A Group-Based Approach to Measuring Polarization

Isaac D. Mehlhaff*

March 9, 2022

Abstract

Political polarization is a key concern in many important topics within the social sciences, yet the conceptual understanding and quantitative measurement of polarization are often unaligned. The rapidly expanding literature on polarization emphasizes two key dynamics: distance and concentration of a distribution's component groups. To capture these dynamics, I introduce the cluster-polarization coefficient (CPC), a measure of multimodal data structuration that scales to high-dimensional analysis and enables comparison across diverse spatiotemporal contexts. I present simulations and validation exercises to show that the CPC predicts distributional polarization with greater accuracy than current measures and I demonstrate its use as a post-processing technique by examining American elite polarization in comparative perspective.

1 Introduction

Political polarization has become an indispensable concept in the social sciences, with scholars, pundits, and even average citizens adding the term to their working vocabulary. Yet despite the proliferation of research on polarization, analysts often arrive at different conclusions with respect to its effect on key variables of interest, and they enthusiastically debate whether polarization is even

*The University of North Carolina at Chapel Hill; mehlhaff@live.unc.edu; word count: 9,757.

occurring at all (Abramowitz and Saunders 2008; DiMaggio, Evans, and Bryson 1996; Fiorina 2011). I suggest that much of the disagreement surrounding the identification, causes, and consequences of polarization stems in part from a lack of valid, generalizable, and widely employed measurement techniques. Polarization researchers have largely settled on a conceptual understanding that emphasizes two dynamics—intergroup heterogeneity and intragroup homogeneity—yet this theoretical foundation is rarely translated into empirical application, with analysts employing a wide variety of measures that capture, at best, only one of these dynamics. This incongruence between concept and measure results in empirical findings that sometimes contradict one another and may not reflect the concept being investigated.

Further, even a nominally valid measure may be analytically limiting if it does not enable comparison over time or space. Scholars often constrain their studies to the overall left-right dimension even though this unidimensional approach may not be appropriate for comparing political systems in different regions of the world or in countries with different levels of socioeconomic development.[1] In Europe, for instance, the left-right dimension has been widely recognized as the most salient, if not the most important (Fuchs and Klingemann 1990; Knutsen 1988; Volkens and Klingemann 2002). In Latin America, however, researchers have shown that voters do not consistently place themselves on such a spectrum and do not view politics through the left-right ideological lens to the extent that voters in more highly developed democracies do (Saiegh 2015; Zechmeister 2015; Zechmeister and Corral 2013). An ideal measure of polarization should enable comparison by employing data that more accurately correspond to the character of contestation in a given political system.

Aiming for a balance between analytical precision and ease of implementation, I offer a path forward by presenting the cluster-polarization coefficient (CPC). The CPC provides a widely applicable measure of multimodal data structuration that explicitly models intergroup heterogeneity and intragroup homogeneity, placing front and center the interpretation of polarization as a group-based phenomenon. Simulations show the CPC's accuracy in both uni- and multidimensional settings

[1] On the multidimensional nature of political polarization, see Bermeo (2003) and Tomz and Van Houweling (2008).

when compared to extant measurement strategies, and it performs well in convergent and construct validation exercises. I demonstrate its use as a post-data-processing technique by evaluating sets of ideological ideal point estimates recovered from Twitter data across several European and North American countries (Barberá 2015). To make the measurement procedure widely accessible to researchers and practitioners, I provide an open-source R package. In addition to calculating the CPC with researcher-specified group memberships, this package contains support for hierarchical, k-means, partitioning-around-medoids, and density-based clustering, making the measure easily applicable to a wide variety of data structures.

2 Dynamics of Political Polarization

The academic literature on political polarization is diverse, with important sub-literatures that address party system polarization on the one hand and societal, ideological, or mass polarization on the other. Research on party system polarization primarily focuses on the distance between parties, party families, or coalitions on various programmatic dimensions or on the overall left-right scale (Klingemann 2005; Mainwaring and Scully 1995; Sartori 1976). Analysts concerned with societal, ideological, or mass polarization are likewise concerned with an element of distance between individuals or sub-groups, but they also add a second dynamic to their conceptual understanding of polarization: the degree of concentration within social or political groups, or the extent to which their members cluster tightly together. In this sub-literature, polarization is considered to increase as the positions of social or political groups grow farther apart from one another or as they become more internally homogeneous (see, for example, Esteban and Ray 1994; Rehm and Reilly 2010; Ura and Ellis 2012). These dual dynamics—distance between and concentration of component groups—characterize polarization across a wide variety of applications, ranging from issue polarization (Baldassarri and Bearman 2007; DiMaggio, Evans, and Bryson 1996) to elite (Druckman, Peterson, and Slothuus 2013; Levendusky 2009) and even geographic polarization (Motyl 2016; Nall 2015). This is the theoretical understanding of polarization with which I am primarily concerned. The measure I propose below is specifically designed to capture these two dynamics, and it may not be suitable for analysis of party system polarization, which typically affords fewer data points and does not place theoretical emphasis on the concentration dynamic.

The challenge for measuring polarization in real-world data is that these two dynamics are occurring simultaneously and each may increase or decrease over time independent of the other. For example, if political parties take opposite and increasingly extreme positions on a particular issue while intra-party disagreement on those positions also rises, polarization may increase, but its rate of increase will be slower than if the members of those increasingly extreme parties were unified on the party's position. To translate the conceptual understanding of polarization into quantitative terms, a measure must incorporate both dynamics.

3 Cluster-Polarization Coefficient

I offer an approach to capturing these group dynamics, which I call the cluster-polarization coefficient (CPC). This measure is centrally concerned with groups and how dynamics within and between those groups manifest in polarization. Citizens and political elites hold multifaceted social and political opinions, exist in a political climate that is similarly multifaceted, and cluster together with other citizens and elites who hold similar social and political opinions. I explicitly model these grouping patterns by decomposing the total variation in clustered data into components corresponding to the distance between clusters and the concentration of each cluster. I also develop this measure with an eye toward enabling spatiotemporal comparison, as intra-state political dynamics, the number and nature of important sociopolitical cleavages, and the size and number of political coalitions may vary across countries or within countries over time.

3.1 CPC Derivation

To explain the logic underlying the CPC, I briefly turn to a similar, more familiar statistic: the coefficient of determination (R^2).[2] Researchers often wish to know the extent to which a dependent variable is correlated with a certain explanatory variable, holding constant a host of control variables. This, in simplified terms, is the optimization problem of ordinary least squares regression (OLS). In solving this optimization problem, we express the total variance of a dependent variable (TSS) as the sum of the variance explained by a vector of covariates (ESS) and the residual variance left unexplained (RSS), as in (1). OLS is then the optimization strategy used to minimize RSS, indirectly maximizing ESS. Dividing by TSS and solving for the ESS term gives an expression for the proportion of variance in the dependent variable explained by the vector of covariates—more commonly referred to as R^2:

$$\begin{aligned} TSS &= ESS + RSS, \\ \rightarrow R^2 &= 1 - \frac{RSS}{TSS}. \end{aligned} \tag{1}$$

I set up the CPC derivation in a similar manner to the R^2 derivation, decomposing the total variance of clustered data (TSS) in (2) into components directly corresponding to the two dynamics of polarization: the variance accounted for between the clusters (BSS, corresponding to distance) and the variance accounted for within all clusters ($TWSS$, corresponding to concentration). Dividing by TSS and solving for the BSS term gives an expression for the proportion of the total variance accounted for by the between-cluster variance—what I call the cluster-polarization coefficient (CPC):[3]

[2] On the coefficient of determination, see Barrett (1974), Cornell and Berger (1987), and Renaud and Victoria-Feser (2010), among others.

[3] Expressed in this way, the CPC also appears related—though not identical—to a one-way ANOVA F-statistic. This is a useful similarity for deriving properties of the measure, which I show in the Supplementary Materials.

$$
\begin{aligned}
&TSS = BSS + TWSS, \\
&\rightarrow CPC = 1 - \frac{TWSS}{BSS + TWSS} = \frac{BSS}{BSS + TWSS}.
\end{aligned} \tag{2}
$$

More formally, I compute three terms in (3): the total sum of squares TSS, the between-cluster sum of squares BSS, and the total within-cluster sum of squares $TWSS$, where each individual i in cluster k holds a position on dimension j:

$$
\begin{aligned}
TSS &= \sum_{i=1}^{n_i} \sum_{j=1}^{n_j} (x_{i[j]} - \mu_j)^2, \\
BSS &= \sum_{k=1}^{n_k} \sum_{j=1}^{n_j} (\mu_{k[j]} - \mu_j)^2, \\
TWSS &= \sum_{k=1}^{n_k} \sum_{i \in C_k} \sum_{j=1}^{n_j} (x_{i[k][j]} - \mu_{k[j]})^2.
\end{aligned} \tag{3}
$$

Mirroring the algebraic procedures in (2), I arrive at formal expressions of the variance of clustered data and of the CPC in (4):[4]

[4]For formal proofs of the expressions upon which this approach is based, see Fisher (1928) and Wishart (1931).

$$
\begin{aligned}
TSS &= BSS + TWSS, \\
\rightarrow \sum_{i=1}^{n_i}\sum_{j=1}^{n_j}(x_{i[j]} - \mu_j)^2 &= \sum_{k=1}^{n_k}\sum_{j=1}^{n_j}(\mu_{k[j]} - \mu_j)^2 + \sum_{k=1}^{n_k}\sum_{i \in C_k}\sum_{j=1}^{n_j}(x_{i[k][j]} - \mu_{k[j]})^2, \\
\rightarrow CPC &= 1 - \frac{\sum_{k=1}^{n_k}\sum_{i \in C_k}\sum_{j=1}^{n_j}(x_{i[k][j]} - \mu_{k[j]})^2}{\sum_{k=1}^{n_k}\sum_{j=1}^{n_j}(\mu_{k[j]} - \mu_j)^2 + \sum_{k=1}^{n_k}\sum_{i \in C_k}\sum_{j=1}^{n_j}(x_{i[k][j]} - \mu_{k[j]})^2}, \\
&= \frac{\sum_{k=1}^{n_k}\sum_{j=1}^{n_j}(\mu_{k[j]} - \mu_j)^2}{\sum_{k=1}^{n_k}\sum_{j=1}^{n_j}(\mu_{k[j]} - \mu_j)^2 + \sum_{k=1}^{n_k}\sum_{i \in C_k}\sum_{j=1}^{n_j}(x_{i[k][j]} - \mu_{k[j]})^2}, \\
&\forall \quad i \in (1, \ldots, n_i), \quad j \in (1, \ldots, n_j), \quad k \in (1, \ldots, n_k).
\end{aligned} \tag{4}
$$

The CPC thus possesses two desirable properties: It is naturally bounded on the interval $[0,1]$ and it takes into account both dynamics of polarization. The CPC increases when the distance between groups increases or when groups become more tightly concentrated around their collective ideal point, but the rate of those increases depends on the relative levels of BSS and $TWSS$.

Reexamining (2) offers some insight into these varying rates of change. If each group is extremely homogeneous (high $TWSS$), further pulling the groups apart (increasing BSS) will have minimal impact on overall polarization. The converse is also true: Bringing two groups closer together will decrease the level of polarization, but it will decrease much more slowly if group members still cling tightly to their own group's position than if group members are more diffuse and willing to intermingle among the groups. By contrast, if the groups are already very far apart (high BSS), making them even more homogeneous (increasing $TWSS$) will not have a dramatic effect on overall polarization. More intuitively, allowing group members to deviate from the group's collective ideal point will decrease the level of polarization, but it will decrease much more slowly if the groups are extremely far apart and there is virtually no opportunity for overlap between them. In sum, the CPC exhibits decreasing returns to scale; when a distribution displays high BSS or low $TWSS$ (or both), further changes to either of those variables will have a minimal impact on overall polarization. In the Supplementary Materials, I derive a sampling distribution and set of statistical properties for the CPC and briefly examine additional implications.

3.2 Adjusted CPC Derivation

While the CPC as derived above adheres closely to the minimal definition of polarization, it nevertheless exhibits behavior that complicates its use for comparison across contexts. In particular, being based on sums of squares makes it monotonic increasing with n. To make the CPC more generalizable to contexts with varying numbers of observations, variables, and clusters, I incorporate corrections for lost degrees of freedom into the key variance expressions in (5) and derive the adjusted CPC in (6).[5] For consistency, all calculations in this paper use the adjusted CPC.

$$
\begin{aligned}
TSS_{adj} &= \frac{\sum_{i=1}^{n_i}\sum_{j=1}^{n_j}(x_{i[j]}-\mu_j)^2}{n_i-n_j}, \\
TWSS_{adj} &= \frac{\sum_{k=1}^{n_k}\sum_{i\in C_k}\sum_{j=1}^{n_j}(x_{i[k][j]}-\mu_{k[j]})^2}{n_i-n_jn_k}.
\end{aligned}
\tag{5}
$$

[5]There is a long-standing debate surrounding the calculation of degrees of freedom for adjusted R^2 (for a review, see Yin and Fan 2001). I adopt the Wherry/Ezekiel formula (Wherry 1931)—perhaps the most popular formula and the one commonly used in statistical software—and incorporate an additional correction for n_k. The Supplementary Materials display the behavior of the CPC and adjusted CPC with varying numbers of dimensions and clusters.

$$
\begin{aligned}
CPC_{adj} &= 1 - \frac{TWSS_{adj}}{TSS_{adj}}, \\
\rightarrow CPC_{adj} &= 1 - \frac{\frac{\sum_{k=1}^{n_k}\sum_{i\in C_k}\sum_{j=1}^{n_j}(x_{i[k][j]}-\mu_{k[j]})^2}{n_i - n_j n_k}}{\frac{\sum_{i=1}^{n_i}\sum_{j=1}^{n_j}(x_{i[j]}-\mu_j)^2}{n_i-n_j}}, \\
&= 1 - \frac{\sum_{k=1}^{n_k}\sum_{i\in C_k}\sum_{j=1}^{n_j}(x_{i[k][j]}-\mu_{k[j]})^2}{\sum_{i=1}^{n_i}\sum_{j=1}^{n_j}(x_{i[j]}-\mu_j)^2}\frac{n_i-n_j}{n_i-n_j n_k}, \\
&= 1-(1-CPC)\frac{n_i-n_j}{n_i-n_j n_k}, \\
&\forall \quad i\in(1,...,n_i), \quad j\in(1,...,n_j), \quad k\in(1,...,n_k).
\end{aligned} \tag{6}
$$

3.3 Assumptions and Limitations

As with any method, researchers should be mindful of assumptions made in the calculation of the CPC and how those assumptions impose limitations on its use. First, the measure assumes that observations have been assigned to groups accurately. When there is a theoretically meaningful basis for such assignment (e.g. political party), this criterion is easily satisfied. In cases where *a priori* assignment is difficult or data is unlabeled, a wide variety of clustering methods may be used to assign observations to groups (Duda, Hart, and Stork 2001; MacKay 2003), and the CPC can then be calculated based on those cluster memberships.[6] I use such a strategy in the validation and application sections below.[7]

On a related issue, the CPC (and all parametric clustering methods, if they are used) also assumes that the number of groups n_k is accurately identified. When the data naturally produces easily identifiable clusters (e.g. liberals and conservatives), selecting this value is simple. In cases

[6]The associated R package contains support for hierarchical, k-means, partitioning-around-medoids, and density-based clustering.

[7]Readers wishing to see a visual demonstration of clustering and how it connects intuitively to polarization may refer to the Supplementary Materials.

where it is less clear, however, researchers may take advantage of several techniques to identify clusters in data (Celebi, Kingravi, and Vela 2013; Tzortzis and Likas 2014; Yoder and Priebe 2017).[8] The simplest and most common involves repeatedly running a clustering algorithm with an increasingly large number of clusters and selecting n_k based on some predefined criterion (for a review of many such criteria, see Jain 2010). Alternatively, researchers may bypass the issue of selecting n_k *a priori* and use a nonparametric method such as a density-based algorithm to uncover the appropriate number of clusters and to which cluster each observation belongs (Ester, Kriegel, and Xu 1996). In the Supplementary Materials, I provide an extended discussion of clustering techniques and the determination of n_k.

Third, constructing the CPC using sums of squares imposes an important limitation: It is sensitive to extreme outliers. Consider how this measure would respond to adding observations far out in the tails of a distribution. Such observations would lead to rapid increases in $TWSS$, likely overwhelming any increase in BSS that would result from small changes to the group's centroid. As a consequence, the CPC will rapidly decrease and remain relatively insensitive to any further changes in either variable, for the same reasons previously discussed. Scaling variables before calculating polarization estimates, however, can at least partially ameliorate this problem. I demonstrate this in the Supplementary Materials using heavy-tailed log-normal distributions.

Fourth, whether or not a clustering method is used, the CPC requires numeric, ordered data and cannot be calculated with categorical or unordered data.

Finally, even though it varies on the domain $[0,1]$, the CPC does not necesarily have a substantively meaningful interpretation. To give one example, the correct interpretation of a CPC of 0.5 would be "fifty percent of the total variance in the distribution is captured by the between-cluster variance," not "the distribution is fifty percent polarized." For this reason, researchers may wish to scale CPC estimates along with all other real-valued variables before fitting models or conducting inference, thereby enabling interpretation in terms of standard deviations.

[8] Applied researchers should also be sensitive to the possibility that n_k may be endogenous to polarization, changing over time as a consequence of exceptionally high or low polarization levels.

4 Simulation Evidence

4.1 Set-Up

In this section, I present evidence for the efficacy of the CPC by simulating both univariate and bivariate data using Gaussian mixture distributions. The purpose of this simulation exercise is to evaluate—in a controlled environment—the extent to which the CPC captures the two dynamics of polarization, and whether it does so better than preexisting measures. Gaussian mixtures are uniquely suited for this purpose because they provide a straightforward method for mimicking distributional polarization. Each component of a Gaussian mixture is parameterized by a location parameter μ and scale parameter σ, which neatly correspond to the two dynamics of polarization: distance and concentration, respectively.

Figure 1 presents a visualization of this intuition. This figure displays kernel density plots for simulated univariate data with two components. The plots are arranged such that the least polarized distributions fall at the top left and the most polarized distributions fall at the bottom right, and component parameters are provided by the plot labels along the top and right axes. Consider what happens to these component parameters as we move from a less polarized to a more polarized distribution. Moving from left to right across the rows of the facet plot, for example, the distributions become more polarized as the difference between component means increases and the components grow farther apart. Likewise, moving from top to bottom along the columns of the facet plot, the distributions become more polarized as component standard deviations decrease and the components grow more compact.

By manipulating these component parameters, therefore, I can generate mixture distributions with varying levels of polarization and estimate those levels using the CPC and other extant measures. I focus here on comparing the CPC to the three most popular strategies for measuring polarization in the political science literature: difference-in-means (e.g. Fiorina 2011), variance

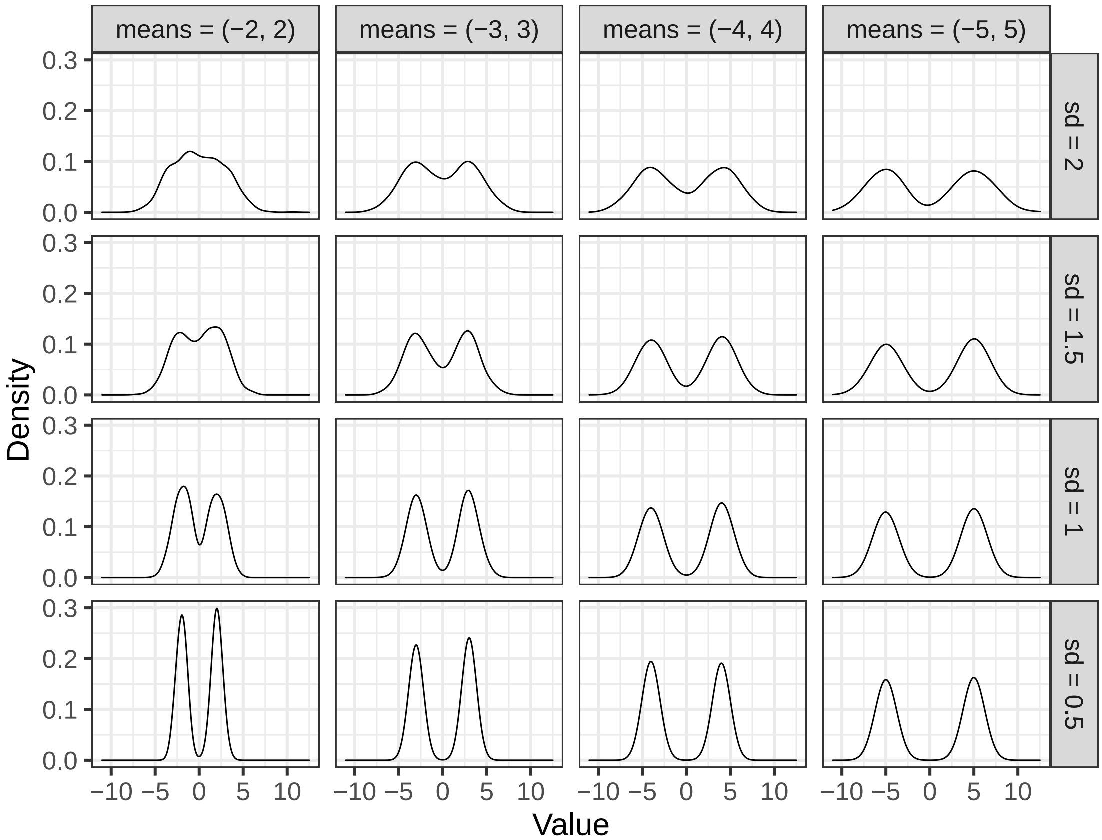


Figure 1: Visualization of Simulation Set-Up. Simulated Gaussian mixture distributions with $\mu_{global} = 0$; rows represent diverging means with standard deviations held constant and columns represent decreasing standard deviations with means held constant; thus, the least polarized distributions appear at top left and the most polarized distributions appear at bottom right.

(e.g. DiMaggio, Evans, and Bryson 1996), and kurtosis (e.g. Baldassarri and Bearman 2007).[9] To identify the accuracy of each polarization measure, I examine the distance and concentration dynamics separately. To simulate polarization as a result of increasing distance, I execute a four-step simulation exercise:

1. Fix component standard deviations at a range of values $\sigma \in \{0.5, 1, 1.5, 2\}$.[10] For identification, I use the same σ for each component and maintain a global mean of zero.

[9]See the Supplementary Materials for a critical review.

[10]As seen in Figure 1, even this relatively short range of values is sufficient to generate distributions ranging from unimodal to distinctly bimodal.

2. For each component standard deviation σ, select 1,000 values of μ as independent draws from $\mathrm{U}(2,5)$.

3. Take 1,000 independent draws from a Gaussian mixture parameterized by $\mathrm{N}(-\mu,\sigma;\mu,\sigma)$.

4. Apply each polarization measure to the resulting distribution.

The result of this procedure is 1,000 distributions, each with $N = 1000$, with which to evaluate the performance of each polarization measure. To simulate polarization as a result of increasing concentration, I execute a similar four-step simulation exercise:

1. Fix component means at a range of values $\mu \in \{2,3,4,5\}$.[11] For identification, I use the same absolute value of μ for each component and maintain a global mean of zero.

2. For each component mean μ, select 1,000 values of σ as independent draws from $\mathrm{U}(0.5,2)$.

3. Take 1,000 independent draws from a Gaussian mixture parameterized by $\mathrm{N}(-\mu,\sigma;\mu,\sigma)$.

4. Apply each polarization measure to the resulting distribution.

The result of this procedure is 1,000 distributions, each with $N = 1000$, with which to evaluate the performance of each polarization measure. For identification, I use equal component weights across all simulated distributions.[12]

Using these simulation frameworks, I evaluate the performance of the adjusted CPC relative to difference, variance, and kurtosis in both univariate and bivariate contexts.[13] Pursuant to the two definitional characteristics of polarization, an appropriate measure should indicate higher polarization when the distance between component means increases or when the standard deviation of each component decreases. Because polarization can occur around more than two poles, especially

[11] Again, as seen in Figure 1, even this relatively short range of values is sufficient to generate distributions ranging from unimodal to distinctly bimodal.

[12] Additional simulations in the Supplementary Materials investigate how the CPC changes in response to varying component weights.

[13] For bivariate data, I calculate difference by taking the average Euclidean distance between all component means, I calculate variance using the sum of squared Euclidean distances, and I calculate kurtosis using Mardia's multivariate kurtosis (Mardia 1970).

in multiparty systems, I conduct these procedures for distributions with two, three, and four components.[14] For all simulations, I calculate the adjusted CPC using true group memberships, which are known from the data randomization procedure. By using true group memberships instead of estimating them using a clustering algorithm, we can be sure that any advantages or disadvantages uncovered in the simulation results are attributable to the CPC itself and not to a clustering method being well- or ill-suited to this particular data structure.

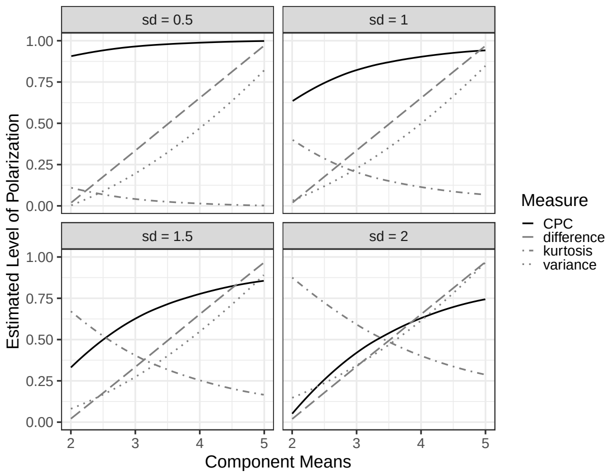
4.2 Results

I evaluate each simulated distribution using all four types of measurements and present the results in two ways. First, Figures 2 and 3 present the raw polarization estimates as a function of the randomized parameters for two-component simulations with univariate and bivariate data, respectively.[15] All measures are scaled to $[0,1]$ to enable comparison and plotted using locally estimated scatterplot smoothing (LOESS). A measure performing in line with theoretical expectations would register a positive slope in plot (a) and a negative slope in plot (b). However, as previously discussed, the magnitude of those slopes and the absolute level of estimated polarization should differ depending on the fixed parameter. For example, the sets of distributions with fixed $\sigma = 0.5$ or fixed $\mu = (-5,5)$ are more polarized on average than the distributions with a greater fixed σ or fixed μ parameters that are closer together. As a result, polarization estimates should generally be higher for those distributions and less sensitive to the value of the randomized parameter.

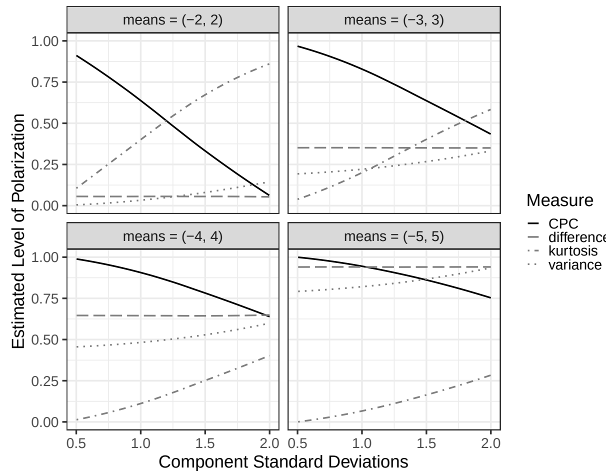
The results presented in Figures 2 and 3 generally align with expectations. Looking first at plot (a), slopes for difference, variance, and the adjusted CPC carry the expected sign. The magnitude of the distance and variance slopes, however, is relatively constant regardless of the fixed parameter, and the absolute level of estimated polarization appears similar. For example, with random component means of $(-5,5)$ at the far right hand side of each facet, difference and variance output almost identical polarization estimates regardless of whether component standard deviations

[14]For three and four components, I calculate the difference score by taking the average distance between all component means.

[15]The Supplementary Materials contain results of three- and four-component simulations.

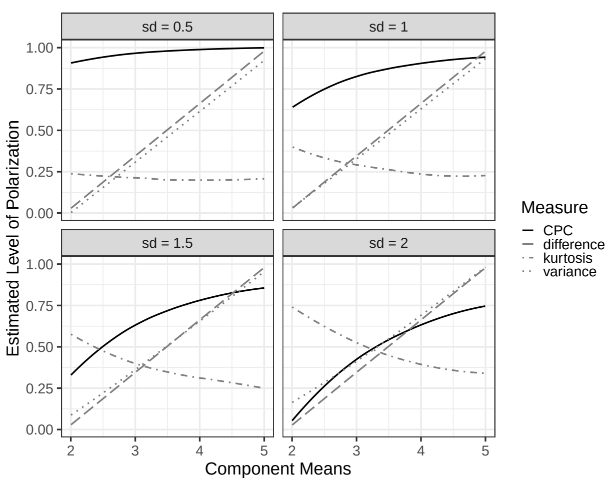


(a) Random means (distance)

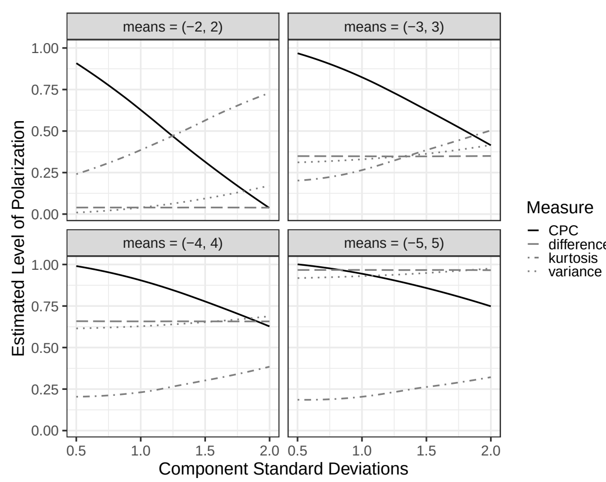


(b) Random standard deviations (concentration)

Figure 2: Univariate Polarization Estimates with Two Components. Results from univariate simulations of polarization measures with two components, showing estimated level of polarization for a randomly varying distribution parameter, holding the other parameter constant. All measures scaled to $[0,1]$ to enable comparison and plotted using LOESS.



(a) Random means (distance)



(b) Random standard deviations (concentration)

Figure 3: Bivariate Polarization Estimates with Two Components. Results from bivariate simulations of polarization measures with two components, showing estimated level of polarization for a randomly varying distribution parameter, holding the other parameter constant. All measures scaled to $[0,1]$ to enable comparison and plotted using LOESS.

are 0.5, 2, or anywhere in between. The CPC, on the other hand, appears more sensitive to those fixed parameters and displays intercepts and slope magnitudes more in line with expectations. The insensitivity of difference and variance to component standard deviations can be seen more clearly in plot (b). While the adjusted CPC again performs as expected, difference and variance appear as nearly flat lines, although they do output higher polarization estimates when the difference between fixed component means grows larger. Across all simulations, the raw value of kurtosis appears inversely related to the simulated level of polarization. I take this into account when evaluating each measure's performance more formally below.

Understanding how raw polarization estimates track with distributional characteristics is valuable, but it complicates a formal evaluation of a measure's effectiveness because the estimated level of polarization (the output of each measure) and the parameters controlling the simulated level of polarization (standard deviation or distance between means) are different quantities and are on different scales. Moreover, we do not have information about the "true" level of distributional polarization—estimating such quantities is the very goal of this measurement approach.

By holding all other distributional characteristics constant and randomly varying only component means and standard deviations, however, we do have information about each distribution's level of polarization *relative* to every other distribution. For example, the simulation to assess the distance dynamic holds standard deviations constant and randomly varies component means. Randomly generated means that are further apart will generate a distribution that is, in theory, more polarized. The result of the simulation, then, is 1,000 distributions that randomly vary in their level of polarization, and those relative levels of polarization can be identified by the relative value of the random component means. I therefore follow the approach taken by Lupu, Selios, and Warner (2017) and use the estimated polarization from each measure to rank order the distributions and compare those rankings to the true rank order recovered from the randomized parameters,[16] with a higher rank indicating a greater level of polarization.[17]

		Univariate				Bivariate			
		Difference	Variance	Kurtosis	CPC	Difference	Variance	Kurtosis	CPC
		$k=2$							
RMSE	Distance	**14.67**	15.38	39.04	30.73	**10.59**	11.05	281	22
	Concentration	406.65	572.36	34.93	**24.01**	406.15	569.16	234.02	**16.79**
	Cverall	287.73	404.87	37.05	**27.58**	287.29	402.53	258.58	**19.57**
MAE	Distance	**10.73**	11.2	29.83	23.84	**7.68**	7.9	212.63	17.16
	Concentration	335.05	496.49	23.53	**17.54**	334.47	494.54	176.58	**12.28**
	Overall	172.89	253.84	26.68	**20.69**	171.07	251.22	194.61	**14.72**
		$k=3$							
RMSE	Distance	17.7	**19.03**	70.7	32.33	16.39	**12.48**	38.61	21.91
	Concentration	400.4	573.84	59.67	**25.41**	409.82	570.81	35.92	**18.08**
	Overall	283.4	405.99	65.42	**29.07**	290.02	403.72	37.29	**20.08**
MAE	Distance	**12.88**	13.64	53.27	25.13	11.99	**8.99**	29.73	17.12
	Concentration	329.82	497.46	40.68	**18.41**	338.37	495.52	24.56	**13.27**
	Overall	171.35	255.55	46.98	**21.77**	175.18	252.26	27.15	**15.2**
		$k=4$							
RMSE	Distance	**10.22**	10.39	103.02	28.95	15.17	**11.4**	36.87	21.81
	Concentration	403.64	567.64	67.85	**21.5**	433.29	569.87	31.98	**17.35**
	Overall	285.51	401.45	87.23	**25.5**	306.57	403.04	34.51	**19.71**
MAE	Distance	**7.44**	7.54	76.73	22.38	11	**8.15**	28.25	16.91
	Concentration	332.94	493.01	52.3	**15.91**	361.31	494.82	22.37	**12.73**
	Overall	170.19	250.27	64.52	**19.15**	186.16	251.49	25.31	**14.82**

Table 1: Error of Distribution Rankings in Simulation. Root mean squared error and mean absolute error calculated for univariate and bivariate simulations with two, three, and four components; bolded values denote measure with lowest error in each category.

Table 1 reports the root mean squared error and mean absolute error for all four measures across all simulations.[18] Bolded values represent the best-performing measure in each category. Examining these results, a clear pattern emerges. Difference and variance register the lowest error

[16]One downside to this rank order approach is that it coerces all polarization measurements to an evenly spaced scale, thereby requiring an assumption of cardinality. However, when considered in conjunction with Figures 2 and 3, I believe the benefit from evaluating the magnitude of error introduced by each measure is greater than the cost imposed by this assumption.

[17]Figures 2 and 3 suggest that lower kurtosis indicates higher polarization, so I attempt to give kurtosis the benefit of the doubt by using its inverse to calculate polarization.

[18]Results are also presented graphically in the Supplementary Materials.

rates on the distance dynamic, regardless of the number of variables or components. This makes intuitive sense; holding the standard deviations of a mixture distribution constant and pulling the component means in opposite directions should, by definition, result in wider differences between means and higher variance for the distribution as a whole. By contrast, kurtosis registers its lowest error rates on the concentration dynamic, but rarely outperforms the other measures. Moreover, it returns inconsistent results, performing well in certain environments yet almost completely falling apart in others. The CPC performs as expected on both dynamics and often improves slightly in higher dimensions and multimodal distributions, which enables comparison across a wider variety of contexts. Finally, although it is necessary to examine each dynamic in isolation, real-world data rarely holds one dynamic constant while manipulating the other. As previously emphasized, a crucial characteristic of a polarization measure is the ability to account for both dynamics as they shift simultaneously. When looking at each measure's respective overall performance, the CPC appears to perform best regardless of the number of variables or components, and often by a drastic margin.[19]

5 Validation

Adcock and Collier (2001) detail three types of validation for quantitative (and qualitative) analysis: content, convergent, and construct. Content validation assesses whether a measure captures the full content of the quantity of interest by including all key elements of the concept and omitting any irrelevant ones. The simulation evidence above acts as a content validation exercise, showing that the CPC captures both dynamics of polarization and is frequently more accurate in doing so than existing measures. Convergent (or criterion) validation assesses whether a measure is empirically related to alternative measures of the same quantity of interest. I pursue this type of validation below by showing that CPC estimates of polarization in Congressional ideal points correlate at reasonable levels with extant measures of polarization. Finally, construct validation

[19]Overall performance is calculated by simply combining results from the distance and concentration simulations and calculating error statistics over the entire body of evidence.

assesses whether a measure is empirically related to measurements of another, distinct concept believed to be causally associated with the quantity of interest. I also pursue this type of validation below by showing that CPC estimates of polarization in Congressional ideal points correlate with indicators of economic inequality. Overall, I find substantial evidence that the CPC offers valid estimates of polarization in real-world data.

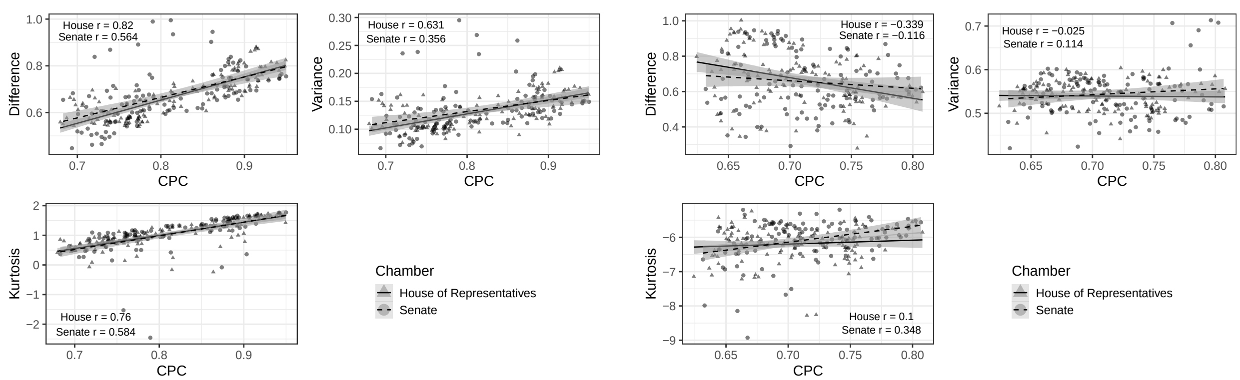
5.1 Data

Like several previous studies investigating polarization in the United States (Hare and Poole 2014; Hetherington 2009; Hirano et al. 2010), I calculate Congressional polarization using DW-NOMINATE data (Lewis et al. 2021). DW-NOMINATE uses a scaling procedure to estimate the ideology of individual legislators, placing them in a latent space on the basis of their role call votes (Poole 1998; Poole and Rosenthal 1985).[20] These estimates allow for a two-dimensional latent space, with the first dimension generally capturing economic issues and the second capturing all others. Having two sets of ideal points for each legislator allows me to demonstrate the CPC's advantage over other measures in multiple dimensions. Using these ideal point estimates also presents an opportunity to reiterate that the CPC is a post-processing technique to recover the polarization of data that has already been processed and scaled appropriately; it does not estimate other quantities of interest—such as ideological ideal points—from raw data. For construct validation, I also examine the polarization of this Congressional data against two common operationalizations of economic inequality: the Gini coefficient for gross household income and the share of total gross income held by individuals in the top percentile. These time series extend back as far as the 63rd Congress (1913-1915) and are compiled by Atkinson et al. (2017).

[20]For critiques of DW-NOMINATE and the use of roll call voting to estimate ideology, see Caughey and Schickler (2016), Lee (2015), and Roberts (2007).

5.2 Convergent Validation

To conduct convergent validation, I use all four types of measures examined in this paper to estimate the level of polarization among legislators within each Congressional chamber and across time. Because the United States has fielded two major political parties for the vast majority of its history (e.g. Abramson et al. 2000),[21] I calculate adjusted CPC estimates with two groups, with group assignments derived from k-means clustering.[22] Because I discussed and presented evidence above that existing measures of polarization are suboptimal, a convergent validation exercise may seem to carry less weight. Indeed, with the simulation results in mind, it is not clear that correlations between the CPC and existing measures should be particularly high. Nevertheless, assuming that existing measures do provide at least some approximation of polarization, their correlation with the CPC should be positive.



(a) First NOMINATE dimension (b) Both NOMINATE dimensions

Figure 4: Correlations Between Adjusted CPC and Extant Polarization Measures Within Chambers and Across Congresses 1-116. Shaded areas represent 95% confidence intervals.

Figure 4 plots CPC estimates against each other measure and displays Pearson's correlation coefficient for the two measures presented on each facet. Estimates are calculated separately for each chamber and trend lines give 95% confidence intervals. I begin with just the first dimension of ideological ideal points, for which polarization estimates are shown in plot (a). Results seem

[21]In the Supplementary Materials, I briefly discuss periods of ideological polarization and convergence in American history and show that the CPC recovers estimates in line with those historical trends.

[22]Assumptions of k-means clustering are referenced in the following section.

to corroborate expectations. The CPC displays positive correlations with each alternate measure, with the magnitude of those correlations displaying a fairly wide range of 0.356 to 0.82. With the exception of that minimum, however, all correlations are at middling to high levels. This suggests that although the CPC is empirically related to these other measures and that they measure the same basic concept, it is still capturing meaningfully different information and is not necessarily interchangeable with alternative measures.

Adding the second dimension of ideal points, however, changes the picture. These results are shown in Figure 4, plot (b). Whereas all polarization measures generally point in the same direction when taking into account just one dimension of ideology, they appear only weakly related to the CPC when considering both dimensions. Difference-in-means estimates are even inversely correlated with CPC estimates. This suggests one of two conclusions: Either current measures or the CPC struggle to capture polarization on more than one variable. Simulation results above point toward the former, and the construct validation exercise below provides further evidence that these underwhelming correlations in two dimensions extend from poor performance by current meausures.

5.3 Construct Validation

To conduct construct validation, I turn to the connection between economic inequality and political polarization, with respect to the United States Congress in particular. As inequality and polarization in the United States have coevolved over the last few decades, scholars have contributed to an influential literature positing a link between the two (Barber and McCarty 2015; Bonica et al. 2015; Stewart, McCarty, and Bryson 2020). McCarty, Poole, and Rosenthal (2006) argue that rising inequality since the 1970s has tracked closely with polarization at both the elite and mass levels. Other scholars suggest this relationship may extend to other contexts. Gelman (2009) and Rehm (2011) provide evidence at the level of individual party identification and voting behavior, Garand (2010) contends that similar dynamics work at the state level, and several comparative studies argue that inequality is strongly associated with polarization cross-nationally

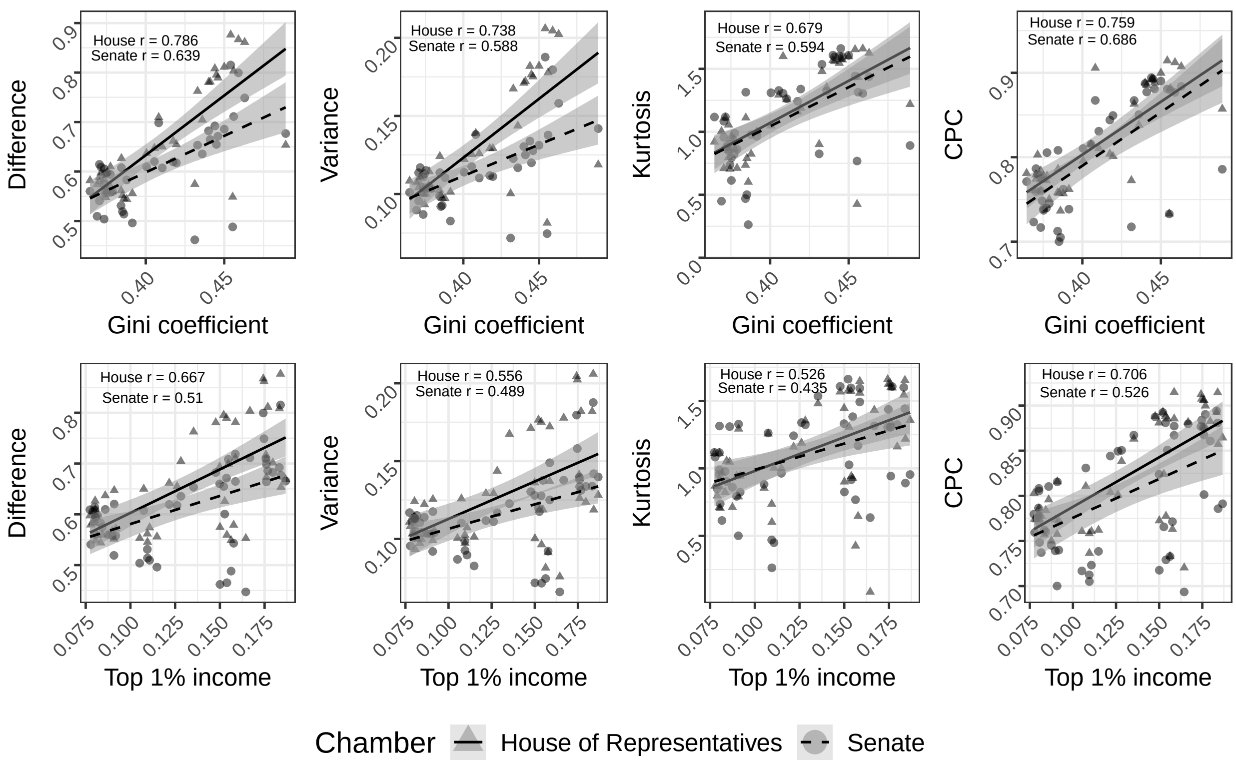


Figure 5: Correlations Between Adjusted CPC and Economic Inequality Within Chambers and Across Congresses 63-114. Polarization calculated using NOMINATE first-dimension ideology estimates. Shaded areas represent 95% confidence intervals.

(Grechyna 2016; Gunderson forthcoming; Pontusson and Rueda 2008). Assuming these previous works have accurately identified an association between economic inequality and Congressional polarization, the CPC should correlate strongly with operationalizations of inequality when applied to legislator ideal points.

I again begin with just the first dimension of ideal point estimates. Figure 5 plots each measure's polarization estimates against the Gini coefficient and percent of income earned by the top 1% of earners during the corresponding congress. Estimates are calculated separately for each chamber and trend lines give 95% confidence intervals. Pearson's correlation coefficient for each chamber is presented on each facet. Evidence aligns with expectations. All polarization measures are positively correlated with both operationalizations of economic inequality and, with one exception, the CPC is correlated at equal or greater magnitude than other measures across all combinations of chamber and operationalization.

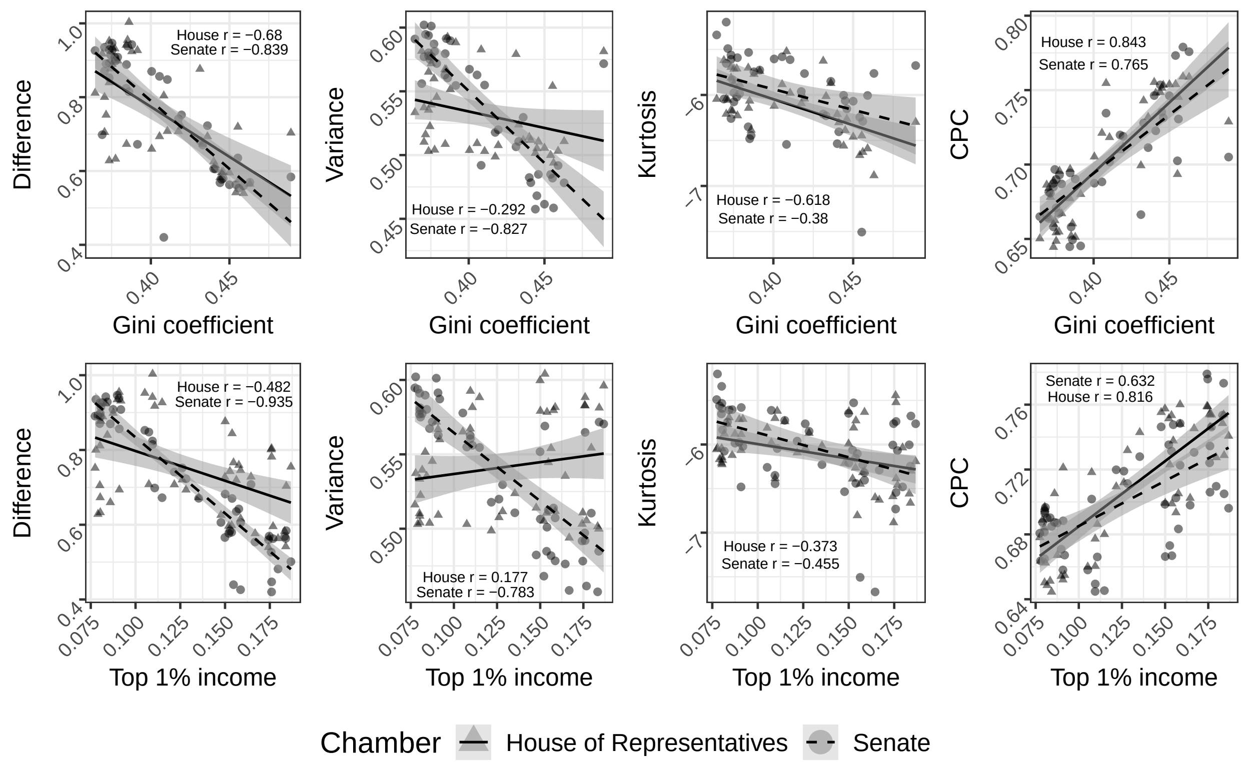


Figure 6: Correlations Between Adjusted CPC and Economic Inequality Within Chambers and Across Congresses 63-114. Polarization calculated using two-dimensional NOMINATE ideology estimates. Shaded areas represent 95% confidence intervals.

Recall, however, that in the convergent validation exercise above, the correlation between the CPC and other measures fell apart after adding the second dimension of ideological estimates. Relating those two-dimensional polarization estimates to the inequality data provides an opportunity to disentangle whether those low correlations are due to poor performance by current measures or the CPC. Figure 6 presents these results and clearly indicates that current measures struggle in this two-dimensional problem. Most correlations between polarization and inequality are now negative. In some cases, results differ drastically across chambers. The CPC is the only measure of the four which maintains a positive correlation after adding the second dimension of ideology, and three of the four correlations even increase in magnitude compared to the one-dimensional case.

In sum, I find evidence that the CPC offers valid estimates of polarization when applied to real-world data. Moreover, it is the only measure that performs well in construct validation in more than dimension. In some cases, current measures of polarization merely provide an imperfect fit with the theoretical concept of interest. But in other cases, such as the two-dimensional ideology data used in the validation exercises here, current measures may provide spurious or inaccurate results.

6 American Polarization: Exceptional or Nothing Special?

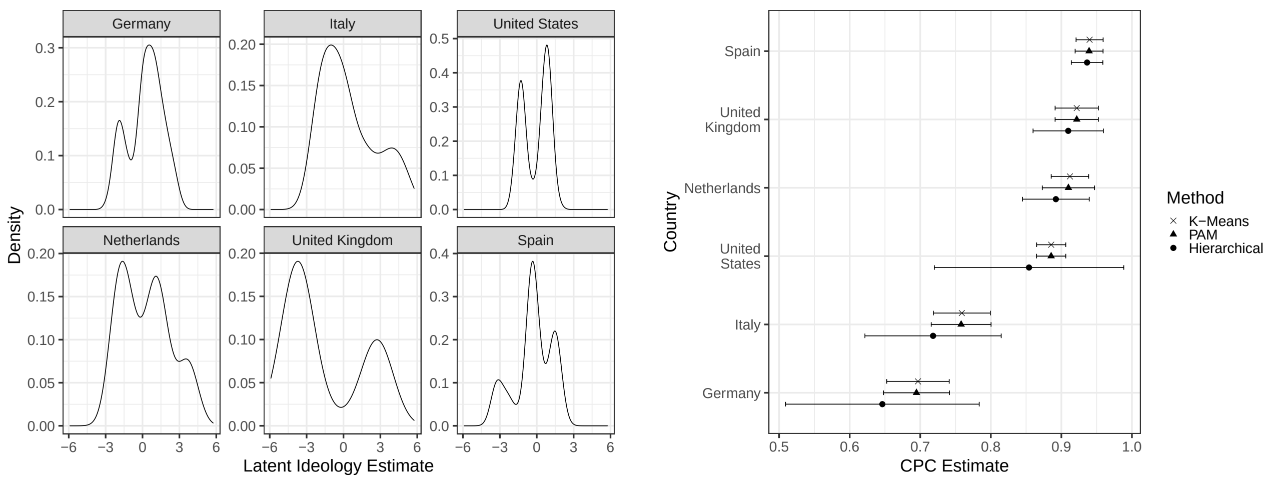
I now broaden my scope and examine elite polarization not just in the United States, but in a set of six industrially developed, consolidated democracies. The subject of elite polarization has enjoyed substantial scholarly attention in the United States (Davis and Dunaway 2016; Hetherington 2001; Robison and Mullinix 2016), but concerns about ideological divisions have also been raised in a wide range of developed (Adams, Green, and Milazzo 2012; Vegetti 2014) and developing countries (Bruhn and Greene 2007; Singer 2016). How does elite polarization in the United States compare to these other contexts? Does its level of severity mirror its level of attention in the media and political science literature or are pundits and scholars lavishing a great deal of concern on a relatively average case? Addressing these questions raises a key methodological challenge, namely, how to measure polarization across political systems with varying numbers of parties.[23]

To demonstrate how the CPC can be used to address this challenge, I apply it to elite ideological ideal points from Western Europe and North America. These latent ideal points are estimated by Barberá (2015) using Twitter data from 2012 and 2013. This application provides a simple and straightforward demonstration of how the CPC enables comparison across cases, as each country differs in the extent to which their latent elite groups are internally homogeneous and externally heterogeneous, and even in how many groups exist within each country. These differences can be seen clearly in Figure 7, plot (a).

I use three different clustering methods to recover group memberships and compare estimates across those methods. K-means (Celebi, Kingravi, and Vela 2013) and partitioning-around-medoids (PAM) (Reynolds et al. 2006) assume cluster sizes are approximately equal, all obser-

[23]Some recent studies grapple with this challenge in the case of mass polarization (Gidron, Adams, and Horne 2020; Reiljan 2020; Wagner 2021).

vations take values on every variable, and variables are on the same numeric scale. K-means additionally assumes spherical clusters and cardinality of data. Hierarchical clustering (Murtagh 1983), by comparison, requires few assumptions, although some variants still assume spherical clusters and some dissimilarity measures assume variables are uncorrelated within clusters. As previously noted, using clustering methods requires the *a priori* specification of how many clusters exist in each country. In this case, the selection of n_k is simple, as Figure 7, plot (a) shows either two or three distinct clusters in each country. To confirm these selections and demonstrate methods for the selection of n_k in more difficult cases, I generate elbow plots and silhouette scores for each country. These plots are displayed and discussed in the Supplementary Materials.



(a) Kernel density plots of ideal points (b) Ideological polarization estimates

Figure 7: Estimates of Polarization in Elite Twitter Data. Plot (a) presents kernel density estimates for each country, with facets ordered by level of polarization. Plot (b) presents adjusted CPC estimates using three clustering methods with 95% confidence intervals, calculated using a case-resampling bootstrap.

Figure 7, plot (b) displays the adjusted CPC with 95% confidence intervals for each country,[24] estimated using each of these clustering methods to make group assignments. Hierarchical clustering produces much wider standard errors in each country compared to k-means or PAM—unsurprising given that it is not the ideal method to use with this data—but the three methods produce very similar polarization estimates. Results suggest that the United States is not, in fact,

[24]Confidence intervals in Figure 7, plot (b) are calculated using a case-resampling bootstrap.

exceptional in its level of elite polarization—at least in comparison to this small sample. CPC estimates in the United States are statistically indistinguishable from those in the Netherlands or United Kingdom, although there is a clear gap between those countries and less-polarized Italy and Germany. Exhibiting higher levels of polarization than the United States are Spain, the United Kingdom, and the Netherlands. At the time of data collection, these countries were enduring vigorous debate over austerity measures in the wake of the Euro crisis (Miley 2017), beginning down the path toward Brexit (Hobolt, Leeper, and Tilley forthcoming), and navigating growing tensions surrounding immigration and ethnic integration (Oosterwaal and Torenvlied 2010), respectively.

7 Discussion

Zooming out and examining the CPC's performance more broadly, I emphasize two additional benefits of the measure for empirical research. First, because the CPC is ultimately a measure of multimodal data structuration, it can be used to gain information about not only polarization, but also other quantities of interest. One example is conditional party government, which mirrors the definition of polarization presented here by emphasizing the degree of preference homogeneity *within* parties and the degree of preference conflict *between* parties (Aldrich 2011; Rohde 1991). Studies building on the conditional party government theory have utilized a wide range of measures to capture a holistic picture of its implications, but a summary measure similar to the one the CPC provides has remained elusive.

Second, the CPC can often obviate the need for dimensionality reduction techniques or be used in tandem with them. Political polarization is typically a multidimensional phenomenon (Bermeo 2003; Tomz and Van Houweling 2008), yet scholars often restrict their analysis to a single left-right dimension or use various methods to recover a single dimension from multivariate data. Because the CPC scales to a high-dimensional space, it can take into account numerous variables without needing to use item response (IRT) models or create additive indices. Alternatively, IRT models can recover a single dimension and the CPC can be used to measure the degree of polarization in the latent distribution. Multidimensional IRT models have also proliferated in recent years (Reckase 2009; Wirth and Edwards 2007), in which case the CPC's consistency in higher dimensions again allows it to measure the degree of polarization in the recovered multivariate distribution. In sum, the CPC is a valuable tool to use alongside dimensionality reduction methods because it gives a piece of information about the underlying distributions that these other methods do not, and it is this information that often holds theoretical significance.

References

Abramowitz, Alan I. and Kyle L. Saunders (Apr. 2008). "Is Polarization a Myth?" In: *The Journal of Politics* 70.2, pp. 542–555.

Abramson, Paul R. et al. (Sept. 2000). "Challenges to the American Two-Party System: Evidence from the 1968, 1980, 1992, and 1996 Presidential Elections". In: *Political Research Quarterly* 53.3, pp. 295–522.

Adams, James, Jane Green, and Caitlin Milazzo (Dec. 2012). "Who Moves? Elite and Mass-Level Depolarization in Britain, 1987–2001". In: *Electoral Studies* 31.4, pp. 643–655.

Adcock, Robert and David Collier (Sept. 2001). "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research". In: *American Political Science Review* 95.3, pp. 529–546.

Aldrich, John H. (2011). *Why Parties? The Origin and Transformation of Political Parties in America*. 2nd ed. Chicago: The University of Chicago Press.

Atkinson, Anthony B. et al. (2017). *The Chartbook of Economic Inequality*. URL: `https://www.chartbookofeconomicinequality.com`.

Baldassarri, Delia and Peter Bearman (Oct. 2007). "Dynamics of Political Polarization". In: *American Sociological Review* 72.5, pp. 784–811.

Barber, Michael J. and Nolan McCarty (2015). “Causes and Consequences of Polarization”. In: *Solutions to Political Polarization in America*. Ed. by Nathaniel Persily. New York: Cambridge University Press, pp. 15–58.

Barberá, Pablo (2015). “Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data”. In: *Political Analysis* 23.1, pp. 76–91.

Barrett, James P. (Feb. 1974). “The Coefficient of Determination - Some Limitations”. In: *The American Statistician* 28.1, pp. 19–20.

Bermeo, Nancy (2003). *Ordinary People in Extraordinary Times: The Citizenry and the Breakdown of Democracy*. Princeton, NJ: Princeton University Press.

Bonica, Adam et al. (2015). “Congressional Polarization and Its Connection to Income Inequality: An Update”. In: *American Gridlock: The Sources, Character, and Impact of Political Polarization*. Ed. by James A. Thurber and Antoine Yoshinaka. New York: Cambridge University Press, pp. 357–377.

Bruhn, Kathleen and Kenneth F. Greene (Jan. 2007). “Elite Polarization Meets Mass Moderation in Mexico’s 2006 Elections”. In: *PS: Political Science & Politics* 40.1, pp. 33–38.

Caughey, Devin and Eric Schickler (Oct. 2016). “Substance and Change in Congressional Ideology: NOMINATE and Its Alternatives”. In: *Studies in American Political Development* 30.2, pp. 128–146.

Celebi, M. Emre, Hassan A. Kingravi, and Patricio A. Vela (Jan. 2013). “A Comparative Study of Efficient Initialization Methods for the K-Means Clustering Algorithm”. In: *Expert Systems with Applications* 40.1, pp. 200–210.

Cornell, J. A. and R. D. Berger (1987). “Factors That Influence the Value of the Coefficient of Determination in Simple Linear and Nonlinear Regression Models”. In: *Phytopathology* 77.1, pp. 63–70.

Davis, Nicholas T. and Johanna L. Dunaway (2016). “Party Polarization, Media Choice, and Mass Partisan-Ideological Sorting”. In: *Public Opinion Quarterly* 80.S1, pp. 272–297.

DiMaggio, Paul, John Evans, and Bethany Bryson (Nov. 1996). “Have American’s Social Attitudes Become More Polarized?” In: *American Journal of Sociology* 102.3, pp. 690–755.

Druckman, James N., Erik Peterson, and Rune Slothuus (Feb. 2013). “How Elite Partisan Polarization Affects Public Opinion Formation”. In: *American Political Science Review* 107.1, pp. 57–79.

Duda, Richard O., Peter E. Hart, and David G. Stork (2001). *Pattern Classification*. 2nd ed. New York: John Wiley & Sons.

Esteban, Joan-María and Debraj Ray (July 1994). “On the Measurement of Polarization”. In: *Econometrica* 62.4, pp. 819–851.

Ester, Martin, Hans-Peter Kriegel, and Xiaowei Xu (1996). “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise”. In: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96). 226-231.

Fiorina, Morris P. (2011). *Culture War? The Myth of a Polarized America*. In collab. with Samuel J. Abrams and Jeremy C. Pope. 3rd ed. London: Pearson.

Fisher, R. A. (Dec. 1928). “The General Sampling Distribution of the Multiple Correlation Coefficient”. In: *Proceedings of the Royal Society of London, Series A, Containing Papers of a Mathematical and Physical Character* 121.788, pp. 654–673.

Fuchs, Dieter and Hans-Dieter Klingemann (1990). “The Left-Right Schema”. In: *Continuities in Political Action: A Longitudinal Study of Political Orientations in Three Western Democracies*. Ed. by M. Kent Jennings and Jan W. van Deth. Berlin: Walter de Gruyter, pp. 203–234.

Garand, James C. (Oct. 2010). “Income Inequality, Party Polarization, and Roll-Call Voting in the U.S. Senate”. In: *The Journal of Politics* 72.4, pp. 1109–1128.

Gelman, Andrew (2009). *Red State, Blue State, Rich State, Poor State: Why Americans Vote the Way They Do*. 2nd ed. Princeton, NJ: Princeton University Press.

Gidron, Noam, James Adams, and Will Horne (2020). *American Affective Polarization in Comparative Perspective*. New York: Cambridge University Press.

Grechyna, Daryna (July 2016). “On the Determinants of Political Polarization”. In: *Economics Letters* 144, pp. 10–14.

Gunderson, Jacob R. (forthcoming). “When Does Income Inequality Cause Polarization?” In: *British Journal of Political Science*.

Hare, Christopher and Keith T. Poole (July 2014). “The Polarization of Contemporary American Politics”. In: *Polity* 46.3, pp. 411–429.

Hetherington, Marc J. (Sept. 2001). “Resurgent Mass Partisanship: The Role of Elite Polarization”. In: *American Political Science Review* 95.3, pp. 619–631.

— (Apr. 2009). “Putting Polarization in Perspective”. In: *British Journal of Political Science* 39.2, pp. 413–448.

Hirano, Shigeo et al. (Aug. 2010). “Primary Elections and Partisan Polarization in the U.S. Congress”. In: *Quarterly Journal of Political Science* 5.2, pp. 169–191.

Hobolt, Sara B., Thomas J. Leeper, and James Tilley (forthcoming). “Divided by the Vote: Affective Polarization in the Wake of the Brexit Referendum”. In: *British Journal of Political Science*.

Jain, Anil K. (June 2010). “Data Clustering: 50 Years Beyond K-Means”. In: *Pattern Recognition Letters* 31.8, pp. 651–666.

Klingemann, Hans-Dieter (2005). “Political Parties and Party Systems”. In: *The European Voter: A Comparative Study of Modern Democracies*. Ed. by Jacques Thomassen. Oxford, UK: Oxford University Press, pp. 22–63.

Knutsen, Oddbjørn (July 1988). “The Impact of Structural and Ideological Party Cleavages in West European Democracies: A Comparative Empirical Analysis”. In: *British Journal of Political Science* 18.3, pp. 323–352.

Lee, Frances E. (2015). “How Party Polarization Affects Governance”. In: *Annual Review of Political Science* 18, pp. 261–282.

Levendusky, Matthew (2009). *The Partisan Sort: How Liberals Became Democrats and Conservatives Became Republicans*. Chicago: University of Chicago Press.

Lewis, Jeffrey B. et al. (2021). *Voteview: Congressional Roll-Call Votes*. URL: https://voteview.com/.

Lupu, Noam, Lucía Selios, and Zach Warner (Jan. 2017). "A New Measure of Congruence: The Earth Mover's Distance". In: *Political Analysis* 25.1, pp. 95–113.

MacKay, David J.C. (2003). *Information Theory, Inference, and Learning Algorithms*. Cambridge, UK: Cambridge University Press.

Mainwaring, Scott and Timothy R. Scully, eds. (1995). *Building Democratic Institutions: Party Systems in Latin America*. Stanford, CA: Stanford University Press.

Mardia, K. V. (Dec. 1970). "Measures of Multivariate Skewness and Kurtosis with Applications". In: *Biometrika* 57.3, pp. 519–530.

McCarty, Nolan, Keith T. Poole, and Howard Rosenthal (2006). *Polarized America: The Dance of Ideology and Unequal Riches*. 2nd ed. Cambridge, MA: The MIT Press.

Miley, Thomas Jeffrey (Apr. 2017). "Austerity Politics and Constitutional Crisis in Spain". In: *European Politics and Society* 18.2, pp. 263–283.

Motyl, Matt (2016). "Liberals and Conservatives Are (Geographically) Dividing". In: *Social Psychology of Political Polarization*. Ed. by Piercarlo Valdesolo and Jesse Graham. New York: Routledge, pp. 7–37.

Murtagh, F. (Nov. 1983). "A Survey of Recent Advances in Hierarchical Clustering Algorithms". In: *The Computer Journal* 26.4, pp. 354–359.

Nall, Clayton (Apr. 2015). "The Political Consequences of Spatial Policies: How Interstate Highways Facilitated Geographic Polarization". In: *The Journal of Politics* 77.2, pp. 394–406.

Oosterwaal, Annemarije and René Torenvlied (Mar. 2010). "Politics Divided from Society? Three Explanations for Trends in Societal and Political Polarisation in the Netherlands". In: *West European Politics* 33.2, pp. 258–279.

Pontusson, Jonas and David Rueda (2008). "Inequality as a Source of Political Polarization: A Comparative Analysis of Twelve OECD Countries". In: *Democracy, Inequality, and Represen-*

tation in Comparative Perspective. Ed. by Pablo Beramendi and Christopher J. Anderson. New York: Russell Sage Foundation, pp. 312–353.

Poole, Keith T. (July 1998). “Recovering a Basic Space From a Set of Issue Scales”. In: *American Journal of Political Science* 42.3, p. 954.

Poole, Keith T. and Howard Rosenthal (May 1985). “A Spatial Model for Legislative Roll Call Analysis”. In: *American Journal of Political Science* 29.2, p. 357.

Reckase, Mark D. (2009). *Multidimensional Item Response Theory*. Statistics for Social and Behavioral Sciences. New York: Springer.

Rehm, Philipp (Apr. 2011). “Risk Inequality and the Polarized American Electorate”. In: *British Journal of Political Science* 41.2, pp. 363–387.

Rehm, Philipp and Timothy Reilly (Mar. 2010). “United We Stand: Constituency Homogeneity and Comparative Party Polarization”. In: *Electoral Studies* 29.1, pp. 40–53.

Reiljan, Andres (May 2020). “”Fear and Loathing across Party Lines” (Also) in Europe: Affective Polarisation in European Party Systems”. In: *European Journal of Political Research* 59.2, pp. 376–396.

Renaud, Olivier and Maria-Pia Victoria-Feser (July 2010). “A Robust Coefficient of Determination for Regression”. In: *Journal of Statistical Planning and Inference* 140.7, pp. 1852–1862.

Reynolds, A. P. et al. (Dec. 2006). “Clustering Rules: A Comparison of Partitioning and Hierarchical Clustering Algorithms”. In: *Journal of Mathematical Modelling and Algorithms* 5.4, pp. 475–504.

Roberts, Jason M. (Aug. 2007). “The Statistical Analysis Of Roll-Call Data: A Cautionary Tale”. In: *Legislative Studies Quarterly* 32.3, pp. 341–360.

Robison, Joshua and Kevin J. Mullinix (Apr. 2016). “Elite Polarization and Public Opinion: How Polarization Is Communicated and Its Effects”. In: *Political Communication* 33.2, pp. 261–282.

Rohde, David W. (1991). *Parties and Leaders in the Postreform House*. Chicago: The University of Chicago Press.

Saiegh, Sebastián M. (2015). “Using Joint Scaling Methods to Study Ideology and Representation: Evidence from Latin America”. In: *Political Analysis* 23.3, pp. 363–384.

Sartori, Giovanni (1976). *Parties and Party Systems: A Framework for Analysis*. Cambridge, UK: Cambridge University Press.

Singer, Matthew (2016). “Elite Polarization and the Electoral Impact of Left-Right Placements: Evidence from Latin America, 1995–2009”. In: *Latin American Research Review* 51.2, pp. 174–194.

Stewart, Alexander J., Nolan McCarty, and Joanna J. Bryson (Dec. 2020). “Polarization under Rising Inequality and Economic Decline”. In: *Science Advances* 6.

Tomz, Michael and Robert P. Van Houweling (Aug. 2008). “Candidate Positioning and Voter Choice”. In: *American Political Science Review* 102.3, pp. 303–318.

Tzortzis, Grigorios and Aristidis Likas (July 2014). “The MinMax K-Means Clustering Algorithm”. In: *Pattern Recognition* 47.7, pp. 2505–2516.

Ura, Joseph Daniel and Christopher R. Ellis (Jan. 2012). “Partisan Moods: Polarization and the Dynamics of Mass Party Preferences”. In: *The Journal of Politics* 74.1, pp. 277–291.

Vegetti, Federico (Sept. 2014). “From Political Conflict to Partisan Evaluations: How Citizens Assess Party Ideology and Competence in Polarized Elections”. In: *Electoral Studies* 35, pp. 230–241.

Volkens, Andrea and Hans-Dieter Klingemann (2002). “Parties, Ideologies and Issues: Stability and Change in Fifteen European Party Systems 1945–1998”. In: *Political Parties in the New Europe: Political and Analytical Challenges*. Ed. by Kurt Richard Luther and Ferdinand Müller-Rommel. Oxford, UK: Oxford University Press, pp. 143–168.

Wagner, Markus (Feb. 2021). “Affective Polarization in Multiparty Systems”. In: *Electoral Studies* 69.

Wherry, R. J. (Nov. 1931). “A New Formula for Predicting the Shrinkage of the Coefficient of Multiple Correlation”. In: *The Annals of Mathematical Statistics* 2.4, pp. 440–457.

Wirth, R. J. and Michael C. Edwards (Mar. 2007). “Item Factor Analysis: Current Approaches and Future Directions”. In: *Psychological Methods* 12.1, pp. 58–79.

Wishart, J. (May 1931). “The Mean and Second Moment Coefficient of the Multiple Correlation Coefficient, in Samples from a Normal Population”. In: *Biometrika* 22.3, pp. 353–376.

Yin, Ping and Xitao Fan (Jan. 2001). “Estimating R2 Shrinkage in Multiple Regression: A Comparison of Different Analytical Methods”. In: *The Journal of Experimental Education* 69.2, pp. 203–224.

Yoder, Jordan and Carey E. Priebe (2017). “Semi-Supervised K-Means++”. In: *Journal of Statistical Computation and Simulation* 87.13, pp. 2597–2608.

Zechmeister, Elizabeth J. (2015). “Left-Right Identifications and the Latin American Voter”. In: *The Latin American Voter: Pursuing Representation and Accountability in Challenging Contexts*. Ed. by Ryan E. Carlin, Matthew M. Singer, and Elizabeth J. Zechmeister. Ann Arbor, MI: University of Michigan Press, pp. 195–225.

Zechmeister, Elizabeth J. and Margarita Corral (June 2013). “Individual and Contextual Constraints on Ideological Labels in Latin America”. In: *Comparative Political Studies* 46.6, pp. 675–701.