PSRM

Making countries small: The nationalization of districts in the United States

Ignacio Lago [ID]

Universitat Pompeu Fabra, Barcelona, Spain
Corresponding author. Email: ignacio.lago@upf.edu

Abstract

I rely on data from 31,754 electoral districts in the United States from 1834 until 2016 to explore how the nationalization of politics occurs within districts. I argue that in the early stages of the American democracy local concerns were more prominent in the distant districts from the capital city than in the nearby districts, and therefore the number of parties was greater in the former than in the latter. However, these differences vanished after the New Deal, when authority was centralized. Nationalization reduced the number of parties everywhere, but above all in the most distant district from Washington, D.C.

Keywords: Country size; decentralization; district; nationalization; United States

The number of national parties over time in electoral democracies has been shown to be affected by two large-scale phenomena, namely the nationalization of voting behavior since the first competitive elections in the 1850s and the degree of (de)centralization of economic and political power. First, nationalization led to the replacement of the highly localized and territorialized politics in the early phases of electoral competition in the nineteenth century with national electoral alignments (Caramani, 2004, 1) and accordingly to a reduction in the number of national parties. The argument is from voters' decision making to party system effects. Second, as national governments centralize power and make policies that affect local areas, the number of national parties reduces because candidates have greater incentives to associate with national organizations, and voters have greater incentives to abandon locally competitive but nationally noncompetitive parties (Chhibber and Kollman, 2004; see also Kollman and Worthington, 2021). The argument is from (de)centralization to the nationalization of the vote and, in contrast with Caramani's argument, it is not temporally bound.

Although Caramani and Chhibber and Kollman rightly point to the key influence of nationalization and the decentralization of power on the number of national parties, their arguments might be underspecified. Other than stating that nationalization and centralization reduce the number of parties, the conditions under which this reductive effect occurs within countries are not indicated.

I argue that the nationalization of politics within countries is driven by the spatial distribution of a country's population relative to the capital city. The geographical distance from districts to the capital city affects the speed and effect of nationalization. More specifically, districts located closer to the stronghold of government are expected to have fewer parties than distant districts in the early years of democracy. The increasing importance of the national government and the decreasing role of local factors as time goes by will mainly affect parties and voters in the distant districts.

This interest in the role played by geography in explaining electoral politics is far from new. Political engineers in early democracies, in particular the founding fathers of the United States,

were already aware of the link between geographic distances and governance. In *Federalist* paper 14, James Madison argued as follows:

> As the natural limit of a democracy is that distance from the central point which will just permit the most remote citizens to assemble as often as their public functions demand, and will include no greater number than can joint in those functions; so the natural limit of a republic is that distance from the centre which will barely allow the representatives to meet as often as may be necessary for the administration of public affairs. (Hamilton *et al.*, 2009, 68)

The effect of geographic distance to the capital city on nationalization across districts within countries is shaped by the actions of parties, voters, and national governments. First, the creation of national parties entails the attraction of enough local notables to spread party labels across geographical areas (Chhibber and Kollman, 2004, 82–83). The formation of national parties is not a "bottom-up" process, with a trajectory of increasing size and scope of party aggregation, but a "top-down" story. In Canada, India, the United Kingdom, and the United States, for instance, the organizational structures of the first political parties with influence on national government policy were centered in the national capitals (Chhibber and Kollman, 2004, chapter 4). As party linkage was gradually extended from the capital cities to outlying geographical areas, local concerns were expected to be incorporated into parties' platforms first in the nearby districts to capital cities and finally in the most distant districts. In other words, party linkage is not expected to be simultaneous in all districts within countries: the more distant districts are from the capital city, the later the party linkage will be.

Second, in the early stages of democracies in the nineteenth century, with poor transportation and communication systems and limited literacy, the flow of information among stakeholders at the national and local levels will be worse as we move away from the capital city. Voters in distant districts from the stronghold of government will be very poorly informed about the national politics, and therefore they will be more concerned about local issues than voters in the closest districts to the capital city.

Finally, the *homogenization* of politics as a consequence of the general integration of societies (Caramani, 2004) is expected to start in the geographical areas closer to the capital city. The expansion of state administration is gradual, and therefore territorial diversity in the course of the nineteenth and twentieth centuries is progressively reduced. The incentives to coordinate on common party labels (Chhibber and Kollman, 1998, 2004) will be first faced by voters and candidates in the districts closest to the capital city as the national government becomes more important for them than in the furthest areas in the country.

When examining the nationalization of politics within countries over time, two expectations emerge on the basis of the previous discussion. On the one hand, in the early stages of democracy in the nineteenth century, local concerns are expected to be more diverse and prominent as we move away from the capital city. As party system fragmentation goes hand in hand with the number of dimensions that are salient to political actors (Stoll, 2011), the (effective) number of parties at the district level should be greater in the furthest areas from the capital city than in the closest areas. Not surprisingly, more provincialization leads to greater party system fragmentation at the district level (Chhibber and Kollman, 2004, chapter 6). Therefore, *in the early elections, the more distant from the capital city a district is, the greater the number of local parties*. On the other hand, the nationalization of politics means that local problems are replaced with national issues. The expectation is that those districts where local issues were more prominent in the early stages of democracies will be the most affected by the nationalization process. Given that the furthest districts from the capital city are expected to have more local parties, nationalization within districts should be *centripetal*: *the furthest districts from the capital city will be those experiencing a higher reduction in party system fragmentation over time*.

The degree to which party systems are nationalized has been shown to affect the delivery of public goods (Hicken *et al.*, 2016). The implication of my argument is that what matters is not only the degree of nationalization but also how it occurs within districts. My research contributes to the literature on political parties by showing that there is a relationship between the location of capital cities and the geographic expansion of national parties. Nationalization is faster in those districts that are located closer to the capital city. Thus nationalization is *heterogeneous* across districts within countries.

Explaining the nationalization of elections within districts

Sample

To examine how the nationalization of elections occurs within districts, I engaged in a large data collection effort. I built a sample including district-election data for the United States Congress elections from 1834 to 2016.[1] The sample includes 31,574 districts. Four features make the United States the ideal case to show that nationalization within electoral districts is not homogenous: (i) it is the oldest democracy (Boix *et al.*, 2013), and electoral data are available since the 1830s; (ii) it is the fourth largest country in the world (FAO, 2021), and therefore there is a huge variation in the distance from the districts to the capital city; (iii) it has experienced no changes in institutional arrangements affecting the nationalization of electoral politics such as the electoral system; and (iv) in the United States there is a clear pattern of party aggregation explained by the centralization of authority. As Chhibber and Kollman (1998, 2004, chapter 7) show, the major change occurred in the 1920s and 1930s, when the effective number of parties dropped to two and stayed there. This aggregation pattern was driven by the New Deal in the 1930s and the introduction of the national income tax in 1913.[2]

Dependent variable

To examine the number of parties at the district level, I used the conventional effective number of electoral parties (ENP_cst) by Laakso and Taagepera (1979):

$$ENP_{ij} = 1/ \sum_{i=1}^{n} p_{ij}^2$$

where *j* is a given election and *p* is the proportion of popular votes received by party *i*. The source is the *Constituency-Level Elections Archive* (CLEA) (Kollman *et al.*, 2019).

Explanatory variables

The key explanatory variable is (the log of) the *distance* in kilometers (plus 1) in a straight line from the district centroids to Washington, D.C., according to Google Maps. Shapefiles for congressional districts are from Lewis *et al.* (2021), as available at https://cdmaps.polisci.ucla.edu.[3] As the effect of *distance* on *ENP_cst* is expected to change over time, the statistical model will be

---

[1]There are no available data before 1834.

[2]The New Deal enacted by Roosevelt from 1933 to 1939 had a very significant impact on the location of government activity within the federal structure of U.S. governments (Fishback & Wallis, 2013). First, the New Deal strengthened the role of the federal and state governments at the expense of local governments. Second, the national government began to give grants to state and local governments to aid them in dealing with poverty and unemployment. Third, regulatory programs that strongly influenced economic activity were more likely to be nationally administered. Fourth, the spending programs tended to reduce the ability of the national administrator to treat states differently.

[3]Alaska and Hawaii are excluded from the analysis.

estimated in different time periods. I have two expectations. First, the coefficient on *distance* should be positive and statistically significant before the New Deal. Given that distant states were progressively joining the union in the second half of the XIX century, the effect of *distance* is expected to be weaker in the 1850s than immediately before the New Deal due to the higher statistical power in the latter. Second, after the New Deal, the coefficient on *distance* should not be statistically significant: the number of district parties should converge due to the centralization of power.

I control for the state population (in millions)[4] and the number of valid votes (in millions) in every district and election. As the heterogeneity of preferences increases as states or districts are larger (Alesina and Spolaore, 2003, 4), highly populated states and districts are expected to have more parties than states and districts with lower population. Thus the expected effect of the variables is positive. The sources are Manson *et al.* (2020) and Kollman *et al.* (2019), respectively.

Finally, no story about party politics in the United States can ignore regional differences, especially in the South. As Chhibber and Kollman (2004, chapter 7) show, the South was predominantly a one-party region from the 1890s to the 1970s. All the analyses have been conducted including and excluding the South.[5]

The empirical analysis proceeds in two steps. The first step shows the decreasing impact of *distance* to the national capital over time. The specification is as follows:

$$ENPcst_{ij} = \beta_0 i + \beta_1 Distance_{ij} + \beta_2 State\ Population_j + \beta_3 Valid\ Votes_{ij}$$
$$+ \beta_4 State\ Dummies_j + \varepsilon_{ij} \tag{1}$$

However, a not significant $\beta_1 Distance_{ij}$ after the New Deal is compatible with three different patterns: the number of parties decreased both in the nearby and the distant districts, but more in the latter; the number of parties only decreased in the distant districts; and the number of parties increased in the nearby districts, while it did not change in the distant districts. In order to disentangle this crucial question, I have used the current time zones in the U.S. and divided the sample into those districts included in the GMT-5 (i.e., the closest districts to Washington, D.C.), the GMT-6, GMT-7, and the GMT-8 (i.e., the furthest districts from Washington, D.C.) time zones[6] and estimated the following model separately for them:

$$ENPcst_{ij} = \beta_0 i + \beta_1 Election\ Year_{ij} + \beta_2 State\ Population_j + \beta_3 Valid\ Votes_{ij}$$
$$+ \beta_4 State\ Dummies_j + \varepsilon_{ij} \tag{2}$$

The parameter $\beta_1 Election\ Year_{ij}$ captures how the number of parties changes over time in the districts depending on their distance to Washington, D.C.

The descriptive statistics are displayed in Table 1.

---

[4]For each election, the data refer to the first year of the decade.

[5]The "non-South" analyses exclude elections held in Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Texas, and Virginia.

[6]The states included in each time zone are the following: GMT-5, Connecticut, Delaware, District of Columbia, Florida, Georgia, Indiana, Kentucky, Maine, Maryland, Massachusetts, Michigan, New Hampshire, New Jersey, New York, North Carolina, Ohio, Pennsylvania, Rhode Island, South Carolina, Tennessee, Vermont, Virginia, West Virginia; GMT-6, Alabama, Arkansas, Illinois, Iowa, Kansas, Louisiana, Minnesota, Mississippi, Missouri, Nebraska, North Dakota, South Dakota, Oklahoma, Texas, Wisconsin; GMT-7: Arizona, Colorado, Idaho, Montana, New Mexico, Utah, Wyoming; GMT-8: California, Nevada, Oregon, Washington (state).

Table 1. Descriptive statistics[a]

	All				Non-South			
	Min	Max	Mean	Std. dev.	Min	Max	Mean	Std. dev.
ENP_cst	1	4.74	1.90	0.37	1	4.74	1.97	0.33
(log of) Distance in km	0	8.28	6.63	0.94	0	8.28	6.61	0.95
State population (in millions)	0.04	37.25	6.14	6.61	0.04	37.25	6.42	6.85
Valid votes (in millions)	0.0006	20.77	0.11	0.25	0.001	20.77	0.12	0.28

[a]The descriptive statistics correspond to the Models for 1834–2016 in Tables A1 (31,574 observations) and A2 (24,631 observations).
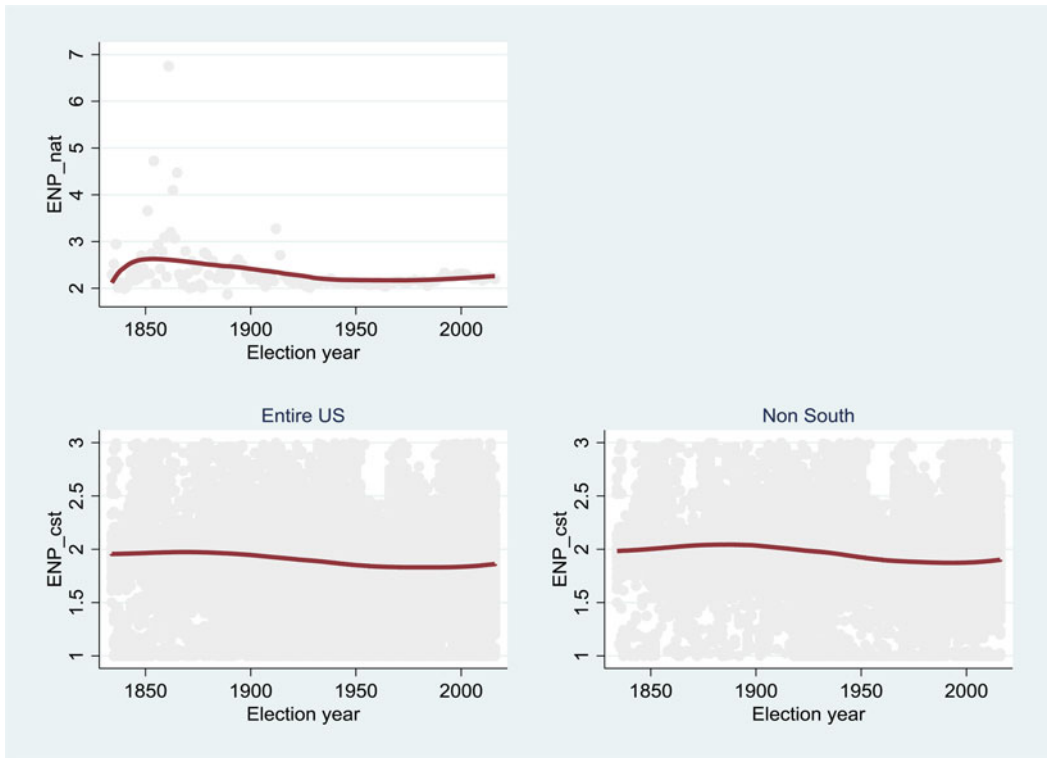
Figure 1. Nationalization in the United States.

Methods

Apart from the availability of data, the number and boundaries of electoral districts in the United States since the 1830s have substantially changed across elections, and this prevents me from using pooled time series cross-sectional data. The models are run using the least square dummy variable estimator (LSDV). As the number of parties varies across states and the coefficient on the *distance* from districts to Washington, D.C., is correlated with states, state fixed effects are included in the models to remove unobserved heterogeneity across districts at the level of states. The state fixed effects also capture to some extent the fact that American party systems are collections of state party systems embedded in regional political movements. Similarly, because there are many observations for every state, standard errors are clustered by states.

As a changing relationship over time between the number of parties at the district level and the *distance* to the national capital is hypothesized, a rolling regression with a 25-year recursive
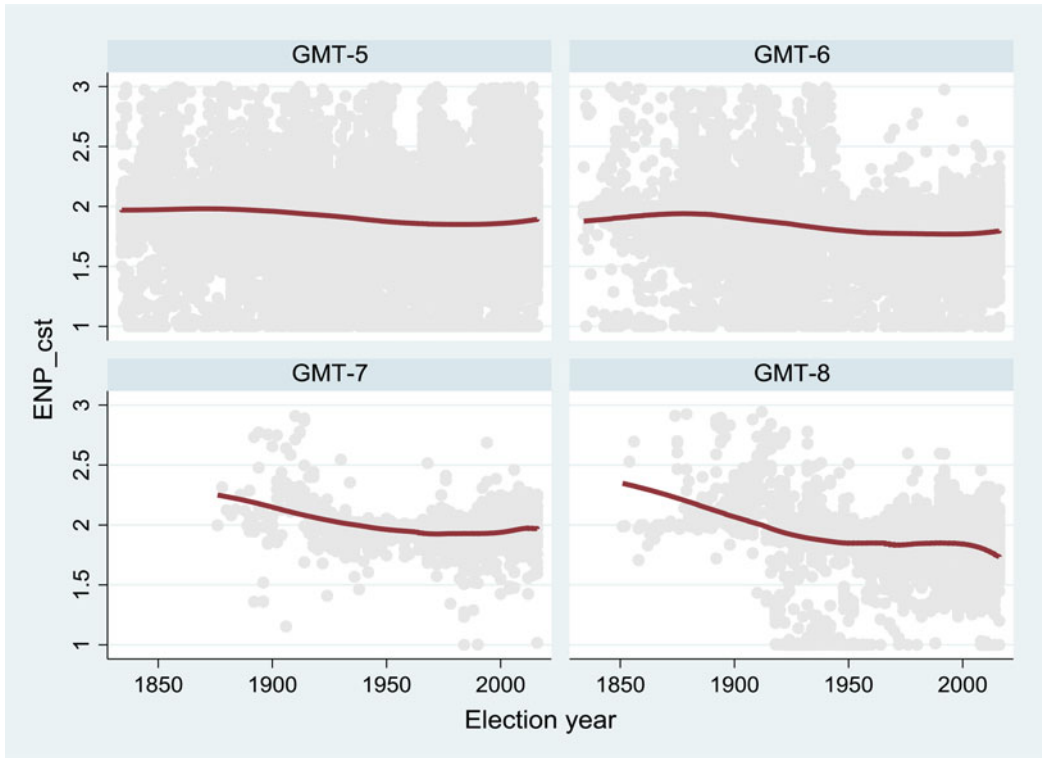
Figure 2. Nationalization in the United States and distance to Washington D.C.

window has been run. I first run specification (1) for the period 1834–1850, and then this starting period is held fixed, while the ending period advances and the window size grows.[7] A smaller recursive window does not appreciably change the results, but it makes the figures more difficult to examine.[8] This rolling regression allows us to examine how the effect of *distance* varies over time.

Results

The lowess curves (bwidth = 0.8) in Figure 1 show the changes over time in the effective number of electoral parties at the national and district levels in the United States. Roughly speaking, when examining the national level (upper panel), the number of parties settles to two in the 1920s and remains stable since then. The average number of parties at the district level has been near two since the first democratic elections in the nineteenth century and it slightly dips below two in the 1920s (lower left panel). The story is basically the same when excluding the South (lower right panel).[9]

To see how *ENP_cst* changes over time and across space, I have used the current time zones in the U.S. and divided the sample into those districts included in the four time zones. The lowess curves (bwidth = 0.8) in Figure 2 show, first, that in distant districts (in particular, those in the GMT-8 zone) the number of parties in the first elections is higher than in the nearby districts (in particular, those in the GMT-5 zone), and, second, that the more distant from the capital city districts are, the greater the decline in the number of parties.

---

[7]In the last period, 2000–2016, the window has 16 years.

[8]The results are qualitatively the same when using a 10-year recursive window.

[9]Those districts where the effective number of electoral parties is above three have been dropped (about 1 percent) for ease of interpretation of the figures.
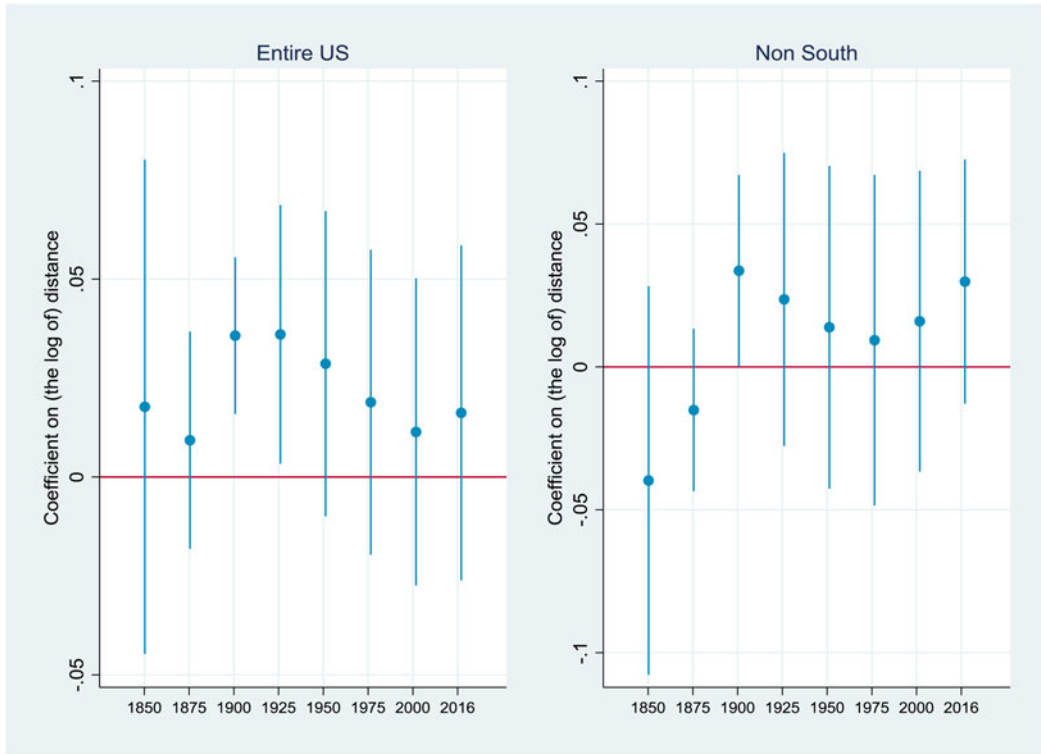
Figure 3. The effect of the distance to Washington, D.C. on *ENP_cst*.

The hypothesis about the varying effect of the *distance* to Washington, D.C. over time on the number of parties at the district level is tested through the estimation of (1). For ease of interpretation, Figure 3 only shows the coefficient on (the log of) the *distance* to Washington, D.C. at the 95 percent confidence level across the 25-year recursive windows. The results of the eight regressions are shown in Tables A1 and A2 in the Appendix. When examining all districts (left panel), the coefficient is positive and statistically significant at the 0.01 and 0.05 percent levels, respectively, in the 1834–1900 and 1834–1925 periods, that is, in the years prior to the New Deal. After the New Deal, the coefficient is no longer statistically significant. If the South is excluded (right panel of Figure 3), the story is quite similar, but the coefficient on (the log of) the *Distance* to Washington, D.C. is only statistically significant (at the 0.05 percent level) in the 1834–1900 period.

These results, however, might be affected by the fact that the within-district coordination occurs much later in the larger and more recent states admitted to the union than in the smaller and older states. I have replicated the previous analysis, including in the sample only the 31 states that joined the union before 1851. The reason for starting in 1851 is that the furthest state from Washington, D.C., California, entered the union in 1850, and therefore the range of (the log of) distance to the capital city does not change since then given that Alaska and Hawaii are not included in the analysis. The rolling regression starts in 1850, and a 25-year recursive window is used again. As can be seen in the left panel of Figure 4, the results are qualitatively the same. The coefficient on *distance* is again positive and statistically significant in the 1850–1900 and 1850–1925 periods. When excluding the South in the right panel, the 1850–1900 period is the only period where the distance to the capital city makes a difference. The results of the regressions are shown in Tables A3 and A4 in the Appendix.
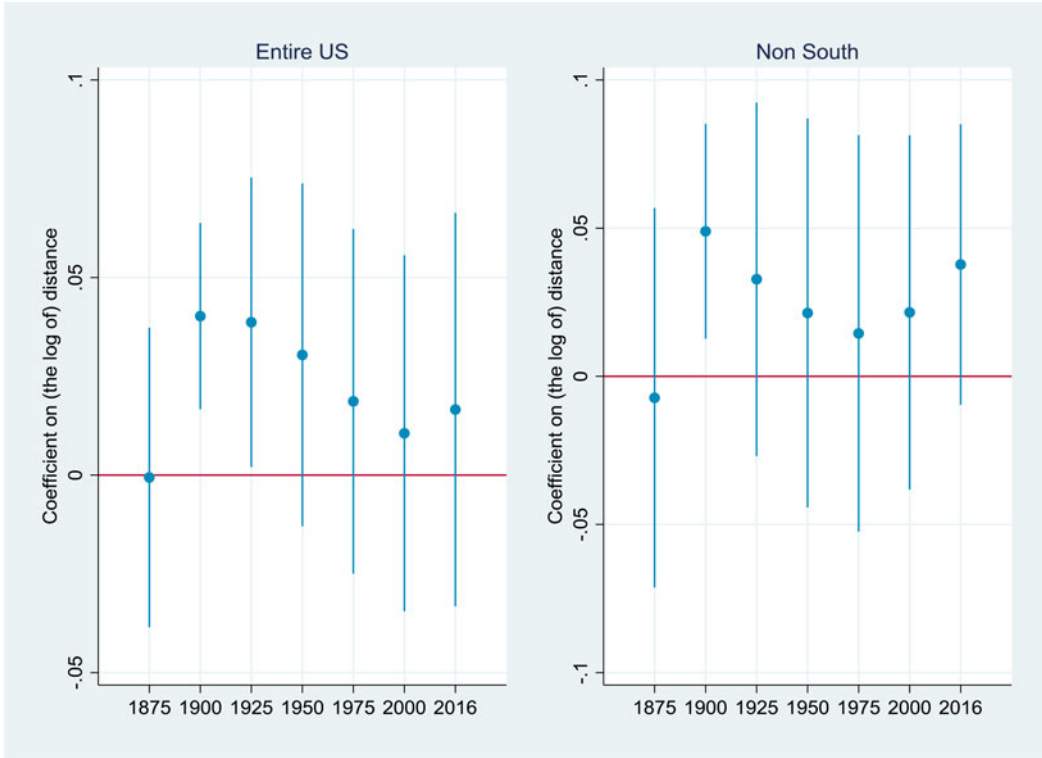
Figure 4. The effect of the distance to Washington, D.C. on *ENP_cst* with a restricted sample.

Table 2. The effect of distance on the number of parties (Entire U.S.)

	Models			
	GMT-5	GMT-6	GMT-7	GMT-8
Election year	−0.0020*** (0.00037)	−0.0027*** (0.00056)	−0.0038** (0.00107)	−0.0072** (0.00160)
State population (in millions)	0.018*** (0.0059)	0.022** (0.0089)	0.047*** (0.0116)	0.015** (0.0042)
Valid votes (in millions)	0.082** (0.036)	0.046 (0.028)	0.061 (0.27)	1.24*** (0.10)
State dummies	YES	YES	YES	YES
Constant	5.91*** (0.71)	6.67*** (1.08)	9.23*** (2.04)	15.54*** (3.05)
Number of observations	19,198	8776	857	2743
R^2	0.21	0.27	0.14	0.15

Note: Shown are OLS coefficients with standard errors clustered by state in parentheses.
***$p < 0.01$; **$p < 0.05$.

The second hypothesis about the centripetal nationalization within districts is examined in Table 2. Controlling for the state population, the number of voters in the districts, and the state dummies, I am interested in seeing whether the slope in party system fragmentation over time differs across districts depending on their distance to Washington, D.C. As can be seen, in the four subsamples the effective number of electoral parties drops as time goes by: the variable identifying the election years is negative and statistically significant at the 0.05 percent level or better. Interestingly, the negative coefficient for the states included in every time zone monotonically increases as we move away from the capital city. The coefficient in the GMT-8 time zone is

3.6 times bigger than in the GMT-5 time zone. When excluding the South (see Table A5 in the Appendix), the results do not change appreciably.

Finally, my argument primarily captures between-state differences rather than within-state differences and therefore using state fixed effects could be misleading. In Table A6 in the Appendix I ran model (1) using the state average effective number of electoral parties as dependent variable and the *distance* in hundreds of kilometers (plus 1) in a straight line from the state capitals to Washington, D.C., according to Google Maps. State fixed effects are not included. The results are qualitatively the same.

Conclusions

An important omission in the literature about political parties and voting behavior is how the effect of the nationalization of politics and the (de)centralization of economic and political power occurs within districts. Using data from the United States, I have argued that in the early stages of electoral democracies the distant districts from the capital city should have more parties than the nearby districts due to the greater saliency of local issues in the former. The effect and speed of nationalization within districts over time depends on the geographical distance from districts to the capital city. In the short term, nationalization is faster in the districts closest to Washington, D.C., while in the long term nationalization also affects distant districts. The consequence is that the number of local parties across districts will converge as time goes by.

References

Alesina A and Spolaore E (2003) *The Size of Nations.* Cambridge: MIT Press.

Boix C, Miller MK and Rosato S (2013) A complete data set of political regimes, 1800–2007. *Comparative Political Studies* **46**, 1523–1554.

Caramani D (2004). *The Nationalization of Politics: The Formation of National Electorates and Party Systems in Western Europe.* Cambridge: Cambridge University Press.

Chhibber P and Kollman K (1998) Party aggregation and the number of parties in India and the United States. *American Political Science Review* **92**, 329–342.

Chhibber PK and Kollman K (2004) *The Formation of National Party Systems: Federalism and Party Competition in Canada, Great Britain, India, and the United States.* Princeton, NJ: Princeton University Press.

FAO (2021) FAOSTAT land use domain. Available at http://www.fao.org/faostat/en/#data/RL. FAO.

Fishback PV and Wallis JJ (2013) What was new about the new deal? In Crafts N and Fearon P (eds), *The Great Depression of the 1930s: Lessons for Today.* Oxford, England: Oxford University Press, pp. 290–327.

Hamilton A, Madison J and Jay J (2009) *The federalist papers.* Edited and with an introduction by Ian Shapiro; with essays by John Dunn, Donald L. Horowitz, and Eileen Hunt Botting. Yale University Press.

Hicken A, Kollman K and Simmons JW (2016) Party system nationalization and the provision of public health services. *Political Science Research and Methods* **4**, 573–594.

Kollman K and Worthington ABH (2021) Party systems and political centralization and decentralization. *Journal of Politics* **83**, 1178–1183.

Kollman K, Hicken A, Caramani D, Backer D and Lublin D (2019) *Constituency-level elections archive.* Center for Political Studies, University of Michigan. Available at http://www.electiondataarchive.org.

Laakso M and Taagepera R (1979) Effective number of parties: a measure with application to west Europe. *Comparative Political Studies* **12**, 3–27.

Lewis JB, DeVine B, Pritcher L and Martis KC (2021) *United States congressional district shapefiles.* Retrieved October 30, 2021, from https://dmaps.polisci.ucla.edu.

Manson SM, Kugler TA, Schroeder J, Van Riper D and Ruggles S (2020) *IPUMS International Historical Geographic Information System: Version* 1. Minneapolis, MN: IPUMS.

Stoll H (2011) Dimensionality and the number of parties in legislative elections. *Party Politics* **17**, 405–429.