

# Nationalization of U.S. Local Elections

Francesca Tang\*

## Abstract

Scholars have documented an upward trajectory of the nationalization of the American electorate, but frequently overlook local elections. The degree to which local elections have nationalized remains unclear. Furthermore, existing tools for measuring electoral cannot be appropriately applied to local elections. I develop a measure of nationalization based on a k-means clustering algorithm and show that nationalization trends differ between subnational and local elections: congressional and gubernatorial elections have markedly nationalized but local elections have not followed the same rising trend and exhibit considerably lower degrees of nationalization. Conclusions made exclusively based on subnational elections, elections that are federal or elected at the state-level, are not complete and political behavior has not converged across local units to the extent previously postulated. To carry out my analysis and satisfy the need for centralized and comprehensive local election data, I also present an aggregate dataset, including presidential election returns by municipality that covers 2000 to 2020 and more than 16,000 municipalities.

**Keywords:** Nationalization, elections, local politics, machine learning, clustering.  
Word count: 11,292

---

\*Department of Politics, Princeton University, Princeton, NJ (email: frtang@princeton.edu).

# 1 Introduction

The past few decades have been marked by the increasing nationalization of the American electorate, defined as the growing convergence or uniformity of political engagement across geographical units (Jacobson 2015; Hopkins 2018). More specifically, the rise in electoral nationalization points to a growing convergence of presidential voting behavior and subpresidential voting behavior. This trend suggests that national winds are permeating subnational environments that were once protected from such influences (Stokes 1967; Sundquist 1973).

In what follows, I distinguish between two types of elections: non-federal elections that elect representatives below the state and elections that are either federal or elect representatives at the state-level. The first category, which I will refer to as “local elections”, includes mayoral and state legislative elections. The second type of elections, which I will refer to as “subnational elections”, includes U.S. House, Senate and gubernatorial elections. Local and subnational elections are two kinds of subpresidential elections. There are several factors that separate these two categories of elections, the first and most important being features about the elections themselves: unlike local elections, subnational elections have higher turnout (Hajnal 2009), fewer uncontested races, and are more often concurrent with presidential elections. Other factors include salience to voters and media coverage, and are heavily covered by national news media (Angelucci, Cage, and Sinkinson 2023+; Hayes and Lawless 2018). Surveys show that approximately two thirds of American can name their governor, a third of Americans can name their representative<sup>1</sup>, while only 20% can name their state legislator<sup>2</sup>. The decline of local news and redirection of resources from state and local news media also contribute to this difference. Although mayors of cities like New York and Chicago may receive large amounts of media coverage and have very broad name-recognition, most municipalities in America do not fall under this category.

---

<sup>1</sup>Survey by Haven Insights LLC in 2018.

<sup>2</sup>Survey by Johns Hopkins in 2018.

The overwhelming consensus on rising nationalization has largely been founded upon subnational elections, namely House and Senate (Amlani and Algara 2021; Carson, Sievert, and Williamson 2020). But these subnational elections are not representative of all American elections. The characteristics of these higher-office subnational elections naturally lend to greater alignment between their election outcomes and those of presidential elections. This suggests that statements about nationalization solely based on subnational elections may be inadequate. Whether purposefully or not, the current discourse on nationalization and polarization has largely left out local elections. If electoral nationalization is to be taken as the uniformity of election outcomes across geographical units, I argue that local elections need to be added to the discussion and measurement of nationalization in order to reach more accurate and comprehensive conclusions.

Since local elections have been neglected in much of the nationalization research, current measures of the convergence of presidential and subpresidential votes have generally sufficed. The most common approach involves finding the linear correlation between the two votes or using a linear regression model. This approach only carries significance and produces easily-interpretable results if the data satisfies certain assumptions and is linearly related. While House and Senate votes are frequently linearly correlated with their corresponding presidential votes, local election votes are rarely linearly correlated and appear to exhibit much more complex relationships (see Appendix C for an example). With data that is non-linearly related, the interpretation of estimated coefficients becomes more difficult and hard to compare across different elections. For instance, in cases where coefficients are insignificant, one could not conclude that the election with the larger coefficient has a higher level of nationalization. Even if they are significant, this comparison of the magnitude of the coefficients is not evident either: a 0.8 correlation does not necessarily imply double the amount of nationalization than a 0.4 correlation. Gauging and comparing exactly how nationalization of various subpresidential elections has evolved over time requires a more

appropriate measure that will allow us to summarize the difference in voting behavior even when the data is nonlinear.

I propose a new measure of nationalization that quantifies the similarity between voting behavior in presidential elections and any other subpresidential election, including local elections. This measure relies on a simple k-means clustering algorithm that partitions units of analysis (congressional district, county, municipality, etc.) based on the margin of victory of a given party, allowing nationalization to be compared across different subpresidential elections and accounting for nonlinear relationships inherent in voting behavior. As a complement to my measure, I also examine the distribution of the differences in margin of victories between presidential and subpresidential elections. This allows for a more detailed study of the differences in voting outcomes, but also poses problems such as choosing the percentiles used to compare the variation in convergence, drawing conclusions from statistically different distributions, and deriving a single measure of nationalization for each election. Instead of setting arbitrary thresholds and making decisions by eye, I use a nonparametric algorithm that conveniently summarizes data without losing valuable information. Nevertheless, I present this alternative approach to supplement my proposed measure.

It is necessary to emphasize that although it is a type of nationalization, electoral nationalization does not necessarily imply or advocate for the causal, top-down effect of the presidential vote on the subpresidential vote. There are highly likely other factors such as partisanship that are involved in producing a particular trend in the nationalization of elections. I focus on the measurement of the convergence in voting outcomes at the two levels, not on whether the presidential preference is representative of national partisanship and dictating down ballot votes.

In order to address the question of the extent to which nationalization has trickled down to state and local politics, I present a dataset that is comprised of three categories: presidential election returns, subnational election returns (congressional and gubernatorial), and

local election returns (mayoral and limited state legislative). The biggest contribution of this dataset is the presidential election results at the municipal level, as the large majority of states does not aggregate results by municipality. This dataset covers 2000 to 2020 and contains 16,678 municipalities. Warshaw and Tausanovitch (2013) compile a similar dataset of presidential vote estimates at the municipal level but their dataset does not include elections prior to 2008, contains fewer municipalities, and does not include total votes. Although not the original vote share numbers, my dataset provides a very close estimate of how a considerable set of municipalities voted in presidential elections.

Using this dataset and measure of electoral nationalization, I offer a precise and unified method that contributes to the existing literature in three major ways. First, although I identify evidence that electoral nationalization has indeed been on a steady increasing trajectory in the last few decades for House, Senate and gubernatorial elections, nationalization of local elections is much lower and has not seen a significant uptick in recent elections. This disparity suggests that voters are not voting the same way in subnational elections when compared to local elections: in the latter case, national partisanship and agendas have not prevailed to the same degree. The inclusion of local election results casts doubt on the pervasiveness of nationalization beyond subnational elections, possibly pointing to a degree of separation from national influences and less heightened partisanship at the local level. The second contribution is a nonparametric and data-driven measure of electoral nationalization that uses a k-means clustering algorithm. This measure is a more appropriate choice than the status quo approach of ordinary least squares or linear correlation because it addresses the relevant issue of nonlinear relationships between presidential and subpresidential electoral outcomes and summarizes the relationships so that direct comparisons of nationalization can be made across elections, over time. I also supplement this measure with observations of the distributions of the absolute difference in presidential and subpresidential margin of victories, reaching similar conclusions. Finally, I find that election timing plays a more central role

in the reduced level nationalization of local elections than the heavily nonpartisan nature of mayoral elections.

This paper proceeds as follows. First, I discuss the literature on electoral nationalization and local politics. I then outline the multi-pronged process of data collection. Next, I define my measure of electoral nationalization and subsequently, present the results for various subpresidential elections. Finally, I conclude and discuss the implications of the results.

## Nationalization of U.S. Elections

Early work on how presidential evaluations affect downstream voting behavior (Carsey and Wright 1998; Simon 1989) do not frame these effects in the context of nationalization but these links have been established more recently in studies such as Abramowitz and Webster (2016); Jacobson (2015); Sievert and McKee (2019). The twentieth century saw peaks and troughs of nationalization in American electoral politics, with the current period situated at an all-time-high. The 1960s and 1970s experienced a peak in candidate-centered voting, and local elections were heavily focused on local issues and candidates (Bartels 2000; Green, Palmquist, and Schickler 2002). What many call the decline of parties, or a partisan dealignment, and a shift away from the Michigan model (a model that argues for the driving force of party identification behind voting Campbell et al. 1980), these two decades marked a period of lower levels of straight-ticket voting and party loyalty for both the Democratic and Republican party. Earlier works such as Kawato (1987) and William Claggett and Zingale (1984) that examine nationalization in terms of partisan support across various geographical subunits do not find clear evidence of growing nationalization at the time of writing.

However, the 1980s marked a turning point in the reconfiguration of American politics. For instance, there was a rise in national party strength as well as an increase in markers of partisanship such as negative partisanship and ideological polarization (Huckshorn et al.

1986; Ceaser and Saldin 2005; Abramowitz and Webster 2016). Bartels (2000) offers evidence that party loyalty in presidential and U.S. House elections has been on the rise and Jacobson and Carson (2016) document the increasingly similar electoral outcomes of presidential and House and Senate races. Some have even posited that Reagan’s presidency consolidated the Republican Party while the Democratic Party positioned itself as the cohesive, liberal alternative. Bartels (1998) suggests that the period immediately after Reagan’s presidency marked a long-term nationalization of the American political landscape. In general, there is little disagreement about the rise of electoral nationalization, and nationalization as a whole, in modern American elections.

If these observations of rising nationalization of U.S. elections are valid, they are concerning for several reasons. First, they run counter to the original vision of federalism held by the framers of the Constitution who emphasized the authority of local governments over local policies. They expected a citizen’s attachment to their state government to take precedence over their attachment to the national government as state and local governments have an advantage in capturing their loyalties by addressing local concerns that affect their daily lives (Alexander Hamilton and Jay 1788). Second, the accountability of local representatives is weakened if their electoral success is largely tied to national issues and agendas rather than their own performance. At the same time, local officials have much less authority over national issues and with reduced incentives to engage in local issues, this comes at the detriment to the well-being of their constituents. Third, increased levels of nationalization can also lead to centralized, uniform party platforms and policy agendas across state and local governments, reducing variation and regional differences that previously existed. Finally, some studies have also made the link between increased nationalization and greater polarization as well as the changing landscape of local news (Martin and McCrain 2019; Moskowitz 2021).

# Nationalization of Local Elections

Much of the current literature on nationalization does not engage heavily with elections at the very local level. Conclusions about electoral nationalization are drawn almost exclusively from congressional or occasionally state-level races even though its definition hinges on uniformity across local units. The study of local government elections has seen a surge in the last decade, but the understanding of local elections is still somewhat limited and the methodologies used lag behind those used in general American Politics research; research on local elections has heavily relied on either qualitative studies or exclusively on large urban cities. More importantly, little methodological advancements have been made on this front in order to better understand the relationship between local and national politics. This is largely due to the lack of centralized, comprehensive and clean data, something that some scholars are working to remedy.

A divide between local politics research and the rest of mainstream American politics research still exists, especially due to the traditionally more qualitative nature of urban politics research (Anzia 2021). Although more recently, there has been an increasing body of research that puts state and local politics at the center of political activity and party mobilization due to various national trends like partisanship (Grumbach 2022; Hertel-Fernandez 2019). Scholars have found that local elections tend to have much lower turnout, less media coverage, and are often nonpartisan (Schaffner, Streb, and Wright 2001; Hajnal 2009; Coppock and Kirkland 2018; Warshaw 2019). This can be partially explained by the fact that many municipal elections do not take place concurrently with national elections and in general, off-cycle elections receive less voter interest and exhibit lower incumbency advantage (de Benedictis-Kessner 2017).

Only recently have partisanship and ideology, previously exclusive to national and state politics, permeated local and urban politics research; there has been an increasing amount of evidence that partisanship has seeped into local politics and that the partisan prefer-



ence or ideology of voters is an important predictor of voting behavior (Hajnal and Troun-  
stine 2014; Boudreau, Elmendorf, and MacKenzie 2015; Einstein and Glick 2018; Warshaw,  
de Benedictis-Kessner, and Velez 2022). From these studies, one may expect to find signs  
of growing nationalization at the local level as well. As voters may be less knowledgeable  
about or invested in policy positions of state and local candidates, given that these are low-  
information elections, they may become more inclined to use party identification to pick be-  
tween candidates when possible. In fact, Rogers (2016) claims that state legislative elections  
specifically are heavily influenced by sentiments regarding the president and the national  
parties. Rogers (2023*a,b*) also establish the link between the drop in incumbency advantage  
and the increase in correlation between the vote shares of presidential and state legislative  
elections, albeit still a smaller correlation than that of presidential and House elections. Hol-  
liday (2023) also finds that partisanship is a dominant force even in down-ballot partisan  
races in Maricopa County, Arizona, but that they are not necessarily as “presidentialized”.  
These results suggest that polarization and national issues could be playing a role in local  
or lower salience elections, where voters are tapping into their party identities, preferences  
and loyalties when voting even when these signals are weaker than those in higher-office  
subnational elections.

There has also been evidence from sources other than election data, demonstrating that  
local representatives are still rooted in their local agendas. One of the few studies that  
attempts to capture nationalization at the mayoral level is Das et al. (2022), who analyze  
tweets from mayors and find that the average American mayor is relatively unaffected by  
national rhetoric and mostly focuses on the needs of their local constituents. Local news  
media has also been found to attenuate some effects of nationalization (Moskowitz 2021).  
More recently, Kuriwaki (2023) uses cast vote records from South Carolina to show that  
the rate of ticket splitting is higher for local elections like county council or sheriff than for  
Congress. But there are clear drawbacks from using cast vote records, namely that very few

states offer this data and ticket splitting can only account for nationalization in concurrent elections. In general, the measurement of nationalization using electoral returns is largely absent in scholarship of local elections.

A definitive measure of the convergence of presidential and subpresidential voting behavior, that can be applied to a wide range of elections, still does not exist. As stated in Hopkins (2018), nationalization research in the United States is still limited and has barely scratched the surface of state and local elections. Since local elections have been mostly ignored in the discussion surrounding nationalization, it is still unclear how localized voting behavior is and what mechanisms could be driving any potential differences. As local elections differ significantly from other elections, I expect to find contrasting nationalization trends between these elections.

## Data and Dataset Construction

To accomplish the goal of rigorously quantifying nationalization of various levels of elections as described above, a set of electoral returns of presidential, House, Senate, gubernatorial, and local elections is needed. To consistently relate subpresidential elections to presidential elections, the same unit of comparison is used, i.e. county, congressional district, or municipality. Therefore, I collect data for House and presidential elections at the congressional district level; Senate and presidential elections at the county level; gubernatorial and presidential elections at the county level; gubernatorial and presidential elections at the congressional district level; mayoral and presidential elections at the municipal level; and state legislative and presidential elections at the state legislative district level. Ideally, nationalization would be measured for a set of presidential and subpresidential elections based on the corresponding electoral unit, but for governors and senators, who are elected at the state-level, having only fifty states would be too small of a sample size. I thus use election

returns at the county level.

As a result of assessing the convergence or divergence of voting behavior, the nationalization measured using these units (congressional districts, counties, municipalities, and state legislative districts) will also reflect a similar convergence or divergence of representation. If voting precincts were used uniformly for all measurements of nationalization then that would more purely represent voting behavior but collecting precinct-level data for all elections would require a much larger endeavor.

Putting all these parts together requires several different data sources. From CQ Press' Voting and Elections Collection, I collect presidential, gubernatorial, and Senate election results at the county level as well as House elections at the congressional district level. As for the presidential and gubernatorial electoral returns at the congressional district level, I use data compiled by myself and others.<sup>3</sup> These approximate returns are obtained by using county-level returns and allocating votes proportionally in each state according to the proportion of population of the county that belongs to each congressional district. Building upon Warshaw, de Benedictis-Kessner, and Velez (2022)'s local election dataset of U.S. mayoral elections of municipalities with a population of 50,000 or more residents, I fill in missing municipalities, party affiliations, and results from the most recent elections in 2022 where possible. I also record whether the mayoral election is a nonpartisan or partisan race (whether party affiliation appears on the ballot) and party affiliations of candidates who run in nominally nonpartisan races. This latter variable is relevant because despite a large quantity of local races being nominally nonpartisan, affiliations of candidates are often known to voters or made evident during the campaign (Coppock and Kirkland 2018). The party affiliations of mayoral candidates in nonpartisan races is found by combing through candidate websites, local news media, and other relevant platforms (Gerber and Hopkins 2011, do something similar). The final dataset of mayoral elections contains 630 municipalities and

---

<sup>3</sup>Christian Baehr, Dahyun Choi, and Rocío Titiunik, Department of Politics, Princeton University

5084 races between 2000 to 2022, accounting for around 37% of the U.S. population. Table 1 displays the breakdown of municipalities included based on population and other variables. Due to restraints in data, the municipalities included in the dataset introduce a few biases but the dataset still offers a decent variety of small to large towns from all fifty states. Although nonpartisan elections represent 76% of the total races, elections where at least one known Democrat is running constitute 67% of the municipalities included. As expected, there is a large minority of races that are uncontested (19%) and the large majority of races are non-concurrent elections (91%), two features that are not characteristic of non-local elections.

Table 1: **Dataset of U.S. mayoral elections, 1990-2022**

	Number	%	Avg Population	Total Population
All municipalities	630	100	185,061	116,588,526
>250,000	91	14	686,411	62,463,388
>500,000	37	6	1,200,726	44,426,866
<250,000	539	86	100,418	54,125,138
<100,000	334	53	70,998	23,713,202
All races	5,084	100		
Partisan	1,124	22		
Nonpartisan	3,878	76		
Concurrent	442	9		
Non-concurrent	4,642	91		
Same-Year	750	15		
Off-Year	4,334	85		
Uncontested	984	19		
Democrat runs	3,428	67		

Partisan are partisan races where candidates’ party affiliations are written on the ballot and there are partisan primary elections. Concurrent is a mayoral race that was held at the same time as a presidential election and Same-Year is a mayoral race that was held in the same year as a presidential election. Democrat runs refers to a mayoral race where there is at least one Democrat running in the election, regardless of whether the race is nonpartisan or partisan.

The largest contribution of this data collection process is providing a more comprehensive dataset of presidential vote estimation at the municipal level. While this data may be

available for a few select states and the largest urban areas, it is largely missing from the universe of American election data. Warshaw and Tausanovitch (2013) offer a version from 2008 - 2020 but contains fewer municipalities and elections, without including total votes (see Appendix A for a detailed comparison and description of the approximation methodology). Ultimately, the dataset contains presidential election returns for more than 16,000 municipalities for elections from 2000 to 2020 (I only use the 630 municipalities that are also in my mayoral election dataset). It is an addition to the growing body of local election data, allowing researchers to determine how voting behavior observed on the national level relates to that of the local level.

## Methodology and Nationalization Measure

One of this paper’s main contribution is a single data-driven, nonparametric, and easy-to-implement measure of electoral nationalization that can be easily applied to election returns data. To this end, I cluster the Democratic (or Republican) margin of victory of all relevant units of study  $i = 1, \dots, N$  (county, congressional district, or municipality) using k-means clustering at both the subpresidential level and the presidential level<sup>4</sup>. More explicitly, after choosing  $K$ , for any particular election at time  $t$ , I first cluster on the presidential Democratic margin of victories using k-means clustering for a set of units. Then, I do the same for the subpresidential Democratic margin of victories for the same year  $t$  and units. With these two sets of vote clusters, I measure nationalization of an election at time  $t$  by the proportion of units where the presidential election and subpresidential units are assigned to the same vote cluster. It is important to note that the focus here is not on the optimal number of clusters or the quality of resulting clusters, but rather what a simple k-means clustering algorithm allows for: reducing units to centroids or groups in a nonparametric way and comparing

---

<sup>4</sup>A party’s margin of victory is defined as the party of interest’s vote share minus the vote share of its strongest opponent.

them, without needing to set arbitrary thresholds or make assumptions about the data's distribution.

Algorithm 1 outlines the k-means algorithm in terms of margin of victory. I use margin of victory, and not say the Democratic two-party vote, since it can be easily applied to other countries that do not have a clear two-party system and to local elections that have third-party or independent candidates running. The central goal of k-means is to partition  $N$  observations into  $K$  clusters by minimizing within-cluster variances. The resulting cluster assignments minimize the total cluster variance

$$C^* = \min_C \sum_{k=1}^K A_k \sum_{c_t^{(i)}=k} \|x_{i,t} - \mu_k\|^2, \quad (1)$$

where  $A_k = \sum_{i=1}^N \mathbb{1}\{c_t^{(i)} = k\}$ .  $C^*$  in Equation (1) is the solution that minimizes the within-cluster average dissimilarity of observations from the cluster's centroid.

---

**Algorithm 1** K-means Clustering of Margin of Victory

---

initialize  $\mu_1, \dots, \mu_K \in \mathbb{R}$  randomly

Repeat until convergence

**for every**  $i$  **do**

$c_t^{(i)} := \operatorname{argmin}_k \|x_{i,t} - \mu_k\|^2$

**for every**  $k$  **do**

$\mu_k := \frac{\sum_{i=1}^N \mathbb{1}\{c_t^{(i)}=k\} x_{i,t}}{\sum_{i=1}^N \mathbb{1}\{c_t^{(i)}=k\}}$

**end for**

**end for**

---

Now using k-means clustering, I define the level of nationalization  $\gamma_t$  at election  $t$  as the share of units partitioned to the same vote cluster at the subpresidential level and the presidential level:

$$\gamma_t = \frac{|\{i = 1, \dots, N | L_{i,t} = P_{i,t}\}|}{N}, \quad (2)$$

where  $L_{i,t}$  is unit  $i$ 's cluster at the subpresidential level and  $P_{i,t}$  is unit  $i$ 's cluster at the presidential level.  $L_{i,t}$  and  $P_{i,t}$  are obtained for each  $i$  using the k-means clustering algo-

rithm. Given a set of a party's margin of victory for all units  $i = 1, \dots, N$  in an election  $t$ ,  $\{x_{1,t}, \dots, x_{N,t}\}$ , the algorithm is outlined in Algorithm 1. Essentially, Algorithm 1 is performed twice: once on the subpresidential elections and another at the presidential elections, using the same unit of analysis both times.

Since  $x_{i,t}$  and  $\mu_1, \dots, \mu_K$  are one-dimensional, each cluster  $C_t^k := \{i | c_t^{(i)} = k, i = 1, \dots, N\}$  can be directly interpreted based on their average Democratic (or Republican) margin of victories  $\mu_1, \dots, \mu_K$ . It is important to note that the labels of the clusters for the subpresidential elections  $L_{i,t}$  and those for the presidential elections  $P_{i,t}$  for each election  $t$  correspond to the same cluster.<sup>5</sup> To know whether  $P_{i,t}$  equals  $L_{i,t}$ , the vote clusters at the presidential and local level must be aligned. More explicitly, I calculate the resulting centroids of the vote clusters for the presidential and subpresidential elections:  $\mu_{1,P}, \dots, \mu_{K,P}$  and  $\mu_{1,L}, \dots, \mu_{K,L}$ , respectively. Without loss of generality, I label the resulting clusters ordinally according to  $\mu_{1,P} \leq \mu_{2,P} \leq \dots \leq \mu_{K,P}$  and  $\mu_{1,L} \leq \mu_{2,L} \leq \dots \leq \mu_{K,L}$ , where the same  $k = 1, \dots, K$  corresponds to the same voting cluster by mean Democratic margin of victory. However, the specific membership of these voting clusters is not of interest.

Once the clusters  $C_t^k, k = 1, \dots, K$  are found via Algorithm 1 for both the subpresidential and presidential voting returns for all units  $i = 1, \dots, N$ , the nationalization measure  $\gamma_t$  for election  $t$  is calculated using Equation 2. But if we consider a baseline of “zero” nationalization as each unit being randomly assigned to a voting cluster for both elections, we can rewrite Equation 2 as

$$\gamma_t^* = \frac{|\{i = 1, \dots, N | L_{i,t} = P_{i,t}\}|}{N} - \frac{1}{K}. \quad (3)$$

A  $\gamma_t^*$  less than  $\frac{1}{K}$  would suggest that there is a divergence of voting behavior that is non-random. Results in the following section are reported using Equation 3.

---

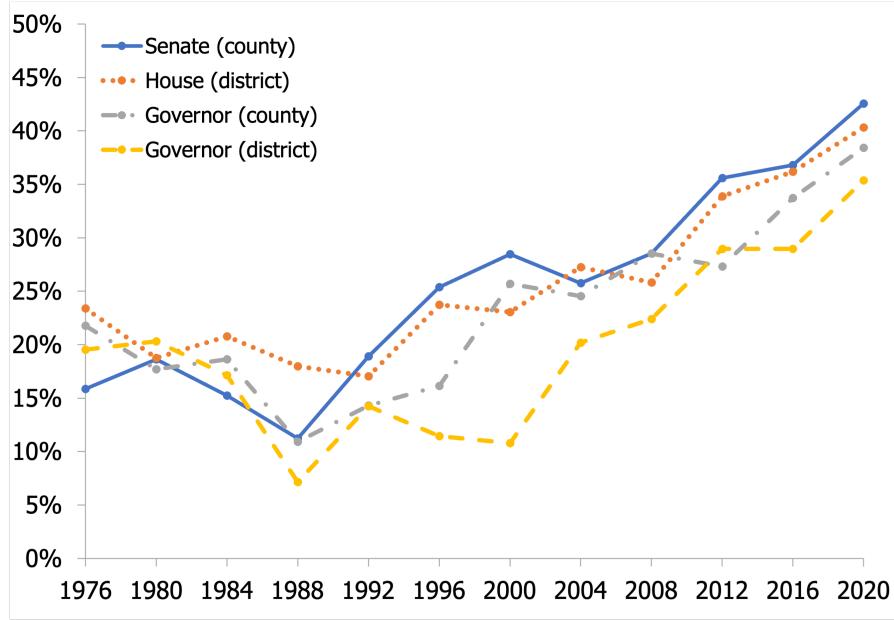
<sup>5</sup>K-means clustering initiates the centroids randomly and the numeric label it assigns to each cluster varies each time so it is important to make sure the label of each cluster corresponds to the same label at the subpresidential versus presidential level.

## Subnational Nationalization Results

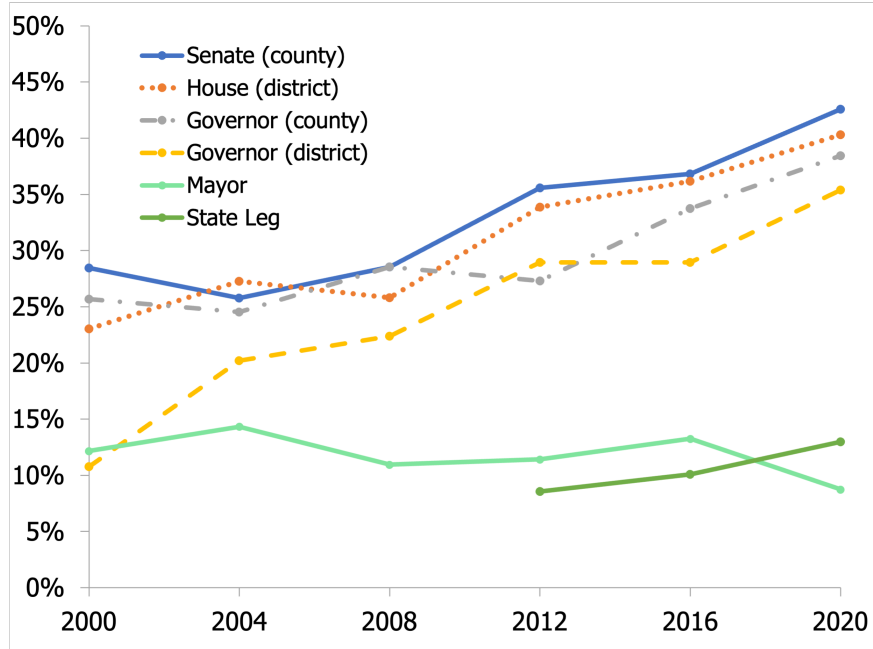
The first part of the results is the nationalization of three subnational elections: House, Senate and gubernatorial elections. Applying Algorithm 1 to the Democratic margin of victory, we obtain  $C^*$  of units that belong to each resulting cluster  $k = 1, \dots, K$ . As mentioned previously, the optimal  $K$  that produces the highest quality clusters is not the priority, but rather a convenient tool that allows us to measure convergence of vote choice. For simplicity, I conduct the analysis for  $K = 2$  and  $K = 3$ ,  $K = 3$  serving as a robustness check. For a more detailed discussion of choice of  $K$  see Appendix D. Table 2 reports the nationalization measure for  $K = 2$  and  $K = 3$  clusters. In both cases,  $\gamma_t$  clearly starts to increase from the late 1980s until 2020, as plotted in Figure 1a for  $K = 2$ . The additional evidence for  $K = 3$  reflects very similar trends as  $K = 2$ . Across the board, the 2020 election resulted in the highest degree of nationalization over the four and a half decades. This is also consistent with what political scientists like Abramowitz and Webster (2016) and Jacobson (2015) have observed about nationalization by examining party loyalty, negative partisanship, and incumbency advantage.

Nationalization appears to be the most pronounced in Senate elections with 56.3% in 2020 for  $K = 3$  and 42.6% for  $K = 2$ . The process of nationalization of Senate elections also witnessed the steepest and most striking increase starting from the 1988 election, as seen in Figure 1a. We also see a slightly smaller magnitude of nationalization, although still substantive in itself, for governor races in 2020, as well as in the few elections before that. In 2020, compared to the 42.6% for Senate races, the nationalization measure for House races is 40.3%, 38.4% for gubernatorial races at the county level, and 35.4% for gubernatorial races at the congressional district level. This has important implications in the varying degrees to which voter choice in congressional and gubernatorial results converge to voter choice in national elections. An obvious explanation is that Senate elections tend to receive more national media coverage, with Senators having longer tenures and higher recognition among





(a) Subnational Elections, for each presidential election in the period 1976-2020.



(b) Subnational and Local Elections, for each presidential election in the period 2000-2020.

Figure 1: Nationalization of Subpresidential Elections,  $K = 2$ .

the general public who are not their constituents, compared to House elections. House races can vary heavily in the amount of national attention they receive and some candidates have a reputation that goes beyond their congressional district. Many congressional election campaigns center their platform and messaging on more nationally-oriented issues compared to gubernatorial campaigns, which can partially elucidate the generally lower level of nationalization in recent elections. Nevertheless, nationalization measured for these three elections is not too distinct.

Table 2: **Nationalization of Subnational Elections, 1976-2020.**

<i>Unit</i>	Election Year	Clusters	Senate <i>County</i>	House <i>CD</i>	Governor <i>County</i>	Governor <i>CD</i>
	1976	2	0.159	0.234	0.218	0.195
	1980	2	0.186	0.188	0.177	0.203
	1984	2	0.152	0.208	0.186	0.171
	1988	2	0.112	0.180	0.109	0.071
	1992	2	0.189	0.171	0.143	0.142
	1996	2	0.254	0.237	0.162	0.114
	2000	2	0.285	0.230	0.257	0.108
	2004	2	0.258	0.273	0.245	0.202
	2008	2	0.285	0.258	0.285	0.224
	2012	2	0.356	0.339	0.273	0.290
	2016	2	0.368	0.362	0.337	0.290
	2020	2	0.426	0.403	0.384	0.354
	1976	3	0.184	0.204	0.162	0.159
	1980	3	0.168	0.218	0.119	0.167
	1984	3	0.214	0.202	0.177	0.060
	1988	3	0.125	0.210	0.073	0.060
	1992	3	0.233	0.186	0.159	0.124
	1996	3	0.273	0.277	0.156	0.072
	2000	3	0.374	0.250	0.260	0.229
	2004	3	0.295	0.172	0.249	0.344
	2008	3	0.332	0.245	0.341	0.256
	2012	3	0.486	0.190	0.357	0.363
	2016	3	0.437	0.098	0.388	0.351
	2020	3	0.563	0.480	0.480	0.421

Measure of nationalization in Senate, House, and Gubernatorial elections, defined as in (3), where Clusters is the number of clusters  $K$  used in Algorithm 1. CD refers to congressional district.

Table 3: **Nationalization of Mayoral Elections, 2000-2020.**

<i>Unit</i>	Election Year	Clusters	Mayor <i>Municipality</i>
	2000	2	0.122
	2004	2	0.143
	2008	2	0.110
	2012	2	0.114
	2016	2	0.133
	2020	2	0.088
	2000	3	0.112
	2004	3	0.134
	2008	3	0.048
	2012	3	0.056
	2016	3	0.100
	2020	3	0.077

Measure of nationalization in mayoral elections, as defined in (3), where Clusters is the number of clusters  $K$  used in Algorithm 1.

## Local Election Nationalization Results

### Mayoral Elections

Turning to local elections, Table 3 reports the results for mayoral elections from 2000 to 2020. As shown in Figure 1b, nationalization of mayoral elections remains distinctly lower than that of the other subnational elections over the last two decades, save for 2000 where gubernatorial races were approximately at the same level. For example, in 2020, mayoral election nationalization is measured to be at 7.7% for  $K = 3$  and 8.8% for  $K = 2$ , which are both significantly lower than the 2020 values for Senate, House, and gubernatorial races. There is also no upward trajectory over this period. This stark contrast suggests that mayoral elections have not nationalized to the degree as House, Senate and gubernatorial races and have maintained a lower, relatively more stable course. Despite studies that demonstrate the growing relevance of national partisanship in local politics and elections, voters are

not turning to party loyalties and affinities in mayoral elections as heavily and electoral nationalization at the more national level cannot be generalized to the local level.

## State Legislative Elections

Aside from mayoral politics, another integral part of local politics is state legislative elections. Despite how its representatives are state legislators, state legislative elections elect candidates neither for federal office nor at the statewide level. These officials have important policy-making responsibilities but Americans are generally relatively much less informed about state politics. For example, it has been shown repeatedly that approximately fewer than 20 percent of Americans can name their state legislators and half do not know whether their state has a one or two-house legislature (Johns Hopkins University, 2018). As a result, state legislative elections also have lower turnout, are frequently uncontested (39% from 2012 to 2020 as seen in Table 4), and receive much less attention from the news media, as these outlets dedicate a greater portion of their reporting to federal and mayoral elections.

But unlike mayoral elections, state legislative elections are partisan and are more likely to be concurrent, specifically 47% of races from 2012 to 2020 as seen in Table 4, which could prompt voters to use their national party affiliations or loyalties when making decisions in these low-information elections. With these considerations in mind, comparing state legislative nationalization to mayoral nationalization could help elucidate the relative effect of certain features that could be contributing to the differences in nationalization, namely the presence of nonpartisan elections.

Following the same basic process as mayoral elections, I collect state legislative election results from Klarner Politics from 2012 to 2022 as well as presidential election results at the state legislative district level from Daily Kos from 2012 to 2020. Due to limited data of presidential election results by state legislative district, this analysis is thus restricted to the period 2012 - 2020. I replicate the analysis with  $\gamma_t^*$  as defined in Equation (3) and the

Table 4: **Dataset of U.S. state legislative elections, 2012-2020**

	Number	%
All races	31,723	100
Same-Year	14,914	47
Off-Year	16,809	53
Uncontested	12,356	39
Democrat runs	24,400	77

Democrat runs refers to a state legislative race where there is at least a Democrat running in the election. All state legislative elections are partisan. State legislative elections are held in November so same-year elections are concurrent elections.

results are presented in Table 5, split by all state legislative elections, State Senate, and State House. During the examined period, nationalization of state legislative elections is very similar to that of mayoral elections: distinctly lower than that of House, Senate and gubernatorial elections.  $K = 2$  results for all state legislative elections is also plotted in Figure 1b. Evidently, the degree to which presidential voting behavior determines subpresidential voting behavior in state legislative elections is found to be much lower.

Moreover, state legislative nationalization is not noticeably higher than mayoral nationalization, and is often even lower, for instance in 2016. Even though they are partisan elections with at least a Democrat and or Republican running in all races, state legislative elections nevertheless exhibit low levels of nationalization. This implies that the presence of partisan ballots is not evoking partisan affiliations to a degree that may lend to higher nationalization. Additionally, removing uncontested races, which accounts for around 39% of all races, does not affect nationalization significantly, much like mayoral elections. The next step is to construct or gather presidential election returns at the state legislative district level prior to

Table 5: **Nationalization of State Legislative Elections, 2012-2020.**

Election Year	Clusters	All	Senate	House
2012	2	0.086	0.126	0.066
2016	2	0.101	0.085	0.099
2020	2	0.130	0.209	0.097
2012	3	0.097	0.114	0.087
2016	3	0.057	0.051	0.059
2020	3	0.118	0.198	0.090

Measure of nationalization in state legislative elections, defined as in (3), where Clusters is the number of clusters  $K$  used in Algorithm 1.

2012 in order to get a better sense of the development of nationalization for a more extensive comparison.

Table 6: **Nationalization of Local Elections, by type of race, 2000-2020**

<i>Type</i>	Election Year	Clusters	Mayor <i>Partisan</i>	Mayor <i>Nonpartisan</i>	Mayor <i>Contested</i>	State Leg <i>Contested</i>
	2000	2	0.101	0.121	0.118	
	2004	2	0.098	0.130	0.153	
	2008	2	0.149	0.118	0.145	
	2012	2	0.188	0.082	0.107	0.061
	2016	2	0.157	0.115	0.164	0.053
	2020	2	0.121	0.049	0.120	0.086
	2000	3	0.196	0.149	0.124	
	2004	3	-0.023	0.095	0.078	
	2008	3	0.010	0.041	0.117	
	2012	3	0.096	0.048	0.076	0.005
	2016	3	0.178	0.105	0.114	0.004
	2020	3	-0.049	0.091	0.172	0.076

Measure of nationalization in mayoral elections, as defined in (3), where Clusters is the number of clusters  $K$  used in Algorithm 1.

## What Distinguishes Local Elections?

To disentangle the potential factors or mechanisms behind the differences between local and other subpresidential elections, Table 6 presents results for different types of mayoral races. The Partisan and Nonpartisan columns separate the municipal elections nationalization analysis into partisan and nonpartisan races, respectively. Nonpartisan race nationalization tracks quite closely with the full sample, suggesting that the nonpartisan nature of mayoral races is not the driving force behind their lower nationalization. These results aligns with the fact that state legislative elections also display low levels of nationalization despite being wholly partisan races. Partisan races exhibit more fluctuations, partially due to the smaller sample size and its makeup of large metropolitan cities. For example, there is a noticeable drop in nationalization of partisan races in 2004 and 2020, as seen in Figure 2a, but not observed in Figure 2b. Note that for the nonpartisan elections, I only use elections where there is at least one Democrat or Republican candidate, party information I obtain by scouring local news outlets and various other websites to calculate the Democratic (or Republican) margin of victory. The Democratic margin of victory used is therefore based on party affiliation that I find for candidates.

For both mayoral and state legislative elections, uncontested elections are a sizeable portion of the sample, a feature that is not found in congressional or gubernatorial elections. These observations are included in the full analysis as uncontested races contain valuable information in terms of measuring nationalization. In a highly nationalized environment, local candidates affiliated with a national party would be more or less incentivized to run in a local race, even challenge an incumbent, if the party is perceived as more favorable in their local unit. If a non-negligible portion of races are uncontested, that suggests a conscious decision candidates make to run or not, given the local dynamics of partisanship identity and loyalty. Thus, leaving uncontested races in the nationalization analysis keeps valuable signals about how local voters perceive the national parties. But to determine if the higher

proportion of uncontested races has any effect on nationalization, I remove races where no Democrat is running. Instead of computing zero for races in which no Democrat runs, they are left out from the analysis. The results are presented in the “Contested” column in Table 6 and plotted in Figure 2c. The level of nationalization is slightly elevated, but not by much in most elections, implying that in mayoral elections voters are voting slightly more in line with presidential elections in races where there is more than one candidate running. This makes sense because in uncontested races, there is no other choice but to either abstain or vote for the one candidate, who may not belong to the party voters prefer at the presidential level. However, this effect is reversed for state legislative elections, where nationalization is lower in contested races. The most likely explanation for the difference is that for uncontested state legislative races, the sole candidate running tends to be from the same party as the dominant party of the district, which means vote choice will much more likely line up at the presidential and state legislative level.

Table 7: **Nationalization of Local Elections, by election timing, 2000-2020**

Year	Clusters	Mayor <i>Non-concurrent</i>	Mayor <i>Concurrent</i>	Mayor <i>Same-Year</i>	Mayor <i>Off-Year</i>	State Leg <i>Same-Year</i>	State Leg <i>Off-Year</i>
2000	2	0.101	0.275	0.251	0.097		
2004	2	0.133	0.104	0.226	0.132		
2008	2	0.100	-0.018	0.275	0.102		
2012	2	0.100	0.029	0.273	0.105	0.075	0.100
2016	2	0.138	0.080	0.181	0.155	0.104	0.095
2020	2	0.065	0.063	0.178	0.060	0.122	0.132
2000	3	0.078	0.279	0.391	0.092		
2004	3	0.125	0.081	0.291	0.141		
2008	3	0.057	0.122	0.160	0.077		
2012	3	0.090	0.171	0.263	0.138	0.105	0.091
2016	3	0.105	0.068	0.274	0.116	0.067	0.059
2020	3	0.086	-0.042	0.232	0.070	0.113	0.103

Measure of nationalization in local elections, as defined in (3), where Clusters is the number of clusters  $K$  used in Algorithm 1. Concurrent is an election that was held at the same time as a presidential election and Same-Year is an election that was held in the same year as a presidential election.



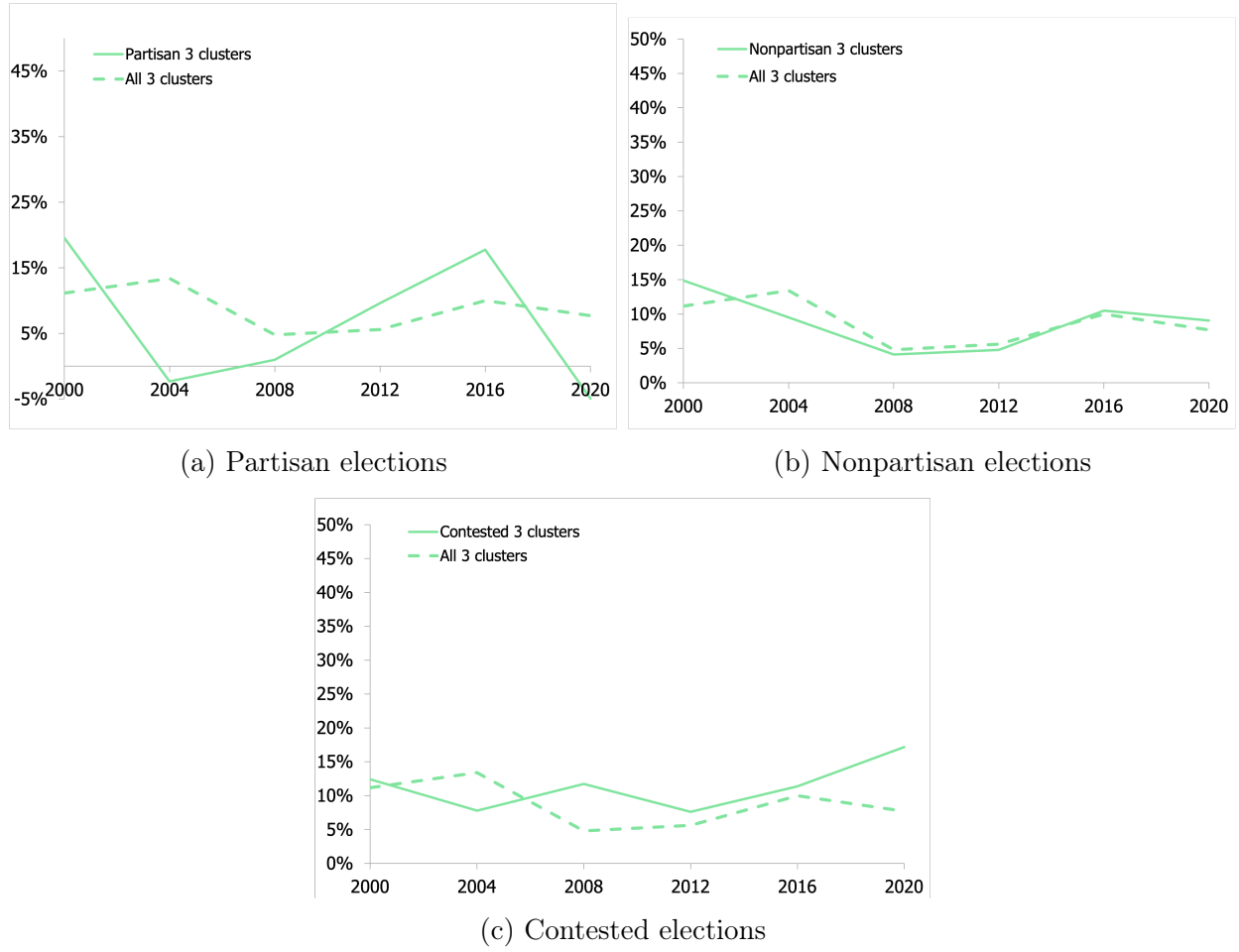


Figure 2: Nationalization of Mayoral Elections, by type of race,  $K = 3$ .

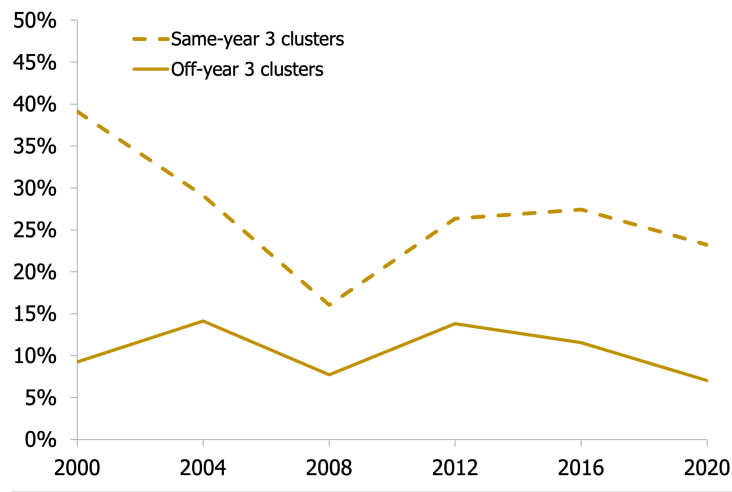


Figure 3: Mayoral elections that happen in the same year as a presidential election vs those that happen in an off-year,  $K = 3$

Another distinguishing factor between local and non-local elections is that many local elections are non-concurrent, meaning that they take place on a different day as presidential elections. One would expect to see more nationalization for concurrent elections as voters are more directly and immediately influenced by their presidential vote choice as opposed to non-concurrent elections. Looking at Table 6, there are no evident patterns of nationalization between the concurrent and non-concurrent (possibly due to the very small sample size of concurrent elections) but the gap between same year and off-year elections is particularly striking, which is plotted in Figure 3. Nationalization of races in the same year, which account for approximately 15% of the whole sample, are consistently at least 9% higher than those in an off-year, with a particularly high level at 39.1% in 2000. This difference is not observed at the other elections, except a slightly higher nationalization of same-year state legislative elections ( $K = 3$ ). Contrasted with the results for partisan and nonpartisan races, these findings point to election timing as a more significant driver of lower local nationalization than the dominance of nonpartisan races. The similarly lower nationalization of state legislative elections also alludes to the less consequential role of partisan elections.

# Distributions of election returns

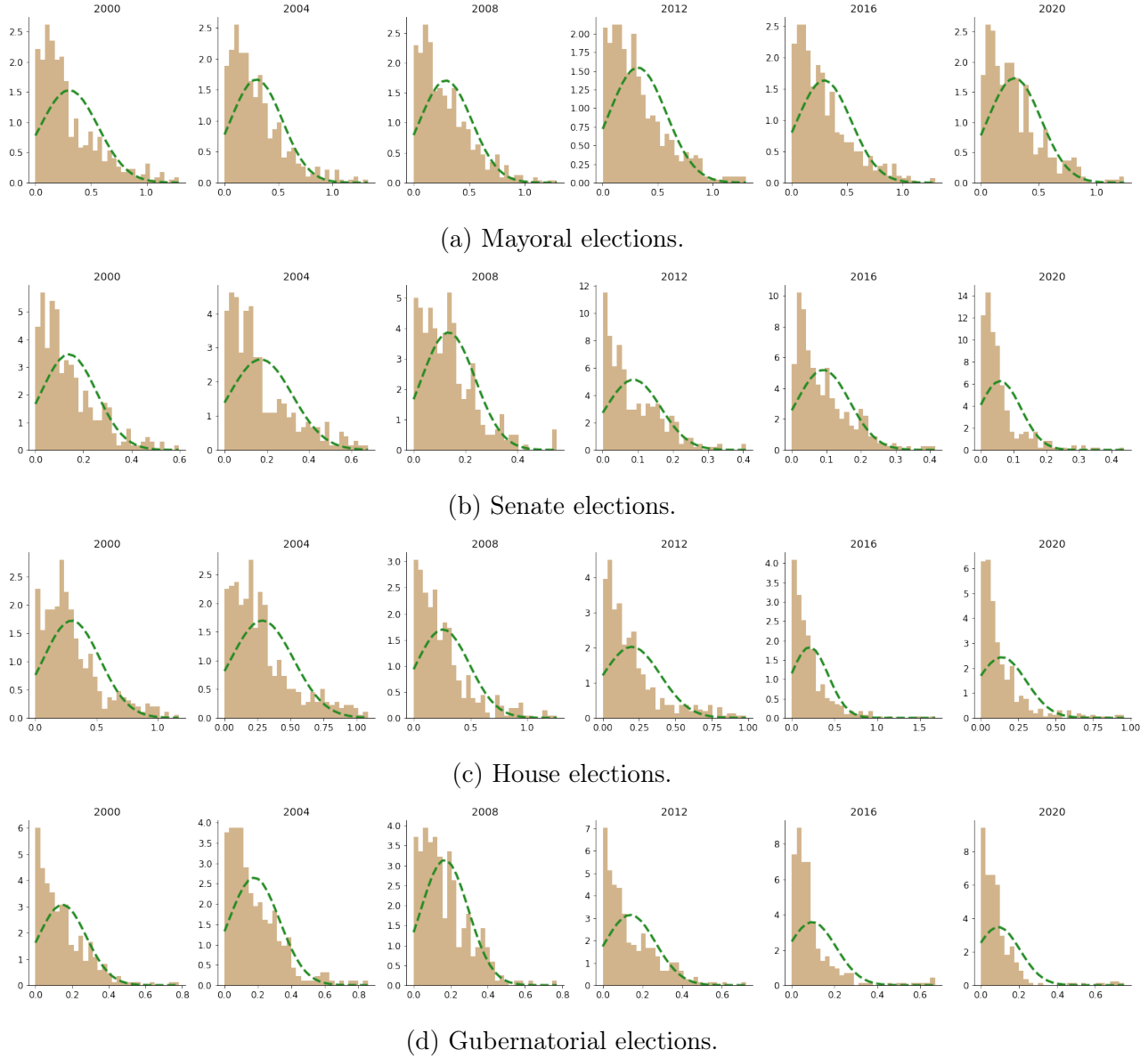


Figure 4: Probability density histograms of the absolute difference between Democratic margin of victory of presidential election and various subpresidential elections, 2000-2020.

To serve as a complementary approach, I directly examine the distributions of the difference in margin of victories to bolster the results obtained using my measure. With the definition of electoral nationalization in mind, if the difference in presidential and subpres-

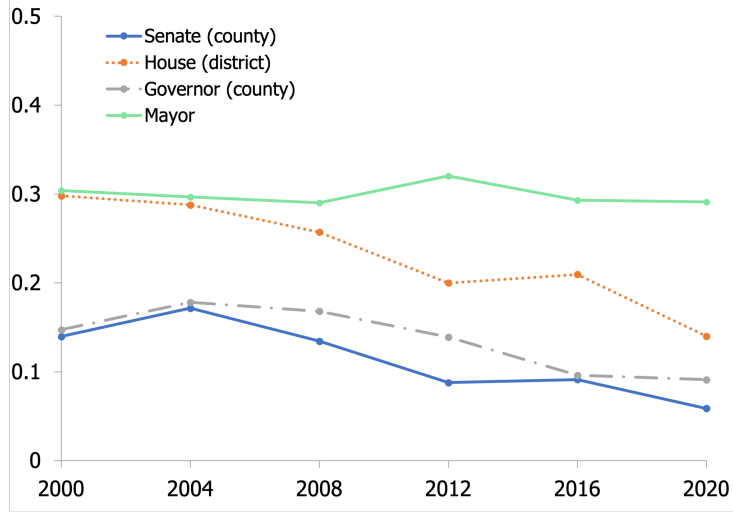


Figure 5: Mean absolute difference between Democratic margin of victory of presidential election and various subpresidential elections, 2000-2020.

idential margin of victories appears to be decreasing in time, then that would suggest a general increase in nationalization. For each pair of presidential and subpresidential elections, I calculate the absolute difference in Democratic margin of victories for each unit. The probability density histograms of the six presidential elections between 2000 and 2020 are presented in Figure 4 and the descriptive features of these distributions are presented in Table 8. Clearly, the same general trends of electoral nationalization prevail. Figure 4a indicates relatively little change in the distribution of the divergence of margin of victories in mayoral elections, juxtaposed against Figure 4b of Senate elections, where the differences become smaller on average and the upper percentiles decrease over time. The stability of the distributions for mayoral elections over time is visually unmistakable here too. For instance, the mean, median, standard deviation, and skew for local elections stay relatively stable over time. On the other hand, there is a sharp drop in the first three metrics and a drastic increase in skew for the non-local elections, becoming more positively skewed. For more recent elections, 2008 and later, Senate, House, and gubernatorial elections all have distributions with larger means and medians, distinctly smaller than that of mayoral elections. The upper range of the distributions decreases over time, implying a greater convergence in the voting

Table 8: **Distribution of the absolute difference between Democratic margin of victory of presidential election and various subpresidential elections, 2000-2020.**

Election	Unit	Year	$\mu$	$\mu_{1/2}$	$P_{75}$	$P_{90}$	$\sigma$	Skew
Mayoral	Municipality	2000	0.304	0.228	0.435	0.677	0.261	1.333
Mayoral	Municipality	2004	0.297	0.231	0.417	0.625	0.240	1.371
Mayoral	Municipality	2008	0.290	0.240	0.420	0.608	0.234	1.184
Mayoral	Municipality	2012	0.320	0.263	0.451	0.698	0.258	1.199
Mayoral	Municipality	2016	0.293	0.233	0.416	0.642	0.244	1.221
Mayoral	Municipality	2020	0.291	0.241	0.399	0.610	0.231	1.221
Senate	County	2000	0.140	0.104	0.204	0.302	0.115	1.230
Senate	County	2004	0.172	0.126	0.250	0.401	0.150	1.206
Senate	County	2008	0.135	0.121	0.185	0.262	0.103	1.264
Senate	County	2012	0.088	0.061	0.140	0.196	0.078	1.160
Senate	County	2016	0.091	0.069	0.130	0.205	0.077	1.282
Senate	County	2020	0.059	0.039	0.068	0.142	0.064	2.461
House	CD	2000	0.298	0.241	0.396	0.664	0.232	1.227
House	CD	2004	0.288	0.227	0.385	0.666	0.235	1.240
House	CD	2008	0.257	0.199	0.334	0.558	0.235	1.612
House	CD	2012	0.200	0.139	0.258	0.476	0.197	1.662
House	CD	2016	0.209	0.143	0.265	0.489	0.218	2.280
House	CD	2020	0.140	0.084	0.182	0.307	0.164	2.440
Governor	County	2000	0.147	0.117	0.218	0.322	0.130	1.443
Governor	County	2004	0.178	0.139	0.262	0.369	0.151	1.462
Governor	County	2008	0.168	0.137	0.233	0.360	0.127	1.092
Governor	County	2012	0.139	0.095	0.208	0.332	0.127	1.295
Governor	County	2016	0.096	0.064	0.111	0.204	0.111	3.001
Governor	County	2020	0.091	0.061	0.111	0.176	0.115	3.560

CD stands for congressional district.

patterns, with the exception of House elections.

In 2000 and 2004, House elections appear to be less nationalized, resembling mayoral elections more than the other elections. But, in 2004, for example, the means and medians for House and mayoral are quite similar yet the distribution for House elections have smaller standard deviations, pointing to greater variability and range. Since my measure still documents a higher level of nationalization for House elections, the clustering algorithm is picking up on some nuances in the differences between the two distributions that distinguish the two

elections in voting behavior. The behavior of the distribution in the upper percentiles implies that House elections still possess congressional districts where the convergence in voting outcomes is greater.

From Figure 5, the plot of the mean absolute differences in margin of victories, we see that the average difference does not drop for mayoral elections, unlike the rest, which emphasizes a non-increasing level of nationalization. But nevertheless, the drop for congressional and gubernatorial elections is not as precipitous and looking solely at the mean neglects other aspects of the distribution of differences. These results bolster the main conclusions obtained by the measure  $\gamma_t$  but also underscore the importance of having a single measure that can encapsulate and represent the trends we see in Figure 4 and Table 8. Instead of manually and arbitrarily determining cutoffs or thresholds for comparison, k-means can make smart decisions that summarize the data in an accurate way.

## Conclusion

Elections in the U.S. have long been dictated by a duopoly of two national political parties that have become increasingly centralized, homogeneous and powerful. The nationalization of the American electorate, along with polarization and partisanship, can and has been reshaping American politics in ways that could reduce accountability, diversity and individualism. However, to gain a better and deeper understanding of this phenomenon, a more appropriate measurement must be developed and applied to subnational elections (federal and or elected at the state-level) but also to local elections. With this quantifiable standard, we can take a pulse on how electoral nationalization has changed over the years and to what extent voting behavior has uniformly converged across geographical units.

I put forth an adaptable measure of electoral nationalization using a k-means clustering algorithm that is appropriate for a variety of subpresidential elections, including local elec-

tions. The results using this measure on congressional and gubernatorial races are congruent with extant literature that tracks the increase in nationalization of U.S. elections starting from the 1980s. In the 2020 general election, there was a notable surge in nationalization to an all-time-high, with Senate elections exhibiting the highest degree of nationalization. However, I do not observe the same upward dynamic in local elections, specifically mayoral and state legislative elections, where nationalization has increased much less in the last two decades and remains at a significantly lower level than that of subnational elections. This may run counter to existing notions that national issues and political organizations have come to dominate even parts of local politics. These findings suggest that voting at the local level is much less impacted by national voting behavior and there are still prevailing idiosyncrasies in voting behavior that cannot be explained by national winds. In addition, electoral timing appears to be a bigger determining factor than partisanship in this lower nationalization of local elections since I do not find any obvious heterogeneity between partisan and nonpartisan races.

A main weakness of existing conclusions regarding a distinct rise in nationalization is that they do not quantitatively take into account local politics. Without the inclusion of local elections, conclusions about the effect and magnitude of nationalization cannot be completely credible or comprehensive. Nationalization is fundamentally a story about the convergence of political behavior throughout state and local units and therefore to leave out local elections is to ignore a crucial part of the definition. I provide an extensive look into nationalization at the subnational and local level via a methodology that is data-driven and easy to implement. A shortcoming to note is that the local election results in this paper is limited as the mayoral election returns only span the last twenty years and the results for state legislative elections only reach back to the last ten years or so. To add to these results, I will explore the possibility of incorporating older election returns and other types of local elections to bolster these results. But the dataset constructed in this study already offers a

fertile resource for more quantitative work to be done in local politics research.

In the future, this measure of nationalization can be applied to political systems in other countries that do not consist of a strict two-party system. It can also be used to shed more light on the relationship between nationalization and polarization or incumbency advantage. Although there has already been headway made in this area of research, there still remains many answered questions about the causes and mechanisms of electoral nationalization and nationalization as a whole.



## References

- Abramowitz, Alan, and Steven Webster. 2016. "The rise of negative partisanship and the nationalization of U.S. elections in the 21st century." *Electoral Studies* 41: 12–22.
- Alexander Hamilton, James Madison, and John Jay. 1788. *The Federalist Papers*. New York, NY: Signet Classics.
- Amlani, Sharif, and Carlos Algara. 2021. "Partisanship & nationalization in American elections: Evidence from presidential, senatorial, & gubernatorial elections in the U.S. counties, 1872–2020." *Electoral Studies* 73.
- Angelucci, Charles, Julia Cage, and Michael Sinkinson. 2023+. "Media Competition and News Diets." *American Economic Journal: Microeconomics* forthcoming.
- Anzia, Sarah F. 2021. "Party and Ideology in American Local Government: An Appraisal." *American Review of Political Science* 24: 133–150.
- Bartels, Larry. 1998. "Panel Attrition and Panel Conditioning in American National Election Studies." *Paper presented at the 1998 meetings of the Society for Political Methodology, San Diego*: Forthcoming.
- Bartels, Larry. 2000. "Partisanship and Voting Behavior, 1952-1996." *American Journal of Political Science* 44: 35–50.
- Boudreau, Cheryl, Christopher S. Elmendorf, and Scott A. MacKenzie. 2015. "Lost in Space? Information Shortcuts, Spatial Voting, and Local Government Representation." *Political Research Quarterly* 68: 843–855.
- Campbell, Angus, Philip E. Converse, Warren E. Miller, and Donald E. Stokes. 1980. *The American Voter*. Chicago, IL: University of Chicago Press.

- Carsey, Thomas M., and Gerald C. Wright. 1998. "State and National Factors in Gubernatorial and Senatorial Elections." *American Journal of Political Science* 42(3): 994–1002.
- Carson, Jamie L., Joel Sievert, and Ryan D. Williamson. 2020. "Nationalization and the Incumbency Advantage." *Political Research Quarterly* 73: 156–168.
- Ceaser, James W., and Robert P. Saldin. 2005. "A New Measure of Party Strength." *Political Research Quarterly* 58: 245–256.
- Coppock, Alexander, and Patricia Kirkland. 2018. "Candidate Choice without Party Labels: New Insights from Conjoint Survey Experiments." *Political Behavior* 40: 571–591.
- Das, Sanmay, Betsy Sinclair, Steven W. Webster, and Hao Yan. 2022. "All (Mayoral) Politics Is Local?" *The Journal of Politics* 84(2).
- de Benedictis-Kessner, Justin. 2017. "Off-Cycle and Out of Office: Election Timing and the Incumbency Advantage." *The Journal of Politics* 80.
- Einstein, Katherine Levine, and David M. Glick. 2018. "Mayors, Partisanship, and Redistribution: Evidence Directly from U.S. Mayors." *Urban Affairs Review* 54: 74–106.
- Ferejohn, John. 1977. "On the Decline of Competition in Congressional Elections." *American Political Science Review* 71: 177–176.
- Gelman, Andrew, and Gary King. 1990. "Estimating Incumbency Advantage Without Bias." *American Journal of Political Science* 34: 1142–1164.
- Gerber, Elisabeth R., and Daniel J. Hopkins. 2011. "When mayors matter: estimating the impact of mayoral partisanship on city policy." *American Journal of Political Science* 55: 326–339.

- Green, Donald, Bradley Palmquist, and Eric Schickler. 2002. *Partisan Hearts and Minds: Political Parties and the Social Identities of Voters*. New Haven, CT: Yale University Press.
- Grumbach, Jacob M. 2022. *Laboratories against Democracy: How National Parties Transformed State Politics*. Princeton, NJ: Princeton University Press.
- Hajnal, Zoltan, and Jessica Trounstine. 2014. “What Underlies Urban Politics? Race, Class, Ideology, Partisanship, and the Urban Vote.” *Urban Affairs Review* 50: 63–99.
- Hajnal, Zoltan L. 2009. *America’s Uneven Democracy: Race, Turnout, and Representation in City Politics*. Cambridge, UK: Cambridge University Press.
- Hayes, Danny, and Jennifer L. Lawless. 2018. “The Decline of Local News and Its Effects: New Evidence from Longitudinal Data.” *The Journal of Politics* 80(1).
- Hertel-Fernandez, Alexander. 2019. *State Capture: How Conservative Activists, Big Businesses, and Wealthy Donors Reshaped the American States – and the Nation*. Oxford, UK: Oxford University Press.
- Holliday, Derek. 2023. “Is All Politics Presidential? Decomposing partisan patterns in U.S. election outcomes across offices, 1972-2020.” : Forthcoming.
- Hopkins, Daniel J. 2018. *The Increasingly United States*. Chicago, IL: University of Chicago Press.
- Huckshorn, Robert J., James L. Gibson, Cornelius P. Cotter, and John F. Bibby. 1986. “Party Integration and Party Organizational Strength.” *The Journal of Politics* 48: 976–991.
- Jacobson, Gary C. 2015. “It’s Nothing Personal: The Decline of the Incumbency Advantage in US House Elections.” *The Journal of Politics* 77.

- Jacobson, Gary C., and Jamie L. Carson. 2016. *The Politics of Congressional Elections*. Lanham, MD: Rowman & Littlefield Publishers.
- Kawato, Sadafumi. 1987. “Nationalization and Partisan Realignment in Congressional Elections.” *American Political Science Review* 81(4): 1235–1250.
- Kritzer, Herbert M., and Robert B. Eubank. 1979. “Presidential Coattails Revisited: Partisanship and Incumbency Effects.” *American Journal of Political Science* 23: 615–626.
- Kuriwaki, Shiro. 2023. “Ticket Splitting in a Nationalized Era.” : Forthcoming.
- Martin, Gregory J., and Joshua McCrain. 2019. “Local News and National Politics.” *American Political Science Review* 113(2).
- Moskowitz, Daniel J. 2021. “Local News, Information, and the Nationalization of U.S. Elections.” *American Political Science Review* 115: 114–129.
- Rogers, Steven. 2016. “National Forces in State Legislative Elections.” *The Annals of the American Academy of Political and Social Science* 667: 207–225.
- Rogers, Steven. 2023a. *Accountability in State Legislatures*. Chicago, IL: University of Chicago Press.
- Rogers, Steven. 2023b. “The Vanishing Incumbency Advantage in State House Elections.” *The Forum* 21(1): 97.
- Rousseeuw, Peter J. 1987. “Silhouettes: a graphical aid to the interpretation and validation of cluster analysis.” *Journal of Computational and Applied Mathematics* 20: 53–65.
- Schaffner, Brian F., Matthew Streb, and Gerald Wright. 2001. “Teams Without Uniforms: The Nonpartisan Ballot in State and Local Elections.” *Political Research Quarterly* 54: 7–30.

- Sievert, Joel, and Seth C. McKee. 2019. "Nationalization in U.S. Senate and Gubernatorial Elections." *American Politics Research* 47: 1055–1080.
- Simon, Dennis M. 1989. "Presidents, Governors, and Electoral Accountability." *The Journal of Politics* 51(2): 286–304.
- Stokes, D.E. 1967. "Parties and the nationalization of electoral forces." *The American party systems: stages of political development*: Forthcoming.
- Sundquist, James L. 1973. *Dynamics of the Party System: Alignment and Realignment of Political Parties in the United State*. Washington, DC: Brookings Institution Press.
- Warshaw, Christopher. 2019. "Local Elections and Representation in the United States." *Annual Review of Political Science* 22: 461–479.
- Warshaw, Christopher, and Chris Tausanovitch. 2013. "Measuring constituent policy preferences in congress, state legislatures, and cities." *The Journal of Politics* 75(2): 330–342.
- Warshaw, Christopher, Justin de Benedictis-Kessner, and Yamil Velez. 2022. "Local Representation in the United States: A New Comprehensive Dataset of Elections." *HKS Faculty Research Working Paper Series RWP22-013*: Forthcoming.
- William Claggett, William Flanigan, and Nancy Zingale. 1984. "Nationalization of the American Electorate." *American Political Science Review* 78(1): 77–91.

# A Presidential Election Returns by Municipality

To put together this dataset, the first step of the process is collecting precinct-level presidential returns for all states from the MIT Election Data Science Lab for 2016 and 2020, Harvard Election Data Archive and Open Elections for 2004 to 2012, and the Federal Elections Project for 2000. I also use individual state and county websites to fill in any gaps when needed, which only leaves missing data for a handful of states in a few elections prior to 2012. To match each voting precinct to the city or town it belongs to, I use TIGER data from the United States Census Bureau that contains geographic entity codes for precincts in every state, collected every ten years. These TIGER files are processed using geographic information systems (GIS) software and require standardizing precinct names to optimize matching to the precinct names in the precinct-level presidential election results. Once each precinct is matched to a specific longitude and latitude, I write an algorithm, to apply to each state, that determines the closest geographical municipality for each precinct subject to a maximum distant constraint based on longitude and latitude coordinates <sup>6</sup>. If the precinct name contains the name of the city or town, then it gets automatically allocated to that city or town.

Table A.1 compares my dataset with Warshaw and Tausanovitch (2013).

Table A.1: **Dataset of presidential votes at municipal level**

	This Paper	Warshaw and Tausanovitch (2013)
Years	2000-2020	2008-2020
No of municipalities in 2020	14,757	6,271
No of municipalities in 2016	15,530	6,336
No of municipalities in 2012	16,678	6,188
No of municipalities in 2008	16,632	6,236

Comparison of the presidential election returns by municipality dataset I compile (This Paper) and the dataset compiled by Warshaw and Tausanovitch (2013).

---

<sup>6</sup>California is the only state that makes available comprehensive municipal-level presidential election results, which I use in lieu.

## B Supplementary figures of electoral nationalization

To verify the trends presented for  $K = 2$ , Figure B.1 plots the nationalization of subpresidential elections for  $K = 3$ . It demonstrates a similar upward trajectory for House, Senate, and governor but a lower, more stable level of nationalization for local elections. This serves as a robustness check, bolstering the main results of this paper.

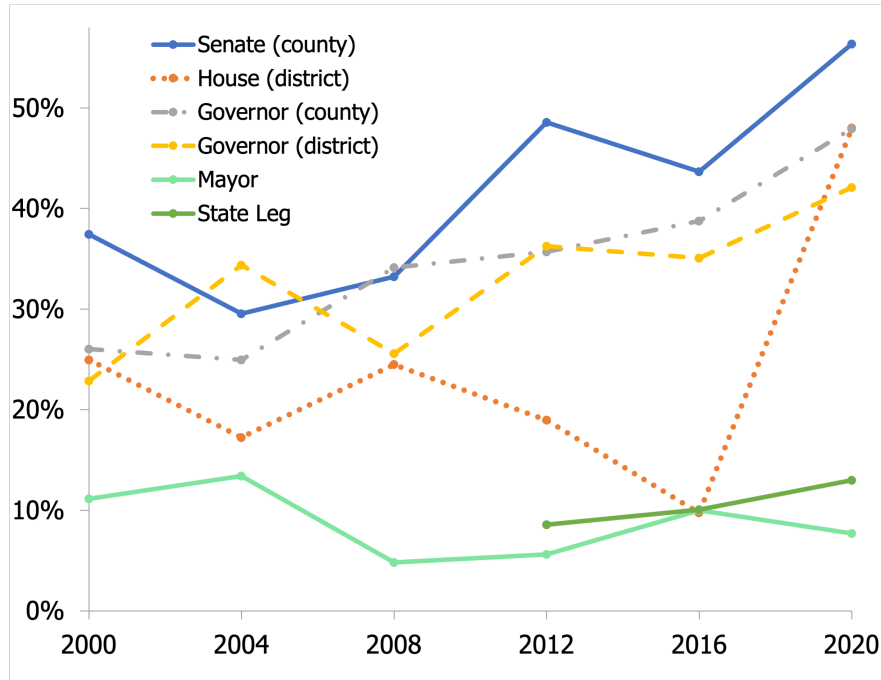
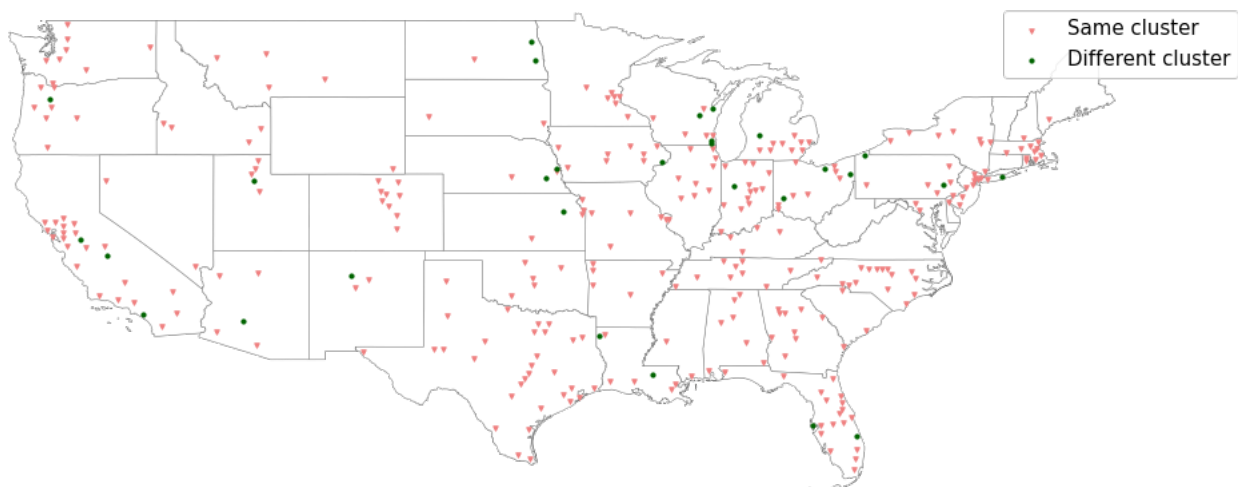
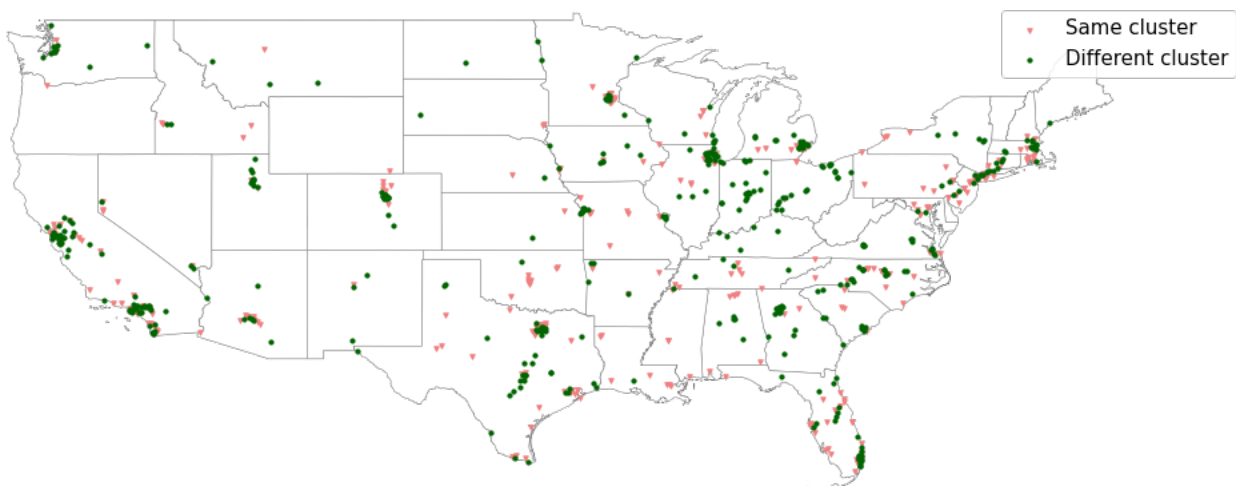


Figure B.1: Nationalization of Subpresidential Elections,  $K = 3$ .

For a geographical representation, Figures B.2a and B.2b visualize the counties and municipalities, respectively, that are assigned to the same voting cluster in the presidential election and the Senate and mayoral elections, respectively. Figure B.2a evidently has very few green counties that belonged to different clusters while Figure B.2b contains green municipalities in most states.



(a) Counties partitioned to the same/different voting clusters at the 2020 presidential and Senate elections.



(b) Municipalities partitioned to the same/different voting clusters at the 2020 presidential and mayoral elections.

Figure B.2: Units in the 2020 election that belong to the same or different voting cluster. The share of the same clusters out of all the units represents  $\gamma_t$ . Geographical coordinates based on centroids of the units.

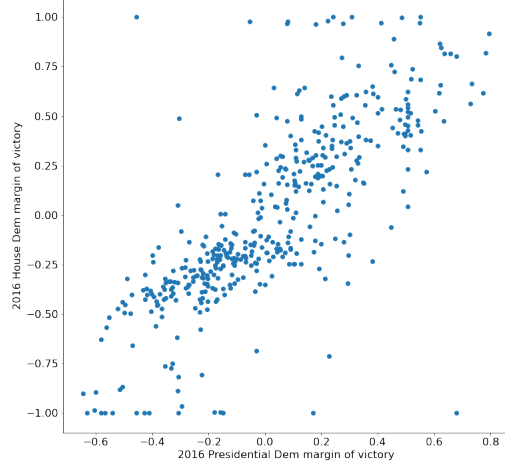


## C The limitations of measuring nationalization with linear models

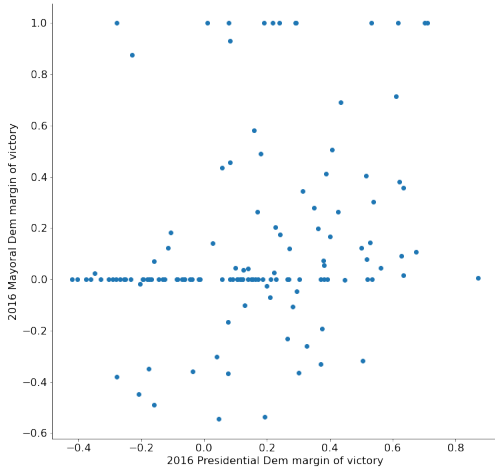
Many existing measures of nationalization rely on ordinary least squares, linear regression, or some version of a linear regression model. As powerful as linear models can be in ascribing the effect of predictors on response variables, it falls victim to problems of multicollinearity, violations of normality assumptions, and sensitivity to outliers. When some linear models become over-specified or contain many variables, this inflates the standard errors of the coefficients, rendering the estimates unstable and difficult to interpret. Arguably more importantly, linear models can only capture linear relationships but the relationship between the presidential vote and the subpresidential vote is not always purely linear.

Figure C.1 plots the presidential vote and the subpresidential vote for the 2016 presidential election as an example. The units (congressional districts) in Figure C.1a do resemble a roughly positively linear relationship for House elections. However, for mayoral elections on the bottom left, Figure C.1b, and state senate elections on the bottom right, Figure C.1c, there is much more complexity. Even if the uncontested races are removed, an ordinary least squares would not be an appropriate choice to model the data in this case. A linear correlation may be enough to explain congressional elections, but a nonlinear model is necessary when measuring nationalization at the local level.

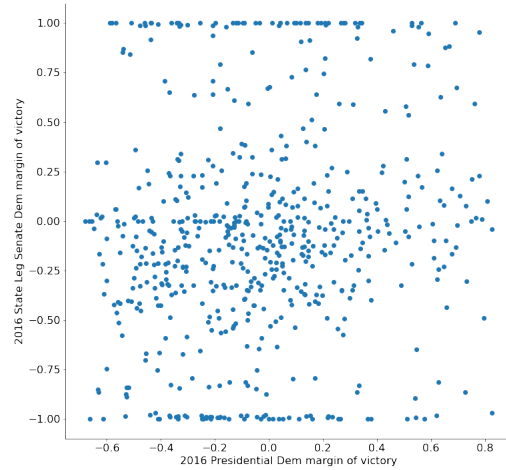
In order to illustrate the inadequacies and complexities in interpretation if an ordinary least squares is applied to the election returns, I replicate the status quo electoral nationalization measure on House, Senate and gubernatorial elections: a regression model following works like Sievert and McKee (2019) where the dependent variable is the Democratic margin of victory and the independent variables are the Democratic presidential margin of victory, incumbency status (1 for Democratic incumbent, -1 for Republican incumbent, and 0 for open seat or otherwise), and incumbent's party (1 if Democrats won the previous election, -1



(a) Presidential Democratic margin of victory vs House Democratic margin of victory at the congressional district level.



(b) Presidential Democratic margin of victory vs Mayoral Democratic margin of victory at the municipal level.



(c) Presidential Democratic margin of victory vs State Senate Democratic margin of victory at the municipal level.

Figure C.1: 2016 presidential election

if Republicans won the previous election). Table C.1 reports results for House, Senate and gubernatorial elections but also for mayoral and state legislative elections using the data I collect. The estimated coefficients are a lot more difficult to interpret when they are not sig-

Table C.1: **Nationalization using regression coefficients**

Year	House	Senate	Gubernatorial	Mayoral	State Leg
1976	0.729* (0.082)	0.504* (0.021)			
1980	0.524* (0.076)	0.354* (0.013)	0.254* (0.029)		
1984	0.341* (0.082)	0.456* (0.024)	0.213* (0.047)		
1988	0.485* (0.077)	0.760* (0.010)	0.237* (0.032)		
1992	0.498* (0.064)	0.660* (0.021)	0.561* (0.030)		
1996	0.559* (0.064)	0.769* (0.017)	0.786* (0.040)		
2000	0.553* (0.056)	0.796* (0.012)	0.457* (0.029)	0.161* (0.177)	
2004	0.545* (0.059)	0.851* (0.018)	0.361* (0.021)	-0.041 (0.131)	
2008	0.493* (0.051)	0.849* (0.012)	0.432* (0.041)	0.098 (0.119)	
2012	0.571* (0.048)	0.809* (0.009)	0.581* (0.026)	-0.034 (0.114)	0.071* (0.018)
2016	0.538* (0.047)	0.769* (0.011)	0.580* (0.028)	-0.071 (0.114)	0.010 (0.017)
2020	0.812* (0.042)	0.920* (0.008)	0.777* (0.027)	0.084 (0.119)	0.098* (0.019)

Measure of nationalization of subpresidential elections using regression coefficient. The dependent variable is the subpresidential election Democratic margin of victory. The main independent variable is the presidential election Democratic margin of victory. The other two predictors are the incumbency status and incumbent's party status. \* denotes statistical significance at the 1% level. Standard errors are reported in parentheses.

nificant and the comparison of nationalization between elections and election is less obvious. Although we see increasing trends in congressional election nationalization and a discrepancy between local and non-local elections, the significance of these regression coefficients and confidence intervals vary greatly, namely for mayoral and state legislative elections. For example, since none of the coefficients for mayoral elections is significant, the only conclu-

sion one can make is that there is no evident linear correlation. However, little can be said about how much less nationalized these elections are compared to the rest and compared between election years. Indeed, the coefficients are lower than that of House, Senate and gubernatorial races, but the difference is immeasurable. Additionally, confidence intervals overlap, making it difficult to ascertain the precise change in nationalization from election to election, whereas this problem does not apply to the methodology outlined in this paper. Much like evidence from Sievert and McKee (2019), coefficient estimates with overlapping confidence intervals cannot be statistically distinguished from one another with confidence. In addition, among the more national elections, there is no clear increase in nationalization until 2016 to 2020, making the argument of a consistent rise in electoral nationalization over several decades less evident. Altering the covariates used in the regression would also change the estimated coefficients, making the interpretation subject to the specificity of the model.

In addition, I perform a Kolmogorov-Smirnov test to investigate whether the presidential and subpresidential vote share distributions are statistically different from one another. However, we see from Table C.2 that the interpretation poses problems once again. The distributions of mayoral elections are significantly different from presidential elections at the municipality level, which is in line with what I find, but we do not see an indicator of increasing convergence for House elections, contrary to what the existing literature demonstrates. Furthermore, the results in Senate only reveal evidence of converging voting distributions in the three most recent elections but the rest seem to imply some sort of divergence. There are unmistakable difficulties in determining and comparing nationalization across types of elections and election years, where no clear trends or single measure can be extracted. A test like this only depicts a single facet of the convergence of voting behavior but overlooks the details, highlighting the necessity and practicality of a measure such as the one I propose.

Table C.2: **Kolmogorov-Smirnov test on Democratic margin of victory of presidential and subpresidential elections, 2000-2020.**

Election	Unit	year	KS stat	p-value
Mayor	Municipality	2020	0.219	0.000
Mayor	Municipality	2016	0.189	0.000
Mayor	Municipality	2012	0.175	0.000
Mayor	Municipality	2008	0.189	0.000
Mayor	Municipality	2004	0.221	0.000
Mayor	Municipality	2000	0.244	0.000
Senate	County	2020	0.074	0.340
Senate	County	2016	0.034	0.993
Senate	County	2012	0.080	0.251
Senate	County	2008	0.172	0.000
Senate	County	2004	0.178	0.000
Senate	County	2000	0.104	0.058
House	CD	2020	0.087	0.047
House	CD	2016	0.138	0.000
House	CD	2012	0.140	0.000
House	CD	2008	0.190	0.000
House	CD	2004	0.215	0.000
House	CD	2000	0.211	0.000

The Kolmogorov-Smirnov (KS) statistic and p-values for presidential and various subpresidential Democratic margin of victories. Reject null hypothesis that distributions are the same if p-value < 0.05. CD stands for congressional district.

## D Choice of $K$

Although, as reiterated, the choice of  $K$  is not an important concern of this paper, I calculate the silhouette score to justify the use of  $K = 2$  and 3 for presenting the main results for electoral nationalization, but the same trends are found with larger  $K$ . There are many methods in determining  $K$  but as I only cluster using one variable, the silhouette score first developed by Rousseeuw (1987) is a sensible choice. Unlike some other scores to evaluate the goodness of clusters, the silhouette score addresses both cohesion (how close units are to their respective centroids) and separation (how far the clusters are from one another).

It is also not sensitive to the scale of the variables and there is no concern about curse of dimensionality since it is just the case of a one-dimensional clustering.

After obtaining the  $K$  clusters using Algorithm 1, for every unit  $i$  in cluster  $C^k$ , let

$$a(i) = \frac{1}{|C^k| - 1} \sum_{j \in C^k, i \neq j} d(i, j) \quad (4)$$

be the average dissimilarity of unit  $i$  to all other units in  $C^k$ , where  $|C^k|$  is the number of units in  $C^k$  and  $d(i, j)$  is the Euclidean distance between units  $i$  and  $j$ .

Next, the smallest mean dissimilarity of unit  $i$  to all points in all clusters except  $C^k$  is denoted by:

$$b(i) = \min_{k \neq l} \frac{1}{|C^l|} \sum_{j \in C^l} d(i, j) \quad (5)$$

Now, the silhouette score is defined as:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))}, \quad (6)$$

if  $|C^k| > 1$  and  $s(i) = 0$ , otherwise. From Equation 6, we can see that  $-1 \leq s(i) \leq 1$ , where 1 indicates well-clustered data and  $-1$  suggests inappropriately-clustered data. The convention is that an average  $s(i)$  above 0.5 is considered strong clustering and anything below 0.25 is considered weak clustering.

Figure D.1 reports the average silhouette scores, according to Equation 6, for the relevant subpresidential elections. Most scores are above 0.5, which indicates strong clusters for the results used in this paper. For most elections, the average score decreases as  $K$  increases for  $K > 2$ . Although  $K = 4$  could be an appropriate choice for some elections, the clusters would be harder to interpret and the same  $K$  needs to be used to compare nationalization  $\gamma_t$  across all elections. In addition, the almost negligible increase in  $s(i)$  from  $K = 3$  to 4 or 5 is not high enough to justify the choice. Therefore,  $K = 2$  would suffice, as well as  $K = 3$

for a robustness check.

Table D.1: **Average Silhouette Scores for subpresidential elections.**

Election	Unit	$K$	Avg SS
Mayor	Municipality	2	0.718
Mayor	Municipality	3	0.651
Mayor	Municipality	4	0.697
Mayor	Municipality	5	0.698
Senate	County	2	0.569
Senate	County	3	0.548
Senate	County	4	0.548
Senate	County	5	0.547
House	CD	2	0.601
House	CD	3	0.590
House	CD	4	0.600
House	CD	5	0.593
Governor	County	2	0.560
Governor	County	3	0.548
Governor	County	4	0.547
Governor	County	5	0.547
State Leg	State Leg District	2	0.553
State Leg	State Leg District	3	0.651
State Leg	State Leg District	4	0.632
State Leg	State Leg District	5	0.637

Avg SS is the average silhouette score across all available elections. CD is congressional district.

## E Nationalization and Incumbency Advantage

An application of this measure of electoral nationalization is to the study of the relationship between nationalization and incumbency advantage. Scholars have reported a decline in incumbency advantage in various U.S. elections over the last few elections. Generally, incumbency advantage is linked to reduced levels of partisanship and a separation of subpresidential elections from presidential elections, resulting in elections more centered on candidates and incumbents with the upper hand (Ferejohn 1977; Kritzer and Eubank 1979).

The recent patterns of decreasing incumbency advantage paint a picture of party divisions and fewer voters who are inclined to defect from their own party loyalties. Hence, Jacobson (2015) and others have made the direct connection between the growing nationalization of congressional elections and diminishing incumbency advantage.

**Table E.1: Nationalization and Incumbency Advantage, 1976-2020.**

Office	Unit	Clusters	Coef	SE	P-value
House	District	2	-1.074	0.568	0.091
Senate	County	2	-1.010	0.666	0.164
Governor	County	2	-0.214	0.377	0.586

Regression of measure of nationalization on incumbency advantage as in Gelman and King (1990). SE represents standard error values.

To provide additional evidence on the relationship between nationalization and incumbency advantage, I calculate the correlation between  $\gamma_t$  and incumbency advantage using Gelman and King (1990)'s incumbency advantage model. The results are reported in Table E.1, where the coefficient between the two variables is negative for all three elections. The lowest and most significant negative correlation is found in the House and these results are overall consistent with works such as Jacobson (2015); Carson, Sievert, and Williamson (2020) who find that incumbency advantage is higher in times of low nationalization and vice versa. Indeed, this validation strengthens the observations made by some scholars and demonstrates the potential of  $\gamma_t$  for other applications.