

The Formation of the American Political Field, 1872-1924

Stephanie L. Mudge, UC-Davis & NYU Abu Dhabi (presenting co-author)

Shahar Zaken, UC-Davis (presenting co-author)

26 February 2025

Short abstract

This article develops a historical account of the formation of the American political field between 1872 and 1924. We deploy a neo-Bourdieuian conception of the political field via a computational analysis of emergence and change in axes of opposition, dominant personas, and the overall scope of American politics using over 20 million pages of historical newspapers. Two novel computational measures show, with striking consistency, that the American political field semantically expanded and then contracted between 1872 and 1924, reaching maximum breadth in 1916. Within this arc were three mini-cycles: 1872-1896, reaching peak breadth in 1892; 1896 to 1912, peaking in 1904; and 1912 to 1924, peaking in 1916. Patterns are broadly consistent with relevant historiography of the period, including literatures on realignments and critical elections. The mini-cyclical pattern is also consistent with the Marxian insight that there is an intimate connection between capitalism's crisis cycles and dynamics of political contention: political field expansion tends to correspond with economic expansion, and *vice versa*. Further analyses focus on the relationship between political formation and geographical/territorial formation; look closely at election years that were dominated by personalities rather than issues, bringing the analytical focus down to the level of the person; and track successive shifts in axes of political opposition alongside the consolidation of an increasingly heavyweight administrative state. Ultimately, this article contributes to political sociology by demonstrating how computational methods, informed by sociological theory and relational sensibilities, can deepen existing understandings of the formation of political fields. More specifically, our aim is to better grasp a volatile period in American political history that, it would seem, has striking similarities with dynamics in the present.

Extended abstract

This article develops a field conception of party politics and puts it to work in a historical and computational account of the relational formation of the American political field between 1872 and 1924—a formative time in American, and indeed global, history. The period saw the consolidation of a Britain-centered gold standard monetary order that would later be recognized by Karl Polanyi for its subordination of community needs and democratic demands to the utopian fiction of the “self-regulating market.” It also saw the end of Reconstruction, American industrial takeoff, and the rise of organized labor—not to mention regressive tendencies that are eerily familiar: repressive racial, anti-socialist, and anti-labor politics; jingoism, imperial expansionism, and nativism; extreme economic inequality, political corruption, monopoly, and protective tariffs—to name a few. American politics’ trajectory in the 1910s and early 1920s defined some of the basic contours of the “short twentieth century” up to the early 1990s (in the formulation of Eric Hobsbawm).

As famously observed by Seymour Martin Lipset (1960), American politics in the short twentieth century took on the characteristics of a “democratic class struggle” contained and channeled by two national parties. The 20th century settlement that Lipset described,

however, required the elimination or evacuation of contention across other axes of political struggle. A series of contentious issues—the form and supply of money; the social, political, and economic rights of women and African Americans; Reconstruction and the political fate of the South; territorial expansion and the economic and political organization of states and localities—all faded as major issues of national contention. The United States became an institutionalized, federated political territory structured by a new architecture of railways stretching from the Atlantic to the Pacific and political and economic power, once concentrated in the Northeast and Midwest, shifted Westward and Southward. The federal state, meanwhile, shifted into an outwardly imperial mode, acquiring new territorial holdings in the Philippines, Puerto Rico and elsewhere. By the mid-1920s a newly articulated empire state-in-the-making unknowingly headed toward the Great Depression and a second World War.

Historiography of this critical period is voluminous but fractured. Predating national polls and surveys, systematic empirical mappings of the prewar political formation of the United States are necessarily partial, tending to focus on certain states or regions, major political personalities, particular public issues, or some combination. But mass print media, a main outlet of partisan and political expression during the time period concerned, offers the possibility of harnessing computational techniques to build a more systematic understanding. This is the project we pursue here.

More specifically, to better understand if and how the volatile politics of the late 19th and early 20th Centuries settled into the pattern Lipset described, we first introduce a theoretical approach that combines Bourdieu's conception of the political field, Mudge's (2018) neo-Gramscian conception of the political party, and a Goffmanian understanding of the development of partisan persons. We situate the political field in its larger, capitalist context, attributing to capital—following Marx—a *de facto* actorhood that exerts a strong gravitational pull in political life.

We put our conception to work via a computational analysis of emergence and change in axes of opposition, dominant personas, and the overall scope of American politics between 1872 and 1924 using a chronologically balanced sample of presidential election-year articles drawn from over 20 million pages of historical newspapers in the *American Stories* dataset (itself derived from the Library of Congress' *Chronicling America* collection). Our sample identified and extracted specifically political articles using a keyword-based approach that detects mentions of political parties active in presidential election years—a strategy consistent with the fact that, other than the state, parties and party oppositions are primary structuring forces in the American political field.

A first analytical task is to develop meaningful, informative, meso-level measures of the field as a whole, conceived as a shifting space of political positions and position-takings. As a first step in this direction, we use word embedding methods to develop indicators of *semantic spread*, or broadening (expansion) and narrowing (contraction), of political meanings.

Word embedding allows researchers to map the “meanings” of words and trace how those meanings change over time. “Meanings” are interpreted relationally, in terms of interpretive semiotics, in which meaning is found in the relationships *between* signs. In computational linguistics, “meanings” refers to *representations* of words captured by word embedding models. Representations take the form of vectors in a high-dimensional space; the position and direction of each vector encodes semantic information derived from the

contextual relationships between words in large text corpora. Words with similar meanings or usage patterns tend to have vectors that are close to each other, making possible the detection of semantic similarity, analogies, and associations (Mikolov et al. 2013). By tracing how associations between words in high dimensional space change over time, researchers can find how the relational “meanings” of words change historically (eg, Hamilton et al 2016; Kutuzov et al. 2018; Kozlowski, Taddy, and Evans 2019).

Our interest in these methods lies in their potential for historically rich descriptive insights into the shifting boundaries and relational structure of American party politics over time, capable of answering field-theoretical questions as to the breadth and scope of the political field, patterns of expansion and contraction, and shifting axes of position and opposition. Our theoretical framework, however, requires extending semantic shift analysis in computational scholarship to-date in order to develop indicators that speak, first, to the matter of the formation of the political field *as a whole*. To this end, we first develop and present two new, complementary indicators, providing a general descriptive picture of the trajectory of the American political field—on the basis of which, following our theoretical framework, we are building out a fuller analysis.

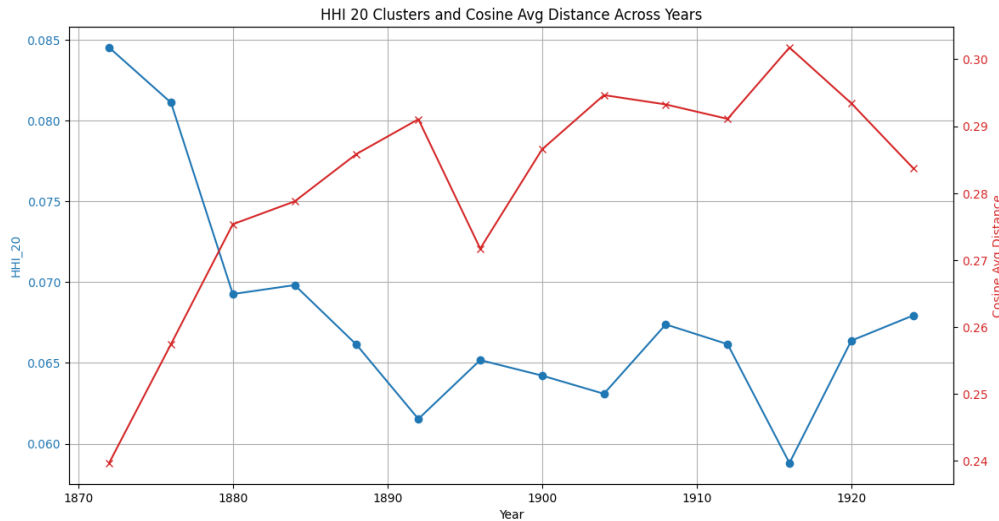
A first of two novel indicators of the overall scope of the political field is *average cosine distance* between political articles for each year. Instead of quantifying changes in the meaning of words, this indicator tracks the average cosine distance between document embeddings so that we can get a picture of the scope, or spread, of the whole semantic terrain of political news. In other words, rather than identifying specific words that change in meaning over time, we quantify the aggregate semantic distance between political articles within and across presidential election years. While semantic shift detection might track, for instance, how the meaning of 'democracy' evolves over a given period, this first indicator shows how the *entire* landscape of political discussion expands, contracts and shifts over time. Practically, we do this by constructing one large word embedding model and situating political newspaper articles in relation to each other using their embedding in the high-dimensional space of the model. Increasing values of our average cosine distance measure indicate field expansion; decreasing values indicate field contraction.

A second indicator of field scope is developed in two steps: it uses K-Means clustering of articles from all years, which tracks unique thematic groupings of political content, in order to then calculate a Herfindahl-Hirschman Index (HHI), which measures the concentration of articles within clusters. For the K-means measure, we cluster all years together instead of separately because comparing across separate clustering operations of different data is arbitrary; it does not produce comparable results because each year has a different data structure. K-means clustering is a method used to partition data into a predetermined number of clusters, where each article is assigned to the cluster with the nearest centroid. The algorithm works iteratively, first initializing random centroids and then refining them by minimizing the within-cluster variance, effectively pulling the centroids toward the densest areas of the data. After inspecting the results of multiple values, we set a predetermined number of clusters at 20. The articles in each cluster are more similar to each other (in terms of their TF-IDF weighted embeddings) than articles in other clusters. Once the clusters were formed, we identified the top TF-IDF words that defined each cluster. We then calculated the distribution of articles across clusters by year and region. Finally, to quantify this distribution, we calculated the HHI. Interpretation of this measure is the

opposite of the average cosine measure: *higher* values of HHI mean *less* diversity in the distribution of articles (lower semantic spread), and *lower* values indicate greater breadth.

Together, these two novel indicators paint a consistent picture of a political field that radically expanded and then partially contracted between 1872 and 1924, reaching a maximum breadth in 1916 (see Figure 1).

Figure 1. The expansion and contraction of the American political field, 1872-1924



Source: Mudge and Zaken 2025

Within this general arc, both indicators—strikingly, we think—suggest three mini-cycles of expansion and contraction in a consistent temporal pattern: an expansion-contraction mini-cycle between 1872-1896, reaching peak breadth in 1892; another mini-cycle from 1896 to 1912, peaking in 1904; and a third mini-cycle from 1912 to 1924, peaking in 1916. We note that the pattern is broadly consistent with relevant historiography of the period, including literatures on realignments and critical elections. The mini-cyclical pattern of field formation is also consistent with the Marxian insight that there is an intimate connection between capitalism's crisis cycles and dynamics of political contention: periods of economic expansion/contraction and political expansion/contraction are, according to our preliminary analysis, strongly and positively correlated.

From here we are developing our analysis in several directions. One considers how semantic spread varied with geographical shifts in news coverage, as Southern states were reincorporated, the country's political boundaries and territories consolidated, and the development of railways extended from coast to coast. Another, which incorporates network methods, is motivated by preliminary findings suggesting that certain presidential election years and not others were heavily defined by political personalities (as opposed to issues). Drawing on the Goffmanian element of our framework, we track the varying centrality of political personas, or *partisan persons*, as part of larger fielding processes. A third focuses on the emergence and collapse of axes of opposition—race, money, trade, women, immigration, labor...—and, in the time period between civil service reform (the Pendleton Act, 1883) and World War I, the consolidation of a federated political field that, anchored by a newly

heavyweight administrative state, defined the parameters of democratic class struggle in the short 20th Century.

Ultimately, this article [contributes to historical](#), political, and economic [sociology by](#) showing [how computational methods](#), informed by sociological theory and relational sensibilities, [can demonstrate](#) and expand existing understandings of [the formation of political fields](#). More specifically, our aim is to deepen our understanding a volatile and critical period in American political history that, it would seem, has striking similarities with dynamics in the present.