify the places missing in at least one Census (a "candidate abandoned settlement"). We refer to these places as "candidates" because the only evidence of abandonment is the absence in at least one decennial census. In our preliminary analysis, we identified approximately 2,600 places missing from at least one Decennial Census (**Figure**). We will then (3) crosscheck each candidate abandonment to ensure it was not affected by other processes (e.g., city annexation). We will accomplish this using satellite imagery from Google Earth, newspaper records from LexisNexis, Wikipedia entries, etc. We will append the comprehensive database of population counts with a flag identifying settlements from which we can find no evidence of occupation.

Each candidate abandoned settlement will be checked using satellite imagery from Google Earth, newspaper records, town histories, etc. as Arenstam Gibbons & Nicholls do for Holland Island (2005). Many of these records exist, for example, Nushagak AK is a fishing village in operation until at least 1936; St. Elmo CO is a former gold mining town in the National Register of Historic Places; and Lisbon MI's official history now reads "The village of Lisbon, whose population once boasted 350 inhabitants, is now deserted… Many feel the decline of Lisbon, the old stagecoach town, was when the railroad came to Sparta." We will systematically scrape text on the web using key search words to develop a historical panel settlement dataset that includes variables that describe the broad characteristics of abandoned and non-abandoned settlements as well as any time-varying political, economic, and climatic shocks each was exposed to over the period of 1890-2020. Our initial cross-checks indicate that some of the drivers of settlement abandonment may include the erosion of mines, fishing reserves, flooding, as well as racial terrorism which we will be able to explicitly test.

In creating this panel database, we can run a first stage regression analysis that captures social mechanisms underlying settlement abandonment. Specifically, we can identify which factors explain most of the variation in a settlement abandonment dependent variable. In analyzing this information, we can better interpret our local average treatment effects (LATE) from settlement abandonment as well as estimate heterogenous treatment effects in how specific types of Displacement Related Stressful Life Events (DR-SLE) impact human longevity.

We will follow best practices as outlined by the US Census Bureau (Alexander, Fisher, and Genadek 2022) and the United Kingdom's Office for National Statistics (ONS) (Clausner et al. 2017) for digitizing historical census data. Notably, we will carry out a validation and evaluation process to ensure correct documentation of the population in each census. Population data for Places are machine-readable for the 1970 Census onward but the population counts for 1890-1960 must converted into a tabular format. We will select a random sample of 3,000 place-years (1.1%) and double-check the input value for these cells. ONS suggests a validation accuracy of 98% (Clausner et al. 2017) while the US Census Bureau's historical data linkages have 90% accuracy (Massey 2017). We will publish the accuracy rates, with corrected values for those discovered to be erroneous, and re-create the database if our accuracy assessment is below 95%.

The National Institutes of Aging funded this project in September 2004 and we anticipate having a database finalized by PAA in 2025.

References

Addyman, P. V. 1989. “The Archaeology of Public Health at York, England.” *World Archaeology* 21(2):244–64. doi: 10.1080/00438243.1989.9980104.

Alexander, J. Trent, Jonathan D. Fisher, and Katie R. Genadek. 2022. “Digitizing Hand-Written Data with Automated Methods: A Pilot Project Using the 1990 US Census.” *Journal of Economic and Social Measurement* 46(2):95–108.

Arkush, Elizabeth. 2008. “War, Chronology, and Causality in the Titicaca Basin.” *Latin American Antiquity* 19(4):339–73. doi: 10.1017/S1045663500004338.

Atherwood, Serge. 2022. “Does a Prolonged Hardship Reduce Life Span? Examining the Longevity of Young Men Who Lived through the 1930s Great Plains Drought.” *Population and Environment* 43(4):530–52.

Black, Richard, Stephen RG Bennett, Sandy M. Thomas, and John R. Beddington. 2011. “Migration as Adaptation.” *Nature* 478(7370):447–49.

Clausner, Christian, Justin Hayes, Apostolos Antonacopoulos, and Stefan Pletschacher. 2017. “Creating a Complete Workflow for Digitising Historical Census Documents: Considerations and Evaluation.” Pp. 83–88 in *Proceedings of the 4th International Workshop on Historical Document Imaging and Processing*, *HIP ’17*. New York, NY, USA: Association for Computing Machinery.

Farbotko, Carol, and Heather Lazrus. 2012. “The First Climate Refugees? Contesting Global Narratives of Climate Change in Tuvalu.” *Global Environmental Change* 22(2):382–90. doi: 10.1016/j.gloenvcha.2011.11.014.

Gill, Richardson B., Paul A. Mayewski, Johan Nyberg, Gerald H. Haug, and Larry C. Peterson. 2007. “Drought and the Maya Collapse.” *Ancient Mesoamerica* 18(2):283–302.

González-Tennant, Edward. 2019. *The Rosewood Massacre: An Archaeology and History of Intersectional Violence*. University Press of Florida.

Hino, Miyuki, Christopher B. Field, and Katharine J. Mach. 2017. “Managed Retreat as a Response to Natural Hazard Risk.” *Nature Climate Change* 7(5):364–70. doi: 10.1038/nclimate3252.

Hori, Makiko, and Mark J. Schafer. 2010. “Social Costs of Displacement in Louisiana after Hurricanes Katrina and Rita.” *Population and Environment* 31(1–3):64–86. doi: 10.1007/s11111-009-0094-0.

Massey, Catherine G. 2017. “Playing with Matches: An Assessment of Accuracy in Linked Historical Data.” *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 50(3):129–43.

McLeman, Robert A. 2011. “Settlement Abandonment in the Context of Global Environmental Change.” *Migration and Global Environmental Change – Review of Drivers of Migration* 21:S108–20. doi: 10.1016/j.gloenvcha.2011.08.004.

Reidmiller, David R., Christopher W. Avery, David R. Easterling, Kenneth E. Kunkel, Kristin LM Lewis, Thomas K. Maycock, and Bradley C. Stewart. 2017. “Impacts, Risks, and Adaptation in the United States: Fourth National Climate Assessment, Volume II.”

Romm, Joseph. 2011. “The next Dust Bowl.” *Nature* 478(7370):450–51.

Siders, AR. 2019. “Managed Retreat in the United States.” *One Earth* 1(2):216–25.

Simms, Jessica RZ, Helen L. Waller, Chris Brunet, and Pamela Jenkins. 2021. “The Long Goodbye on a Disappearing, Ancestral Island: A Just Retreat from Isle de Jean Charles.” *Journal of Environmental Studies and Sciences* 11(3):316–28.

Xie, Hualin, Peng Wang, and Guanrong Yao. 2014. “Exploring the Dynamic Mechanisms of Farmland Abandonment Based on a Spatially Explicit Economic Model for Environmental Sustainability: A Case Study in Jiangxi Province, China.” *Sustainability* 6(3):1260–82.

Xu, Chi, Timothy A. Kohler, Timothy M. Lenton, Jens-Christian Svenning, and Marten Scheffer. 2020. “Future of the Human Climate Niche.” *Proceedings of the National Academy of Sciences* 117(21):11350. doi: 10.1073/pnas.1910114117.