Social Networks and Political Accountability

Victoria Mooers

February 28, 2024

Abstract

Does it matter if your political boundaries reflect your community? In an effort to ensure equal representation, U.S. congressional districts are required to be roughly equal in population. However, congressional district boundaries need not reflect the geography of communities. If voters share political information through their social networks, a community that is more cohesive with its congressional district may be at an advantage: with a common representative, they may be able to more easily keep each other informed about their representative's actions. I study whether the level of match between congressional districts and social networks ("congruence") impacts voters' knowledge of their representatives and, in turn, their representatives' effort in Congress. I find that congruence increases voters' knowledge of their representative: for example, increasing congruence from the minimum to maximum value observed in the data increases the probability that a voter knows their representative's party by 11.8pp, from a mean of 61.7%. However, I find that congruence does not impact representatives' effort, though estimates are noisy. Future versions of this paper will model and measure the impacts of congruence on voter behavior, candidate behavior, and political competition.

1 Introduction

Does it matter if your political boundaries reflect your community? The U.S. Constitution enshrines the principle of one-person-one-vote by requiring that all House of Representatives congressional districts within a state be equal in population. However, a district that satisfies this constraint need not match the geography of the underlying social network. This mismatch may not be benign: the existence of a "public sphere" in which constituents can discuss elections they share has also been postulated as necessary for a well-functioning democracy (Habermas et al. 1992). Constituents who know and can speak with each other may be more successful at learning about the performance of their representatives, and in turn better able to leverage the power of their votes to hold their representatives accountable in delivering favorable policies (Prat and Strömberg 2013; Strömberg 1999; Strömberg 2004a). Conversely, people living in socially fragmented districts (such as heavily gerrymandered ones) may be deprived of the opportunity to learn much from their fellow constituents, as most people in their networks are represented by someone else.

In this paper, I ask: In the U.S., how well do congressional district boundaries reflect existing communities? What impacts does this match (or lack thereof) have on voters' knowledge of their representative and, in turn, their representative's effort in Congress?

Specifically, for each county, I measure the mismatch between social networks and congressional districts, as captured by the share of a county's friends that live in the same congressional district as the county. I refer to this measure as congruence (analogous to Snyder and Strömberg 2010. I calculate congruence by aggregating Facebook's Social Connectedness Index (SCI) (Bailey, Cao, et al. 2018), which uses the Facebook friendship graph to construct the strength of social connectedness between each county pair in the U.S. I study the impact of changes in congruence on voters' survey reports of their familiarity with their representative and on several measures of representative effort, including bill sponsorships, voting behavior, and committee assignments.

Informed voters play an important role in monitoring government, as they can hold politicians accountable for their policies at the voting booth. However, this very role of monitoring may introduce biases in who the government best serves. A substantial literature on models of media coverage and political accountability predicts that more informed voters will be more responsive to policy in their voting decisions and, in turn, representatives will be more responsive to the more informed voters (Prat and Strömberg 2005, Prat and Strömberg 2013, Strömberg 1999, Strömberg 2001, Strömberg 2004a). In these models, there are multiple groups of voters who differ in how informed they are about their representative's actions. The representative wants to get re-elected and can exert effort to deliver public goods to each group. Less informed groups of voters do not change their votes regardless of their representative's success at garnering them more public goods. Accordingly, representatives allocate their effort elsewhere: in particular, towards benefitting the more informed groups of voters, who will change their votes in response to policy outcomes. Consequently, the less informed voters receive lower levels of public goods.

Previous research has shown that the predictions of these models hold true when considering the information provided by traditional media. Exogenous variations in newspaper (Angelucci et al. 2020, Besley and Burgess 2002, Snyder and Strömberg 2010) and radio coverage (Ferraz and Finan 2008, Strömberg 2004b) indeed impact how informed voters are. Further, when voters receive less information about government from traditional media, these same voters receive less funding (Eisensee and Strömberg 2007).

I build on this literature by studying the role of information spread through social networks. Crucially, if social networks are another avenue through which voters gain political knowledge, then voters' level of knowledge can be manipulated through gerrymandering. Gerrymandering is generally discussed in terms of its extensive margin impact of depriving a party of seats they might have won under a more equitably drawn map. However, gerrymandering may be further damaging if communities can be strategically dissected to manipulate their voters' knowledge of their representative, and in turn depressing their success at inducing their representative to serve their needs (an intensive margin impact on representation). With detailed social network data, like the SCI, becoming publicly available for the first time in recent years, it is not implausible that gerrymanderers will use such data to their advantage.[1]

Indeed, previous literature studying political behavior and peer effects (Fafchamps et al. n.d.,Quintelier et al. 2012,Sinclair 2012) has found that peers impact voters' political knowledge, identification, and participation. These studies have generally focused on individual-level interactions, capturing the direct spillovers from one individual to another. This paper instead studies social networks as a whole: I capture how the aggregate impacts of these interactions change when communities are more or less reflected by district boundaries.

Additionally, an expansive literature has studied the impacts of social media and technology access on political knowledge and participation (Enikolopov, Makarin, et al. 2020,Manacorda and Tesei 2020,Di Tella et al. 2021, Guriev et al. 2021) and the responses of politicians (Bessone et al. 2022). In this paper, I am agnostic about whether information is spread online or offline, and instead focus on the role of the geography of connections. Indeed, the SCI is a useful measure of social networks because it closely reflects offline networks (Bailey, Cao, et al. 2018, Kuchler et al. 2022), in part because Facebook usage is relatively even across demographic groups (Auxier and Anderson 2021) and also because Facebook friendships are persistent, accumulating throughout a lifetime. The SCI also captures friendship patterns that might be lost in other proxies for social networks, such as physical proximity. Nonetheless, I test the robustness of my results to alternative proxies for social networks by using data on commuting flows.

In order to identify the causal impact of congruence on outcomes, I utilize changes in congruence driven by district border changes following redistricting. District boundaries are re-drawn primarily every 10 years following the Decennial Census, in order to respect the requirement that each House representative represent roughly the same number of people. Most districts change in this process: following each of the 2010 Census

[1]For example, Bouton et al. 2023 find evidence that gerrymanderers already strategically account for differential turnout rates.

and 2020 Census, I find that over 85% of counties overlap with districts that experience a border change. I construct congruence for each county in each year of 2006-2023. This allows me to include county and year fixed effects: intuitively, I capture what happens when the same county faces a change in congruence due to redistricting. This addresses the concern that congruence may be correlated with unobserved location-specific characteristics that affect outcomes.

I find that congruence varies substantially across the country, with a mean of 45% but a range of 6% to 74%. Rural counties tend to be more congruent (social ties are inversely predicted by distance and rural counties are generally in larger districts), as are counties with a higher white share of the population and lower educational attainment. I demonstrate that congruence significantly increases voters' familiarity with their representatives (robust to including a variety of controls, to district-by-year fixed effects, to accounting for media markets, and to constructing the network using commuting flows). For example, increasing congruence from the minimum to maximum value observed in the data is associated with an increase the probability that a voter knows their representative's party by 11.8pp, from a mean of 61.7%. Focusing instead on congruence measured using commuting flows, I find similar but smaller impacts; commuting flows tend to represent much denser networks (a mean of 86%) and as such may not capture the role of more geographically distant ties. Notably, congruence generally does not impact placebo outcomes like familiarity with the governor or senators. Nonetheless, I find no evidence of congruence impacting measures of representative effort, though estimates are noisy.

1.1 Future Steps

In future iterations of this paper, I will study additional measures of voters' political participation and the consequences for electoral competition. In particular, I plan to use a DeGroot social learning model (Degroot 1974) to model the spread of district-relevant information within social networks; I will then combine this with a probabilistic voting model to get predictions on candidate incentives.

In terms of outcomes, as motivated by the model I plan to examine the impact of congruence on voter turnout, representatives' incumbency advantage, and overall political competition, as well as other measures of voters' engagement (e.g., campaign volunteering and campaign donations) and House candidates' responses (e.g., their targeting of political advertising).

The outcomes focused on electoral competition will shed light on whether my finding that congruence impacts voter information plausibly has electoral consequences. In particular, Bouton et al. 2023 demonstrate that heterogeneity in voter turnout alters gerrymanderer's incentives. If the impacts of congruence also lead to, e.g., changes in voter turnout, this would suggest that policymakers should account for the role of social networks in political participation when making laws governing the drawing of district boundaries.

2 Methods

2.1 Congruence

I study the geographical mismatch of social networks and political boundaries. In particular, I focus on the alignment of county-level social networks and congressional districts. Accordingly, I define "congruence" for a given county as the share of that county's friends that are in the same congressional district as the county.

2.1.1 Facebook Social Connectedness Index

For data on social networks, I use the Facebook Social Connectedness Index (SCI) (see Bailey, Cao, et al. 2018 for details), which is as a measure of the strength of social connection between two locations. $\text{SCI}_{i,j}$ is the relative probability of a friendship link between users in two geographic units i and j:

$$\text{SCI}_{i,j} = \frac{\text{Friendship Links}_{i,j}}{\text{Facebook Users}_i \times \text{Facebook Users}_j}$$

That is, the SCI is the number of friendship links between the two locations, normalized by the total number of possible connections between them.[2] I use the SCI for U.S. county-county pairs from the October 2021 snapshot.[3]

2.1.2 Aggregating SCI to Construct Congruence

Whereas the SCI gives the strength of social connection between two counties, congruence is the share of a county's friends that are in the same congressional district as the county. Accordingly, to construct congruence I need to appropriately aggregate the SCI.

Now, let $i \in J \subset K$, where i is a county, J is the set of all counties in the same congressional district as county i, and K is the set of all counties in the US. The congruence of county i, i.e., the share of county i's

[2] This makes the SCI independent of differences in the number of users in each location. Double the SCI means that it is twice as likely for a user in i to be friends with a user in j. The SCI is scaled from 1 to 1,000,000,000, areas with particularly small populations are removed, and noise is added to preserve privacy.

[3] There are 3,136 counties in the data, and each county appears in a pair with every other county (including itself). As such, there are 9,834,496 county-county pairs. The SCI only includes users who have interacted with (including simply logging into) any of Meta's apps in the 30 days prior to the snapshot. Locations are assigned based on users' provided information (such as their stated city) and their device connection information.

links[4] that are to people in counties in the same district, is then

$$\text{Congruence}_i = \frac{\text{Links}_{i,J}}{\text{Links}_{i,K}} = \frac{\sum_{j \in J} \text{Links}_{i,j}}{\sum_{k \in K} \text{Links}_{i,k}}$$

However, this does not follow immediately from the SCI data. In particular, we could find each $\text{Links}_{i,j}$ as

$$\text{Links}_{i,j} = \text{SCI}_{i,j} \times \text{Facebook Users}_i \times \text{Facebook Users}_j$$

but the number of Facebook users in each county is not made available, so this is not possible to back out. Bailey, Gupta, et al. 2020 argue that we can substitute the population of an area for the number of Facebook users. This requires the assumption that Facebook usage rates are the same across counties. Indeed, the share of the population that uses Facebook does not vary substantially across demographic groups: in 2021, Facebook usage rates among American adults varied slightly between urban and suburban (70%) and rural areas (67%); when sliced by race, income, and education, usage rates varied between 61% and 74% (Auxier and Anderson 2021). The largest gaps emerge by age, with the lowest usage rates among 65+ year-olds (50%) and the highest usage rates among 30-49 year-olds (almost 80%); use among 18-29 year-olds reflects the national average at 70%. There are also minimal differences in Facebook usage rates by political party (Vogels et al. 2021).[5] Nonetheless, the assumption of equal usage rates masks possible heterogeneity across counties, and this limitation should be kept in mind throughout.

Replacing the number of Facebook users in a given county with the county's population and re-arranging, we get

$$\text{Congruence}_i = \frac{\sum_{j \in J} \text{Links}_{i,j}}{\sum_{k \in K} \text{Links}_{i,k}} = \frac{\sum_{j \in J} (\text{SCI}_{i,j} \times \text{Pop}_j)}{\sum_{k \in K} (\text{SCI}_{i,k} \times \text{Pop}_k)}$$

which is feasible to calculate. I use the U.S. Census Bureau's 2020 Census Redistricting Data block-level counts to find the population of each county.

Some counties overlap with multiple congressional districts; for these counties, I use the population-weighted average congruence:

$$\text{Congruence}_i = \sum_{d \in D(i)} \frac{n_{i,d}}{n_i} \text{Congruence}_{i,d}$$

[4]For convenience, I refer to "Friendship Links" as simply "Links" from here on.

[5]Respondents were asked whether they ever use each social media platform. Facebook usage rates rose until 2016, and remained stable at around 70% of U.S. adults from then until at least 2021. As of 2021, Facebook was the social media platform with the least heterogeneity in usage rates by age. While 18-29 year-olds were the heaviest users of all other platforms, their Facebook use was only exceeded by their use of YouTube (95%) and Instagram (71%) (Auxier and Anderson 2021).

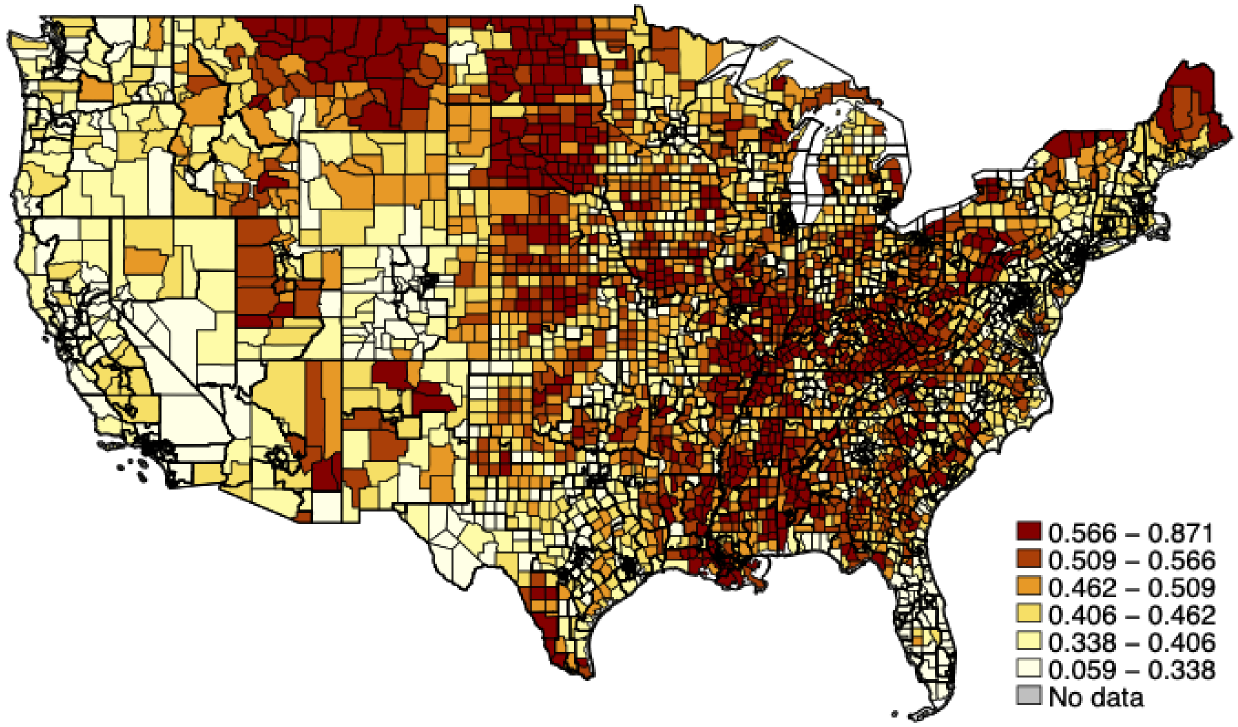

Figure 1: Unadjusted 2020 County Congruence with Congressional Districts

where $D(i)$ is the set of districts county i has land in, and $\frac{n_{i,d}}{n_i}$ is the share of the population of county i that lives in district d.

2.1.3 Summary Statistics and Predictors of Congruence

Congruence varies substantially across the continental U.S. Figure 1 shows congruence by county (counties are demarcated by thin lines, and congressional districts by thick lines).[6] The color bins in the figure each contain an equal number of counties, and hence represent different ranges of congruence values. The large bulk of counties have congruence between 35% and 55%.[7] In particular, among the continental 48 states, mean congruence is 45% with a standard deviation of 11.5pp; minimum congruence is 6% and maximum is 74%.[8]

[6]Throughout this paper, I focus on the 48 contiguous states – i.e., in results I exclude Alaska, Hawaii, Washington, D.C., and territories. I exclude Alaska and Hawaii since they are likely to have distinctly different patterns of connections to the rest of the country. I exclude D.C. and the territories because their representatives are non-voting members. However, all counties as well as foreign friendships are included for calculating the scaled total number of friends (denominator) for each county. Results are robust to not adjusting for foreign friendships.

[7]Appendix Figure 4 is the same figure, using evenly spaced bins. Appendix Figures 5 and 6 zoom in on individual states.

[8]Loving County, TX has the lowest congruence. It is a county on New Mexico's southern border, and it is the least populous county in the United States, with 64 residents as per the 2020 Census. Clay County, KY has the highest congruence. It is in southeastern Kentucky, central to the Appalachian region, and has a population of 20,345 as per the 2020 Census.

Some of the variation in congruence is clearly driven by how district boundaries are drawn. As might be expected given that social networks tend to follow state boundaries (Bailey, Cao, et al. 2018), counties in single district states[9] have higher congruence on average (53%).[10] Gerrymandering can also determine congruence, particularly if characteristics targeted by a gerrymanderer (e.g., political affiliation, race) are also predictive of social ties. Gerrymandering can increase or decrease congruence, depending on whether a community is "packed" into a district or "cracked" between them. For example, the counties that overlap with Texas's 15th congressional district – a district nicknamed the "fajita strip" for its notorious long and skinny shape (Solomon 2021; Zak 2022) and reflective of Texas's history of gerrymandering (*Redistricting Report Card - Texas* 2023) – taken together actually have about average congruence for Texas (39% vs 41% for all of Texas), but individually vary from having low congruence (24%) to higher than average (50%) (see Figure 6).[11] This variance coincides with critiques that the district groups dramatically different communities under one representative.[12] Lastly, due to the restriction that each congressional district within a state represent roughly the same population (across states averaging about 760,000 in 2020, see Eckman 2021; Whitaker 2017), counties with large populations (including most urban areas) are more likely to be split by a district boundary in order to accommodate this constraint; a one percent increase in the logarithm of county population is associated with a 2.8pp decrease in congruence.

The other determinant of congruence is the geography of social networks. The biggest predictor of social ties is distance (Bailey, Cao, et al. 2018), so counties that are further from a congressional district border will generally have higher congruence. Naturally, this is more likely to occur in geographically large districts, which are necessarily in areas with lower population density (again because each district is meant to have roughly the same population). This leads to higher congruence in rural areas. Simultaneously, urban areas have much more geographically dispersed social networks, because they have strong ties to other urban centers around the country (Bailey, Cao, et al. 2018); this further drives down congruence in urban areas.

Table 1 summarizes how congruence varies with each of these predictors of district boundaries and social

[9] In 2020, Delaware, Montana, North Dakota, South Dakota, Vermont, and Wyoming. These six states contain 215 counties, or 7% of all counties in the data.

[10] There is some heterogeneity: Delaware has below average congruence (38%) as does Wyoming (42%). Vermont has average congruence, while the remaining states have above average congruence (54%-57%).

[11] These counties' average congruence with district 15 alone (ignoring other districts they overlap with) is 37%, and varies from 20% (Guadalupe County, outside San Antonio) to 50% (Jim Hogg County, near to but not on the border with Mexico).

[12] Zak 2022 writes: "The next member of Congress representing the long and skinny 15th Congressional District in Texas will represent communities that stand in stark contrast. At the very northern tip of the 15th District lies New Braunfels...[ranked] the 25th best city to live in America [due to] its fast-growing economy, low cost of living, and optimal health and safety conditions. At the district's southernmost point is Hidalgo County which borders Mexico. It has a poverty rate of almost 30%, nearly three times the national average. The contrasting needs of these areas are more representative of America's diversity than they are typical for one federal lawmaker...In February 2021, South Texan residents pleaded for new districts that more accurately represented their communities in a Texas Senate hearing. 'We need the right to elect people who can focus on our issues and our needs,' said Michael Seifert, a self-employed Migrant Advocate. 'The present election maps, however, do not respect this right or our needs.'"

networks. Table 1 also shows how congruence is associated with socioeconomic, demographic, and political characteristics of counties' populations, as these characteristics are likely to be correlated with the outcomes of interest. County population, population density, and the share of the population that is rural comes from the 2020 Census. Clustering is, for an individual in the county, the average fraction of an individual's friend pairs who are also friends with each other in 2022, from Chetty et al. 2022. Biden vote shares are from MIT Election Data and Science Lab 2021. The remaining demographics coming from the 2015-2019 5-Year ACS.[13]

The determinants of district boundaries – whether the state has a single district, and the county's population – significantly predict congruence. In particular, as expected, counties in single district states are more congruent. Before adding the predictors of social networks and demographic characteristics (column 1), we see that population is negatively correlated with congruence; however, this reverses once we control for the county's population density, the share of the population that is rural, and the level of clustering in the county. That is, once we condition on those three correlates of social networks, a county with a higher population is more congruent. In line with the discussion above, congruence is decreasing in population density and increasing in the share of the county that is rural. However, once we control for demographic characteristics, population density no longer contributes any additional predictive power, except when we restrict to the top 50% most rural counties (these counties have at least 66% rural population, and most of them have 100% rural population). Additionally, the share of the population that is rural becomes negatively correlated with congruence once we control for demographics. Not surprisingly, counties with higher rates of clustering are more congruent in all specifications.

Turning to county demographics, we see that congruence is negatively correlated with the share of the population that was born outside of the U.S. and the share that moved within the last year: these are both indicators of more mobile populations, which intuitively are more likely to have more geographically dispersed networks. Congruence is positively predicted by the share of the population that is white as well as (though more marginally) by the share of the population without a high school degree and the share of the population below the poverty line. The latter two are strongly significantly associated with congruence when we restrict to the 50% least rural counties. Analogously, the share of the population with at least a college degree is negatively associated with congruence. Lastly, conditional on these geographic, network, and demographic features, Biden's vote share in 2020 is decreasing in congruence, particularly when we focus on the 50% most rural counties.

[13]As discussed in Section 2.2, I will measure congruence and outcomes over time. Due to the need to use a non-overlapping sequence of 5-Year ACS estimates, the 2015-2019 5-Year ACS is the one closest to 2020 that I use.

	Dependent Variable: Congruence					
	(1)	(2)	(3)	(4)	(5)	(6)
Single District State	0.050***	0.076***	0.095***	0.096***	0.109***	0.079***
	(0.008)	(0.007)	(0.006)	(0.006)	(0.008)	(0.010)
log(Population)	-0.026***	0.007***	0.011***	0.013***	0.009**	0.021***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.004)	(0.002)
Pop. Density		-12.14***	1.831	2.102	565.2**	-0.452
		(2.348)	(2.178)	(2.173)	(272.0)	(2.018)
% Pop. Rural		0.054***	-0.025***	-0.025***		
		(0.008)	(0.008)	(0.008)		
Clustering		2.911***	2.197***	2.317***	2.441***	2.483***
		(0.107)	(0.111)	(0.114)	(0.152)	(0.186)
% Foreign Born			-0.367***	-0.346***	-0.393***	-0.444***
			(0.040)	(0.040)	(0.077)	(0.048)
% Moved Last Year			-0.255***	-0.267***	-0.023	-0.355***
			(0.040)	(0.040)	(0.065)	(0.054)
% White			0.127***	0.112***	0.090***	0.126***
			(0.012)	(0.012)	(0.018)	(0.017)
% No HS			0.112*	0.067	-0.144*	0.418***
			(0.061)	(0.061)	(0.082)	(0.103)
% College/Grad			-0.549***	-0.485***	-0.519***	-0.382***
			(0.036)	(0.039)	(0.072)	(0.047)
% Below Poverty			0.070*	0.103***	0.054	0.223***
			(0.037)	(0.038)	(0.051)	(0.059)
% Votes for Biden				-0.066***	-0.137***	-0.015
				(0.014)	(0.024)	(0.018)
Sample					Top 50% Rural	Bottom 50% Rural
Observations	3,107	3,055	3,055	3,054	1,523	1,531
R^2	0.150	0.362	0.501	0.504	0.366	0.564

*** p<0.01, **p<0.05, * p<0.1

Notes: "Single District State" is a dummy variable indicating the state has only one congressional district. "Pop. Density" is the population per square meter. "% Foreign Born" is the share of non-citizens and naturalized citizens. "% Moved Last Year" is the share of people who moved within the last year, including those that moved within the same county. "% No HS" is the share of individuals 25 or over who do not hold a high school degree. "% College/Grad" is the share of individuals 25 or older who hold at least a college degree.

Table 1: Predictors of 2020 County Congruence with Congressional Districts

2.1.4 Changes in Congruence over Time

How is congruence changing over time? I examine changes in congruence following the 2010 and 2020 redistricting cycles. As described in more detail in Section 2.2, I re-calculate congruence for each year by holding the social network fixed. I then calculate the change in congruence for a county between 2012 and 2013 (the last year using the pre-2010 Census district boundaries vs. the first year using the post-2010 Census district boundaries) and between 2022 and 2023. In both years, the average change in congruence is nearly zero (0.09pp in 2013, 0.29pp in 2023), with a standard deviation of 5.3pp. In 2013, the biggest drop in congruence was by 29.3pp, while the biggest increase was by 37.4pp. In 2023, the biggest drop was by 26.5pp and the biggest increase was by 38.4pp. Following the 2010 Census, 445 counties (14%) experienced no change in congruence, and following the 2020 Census, 408 counties (13%) experienced no change in congruence. Thus, most counties experience some change in congruence, but very large changes are driven by outliers.

In Table 2, I regress the change in congruence on the same variables as in Table 1, including years 2013 and 2023 in the sample.[14] For 2013, I use 2010 Census estimates of population, population density, and rural share; 2010-2014 5-Year ACS estimates for demographics; and the same clustering measure and Biden vote shares as used for 2020. In Appendix Tables 7 and 8 I show analogous tables for 2013 and 2023 separately.

In general, few of the features considered here predict whether a county will become more or less congruent; indeed, across all specifications the R^2 value remains below 0.01. Most notably, counties with more clustering are more likely to become more congruent (this remains weakly true even when only considering 2013, as seen in Appendix Table 7), while counties with higher shares of college graduates are more likely to become less congruent.

2.2 Redistricting

Congruence measures variation in the match between social networks and congressional district boundaries, but it is not itself exogenous. Congruence is correlated with factors that determine district boundaries, factors that determine social networks, and sociodemographic characteristics. Accordingly, in order to have plausibly exogenous variation in congruence, I need to control for these factors, especially when they are possibly correlated with outcomes of interest: voters' knowledge of their representative and representatives'

[14] Though this time excluding the indicator for a single district, as the congruence for a county in a single district state cannot change under my methodology, unless the state were to gain an additional county. Because I re-calculate congruence holding the social network fixed, changes in district boundaries are the only thing that can create a change in congruence. Because state boundaries are not changing, a county in a single district state does not experience a change in congruence. Following 2020 re-districting, Montana gained a second congressional district, causing congruence to drop for counties across the state.

	Dependent Variable: Change in Congruence					
	(1)	(2)	(3)	(4)	(5)	(6)
log(Population)	-0.001**	0.001	0.001	0.002*	0.001	0.000
	(0.000)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Pop. Density		-0.451	0.002	0.023	-127.5	-0.581
		(1.010)	(1.058)	(1.058)	(138.6)	(0.935)
% Pop. Rural		0.006	0.006	0.005		
		(0.004)	(0.004)	(0.004)		
Clustering		0.113**	0.113**	0.126**	0.082	0.130
		(0.044)	(0.053)	(0.054)	(0.079)	(0.081)
% Foreign Born			0.013	0.016	-0.039	0.025
			(0.018)	(0.018)	(0.040)	(0.020)
% Moved Last Year			0.020	0.019	0.038	-0.011
			(0.019)	(0.019)	(0.034)	(0.023)
% White			0.007	0.006	0.020**	-0.008
			(0.005)	(0.005)	(0.009)	(0.007)
% No HS			-0.002	-0.007	0.017	-0.039
			(0.027)	(0.027)	(0.040)	(0.041)
% College/Grad			-0.040**	-0.033*	-0.085**	-0.014
			(0.018)	(0.019)	(0.037)	(0.022)
% Below Poverty			-0.002	0.002	-0.012	0.023
			(0.017)	(0.017)	(0.026)	(0.025)
% Votes for Biden				-0.007	0.004	-0.013*
				(0.007)	(0.012)	(0.008)
Sample					Top 50% Rural	Bottom 50% Rural
Observations	6,214	6,110	6,110	6,108	3,046	3,062
R^2	0.001	0.003	0.005	0.005	0.007	0.005

*** p<0.01, **p<0.05, * p<0.1

Notes: "Pop. Density" is the population per square meter. "% Foreign Born" is the share of non-citizens and naturalized citizens. "% Moved Last Year" is the share of people who moved within the last year, including those that moved within the same county. "% No HS" is the share of individuals 25 or over who do not hold a high school degree. "% College/Grad" is the share of individuals 25 or older who hold at least a college degree.

Table 2: Changes in County Congruence with Congressional Districts

effort.[15]

In order to capture plausibly exogenous variation in congruence, I measure the impact on outcomes of a change in congruence driven by congressional redistricting. This allows me to use county fixed effects to control for unobserved location-specific confounders. The intuition is: holding the county fixed, what is the impact of an increase (or decrease) in congruence driven by redistricting?[16]

Because the SCI is not intended to be used as a panel, I hold the social network fixed (i.e., I always use the October 2021 snapshot) and I re-calculate congruence with each congressional border change.[17] This relies on the assumption that social networks are slow-changing, which is reasonable given studies like Bailey, Gupta, et al. 2021, (which finds that countries with higher social connectedness trade more, and that this relationship is similar in every year back to 1980) and Kuchler et al. 2022 (which finds that institutional investors are more likely to invest in firms in regions with higher social connectedness to the investor's region, and that this relationship remains at least back to 2007). Additionally, Enke et al. 2023 note that the correlation between years of the SCI is above .99. I construct congruence for each year from 2005 to 2023, i.e. looking at the 109th-117th Congresses.

The general specification (supposing an individual-level outcome) then is

$$y_{ict} = \alpha + \beta_0 \text{Congruence}_{ct} + \kappa_c + \lambda_t + X_{ct}\delta + Z_{ict}\gamma + \varepsilon_{ict} \tag{1}$$

where y_{ict} is the outcome for a given individual i in county c in year t, α is an intercept, Congruence_{ct} is the congruence of county c in year t, κ_c are county fixed effects (to capture time-invariant location-specific confounders), λ_t are year fixed effects (to capture shocks specific to a particular year/election), X_{ct} is a vector of county-by-year controls (to further adjust for things like changing demographics over time), and Z_{ict} is a vector of individual controls. Errors ε_{ict} are clustered at the county level.

[15]In particular, these predictors of congruence are of concern to the extent that they are correlated with outcomes of interest for reasons not driven by the extent to which the boundaries reflect the social network. For example, if congressional districts were drawn randomly and my county suddenly became very interior to the district, it is not obvious that *being interior alone* should matter for the extent to which I am informed about my representative and my representative's effort – except for the fact that it means that more of my social network is likely to lie within the same district. However, it may be that my county can only be so interior to the district because my county is rural and has low population density, and so the district that contains it can be geographically larger; because people in rural areas have different political leanings, interests, and behavior for reasons separate from the match between their social networks and their districts, this is more concerning.

[16]In essence, I am using a two-way fixed effects model with a continuous treatment, where the change in congruence due to a redistricting event is the treatment. In future versions of this paper, I plan to apply estimators from Callaway et al. 2024 or de Chaisemartin et al. 2023, recent working papers that address the issues with two-way fixed effects estimators in this kind of setting.

[17]Specifically, I assume that the share of a county's friends in each other county remains the same. To do this, I also assume that the populations are the same as Decennial Census 2020 populations (otherwise changes in population of one county would affect friendship shares for many other counties). Alternatively, one could assume the SCI of each county pair remains the same, but recalculate the friendship shares every five years (using the 5-year ACS) or every 10 years (using the Decennial Census).

In alternative specifications, I also include district-by-year fixed effects. This can be thought of as controlling for district election-specific factors that impact outcomes for all counties in the district. These can include characteristics of the candidates, scandals, national attention, etc. The intuition for this specification then is: holding the county fixed and controlling for district election-specific shocks, what is the impact of an increase (or decrease) in congruence driven by redistricting? As another alternative, I include state-by-year fixed effects, which can address concerns about substantial changes in congruence being predicted by a state's changing population size (causing the state to gain/lose seats during reapportionment).

When studying outcomes related to representative effort, outcomes are at the representative (i.e., district) level. For these outcomes, I construct instead "district congruence," which is the share of a district's friends that are in that district (see A.1.1). Accordingly, we can no longer use county fixed effects. Further, district fixed effects cannot reliably be used as a substitute, as the geography included within a district usually changes with each redistricting event, and so there is no stable geographic unit over time. Consequently, while redistricting creates additional variation in district congruence, it cannot provide identification for the impact of district congruence on representative effort. I include representative and district controls in order to address likely observable confounders, but the claim for identification relies on assuming no unobservable confounders remain, and is necessarily weaker.

3 Outcomes Data and Descriptive Statistics

3.1 Voters' Information

In order to study the impact of congruence on voters' familiarity with their representatives, I use responses in the Cooperative Election Study (CES) (formerly the Cooperative Congressional Election Study, or CCES; see for example Schaffner, Ansolabehere, and Shih 2023). The CES is a nationally representative survey that has run annually from 2006 to 2022 and ask about topics including demographics, political attitudes, political knowledge, and voting intentions and choices. In federal election years (i.e., all even years), a pre-election survey is conducted from late September to late October, and a post-election survey is conducted in November. In non-federal election years (i.e., all odd years) a single survey is conducted in the fall. I use the pre-election surveys (or single surveys in odd years) for 2006-2022. The CES sample consists of 50,000+ adults in every federal election year since 2010 (>30,000 in 2006 and 2008) and 10,000+ adults in every odd year. I use the CES's cumulative weights, which re-weight observations to make sample sizes comparable across years (see Kuriwaki 2018). The CES includes each respondent's county and congressional district, enabling me to link respondents to county-level congruence measures and to observe responses to questions about each respondent's own representative.

I construct three binary variables to assess how familiar respondents are with their current representative. Brief descriptions of these variables are in in Table 3. Respondents are asked to "Please indicate whether you've heard of this person and if so which party he or she is affiliated with...". They are asked this about their current House representative, both of their senators, and their governor. Respondents can answer "Never Heard of Person", "Republican", "Democrat", "Other Party/Independent", or "Not Sure". The first dummy variable, "Heard of Incumbent", is coded as 0 if the respondent answered "Never Heard of Person" and 1 otherwise.[18] This variable captures whether the respondent claims to have any familiarity with their representative at all: do they even recognize the name? The second dummy variable, "Selected Party", is coded as 0 if the respondent answered "Never Heard of Person" *or* "Not Sure", and 1 otherwise. This variable indicates whether, beyond recognizing the representative, the respondent claims to have some knowledge about them: they claim to know the party the representative belongs to (though they may just be guessing). Lastly, the third dummy variable, "Selected Correct Party", is coded as 1 if the respondent selected the correct party for the incumbent and 0 otherwise. While lucky guesses cannot be ruled out, this variable generally indicates that the respondent at least knows enough about their representative to know what party their representative belongs to. Table 4 shows that, as expected, fewer people select their representative's party than claim to have heard of them, and fewer still select the correct party (though, among those who select a party, the overwhelming majority select the correct party).

As when constructing the congruence measure, I only include respondents in the 48 contiguous states.[19] Additionally, not all counties are represented in every year; in even years, there is at least one respondent from 80-90% of counties, while in odd years about two-thirds of counties have at least one respondent. Because the weighted sample is representative of people living in the U.S. (rather than of U.S. counties) and more people live in urban areas (which tend to have lower congruence), the average respondent's county

[18]In 2006, 2007, and 2009, respondents do not have the option to say "Never Heard of Person" and instead can only say "Not Sure." Consequently, I drop these years in regressions using the "Heard of Incumbent" variable. Note that 2008, 2010, and 2011 still provide observations of this variable prior to the redistricting that follows the 2010 Decennial Census, because the new districts first apply in 2012.

[19]In these 48 states and across all 17 years of the CES, there are 612,085 respondents (552,307 excluding 2006, 2007, and 2009). I exclude missing responses to the candidate party recognition question (<2% of respondents in each year; for most of these cases, the House candidate name is missing in the survey). When including individual demographic controls, I similarly exclude respondents who did not answer the relevant demographic questions. I also exclude a small number of respondents in 2006 and 2007 that are assigned to counties that are not in their state of residence. Lastly, in the 2020 survey, 925 respondents in North Carolina were assigned to incorrect congressional districts, and consequently were shown the candidate names for the wrong district. I exclude these respondents, since they were not asked about their familiarity with their own representative. See Schaffner, Ansolabehere, and Luks 2021.

Variable	Description
Heard of Representative	When shown the name of their current House representative and asked to indicate the party their representative is affiliated with, respondent did not indicate they had "Never Heard of Person" , and they instead chose "Republican", "Democrat", "Other Party/Independent", or "Not Sure". Binary. From Cooperative Election Study.
Selected Party	When shown the name of their current House representative and asked to indicate the party their representative is affiliated with, respondent did not indicate they had "Never Heard of Person" or "Not Sure", and they instead chose "Republican", "Democrat", or "Other Party/Independent". Binary. From Cooperative Election Study.
Selected Correct Party	When shown the name of their current House representative and asked to indicate the party their representative is affiliated with, respondent chose the correct party. Binary. From Cooperative Election Study.
Share Votes Against Leaders	The share of a representative's votes in a given Congress that are opposite the vote of the majority of the representative's party's leadership. Percentage. From the Voteview.
Total Bills Sponsored	The total number of bills a representative sponsored during a given Congress. Count. From Legislative Effectiveness Data.
LES/Benchmark	The ratio of a representative's actual Legislative Effectiveness Score to their predict Legislative Effectiveness Score. Continuous. From Legislative Effectiveness Data.
Distributive Committee	Indicates a majority of a representative's committee assignments were to a distributive committee (Agriculture, National Security or Homeland Security, Resources, Science, Small Business, Transportation & Infrastructure, or Veterans Affairs) during the given Congress. Binary. From Stewart 2021.
No Policy Committee	Indicates the representative did *not* serve on any policy committee (House Judiciary Committee, International Relations, or Foreign Affairs) during the given Congress. Binary. From Stewart 2021.

Table 3: Descriptions for Outcome Variables

Variable	Observations	Mean (%)	SD (pp)
Heard of Incumbent	545,185	93.2	25.2
Selected Party	604,254	68.6	46.4
Selected Correct Party	604,254	61.7	48.6

Table 4: CES Data: Summary Statistics

congruence is slightly lower at 37% (compared to 45% for the average county).[20]

3.2 Representative Effort

In order to measure the impact of congruence on representatives' effort, I construct measures of effort based on representatives' voting behavior, sponsored bills, and committee memberships. Across outcomes, I focus on the 109th-117th Congresses (covering 2005-2022), yielding 3,943 representative-Congress observations.[21]

Representatives who are more responsive to their constituents might be expected to more frequently vote against their party (see the argument in Snyder and Strömberg 2010). The intuition is that representatives pay a cost if they vote against their party's leadership: party leaders have multiple tools to hold their members in line, as they have much control over legislative agenda setting, committee assignments, and the flow of debate on the Senate floor (United States Senate n.d.). Consequently, a representative would only be willing to pay the cost of opposing the party's leadership if the representative believes that they will be commensurately rewarded by their constituents for doing so. I use data on roll call votes from the Voteview: Congressional Roll-Call Votes Database (Lewis et al. 2022). Voteview contains every representative's vote on every roll call vote in each Congress. I label a representative as voting against the party if their vote differs from that of the majority of their party's leadership (as in Snyder and Strömberg 2010), and I construct for each representative the share of all votes during a given two-year Congress in which they voted against their party's leadership.[22] As seen in Table 5, representatives vote against their party's leadership on average 10.6% of the time. The representative at the 10th percentile votes against their leaders in 3.6% of votes, while the representative at the 90th percentile votes against their leaders in 18.3% of votes.[23]

A straightforward measure of representative effort is the number of bills they co-sponsor (another key measure used in Snyder and Strömberg 2010). For information on bill sponsorships, I use the Legislative Effectiveness Data from the Center for Effective Lawmaking (Volden and Wiseman 2023a). This data contains information on the number of bills each representative (co-)sponsored, both in total and categorized by how far the bill got towards becoming law; bills are also categorized by whether they are "commemorative",

[20] The distribution is otherwise similar to the county-level distribution, with a standard deviation of 11pp, a minimum of 9% and a maximum of 74%.

[21] There are 435 voting members of the House of Representatives, and 431 once Alaska and Hawaii are excluded. However, during the two-year span of any given Congress, some representatives exit their seats (due to deaths, resignations, etc.) and others replace them, so there are a larger number of representatives that serve at any point during a Congress.

[22] Alternative definitions – such as flagging votes that differ from the vote of the majority of the party's *members*, or subsetting to close elections – make little difference for the results.

[23] Outliers are rare and often driven by unusual circumstances: for example, the lone representative whose votes were 100% against party leadership in the given Congress died a month into the Congress (Walter B. Jones, Jr. of North Carolina's 3rd Congressional District in the 116th Congress), so this average is over very few votes.

"substantive", or "substantive and significant", which broadly corresponds to the extent to which the bill was a major vote (see Volden and Wiseman 2023b). Here I focus only on the total number of bills sponsored, but separately examining by the stage of the legislative process reached or whether bills are substantive yields similar results. The average representative sponsors 16 bills in a given Congress; the representative at the 10th percentile sponsors 5 bills, while the representative at the 90th percentile sponsors 30 bills.

The Legislative Effectiveness Data also includes several measures of "legislative effectiveness," including a legislative effectiveness score (LES) that is an index that accounts for the number of bills a representative got to each stage of the lawmaking process, weighted by the substantiveness of these bills, and in comparison to all other representatives. A predicted LES is calculated for each representative, based on the representative's seniority, majority party membership, and chairing of committees or sub-committees (Volden and Wiseman 2023b). The ratio of the actual LES to the benchmark LES then gives a measure of whether the representative met expectations, given their characteristics. I focus on this ratio, but the other measures of legislative effectiveness yield similar results. The average representative has a ratio right around 1, indicating performing as expected. The representative at the 10th percentile has a ratio of 0.14, while the representative at the 90th percentile has a ratio of 2.17.

Lastly, whether a representative is able to focus on bills benefitting their district or has their attention diverted to national issues will depend largely on their committee assignments. As again discussed in Snyder and Strömberg 2010, representatives assigned to "distributive" committees (Agriculture, National Security/Homeland Security, Resources, Science, Small Business, Transportation & Infrastructure, Veterans Affairs) have more power over bills that allocate funding to domestic projects, and consequently have more scope to negotiate funding that will benefit their own constituents. Conversely, representatives assigned to "policy" committees (House Judiciary Committee, International Relations, Foreign Affairs) spend substantial time on national-level issues, and this could distract from their ability to focus on their own district. Accordingly, a representative that is heavily monitored by their constituents may try to serve on distributive committees in order to deliver more funding to their districts, while a representative who is less concerned that voters will respond to their policy choices may have less to lose from serving on policy committees. Using Stewart 2021's data on committee assignments, I construct binary variables indicating (1) whether a majority of a representative's committee assignments in a given Congress were on distributive committees, (2) whether a representative served on *any* policy committee in a given Congress (following Snyder and Strömberg 2010). On average, 36.5% of representatives are serving on mostly distributive committees, while 18.5% of representatives are serving on some policy committee.

Variable	Observations	Mean	SD	Min	Max
Share Votes Against Leaders	3,943	10.6%	7.4%	0%	100%
Total Bills Sponsored	3,943	16.1	11.4	0	120
LES/Benchmark	3,943	1.0	1.0	0	13.7
Distributive Committee	3,888	36.5%	48.1%	0%	100%
Policy Committee	3,888	18.5%	38.8%	0%	100%

Table 5: Effort Data: Summary Statistics

4 Results

4.1 Voters' Familiarity with Representatives

First, I examine whether congruence impacts voters' familiarity with their representatives. Voters that are more familiar with their representative may be more effective at monitoring their representatives, if they are also more likely to be aware of their representative's actions in Congress. Indeed, across specifications and outcomes, I estimate positive and significant coefficients for the impact of congruence on voters' familiarity with their representatives.

Table 6 presents results from equation 1 with outcomes "Heard of Incumbent", "Selected Party", and "Selected Correct Party" constructed from the CES data. I estimate coefficients using linear probability models. All five columns include county and year fixed effects, column 2 adds district-by-year fixed effects, column 3 adds individual demographic controls (as reported in the CES), and column 4 adds county-by-year demographic controls (constructed using 5-year ACS estimates and Decennial Census estimates).[24] Congruence is measured on a scale from 0 to 1, and outcome variables are binary. As such, reported estimates give the change in probability of the outcome (measured between 0 and 1) that would result from a 0 to 1 change in congruence.

Across columns (1)-(4), estimated coefficients are positive and significant at at least the 5% level. Estimates generally increase moving right across these columns. The increasing magnitudes as additional observables are controlled for suggests that unobservable confounders may remain. Reassuringly, in order for the actual impact of congruence to be zero or negative, such unobservables would need to bias estimates in the opposite direction of the observables.

Consider the simplest (and most conservative) specification given in column (1). If we assume linear

[24] Note that when including fixed effects, singleton observations are dropped. For example, if there is only one respondent in a county, that county is dropped when including county fixed effects. Similarly, if there is only one respondent in a district in a given year, that district is dropped when including district-by-year fixed effects. Consequently, sample sizes are slightly smaller than those reported in 4.

impacts, an increase in a county's congruence from the minimum value observed in the data (8%) to the maximum value observed (74%) would increase the probability that a respondent in that county has heard of their representative by 6.3pp (recall from Table 4 that the mean is 93.2%). The same change in a county's congruence would increase the probability a respondent in that county selects a party by 16pp (from mean 68.6%) and selects the correct party by 11.8pp (from mean 61.7%).[25]

4.1.1 Event Studies

In order to dig deeper into the dynamic effects of a change in congruence, I run event studies. I focus on the redistricting that followed the 2010 Census, and I focus on even years of the CES survey until 2022 (the last year before the next national redistricting event).[26] The Census was conducted in April 2010, and states needed to draw new congressional district borders in time for the November 2012 elections. Accordingly, the congressional representatives first elected under the new borders assumed office in January 2013. Because the CES questions ask about *current* representative, under this timeline 2012 would be the last year before the "treatment" of a change in congruence occurring in January 2013.The event studies accordingly take the following form:

$$y_{ict} = \kappa_c + \lambda_t + \sum_{\tau=2006}^{\tau=2010} \beta_\tau \Delta\text{Congruence}_c \mathbb{I}(\tau = t) + \sum_{\tau=2014}^{\tau=2022} \beta_\tau \Delta\text{Congruence}_c \mathbb{I}(\tau = t) + X_{ct}\delta + Z_{ict}\gamma + \varepsilon_{ict}$$

where $\Delta\text{Congruence}_c$ is the change in congruence experienced by county c between 2012 and 2013, and the other variables are as in Equation 1.

The event studies show that the change in voter knowledge due to changes in congruence in redistricting most strongly takes effect in the first survey after redistricting (2014). Impacts are relatively stable (or, if anything, decreasing) over time.

[25]Recall from Section 2.1.4 the a one standard deviation change in congruence following redistricting is roughly 5pp. An increase in congruence by 5pp increases the probability of having heard of the incumbent by 0.5pp, of selecting a party by 1.2pp, and of selecting the correct party by 0.9pp. The largest changes in congruence following redistricting are normally by around 30pp. Such an increase in congruence increases the probability of having heard of the incumbent by 2.9pp, of selecting a party by 7.3pp, and of selecting the correct party by 5.4pp.

[26]The odd years have a sample about one-fifth the size of even years. As such, including odd years yields similar but noisier results. The event studies thus reflect voters' knowledge of their current representative shortly before the election that will replace or re-elect that representative.

	(1) County & Year FEs Only	(2) Add District x Year FEs	(3) Add Individual Demographic Controls	(4) Add County-Year Controls
		Heard of Incumbent		
Congruence	0.096**	0.145***	0.156***	0.159***
	(0.038)	(0.043)	(0.043)	(0.043)
	[0.012]	[0.001]	[0.000]	[0.000]
Obs	545,127	545,127	545,127	545,127
R^2	0.032	0.078	0.129	0.129
		Selected Party		
Congruence	0.243***	0.289***	0.339***	0.345***
	(0.078)	(0.074)	(0.069)	(0.069)
	[0.002]	[0.000]	[0.000]	[0.000]
Obs	604,209	604,208	604,208	604,208
R^2	0.042	0.098	0.211	0.211
		Selected Correct Party		
Congruence	0.179**	0.278***	0.332***	0.337***
	(0.084)	(0.078)	(0.072)	(0.072)
	[0.034]	[0.000]	[0.000]	[0.000]
Obs	604,209	604,208	604,208	604,208
R^2	0.044	0.108	0.236	0.236
Ind. Controls			X	X
County x Year Controls				X
FEs	County, Year	County, Year, District x Year	County, Year, District x Year	County, Year, District x Year

*** $p<0.01$, **$p<0.05$, * $p<0.1$

Standard errors clustered at the county level in parentheses. P-values in square brackets.

"Heard of Incumbent" not available in 2006, 2007, or 2009. Individual controls include gender, race, education, age categories, and whether the respondent is affiliated with the same party as their representative. County-by-year controls include population and shares by race, age categories, gender, and county urban population share.

Table 6: Effect of Congruence on Voter Familiarity with Representative

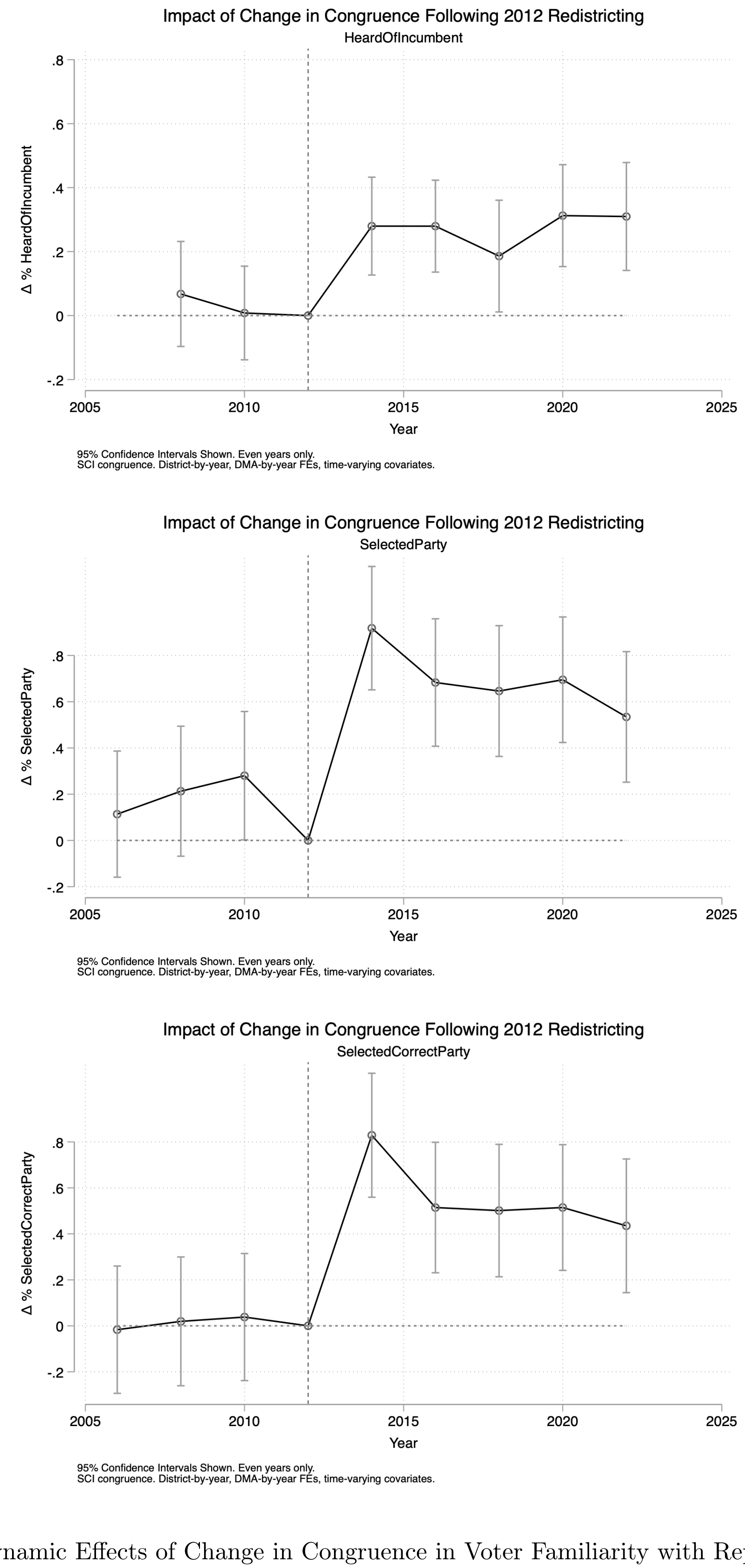

Figure 2: Dynamic Effects of Change in Congruence in Voter Familiarity with Representative

4.2 Representative Effort

Second, I study whether congruence impacts representatives' effort. Having found positive estimates of the effect of congruence on voters' familiarity with their representatives, it is plausible that the more informed voters in more congruent areas are better able to hold their representatives accountable. If so, we should expect to see that representatives in more congruent districts exhibit higher effort, all else equal. Instead, I find insignificant impacts of congruence on representative effort, though estimates are noisy.

As discussed in Section 2.2, I use district-level congruence as the explanatory variable when studying measures of representative effort. Redistricting provides variation in district-level congruence, but no district can experience a change in congruence without its boundaries changing, so we cannot observe what happens when the same district experiences a change in congruence. Accordingly, identification of the representative effort outcomes relies on the assumption that we have controlled for all relevant confounders. To control for district characteristics, I use ACS 1-year estimates of district demographics (averaged across both years of a Congress).[27] Using characteristics reported in the Legislative Effectiveness Data, I control for whether a representative is the Speaker of the House, part of the majority leadership, part of the minority leadership,[28] a freshman, as well as the number of terms the representative has served, the year the representative was first elected, the representative's age, the size of the delegation from the representative's state, and the closeness of the election that elected the representative.

Figure 3 summarizes point estimates for the five measures of representative effort summarized in Table 5, as well as for an index constructed from the five measures. Note that all measures shown in the figure have been standardized to have mean zero and variance one. The indicator for being on a policy committee has been negated, such that for all outcomes a positive value indicates higher effort on the district constituents' behalf. Put differently, points to the right of the vertical red line indicate higher effort, while points to the left indicate lower effort. Accordingly, the index is constructed by taking the average of the five standardized measures.

The bars surrounding points indicate 90%, 95%, and 99% confidence intervals, in order of decreasing thickness. The green (highest) points show the coefficient when district congruence is simply regressed on each outcome, without including any controls. The orange points add representative-by-Congress controls

[27]Controls include population, median income, and shares by race, citizenship, nativity, age categories, gender, education, poverty status, and moves within the past year.

[28]For the Democratic party, the leadership consists of the Speaker, Assistant Speaker, Majority Leader, Majority Whip, Democratic Caucus Chairman and Vice-Chairman, and Democratic Steering Committee Co-Chairmen (eight leaders, as the Democrats had the majority and hence the speaker). For the Republican party, the leadership consists of the Minority Leader; Minority Whip and Deputy Whip; Republican Conference Chairman, Vice-Chairman, and Secretary; and the Republican Policy Committee Chairman (seven leaders).

and district-by-Congress controls. The blue points incorporate state fixed effects and Congress fixed effects, while the maroon points instead use state-by-Congress fixed effects. Lastly, the light turquoise points include congressional district fixed effects, though as discussed this is a spurious concept because the land included in a congressional district changes following redistricting, and so fixed effects will be very noisily measured.[29]

Accordingly, the specifications including state and Congress or state-by-Congress fixed effects (navy and maroon points) indicate my preferred estimates of the impact of congruence on representative effort. Focusing on these specifications, we see that none of the measures are significant at the 90% level. While point estimates are positive for the predicted-to-actual LES ratio, the share of votes against leaders, and the indicator of serving on a majority distributive committees, these are not enough to result in a substantially positive point estimate when looking across all measures in the index. In particular, point estimates for the number of bills sponsored are actually negative.[30]

5 Robustness

I find in my main specification that congruence has a positive effect on voters' familiarity with their representatives. I explore the robustness of this finding by testing whether congruence impacts placebo outcomes, by accounting for media markets, and by constructing an alternative measure of congruence using commuting flows.

5.1 Placebo Outcomes

I test whether congruence impacts voters' familiarity with their governor and senators: because these offices are elected through statewide elections, and consequently congressional district borders are not relevant for them, congruence should not impact them.[31] CES respondents answer similar questions about whether they have heard of and can identify the party of their governor and each of their senators. From these responses, I construct outcome variables analogous to the ones in the main analysis, and which measure whether voters

[29]While many districts mostly stay the same following redistricting (recall that the standard deviation of the change in congruence following redistricting is 5pp), a district that experiences a large change of congruence is precisely one for which the boundaries of the district meaningfully changed.

[30]The specification that includes district fixed effects indicates positive point estimates that are significant at the 1% level across all measures except the number of bills sponsored and the share of votes against leaders, despite very wide confidence intervals. While the meaning of this specification is ambiguous, these estimates are suggestive that more confounders likely remain that obscure the impact of congruence.

[31]These are not perfect placebos: a county that is more congruent with its congressional district may also be more likely to be more congruent with its state (consider a couple rural districts in a rural state). However, this would push me to find significant estimates of the impact of congruence on the placebo outcomes, even if congruence is appropriately capturing the alignment of voters' social networks with their congressional districts.

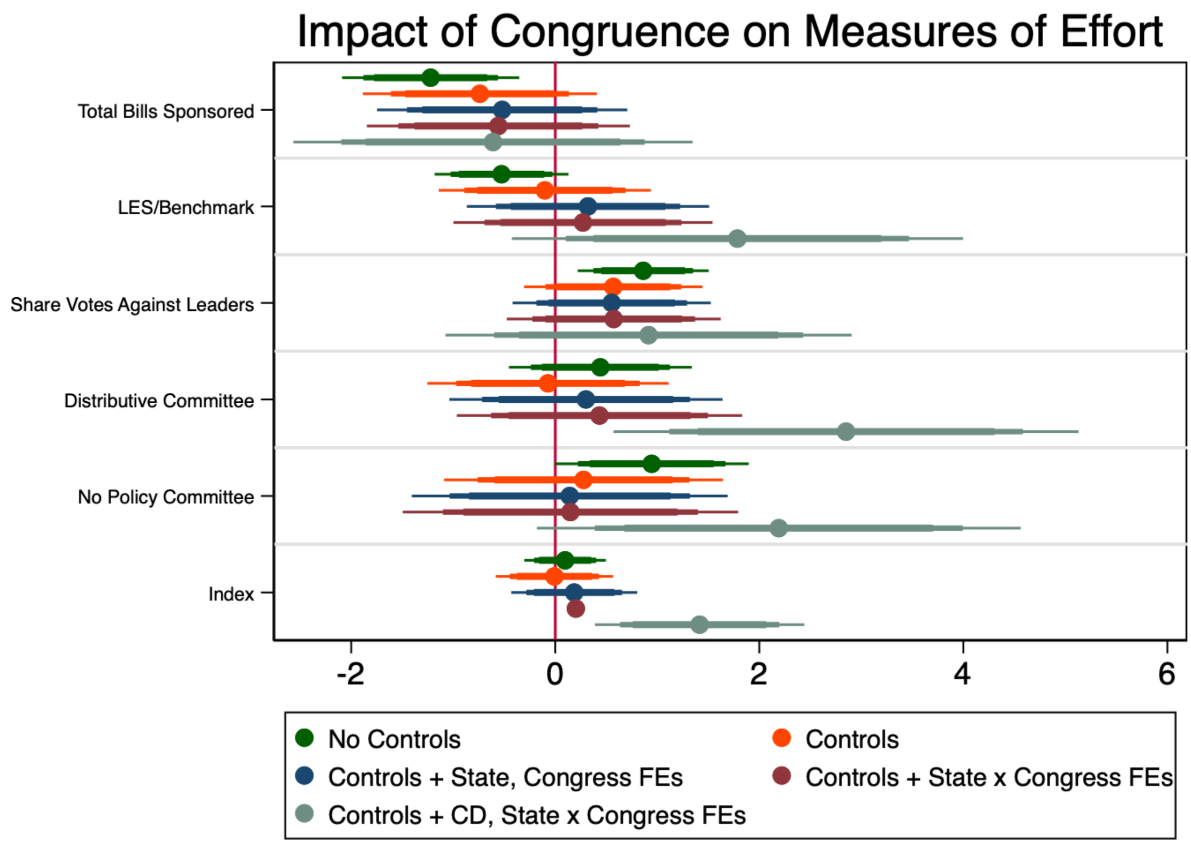

Figure 3: Effect of Congruence on Measures of Representative Effort

have heard of, select a party for, and select the correct party for their governor and their senators. These variables are summarized in Appendix Table 9.

I find scant evidence for an impact of congruence on these nine outcomes. Results are reported in Appendix Tables 10-12.[32]

5.2 Media Markets

One concern is that social networks may simply map closely to existing media markets, and as such that any effect I pick up is simply driven by the role of media in voter information described in prior literature (Snyder and Strömberg 2010,Prat and Strömberg 2005,Enikolopov, Petrova, et al. 2011,Angelucci et al. 2020,Eisensee and Strömberg 2007,Strömberg 2004b). To address this, I use Nielsen Designated Market Area (DMA) regions, which reflect TV and radio markets, and construct a measure of "DMA congruence". This is based on the logic in Snyder and Strömberg 2010 that a media market that is better aligned with a congressional district will produce more content about the district's representative, thereby making voters in that district more informed about their representative.[33] DMAs contain many counties, but DMA borders follow county borders, so no county is in multiple DMAs. For county i in district J, which contains counties j, and in DMA M, which contains counties m, DMA congruence is defined as

$$\text{DMA Congruence}_i = \frac{\sum_{j \in (J \cap M)} \text{Population}_j}{\sum_{m \in M} \text{Population}_m}$$

I then test whether I estimate a significant effect of my original congruence measure after controlling for DMA congruence.[34] Results are shown in Appendix Table 13. In general, results are similar to the original specification.

[32]While point estimates are significant at the 5% level in 2 cases and at the 10% level in 2 cases, over the 36 specifications ran this is not far beyond (though still more than) what is expected from random chance. One of these instances is when considering whether respondents select the correct party for senator 1, and for this estimates become much smaller and insignificant as soon as controls are added. The other three instances occur when considering whether respondents select a party for the governor. This is somewhat more concerning, but it should be noted that (1) estimates again become slightly smaller and less significant as controls are added, and (2) point estimates are negative, which is the opposite direction of the effects found when considering representatives.

[33]Of course, Snyder and Strömberg 2010 actually find no effect of TV and radio market congruence in their paper (though it is constructed somewhat differently), and instead find effects of newspaper market congruence, which I have not yet included here.

[34]DMA congruence and the originally defined county congruence are about 25% correlated.

5.3 Commuting Flows as an Alternative Network Measure

Commuting flows can be used as an alternative measure of social networks: the number of people that commute between two counties reflects patterns of who is regularly physically proximate to each other. Replicating the analysis using commuting flows can shed light on the extent to which the SCI captures "real world" offline networks. I use the 2016 5-Year ACS County-County Commuting Flows, which report the average number of people that commute between two counties, and I construct commuting congruence as the share of a county's commuters that stay within the county's district when commuting. For county i in district J (which contains counties j) and all US counties K (which contains counties k),

$$\text{Commuting Congruence}_i = \frac{\sum_{j \in J} \text{Commuters}_{i,j}}{\sum_{k \in K} \text{Commuters}_{i,k}}$$

In Appendix Table 14, I report results for the effect of commuting congruence on voters' familiarity with their representatives. Estimates are of smaller magnitudes but otherwise are similar: across specifications, commuting congruence has a positive effect on measures of voters' familiarity with their representatives, with significance at at least the 5% level for all but one estimate. I interpret the smaller estimates as reflecting the fact that commuting flows are a rougher approximation of social networks than the SCI. Additionally, the larger effects when using the SCI to construct congruence likely also reflect use of Facebook to share news about representatives.[35]

6 Conclusion

Counties across the U.S. vary substantially in their social cohesiveness with their congressional district – their congruence. While people living in the average county share a representative with about half of the county's friends, this varies from 6% to 74%. I show that congruence increases voters' familiarity with their representative: when a county becomes more congruent due to redistricting, voters are more likely to recognize the name and know the party of their representative. I find similar results regardless of whether I construct the network using the SCI or commuting flows, which strengthens the case that these impacts are not unique to Facebook users.

However, I do not find clear evidence of this translating into an impact of congruence on measures of representative effort. The measures of representative effort are at the representative-congress level, and consequently effects would need to be large to be detected with my estimation strategy. Future research

[35] I have verified that results are again similar if I include DMA congruence in specifications using commuting congruence.

could explore more granular measures (such as county-level federal funding, or analyses of representatives' speech or websites). Additionally, my estimates are limited by the fact that I do not observe exogenous changes in a district's congruence: if the district's congruence changes, then the district itself must have changed

In future versions of this paper, I plan to examine voter turnout and incumbency advantage. Because I find that congruence matters for voters' information, I will study whether this is reflected in changes in behavior at the voting booth. If so, my results will inform policy by providing new evidence on gerrymanderers' incentives and impacts: a gerrymanderer may try to draw district boundaries to strategically manipulate turnout rates (as gerrymanderers already pay attention to turnout when drawing boundaries, shown in Bouton et al. 2023). This evidence is especially important as detailed social network data, like the SCI, has become publicly available for the first time in recent years, and consequently usable for gerrymandering. In order to understand the mechanisms underlying my results, I will also study the impact of congruence on other measures of voters' engagement (e.g., campaign volunteering and campaign donations) as well as on House candidates' responses (e.g., their targeting of political advertising).

References

Angelucci, C., Cagé, J., & Sinkinson, M. (2020). *Media Competition and News Diets* (Working Paper No. 26782). National Bureau of Economic Research. https://doi.org/10.3386/w26782

Auxier, B., & Anderson, M. (2021). Social Media Use in 2021.

Bailey, M., Cao, R., Kuchler, T., Stroebel, J., & Wong, A. (2018). Social Connectedness: Measurement, Determinants, and Effects. *Journal of Economic Perspectives*, *32*(3), 259–280. https://doi.org/10.1257/jep.32.3.259

Bailey, M., Gupta, A., Hillenbrand, S., Kuchler, T., Richmond, R., & Stroebel, J. (2020). International Trade and Social Connectedness, 55.

Bailey, M., Gupta, A., Hillenbrand, S., Kuchler, T., Richmond, R., & Stroebel, J. (2021). International trade and social connectedness. *Journal of International Economics*, *129*, 103418. https://doi.org/10.1016/j.jinteco.2020.103418

Besley, T., & Burgess, R. (2002). The Political Economy of Government Responsiveness: Theory and Evidence from India. *The Quarterly Journal of Economics*, *117*(4), 1415–1451. https://doi.org/10.1162/003355302320935061

Bessone, P., Campante, F., Ferraz, C., & Souza, P. (2022). *Social Media and the Behavior of Politicians: Evidence from Facebook in Brazil* (tech. rep. w30306). National Bureau of Economic Research. Cambridge, MA. https://doi.org/10.3386/w30306

Bouton, L., Genicot, G., Castanheira, M., & Stashko, A. (2023). Pack-Crack-Pack: Gerrymandering with Differential Turnout.

Callaway, B., Goodman-Bacon, A., & Sant'Anna, P. (2024). Difference-in-Differences with a Continuous Treatment. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4716682

Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., Johnston, D., Koenen, M., Laguna-Muggenburg, E., Mudekereza, F., Rutter, T., Thor, N., Townsend, W., Zhang, R., Bailey, M., . . . Wernerfelt, N. (2022). Social capital I: Measurement and associations with economic mobility. *Nature*, *608*(7921), 108–121. https://doi.org/10.1038/s41586-022-04996-4

de Chaisemartin, C., D'Haultfoeuille, X., Pasquier, F., & Vazquez-Bare, G. (2023). Difference-in-Differences Estimators for Treatments Continuously Distributed at Every Period.

Degroot, M. H. (1974). Reaching a Consensus. *Journal of the American Statistical Association*, *69*(345), 118–121. https://doi.org/10.1080/01621459.1974.10480137

Di Tella, R., Gálvez, R. H., & Schargrodsky, E. (2021). *Does Social Media cause Polarization? Evidence from access to Twitter Echo Chambers during the 2019 Argentine Presidential Debate* (Working Paper No. 29458). National Bureau of Economic Research. https://doi.org/10.3386/w29458

Eckman, S. J. (2021). *Apportionment and Redistricting Process for the U.S. House of Representatives* (tech. rep. R45951). Congressional Research Service.

Eisensee, T., & Strömberg, D. (2007). News Droughts, News Floods, and U. S. Disaster Relief. *The Quarterly Journal of Economics*, *122*(2), 693–728. https://doi.org/10.1162/qjec.122.2.693

Enikolopov, R., Makarin, A., & Petrova, M. (2020). Social Media and Protest Participation: Evidence From Russia. *Econometrica*, *88*(4), 1479–1514. https://doi.org/10.3982/ECTA14281

Enikolopov, R., Petrova, M., & Zhuravskaya, E. (2011). Media and Political Persuasion: Evidence from Russia. *American Economic Review*, *101*(7), 3253–3285. https://doi.org/10.1257/aer.101.7.3253

Enke, B., Fisman, R., Freitas, L. M., & Sun, S. (2023). *Universalism and Political Representation: Evidence from the Field* (tech. rep. w31265). National Bureau of Economic Research. Cambridge, MA. https://doi.org/10.3386/w31265

Fafchamps, M., Vaz, A., & Vicente, P. C. (n.d.). Voting and Peer Effects: Experimental Evidence from Mozambique.

Ferraz, C., & Finan, F. (2008). Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes. *The Quarterly Journal of Economics*, *123*(2), 703–745. https://doi.org/10.1162/qjec.2008.123.2.703

Guriev, S., Melnikov, N., & Zhuravskaya, E. (2021). 3G Internet and Confidence in Government. *The Quarterly Journal of Economics*, *136*(4), 2533–2613. https://doi.org/10.1093/qje/qjaa040

Habermas, J., Burger, T., & Lawrence, F. (1992). *The structural transformation of the public sphere: An inquiry into a category of bourgeois society*. The MIT press.

Kuchler, T., Li, Y., Peng, L., Stroebel, J., & Zhou, D. (2022). Social Proximity to Capital: Implications for Investors and Firms (I. Goldstein, Ed.). *The Review of Financial Studies*, *35*(6), 2743–2789. https://doi.org/10.1093/rfs/hhab111

Kuriwaki, S. (2018). Cumulative CES Common Content. https://doi.org/10.7910/DVN/II2DB6

Lewis, J. B., Poole, K., Rosenthal, H., Boche, A., Rudkin, A., & Sonnet, L. (2022). Congressional Roll-Call Votes Database.

Manacorda, M., & Tesei, A. (2020). Liberation Technology: Mobile Phones and Political Mobilization in Africa. *Econometrica*, *88*(2), 533–567. https://doi.org/10.3982/ECTA14392

MIT Election Data and Science Lab. (2021). *County presidential election returns 2000-2020*. https://doi.org/10.7910/DVN/VOQCHQ

Prat, A., & Strömberg, D. (2005). *Commercial Television and Voter Information* (CEPR Discussion Paper No. 4989).

Prat, A., & Strömberg, D. (2013). The Political Economy of Mass Media. In D. Acemoglu, M. Arellano, & E. Dekel (Eds.), *Advances in Economics and Econometrics*. Cambridge University Press.

Quintelier, E., Stolle, D., & Harell, A. (2012). Politics in Peer Groups: Exploring the Causal Relationship between Network Diversity and Political Participation. *Political Research Quarterly*, *65*(4), 868–881. https://doi.org/10.1177/1065912911411099

Redistricting Report Card - Texas 2021 Final Congressional Plan C2193 - Enacted. (2023).

Schaffner, B., Ansolabehere, S., & Luks, S. (2021). Cooperative Election Study Common Content, 2020. https://doi.org/10.7910/DVN/E9N6PH

Schaffner, B., Ansolabehere, S., & Shih, M. (2023). Cooperative Election Study Common Content, 2022. https://doi.org/10.7910/DVN/PR4L8P

Sinclair, B. (2012). *The social citizen: Peer networks and political behavior*. The University of Chicago Press.

Snyder, J. M., & Strömberg, D. (2010). Press Coverage and Political Accountability. *Journal of Political Economy*, *118*(2), 355–408. https://doi.org/10.1086/652903

Solomon, D. (2021). Explaining the Most Bizarrely Shaped Districts in Texas's Proposed Congressional Map. *Texas Monthly*.

Stewart, C. (2021). House Standing Committee Assignments, 103rd - 117th Congress. https://doi.org/10.7910/DVN/XLIHUC

Strömberg, D. (1999). *The Politics of Public Spending* (Doctoral dissertation). Princeton University. Princeton.

Strömberg, D. (2001). Mass media and public policy. *European Economic Review*, *45*(4), 652–663. https://doi.org/10.1016/S0014-2921(01)00106-4

Strömberg, D. (2004a). Mass Media Competition, Political Competition, and Public Policy. *The Review of Economic Studies*, *71*(1), 265–284. https://doi.org/10.1111/0034-6527.00284

Strömberg, D. (2004b). Radio's Impact on Public Spending. *The Quarterly Journal of Economics*, *119*(1), 189–221. https://doi.org/10.1162/003355304772839560

United States Senate. (n.d.). About Parties and Leadership.

Vogels, E., Auxier, B., & Anderson, M. (2021). Partisan differences in social media use show up for some platforms, but not Facebook.

Volden, C., & Wiseman, A. (2023a). Legislative Effectiveness Data from 1973-2022.

Volden, C., & Wiseman, A. (2023b). Methodology.

Whitaker, L. P. (2017). *Congressional Redistricting Law: Background and Recent Court Rulings* (tech. rep. R44798). Congressional Research Service.

Zak, K. (2022). Here's what you need to know about the Texas 'fajita strip' congressional district. *Annenberg Media*.

A Appendix

A.1 Additional Methodology

A.1.1 District Congruence

District congruence is constructed as the share of a district's friends that live in that district. Recall our notation: $i \in J \subset K$, where i is a county, J is the set of all counties in the same congressional district as county i, and K is the set of all counties in the US. Additionally, n_i is the population of county i and $n_{i,d}$ is county i's population living in district d.

Now, let D represent the set of all districts in the U.S., and we will focus on a particular district $d \in D$. Congruence of district d is then:

$$\text{Congruence}_d = \frac{\text{Links}_{d,d}}{\sum_{e \in D} \text{Links}_{d,e}} = \frac{\text{SCI}_{d,d} \times \text{Pop}_d}{\sum_{e \in D} (\text{SCI}_{d,e} \times \text{Pop}_e)}$$

But, district-district SCI is not available. So, to calculate this from the county-county SCI, I first define the following terms:

$$\text{FriendShare}_{i,j} = \frac{\text{Links}_{i,j}}{\sum_k \text{Links}_{i,k}} = \frac{\text{SCI}_{i,j} \times \text{Pop}_j}{\sum_{k \in K} (\text{SCI}_{i,k} \times \text{Pop}_k)}$$

$$\mathbb{I}\{i, j \in d\} = \begin{cases} 1 & \text{if both } i, j \text{ have population in } d \\ 0 & \text{otherwise} \end{cases}$$

Then,

$$\text{Congruence}_d = \sum_{i \in d} \sum_{k \in K} \left(\text{FriendShare}_{i,k} \times \mathbb{I}\{i, k \in d\} \times \frac{\text{Pop}_i}{\text{Pop}_d} \right)$$

A.2 Additional Empirical Results

A.2.1 Maps of Congruence

Figure 4 shows unadjusted congruence across the continental 48 states in 2020, with color bins of equal widths. Most counties have congruence between 20% and 60%, or in the middle two bins.

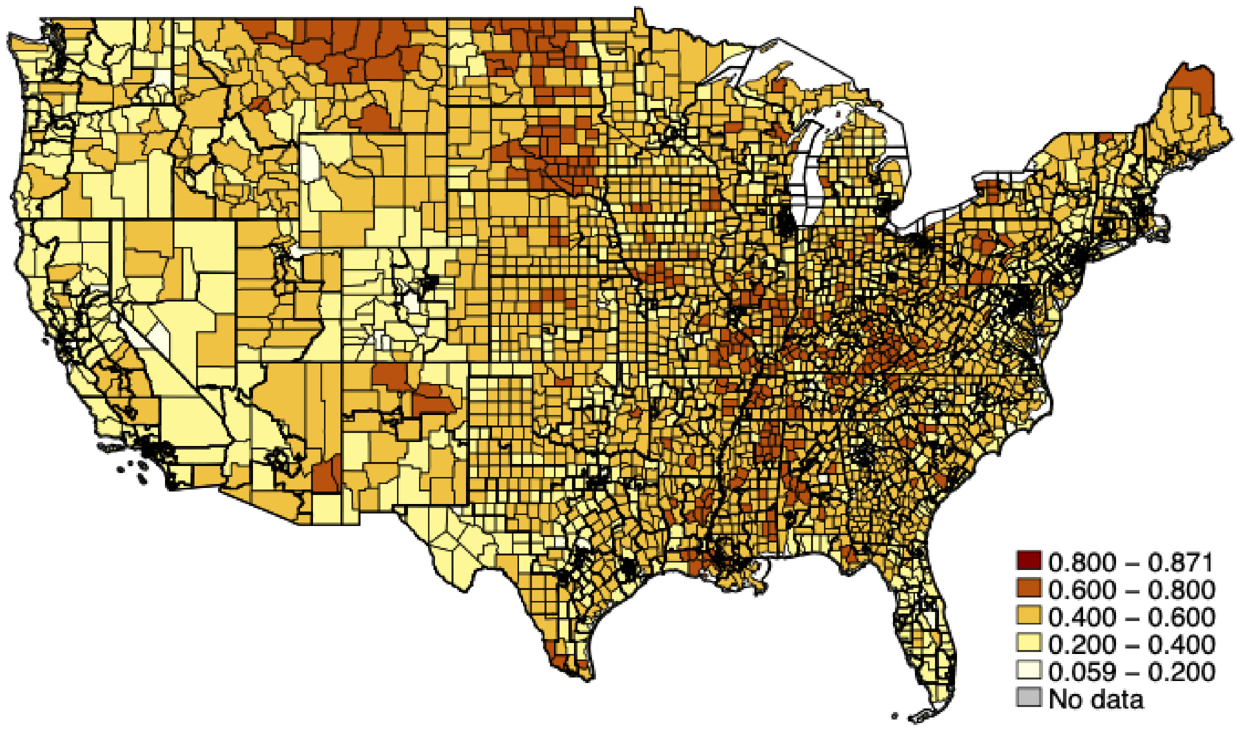

Figure 4: Unadjusted 2020 County Congruence with Congressional Districts - Equal-Spaced Bins

We can look at how congruence varies within individual states. The maps in Figures 5 and 6 show congruence for Florida, Georgia, Illinois, Michigan, Missouri, New York, Texas, and Virginia. Each map groups an equal number of the state's counties in each color bin; the bin boundaries are thus re-calculated for each state, so maps are not comparable between each other, but rather should be used to understand the most and least congruent areas within a state. Note that in each state depicted, the range between the minimum and maximum congruence is at least 40pp: there is substantial variation in congruence within states.

A.2.2 Changes in Congruence

Changes in Congruence Following 2010 Redistricting To motivate the results in Figure 2, below I present evidence that (1) current congruence is not strongly predictive of future changes in congruence, (2) changes in congruence are roughly normally distributed. In the following figures, I define "switchers" as counties that experienced a change in congruence of magnitude at least 1pp going from 2012 to 2013; "stayers" experienced no or small changes in congruence (less than 1pp magnitude). Roughly half of counties

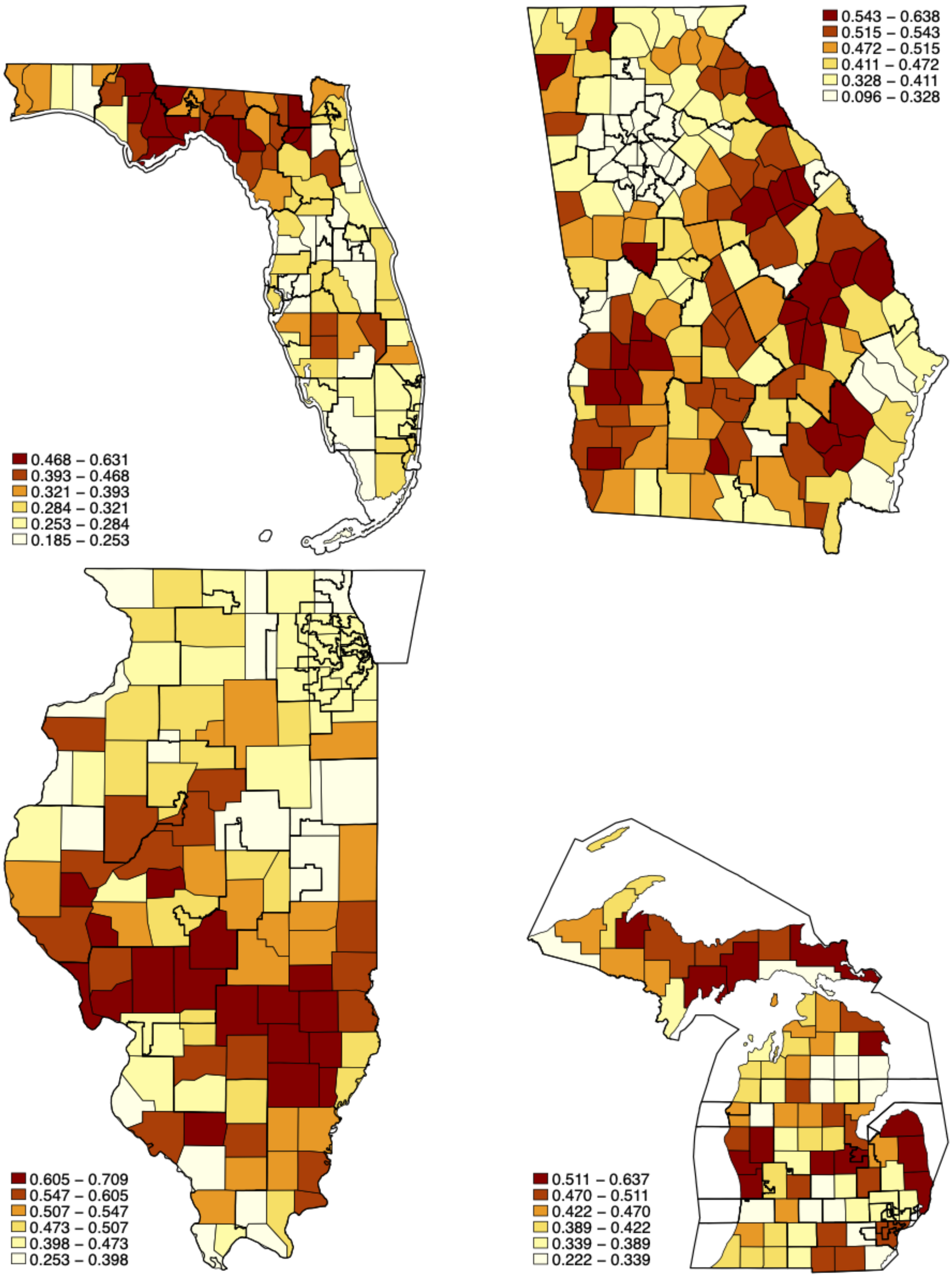

Figure 5: Unadjusted 2020 County Congruence with Congressional Districts - State Maps

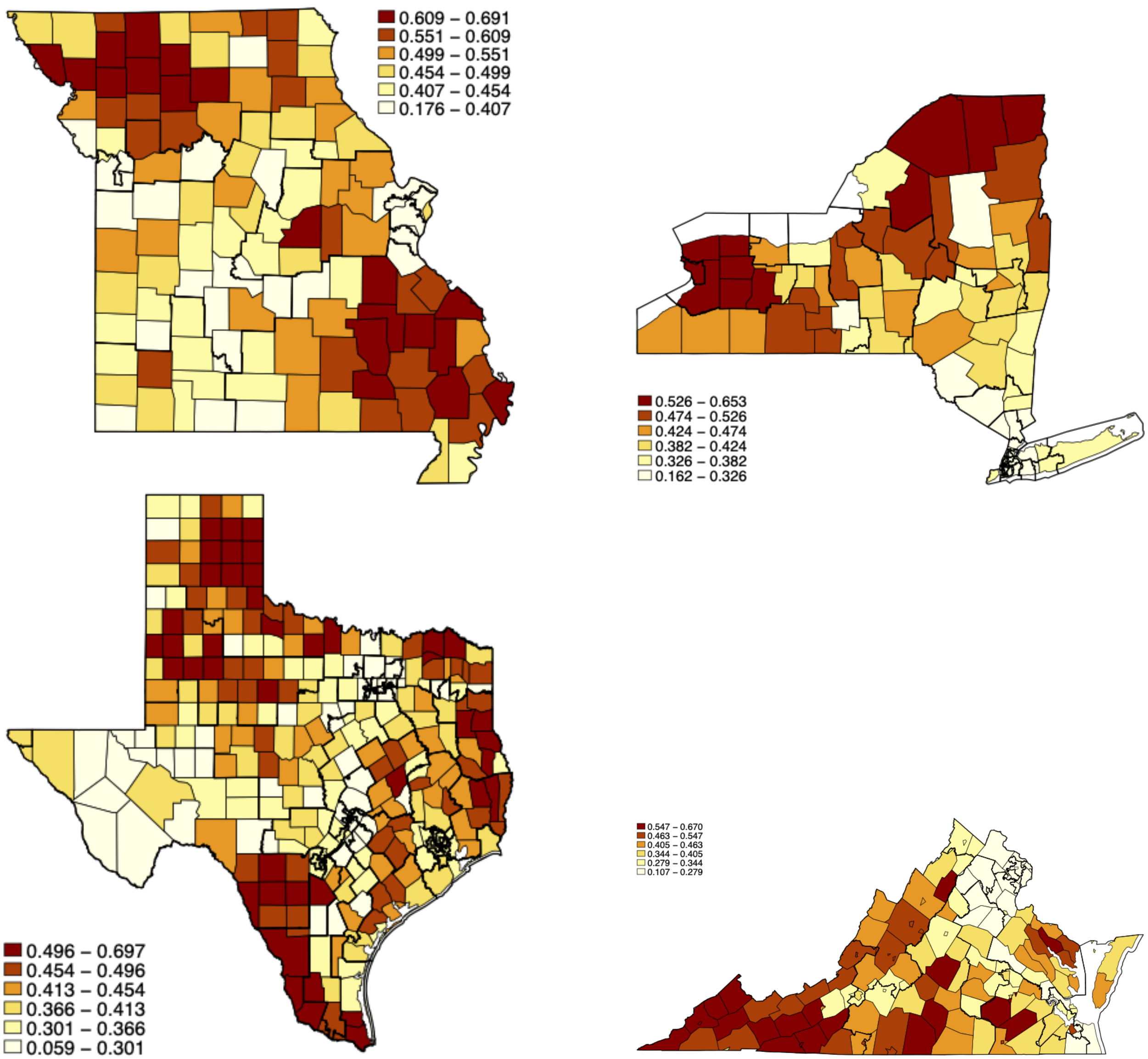

Figure 6: Unadjusted 2020 County Congruence with Congressional Districts - State Maps

	Dependent Variable: Change in Congruence					
	(1)	(2)	(3)	(4)	(5)	(6)
log(Population)	-0.000790	0.000266	0.000844	0.000918	0.00163	-0.00177
	(0.000661)	(0.00116)	(0.00123)	(0.00127)	(0.00256)	(0.00157)
Pop. Density		1.227	1.051	1.042	-19.48	0.0651
		(1.476)	(1.549)	(1.550)	(199.5)	(1.369)
% Pop. Rural		0.00454	0.00754	0.00757		
		(0.00505)	(0.00571)	(0.00572)		
Clustering		0.0603	0.131*	0.134*	0.165	0.0265
		(0.0621)	(0.0766)	(0.0787)	(0.116)	(0.114)
% Foreign Born			0.0280	0.0297	-0.0228	0.0298
			(0.0254)	(0.0260)	(0.0616)	(0.0277)
% Moved Last Year			0.0248	0.0245	0.0732	-0.0277
			(0.0276)	(0.0277)	(0.0498)	(0.0325)
% White			0.00686	0.00649	0.0162	-0.0143
			(0.00776)	(0.00793)	(0.0125)	(0.0104)
% No HS			-0.00155	-0.00326	-0.00968	-0.0199
			(0.0376)	(0.0383)	(0.0570)	(0.0555)
% College/Grad			-0.00240	-0.00112	0.00221	0.00334
			(0.0264)	(0.0281)	(0.0543)	(0.0325)
% Below Poverty			-0.0116	-0.0106	-0.0256	0.0242
			(0.0241)	(0.0247)	(0.0371)	(0.0331)
% Votes for Biden				-0.00179	-0.0144	0.00963
				(0.00966)	(0.0171)	(0.0111)
Sample					Top 50% Rural	Bottom 50% Rural
Observations	3,107	3,055	3,055	3,054	1,523	1,531
R^2	0.000	0.001	0.002	0.002	0.007	0.007

*** p<0.01, **p<0.05, * p<0.1

Notes: "Pop. Density" is the population per square meter. "% Foreign Born" is the share of non-citizens and naturalized citizens. "% Moved Last Year" is the share of people who moved within the last year, including those that moved within the same county. "% No HS" is the share of individuals 25 or over who do not hold a high school degree. "% College/Grad" is the share of individuals 25 or older who hold at least a college degree.

Table 7: Changes in County Congruence with Congressional Districts, 2013

	Dependent Variable: Change in Congruence					
	(1)	(2)	(3)	(4)	(5)	(6)
log(Population)	-0.00121*	0.00116	0.00156	0.00188	0.000595	0.00235
	(0.000636)	(0.00116)	(0.00119)	(0.00123)	(0.00247)	(0.00156)
Pop. Density		-1.978	-0.891	-0.844	-233.8	-1.223
		(1.385)	(1.447)	(1.448)	(192.1)	(1.282)
% Pop. Rural		0.00519	-0.000124	-0.000220		
		(0.00498)	(0.00552)	(0.00552)		
Clustering		0.170***	0.0944	0.113	-0.0155	0.210*
		(0.0630)	(0.0739)	(0.0760)	(0.107)	(0.118)
% Foreign Born			-0.00468	-0.00150	-0.0652	0.0212
			(0.0267)	(0.0269)	(0.0546)	(0.0306)
% Moved Last Year			0.0177	0.0158	0.0190	0.00775
			(0.0266)	(0.0266)	(0.0461)	(0.0343)
% White			0.0151*	0.0128	0.0327**	-0.00159
			(0.00784)	(0.00813)	(0.0127)	(0.0109)
% No HS			0.0296	0.0222	0.0737	-0.0402
			(0.0399)	(0.0405)	(0.0577)	(0.0655)
% College/Grad			-0.0680***	-0.0577**	-0.162***	-0.0269
			(0.0241)	(0.0259)	(0.0510)	(0.0300)
% Below Poverty			0.0163	0.0216	0.00972	0.0188
			(0.0246)	(0.0251)	(0.0362)	(0.0376)
% Votes for Biden				-0.0104	0.0256	-0.0344***
				(0.00962)	(0.0168)	(0.0111)
Sample					Top 50% Rural	Bottom 50% Rural
Observations	3,107	3,055	3,055	3,054	1,523	1,531
R^2	0.001	0.006	0.013	0.014	0.022	0.017

*** p<0.01, **p<0.05, * p<0.1

Notes: "Pop. Density" is the population per square meter. "% Foreign Born" is the share of non-citizens and naturalized citizens. "% Moved Last Year" is the share of people who moved within the last year, including those that moved within the same county. "% No HS" is the share of individuals 25 or over who do not hold a high school degree. "% College/Grad" is the share of individuals 25 or older who hold at least a college degree.

Table 8: Changes in County Congruence with Congressional Districts, 2023

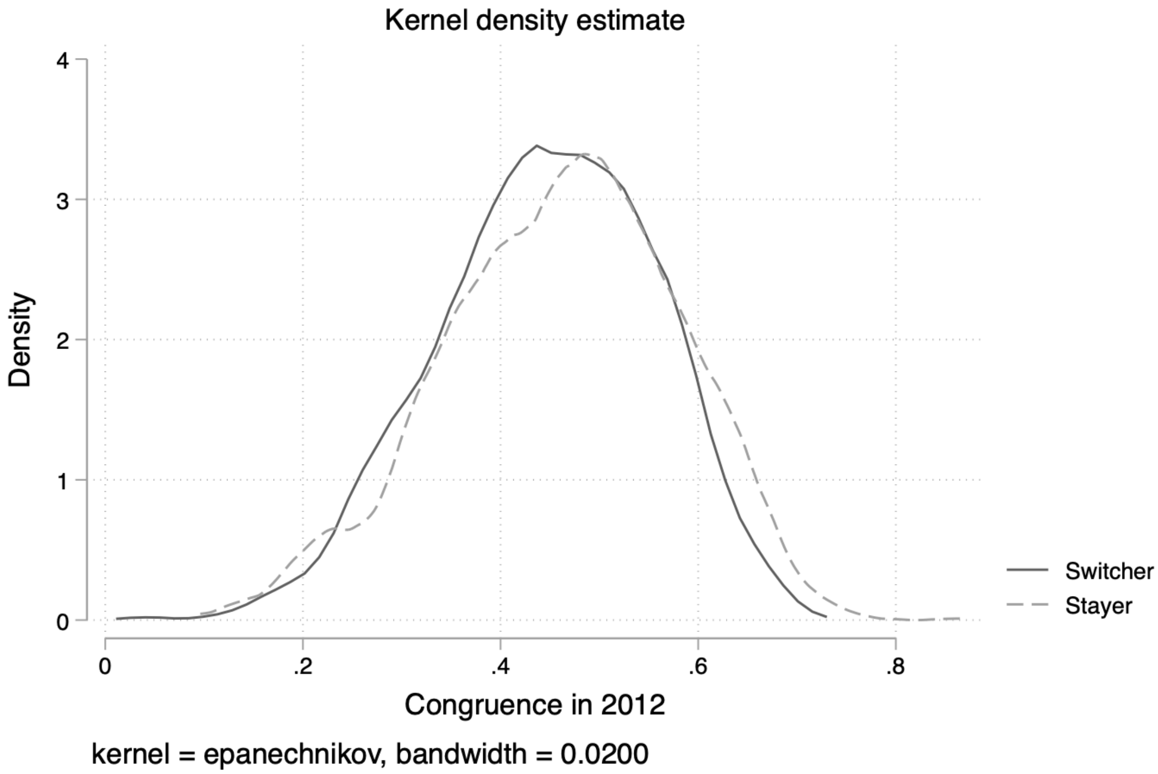

Figure 7: Kernel Density Estimate of 2012 Congruence for Switchers vs. Stayers

are switchers and half are stayers.[36]

Figure 7 shows the distribution of 2012 congruence among switchers (solid line) and stayers (dashed line). The distributions are relatively similar. Figure 8 shows this distributions instead separately as histograms: the y-axis is now the count of counties in each 2pp-wide bin. Here, we see also that the count of counties in each group at a given level of congruence in 2012 is relatively similar.

Figure 9 shows the distribution of *changes* in congruence among switchers; on top of the histogram, a kernel density estimate as well as a normal distribution is plotted. Here, we see that, conditional on experiencing a change in congruence of at least 1pp in magnitude, the changes themselves are close to normally distributed.

[36] Roughly 1/6 of counties (or 1/3 of the "stayers" as defined here) experience changes of congruence <0.001pp in magnitude. The other 2/3 of stayers mostly have changes in congruence between 0.1pp and 1pp in magnitude.

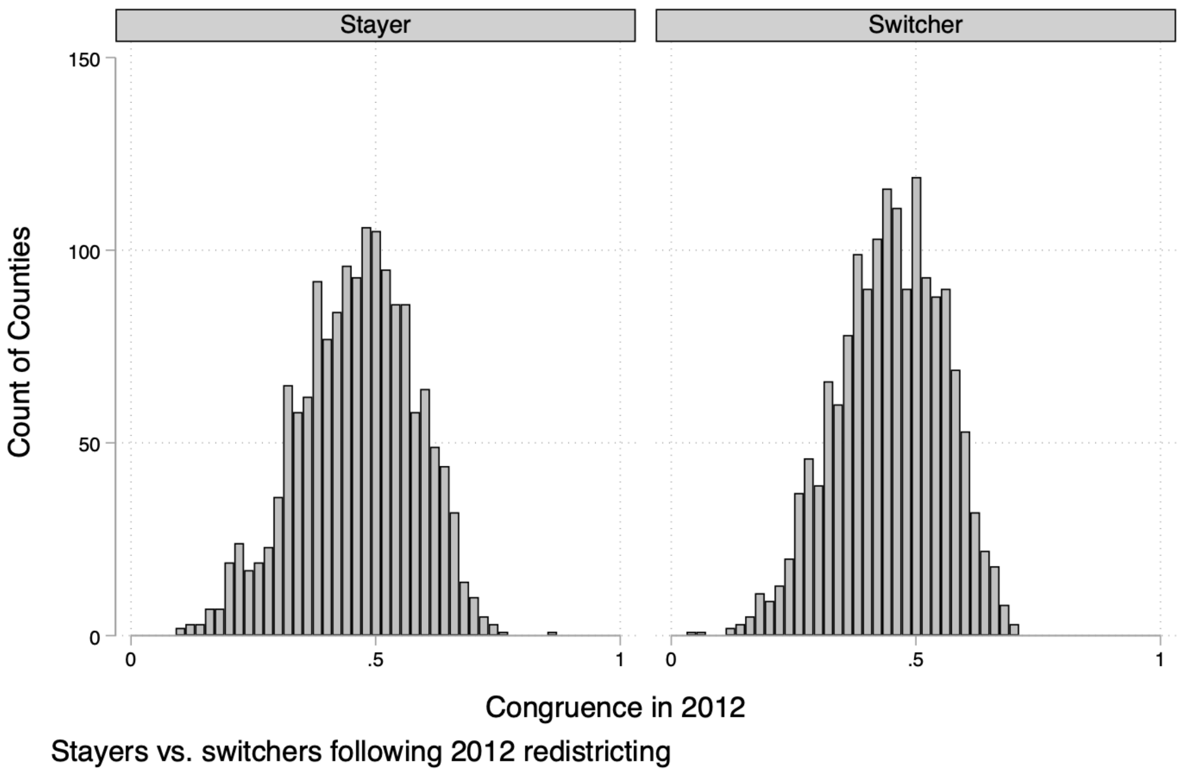

Figure 8: Histograms of 2012 Congruence for Switchers vs. Stayers

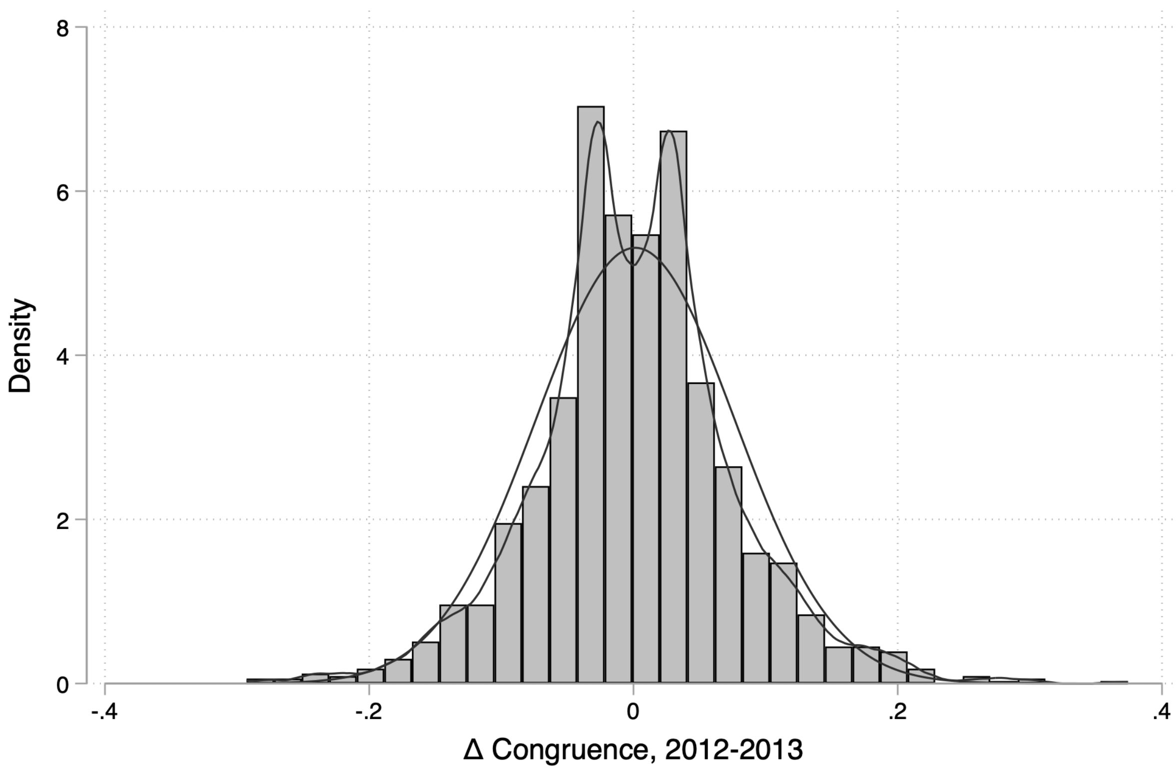

Figure 9: Distribution of Changes in Congruence from 2012-2-13, Among Switchers

Variable	**Observations**	**Mean (%)**	**SD (pp)**
Heard of Governor	549,740	96.3	18.8
Selected Governor Party	608,985	81.3	39.0
Selected Correct Gov. Party	608,985	74.7	43.5
Heard of Senator 1	549,244	94.9	22.0
Selected Senator 1 Party	608,414	75.0	43.3
Selected Correct Sen. 1 Party	608,414	67.7	46.7
Heard of Senator 2	549,246	94.9	22.0
Selected Senator 2 Party	608,402	74.3	43.7
Selected Correct Sen. 2 Party	608,402	66.9	47.0

Table 9: CES Data: Summary Statistics for Placebo Outcomes

A.2.3 Placebo Tests for Voter Information

Table 9 provides summary statistics for the nine outcome variables used for placebo tests. These distributions are generally similar to those for House representatives (Table 4), though respondents are generally more likely to recognize and select the correct party for their Senators and Governor.

The following three tables show the results of the placebo tests. In general, congruence does not significantly predict the placebo outcomes. While not shown here, I have verified that results are similar when instead using commuting congruence or controlling for DMA congruence.

A.2.4 DMA Congruence

Table 13 shows results from specifications that include both the originally defined county congruence measure and DMA congruence. *(Note that this table has a smaller sample size because it does not include 2021 and 2022; to be updated.)*

A.2.5 Commuting Congruence

Table 14 shows results from specifications that construct congruence using commuting flows. *(Note that this table has a smaller sample size in column (3) because that column does not include 2020, 2021, and 2022; to be updated.)*

	(1) County & Year FEs Only	(2) Add District x Year FEs	(3) Add Individual Demographic Controls	(4) Add County-Year Controls
	Heard of Governor			
Congruence	-0.004	-0.004	0.004	0.007
	(0.028)	(0.027)	(0.027)	(0.027)
	[0.897]	[0.873]	[0.872]	[0.793]
Obs	549,682	549,682	549,682	549,682
R^2	0.032	0.083	0.120	0.120
	Selected Governor Party			
Congruence	-0.056	-0.144**	-0.105*	-0.100*
	(0.065)	(0.061)	(0.057)	(0.057)
	[0.393]	[0.017]	[0.064]	[0.078]
Obs	608,940	608,940	608,940	608,940
R^2	0.043	0.100	0.202	0.202
	Selected Correct Governor Party			
Congruence	-0.091	-0.101	-0.055	-0.049
	(0.084)	(0.065)	(0.060)	(0.060)
	[0.278]	[0.117]	[0.358]	[0.413]
Obs	608,940	608,940	608,940	608,940
R^2	0.046	0.139	0.250	0.250
Ind. Controls			X	X
County x Year Controls				X
FEs	County, Year	County, Year, District x Year	County, Year, District x Year	County, Year District x Year

*** p<0.01, **p<0.05, * p<0.1
Standard errors clustered at the county level in parentheses. P-values in square brackets.
"Heard of Governor" not available in 2006, 2007, or 2009. Individual controls include gender, race, education, age categories, and whether the respondent is affiliated with the same party as their representative. County-by-year controls include population and shares by race, age categories, gender, and county urban population share.

Table 10: Effect of Congruence on Voter Familiarity with Governor

	(1) County & Year FEs Only	(2) Add District x Year FEs	(3) Add Individual Demographic Controls	(4) Add County-Year Controls
	Heard of Senator 1			
Congruence	-0.039	0.005	0.016	0.015
	(0.038)	(0.032)	(0.032)	(0.032)
	[0.300]	[0.888]	[0.616]	[0.642]
Obs	549,187	549,187	549,187	549,187
R^2	0.031	0.080	0.136	0.136
	Selected Senator 1 Party			
Congruence	-0.070	-0.062	-0.018	-0.022
	(0.079)	(0.064)	(0.059)	(0.059)
	[0.372]	[0.331]	[0.760]	[0.715]
Obs	608,370	608,370	608,370	608,370
R^2	0.047	0.104	0.229	0.229
	Selected Correct Senator 1 Party			
Congruence	-0.213**	-0.070	-0.022	-0.024
	(0.097)	(0.066)	(0.061)	(0.061)
	[0.028]	[0.291]	[0.714]	[0.696]
Obs	608,370	608,370	608,370	608,370
R^2	0.050	0.132	0.267	0.267
Ind. Controls			X	X
County x Year Controls				X
FEs	County, Year	County, Year, District x Year	County, Year, District x Year	County, Year District x Year

*** p<0.01, **p<0.05, * p<0.1
Standard errors clustered at the county level in parentheses. P-values in square brackets.
"Heard of Senator 1" not available in 2006, 2007, or 2009. Individual controls include gender, race, education, age categories, and whether the respondent is affiliated with the same party as their representative. County-by-year controls include population and shares by race, age categories, gender, and county urban population share.

Table 11: Effect of Congruence on Voter Familiarity with Senator 1

	(1) County & Year FEs Only	(2) Add District x Year FEs	(3) Add Individual Demographic Controls	(4) Add County-Year Controls
		Heard of Senator 2		
Congruence	0.026	0.031	0.043	0.045
	(0.036)	(0.037)	(0.036)	(0.036)
	[0.468]	[0.402]	[0.233]	[0.216]
Obs	549,189	549,189	549,189	549,189
R^2	0.032	0.083	0.132	0.132
		Selected Senator 2 Party		
Congruence	-0.018	0.013	0.059	0.061
	(0.088)	(0.068)	(0.063)	(0.063)
	[0.842]	[0.851]	[0.351]	[0.333]
Obs	608,358	608,358	608,358	608,358
R^2	0.045	0.145	0.213	0.226
		Selected Correct Senator 2 Party		
Congruence	-0.098	0.002	0.051	0.053
	(0.104)	(0.067)	(0.062)	(0.062)
	[0.345]	[0.982]	[0.411]	[0.387]
Obs	608,358	608,358	608,358	608,358
R^2	0.046	0.141	0.261	0.261
Ind. Controls			X	X
County x Year Controls				X
FEs	County, Year	County, Year, District x Year	County, Year, District x Year	County, Year District x Year

*** $p<0.01$, **$p<0.05$, * $p<0.1$
Standard errors clustered at the county level in parentheses. P-values in square brackets.
"Heard of Senator 2" not available in 2006, 2007, or 2009. Individual controls include gender, race, education, age categories, and whether the respondent is affiliated with the same party as their representative. County-by-year controls include population and shares by race, age categories, gender, and county urban population share.

Table 12: Effect of Congruence on Voter Familiarity with Senator 2

	(1) County & Year FEs Only	(2) Add District x Year FEs	(3) Add Individual Demographic Controls	(4) Add County-Year Controls
		Heard of Incumbent		
Congruence	0.067	0.118***	0.133***	0.145***
	(0.041)	(0.046)	(0.045)	(0.047)
	[0.100]	[0.010]	[0.003]	[0.002]
DMA Congruence	0.054***	0.030*	0.028	0.035*
	(0.016)	(0.018)	(0.017)	(0.018)
	[0.001]	[0.096]	[0.102]	[0.054]
Obs	460,951	460,951	460,951	402,066
R^2	0.035	0.083	0.132	0.133
		Selected Party		
Congruence	0.191**	0.214***	0.269***	0.278***
	(0.086)	(0.078)	(0.074)	(0.077)
	[0.027]	[0.006]	[0.000]	[0.000]
DMA Congruence	0.122***	0.093***	0.088***	0.098***
	(0.035)	(0.032)	(0.030)	(0.032)
	[0.000]	[0.003]	[0.004]	[0.002]
Obs	520,034	520,033	520,033	461,154
R^2	0.046	0.103	0.211	0.210
		Selected Correct Party		
Congruence	0.123	0.194**	0.253***	0.280***
	(0.093)	(0.082)	(0.077)	(0.079)
	[0.189]	[0.019]	[0.001]	[0.000]
DMA Congruence	0.131***	0.115***	0.108***	0.121***
	(0.038)	(0.032)	(0.030)	(0.032)
	[0.001]	[0.000]	[0.000]	[0.000]
Obs	520,034	520,033	520,033	461,154
R^2	0.047	0.113	0.236	0.235
Ind. Controls			X	X
County x Year Controls				X
FEs	County, Year	County, Year, District x Year	County, Year, District x Year	County, Year District x Year

*** $p<0.01$, **$p<0.05$, * $p<0.1$
Standard errors clustered at the county level in parentheses. P-values in square brackets.
"Heard of Incumbent" not available in 2006, 2007, or 2009. Individual controls include gender, race, education, age categories, and whether the respondent is affiliated with the same party as their representative. County-by-year controls include population and shares by race, age categories, gender, and county urban population share.

Table 13: Effect of Congruence and DMA Congruence on Voter Familiarity with Representative

	(1) County & Year FEs Only	(2) Add District x Year FEs	(3) Add Individual Demographic Controls	(4) Add County-Year Controls
		Heard of Incumbent		
Commuting Congruence	0.050**	0.048*	0.049**	0.065***
	(0.023)	(0.025)	(0.024)	(0.025)
	[0.032]	[0.050]	[0.042]	[0.008]
Obs	544,910	544,910	544,910	401,874
R^2	0.032	0.078	0.130	0.134
		Selected Party		
Commuting Congruence	0.149***	0.112***	0.132***	0.131***
	(0.039)	(0.037)	(0.033)	(0.036)
	[0.000]	[0.002]	[0.000]	[0.000]
Obs	603,967	603,966	603,966	460,937
R^2	0.042	0.098	0.224	0.225
		Selected Correct Party		
Commuting Congruence	0.135***	0.117***	0.139***	0.150***
	(0.041)	(0.039)	(0.035)	(0.037)
	[0.001]	[0.003]	[0.000]	[0.000]
Obs	603,967	603,966	603,966	460,937
R^2	0.044	0.108	0.251	0.251
Ind. Controls			X	X
County x Year Controls				X
FEs	County, Year	County, Year, District x Year	County, Year, District x Year	County, Year District x Year

*** p<0.01, **p<0.05, * p<0.1
Standard errors clustered at the county level in parentheses. P-values in square brackets.
"Heard of Incumbent" not available in 2006, 2007, or 2009. Individual controls include gender, race, education, age categories, and whether the respondent is affiliated with the same party as their representative. County-by-year controls include population and shares by race, age categories, gender, and county urban population share.

Table 14: Effect of Commuting Congruence on Voter Familiarity with Representative