NBER WORKING PAPER SERIES

PRESS COVERAGE AND POLITICAL ACCOUNTABILITY

James M. Snyder, Jr.
David Strömberg

Working Paper 13878
http://www.nber.org/papers/w13878

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2008

Press Coverage and Political Accountability
James M. Snyder, Jr. and David Strömberg
NBER Working Paper No. 13878
March 2008
JEL No. D72,H5,L82

ABSTRACT

In this paper we estimate the impact of press coverage on citizen knowledge, politicians' actions, and policy. We find that a poor fit between newspaper markets and political districts reduces press coverage of politics. We use variation in this fit due to redistricting to identify the effects of reduced coverage. Exploring the links in the causal chain of media effects -- voter information, politicians' actions and policy -- we find statistically significant and substantively important effects. Voters living in areas with less coverage of their U.S. House representative are less likely to recall their representative's name, and less able to describe and rate them. Congressmen who are less covered by the local press work less for their constituencies: they are less likely to stand witness before congressional hearings, to serve on constituency-oriented committees (perhaps), and to vote against the party line. Finally, this congressional behavior affects policy. Federal spending is lower in areas where there is less press coverage of the local members of congress.

James M. Snyder, Jr.
MIT Department of Economics
E53-457
50 Memorial Drive
Cambridge, MA 02142-1347
and NBER
millett@mit.edu

David Strömberg
IIES
Stockholm University
Stockholm, Sweden
David.Stromberg@iies.su.se

1 Introduction

It has long been recognized that having a press that actively covers politics is essential for democratic governance. Informed voters are better able to hold elected officials accountable for their policy decisions, and most people get their information via the media. In the words of Thomas Jefferson:

> "The functionaries of every government have propensities to command at will the liberty and property of their constituents. There is no safe deposit for these but with the people themselves, nor can they be safe with them without information. Where the press is free, and every man able to read, all is safe."
>
> –Thomas Jefferson to Charles Yancey, 1816. ME 14:384

Frequently, however, an active press is missing from the political stage. At the national level, this is true in many weak democracies and poor countries, often because of political oppression or lack of an economic base to support large mass media. This is true also in mature democracies with a free press, in particular for lower levels of government. One reason is that local politics is boring to many people. Another reason it that there is often a poor fit between media markets and political jurisdictions, making coverage of some jurisdictions too costly. This problem may be increasing with increasing media market size, consolidations in the media industry and a shift from print to electronic media. As the number of active political units within each media market that require media attention increases, reporting, by necessity, becomes more selective and superficial. To counter this trend, regulations such as the Newspaper Preservation Act and the FCC "must carry" laws aim at ensuring local content and programming. For similar reasons, many governments subsidize domestic media.

Despite these important and long standing concerns, empirical evidence of the effects of active media coverage is scarce. One reason is that media coverage is endogenous to most outcome variables of interest. For example, a number of studies have found better governance in countries where the media actively covers politics (Ahrend, 2002, Brunetti and Weder, 2003, Djankov and others, 2003). However, allowing active political coverage is a governmental choice, and more corrupt governments have stronger incentives to silence the press. So, correlations between governance and active political coverage may simply reflect that these variables are jointly determined, as discussed by Besley and Prat (2006). Similarly, while some studies find a correlation between media coverage and voter knowledge, this correlation might simply reflect the fact that news coverage and knowledge are both driven by citizens' intrinsic interest in politics.

One way to more convincingly identify an effect of active media coverage on voters, politicians, and policy, is to find a number of similar political jurisdictions where some factor

causes news coverage of politics to vary, but this factor does not directly affect our outcome variables. The factor we propose is the match, or "congruence," between newspaper markets and U.S. congressional districts. The "economic geography" factors that determine media markets are generally quite different from the "political geography" factors that determine congressional district boundaries. Therefore, it therefore seems a priori reasonable that the match between the two will be haphazard. Below we present some evidence to this effect. In addition, panel features of our data allow us to estimate two specifications that are particularly convincing (at least to us). One investigates the consequences of the changing congruence between media markets and congressional district due to redistricting. The other employs congressional-district-and-year fixed effects, and thus focuses on variation within congressional districts.

The paper follows the chain of media impacts, link by link: congruence increases the newspapers' political coverage which, in turn, makes voters better informed, which increases monitoring and induces politicians to work harder, which, finally, produces better policies (for their constituencies). Figure 1 shows these links, together with simple bivariate graphs of the main relationships of interest.[1] The paper is devoted to investigating whether causal relationships are underlying the patterns in those graphs.

We start by defining the match, or congruence, between newspaper markets and congressional districts in section 2. The measure we use is based on the share of a newspaper's readership that lives in a certain congressional district. In Figure 1(a), the left image shows a perfect match between media markets and congressional districts. Every person in both districts only buys newspapers that only sell in this district. The situation to the right shows a worse match where each person reads a newspaper which only has half of the sales in their district. Intuitively, newspaper coverage of a congressman should be increasing in this readership share. Since more than one newspaper sells in each district, we define congruence as the circulation-weighted average of readership for all newspapers sold in a district.

We then explore whether newspaper coverage is increasing in the readership share in section 3. Using online editions, we counted the number of articles mentioning House representatives in 161 newspapers which, on average, cover 385 districts in each congress from 1991-2002. We find that coverage of representatives is strongly increasing in district readership shares. This can clearly be seen from the graph in Figure 1(b), which plots the share of articles mentioning a given representative in a given newspaper as a function of the share of that newspaper's readers that lives in the representative's district.

Consequently, voters in non-congruent areas are less exposed to news about their rep-

[1] Except for the first graph, each point in the graphs the mean value of the y-variable within each percentile of Congruence. So, for example, the first point in the "Witness appearances" graph shows the mean Congruence and witness appearances in the lowest Congruence percentile.

resentatives. The newspapers sold in their area simply devote less coverage to their House representative. Section 4 explores whether this affects voter information. After analyzing survey responses from the American National Election Studies 1984 to 2004, we find that voters from more congruent areas are better informed about their House representative. More precisely, they are better able to correctly name at least one of the candidates in the House election. They are also more willing to place their representative ideologicially, to rate their representative on a feeling thermometer, and to mention things that they like or dislike about their representative. The graph in Figure 1(c) plots the share that can correctly recall a name against congruence. Since voter information is key to accountability, this relationship is extensively explored. In an extension, we also study how news coverage affects two voting outcomes, participation in elections and the tendency to support the incumbent.

Section 5 then studies effects on representatives' behavior. Some representatives are elected from districts that "fall through the cracks" of some or all local newspapers, and the press is largely silent about them. Others are elected from highly congruent districts with ample coverage and well-informed voters. Such representatives might therefore have a greater incentive to work for their constituency, for example, by considering constituency (rather than party) interests in voting, and by engaging in pork-barrel politics. One useful vehicle in bringing pork to the district is to serve on "constituency" committees with jurisdiction over spending programs, another to appear as a witness before congressional hearings. To analyze these aspects of representative effort, we collected roll-call voting data for all House members 1982-2002, and data on committee assignments and witness appearances for 1982-2004. We find that congressmen from highly congruent districts are more disciplined by their constituencies: their voting conforms less to the party line, they are more likely to stand witness before congressional hearings, and they are, perhaps, more likely to serve on constituency-oriented committees and also less likely to serve on broad policy-oriented committees. Figures1 (d),(e), and (f) show the corresponding bivariate graphs.

If the representatives' efforts in Congress are successful, then we should observe more federal funds flowing into highly congruent districts, and into more congruent counties within districts. To test this, we assembled data on federal expenditure allocations across counties from the Consolidated Federal Funds Report, 1983-2004. The total value of the expenditures we study is about $2,700 per capita (in year 2000 dollar values), or about 10 percent of U.S. GDP. We find that more congruent counties have received higher federal expenditures per capita. Figure 1(g) shows the corresponding bivariate graph.

As mentioned, this paper is closely related to the literature on the political consequences of having a media that actively covers politics. Moreover, our results in section 5 are related to Cohen, Noll, and Zaller (n.d.), who study how the amount of information available to voters affects congressional roll call voting behavior, using an approach which is similar to ours. Consistent with our results, they find evidence that representatives in low-information

districts vote in a more purely partisan manner, while those in high-information districts are more responsive to the underlying ideology of the district. Our work is also related to several paper that use television media market definitions to explore the importance of mass media on the incumbency advantage and campaign finance.[2]

While we study the effects of media content, our results are related to a number of studies of the effect of media penetration and media use. For example, our results in section 4 are closely related to a literature on the effects of mass media use on information.[3] However, while there are numerous studies regressing measures of political knowledge on self-reported media exposure (plus other controls), none of them convincingly identifies a causal relationship. One problem is that people who are interested in politics consume more political news, from a variety of sources, causing political knowledge and news consumption to be positively correlated. The results in section 6 are related to the findings that media penetration has been found to influence government spending (Besley and Burgess, 2002 and Strömberg, 1999, 2004). Our approach is different since we study the effects of news coverage, holding media penetration constant.

Before proceeding, we must briefly discuss why we focus on newspapers and ignore radio and television. The existing evidence indicates that local television stations devote much less coverage to congress than local newspapers (Hess, 1991; Vinson, 2003). In one analysis, we compare the impact of television market structure and newspaper market structure on voter knowledge (section 4), and find that television appears to have no effect on voter knowledge about their congressmen. Less is known about radio, and this will require further investigation.

2 Congruence

The driving force behind all results in this paper is that the number of articles, q_{md}, that a newspaper m writes about a House representative from district d, is strongly increasing in the share of this newspaper's readers that lives in district d, $ReaderShare_{md}$. We present strong evidence of this in the empirical section. For simplicity, assume a proportional relationship, so that

$$q_{md} = \alpha ReaderShare_{md}. \tag{1}$$

[2]See Campbell, Alford and Henry (1984), Niemi, Powell and Bicknell (1986), Stewart and Reynolds (1990), Levy and Squire (2000), Ansolabehere, Snowberg and Snyder (2004).

[3]Examples include Robinson and Levy (1986), Berkowitz and Pritchard (1989), Delli Carpini and Keeter (1989), Robinson and Davis (1990), Weaver and Drew (1993), and Arnold (2004). Mondak (1995) exploits a quasi-experiment – a newspaper strike in Pittsburgh in 1992 that closed the city's two major newspapers for eight months. Unfortunately, he only has responses to vaguely worded questions about respondents' knowledge of local politics.

Let $MarketShare_{mc}$ be newspaper m's market share in county c. We define

$$Congruence_{cd} = \sum_{m=1}^{M} MarketShare_{mc} ReaderShare_{md}. \quad (2)$$

$Congruence_{cd}$ is the average $ReaderShare_{md}$, weighted by market share. It is a proxy for how actively the average newspaper read in county c covers congressional district d, since each newspaper's incentive to cover district d is increasing in $ReaderShare_{md}$. Note that since *Congruence* is defined using market shares, it is not dependent on the total newspaper penetration in the county. The Appendix explains why *Congruence* is the best simple proxy for news impact on voter information, given the relationship in equation 1.

To measure *Congruence*, we combine newspaper sales data with demographic data. Each year, the Audit Bureau of Circulation collects data on each newspaper's circulation in each county, for almost all U.S. newspapers. We have this data for 1982 and for the period 1991-2004. We complemented this with county-circulation data for non-ABC newspapers for 2004 and 1991, and interpolated values in between those years. These were mainly smaller papers, and the data was provided by SRDS.[4] The U.S. Census Bureau collects data on the number of people in each congressional district, by county. We have this data for the censuses of 1980, 1990 and 2000. Since we analyze data from 1982 to 2004, we need to compute $Congruence_{cd}$ for years when we do not have circulation data. This is done by interpolating $Congruence_{cd}$ 1983-1990. The formulas used for computing *Congruence* are shown in the Appendix.

To illustrate how *Congruence* varies, consider the case of Missouri, shown in Figure 2. Newspapers are typically located in the larger cities and towns and sell to a markets around these. The two largest cities in Missouri, Kansas City and St. Louis City both lie on the state border and their markets cross this border. The third and fourth largest cities (disregarding cities which are part of the Kansas City and St. Louis metropolitan areas) are Springfield City, and Columbia City both located in the interior of the state and a congressional district. Finally, the fifth largest city, St. Joseph City, is located on the state border but the area across the border in Nebraska and Kansas are not densely populated.

The congressional district borders of Missouri are drawn with solid lines. First, take a look at the congressional district to the northwest. In the north, *St.Joseph News Press* is the dominant paper. We expect the *St. Joseph News* to mainly cover this congressional district, since 93% of its readers live in this district. Therefore, $Congruence_{cd}$ is high in that area. In the southern part of the district, people read the *The Kansas City Star* with 18% of its readers in this district and thus, $Congruence_{cd}$ is much smaller. *The Columbia Daily*

[4] On average there are about 10,900 observations each year in the ABC data, and about 500 observations in the non-ABC data. There are about 3,000 counties in the U.S., so the average number of observations per county in each year is slightly less than 4.

Tribune sells in the small appendix region to the southeast and only has 5% of its readers in this district. As a result, $Congruence_{cd}$ is low.

The district in the southwest has high congruence. It has two major papers, *The Joplin Globe* (with 80% of its sales in this district) and *The Springfield News Leader* (87%). Note how the eastern counties of this district have high $Congruence_{cd}$, whereas the adjoining county to the east of Springfield has a very low value. Similar strong differences in $Congruence$ between neighboring countries are found in the central and central-northeast parts. These differences are due to the match between congressional districts and the markets of a string of moderately sized papers: *The Sedalia Democrat, The Columbia Daily Tribune, The Post-Tribune/Capital Times, The Mexico Ledger* and *The Hannibal Courier-Post.*

We also construct a measure for congruence of television markets, $TVCongruence$. To identify radio and TV markets, we use the Designated Media Markets (DMAs). A DMA is defined by Nielsen Media Research as all counties whose largest television viewing share is given to stations of that same market area. 210 non-overlapping DMAs cover the entire continental United States, Hawaii and parts of Alaska. Since we do not have viewership data by county, we assume that everyone in a media market only watches broadcasts from that market. For a county c in media market m, $TVCongruence_{cd}$ is thus defined as the population share of media market m that lies in district d. Figure 3 shows how $TVCongruence$ is constructed in Missouri. In the graph to the left, the solid lines are the congressional district boundaries and the fillings identify the media markets. The graph to the right shows the resulting $TVCongruence$.

A key identifying assumption in the empirical investigation is that $Congruence$ is not directly related to variables such as voters' intrinsic interest in politics. We explore this extensively below. For now, note that this seems a priori reasonable. The "economic geography" factors that determine media markets are generally quite different from the "political geography" factors that determine congressional district boundaries.[5] This is quite apparent from Figure 3, showing both media markets and congressional districts. The match between these economic and political geography concerns creates the haphazard patterns in Figures 2 and 3.

[5] Congressional districts boundaries are drawn under the constraints that all districts in each state must have the same population, constraints imposed by courts on racial considerations, partisan considerations, the desire to protect incumbents, and so on. To satisfy the population requirement, congressional district lines are redrawn after the population count of each Census. In our sample period, the congressional lines were redrawn in 1992 based on the 1990 Census and in 2002, based on the 2000 Census.

Newspapers, on the other hand, typically locate in large cities, with strong demand for advertising and news about city affairs. Their county sales depends strongly on the distance between the county and the newspaper seat, and characteristics of the county residents such as their incomes, their level of education and their age structure.

3 Newspaper Coverage

In this section, we examine the main force behind the results in this paper: that the number of articles that a newspaper writes about a House representative is increasing in the estimated fraction of the newspaper's readers that live in the associated congressional district, the *ReaderShare*.

Our data covers the period 1991-2002 (103rd-107th Congresses). The sample consists of 161 newspapers which, on average, cover 385 districts for each congress.[6] We used the NewsLibrary.com web site, which employs a common search engine to search the on-line archives of newspapers, for 142 newspapers. We supplemented this using Lexis/Nexis for 8 newspapers. In addition, we searched 11 newspapers' web sites directly. In all, there are 4,216 observations in our sample, where each observation is a newspaper-district (newspaper-representative) pair in a given year.

We measure relative newspaper coverage as follows: Let q_{mdt} be the number of articles appearing in newspaper m during Congress t that contain both the name of the representative from district d and the word "Congress." [7] Let $q_{mt} = \sum_d q_{mdt}$, where the sum is taken over all districts which contain at least 1% of newspaper m's readers. Then, $CoverShare_{mdt} = q_{mdt}/q_{mt}$. This is the relative share of newspaper m.

The main independent variable of interest is $ReaderShare_{mdt}$. (Note that we now use t to denote time periods – each period is a Congress.) We include several other control variables that are likely to affect the amount of coverage, including indicator variables for party leaders (the Speaker of the House, Majority Leader, Minority Leader, Majority Whip, and Minority Whip); seniority; an indicator variable for freshmen; an indicator for majority party status; an indicator for out-of-state districts; variables for members in scandals; and a variable indicating whether the representative sought higher office (governor or senator) or received a higher appointment towards the end of the term. We also include some district level demographics that might affect coverage, including income and the fraction of people living in urban areas, the share old, the share foreign, and the share blue collar. Summary statistics are in Table 1.

Figure 1(b) shows the basic pattern: a very strong, positive, and approximately linear relationship between *CoverShare* and *ReaderShare*. The bivariate regression line in also shown in the figure. Table 2 shows that the relationship is essentially the same after controlling for other factors that affect coverage.

Most of the controls have the effects that one would expect. Party leaders tend to receive

[6] We do not have data on every newspaper for every year in the sample.

[7] For simplicity, we use two calandar years for each Congress rather than the exact dates the Congress is in session. Thus, for the 103rd Congress we use articles appearing January 1, 1991 through December 31, for the 104th Congress we use articles appearing January 1, 1995 through December 31, 1996, and so on.

more coverage than other representatives. Members who are involved in scandals or who seek higher office also receive more coverage. Three other findings are noteworthy. First, even controlling for *ReaderShare*, newspapers exhibit an in-state bias, covering representatives from their home state more heavily than out-of-state representatives. Second, even controlling for *ReaderShare*, urban representatives receive less coverage than others. This could reflect the fact that urban areas have many powerful local politicians competing for scarce newspaper space – mayors, police chiefs, school superintendents, and even city council members. Local governments provide a broader array of services in urban areas than in suburban and rural areas, and citizens' preferences in urban areas are more heterogeneous, so decisions are more conflictual. As a result, citizens need to pay more attention to local politics, and, since they are busy, they have less time to monitor their U.S. representatives. Finally, it does not appear that newspapers are mainly interested in providing necessary information to help voters vote in the next election (although that might be one consideration, of course). For example, freshmen do not receive more coverage than other representatives, even though they are new and relatively unknown to voters. Similarly, members who are retiring receive just as much coverage – or even slightly more – as those running for re-election.

In quantitative terms, the share of the articles devoted to a representative increases by about 60 percent when the share of the readership increases from 0 to 1. In a different specification, we estimate the elasticity of the number of newspaper articles covering the representative with respect to the reader share to 0.50 (s.e. 0.03). Since the typical newspaper prints approximately 100 articles about a representative each year, this translates into about 50 more articles per year, although, as can be seen from the graph, there is a great deal of variation for any given district readership share.

The very strong relationship between *ReaderShare* and press coverage of representatives will drive our results. People who live in areas where *Congruence* (the market share weighted *ReaderShare*) is high will, on average, be exposed to a considerably larger number of articles about their congressman than people in other areas.

In order to determine what effects this may generate, it is useful to know what newspapers actually report when they cover members of congress. Arnold (2004) studied newspaper coverage of the House representatives in detail. In a representative sample of 25 newspapers, covering equally many House representatives in the 103$^{\text{rd}}$ Congress, he found 8,000 news stories mentioning the candidates. The articles were primarily news stories in reasonably prominent places. Arnold found that the most commonly covered policy-related activities of the congressmen are roll-call voting, advance position taking, and acting as local agents, that is, working actively to acquire constituency benefits. Each of these three categories accounted for about 1,200 news stories. If this is representative for our sample, then newspaper coverage should inform voters of the representatives' ideological positions, how well they act as local agents, and, consequently, help voters rate their representatives.

4 Voter information

To hold their political representatives accountable, voters require information. Information reduces the chances that voters make "mistakes" when casting their ballots, and increases the electoral cost of politicians enacting poor policies, "shirking" and being corrupt. This was the belief of the Framers, and is a standard feature of monitoring models of political agency, see e.g. Persson and Tabellini (2000). A variety of models also predict that policy becomes more distorted in favor of narrow interest groups when voters are less informed, and that in a pluralistic world, policy outcomes favor groups with more informed members, see, e.g., Grossman and Helpman (2001).

Information may also affect citizens' welfare because it is useful in private decisions that depend on government policies, or simply because voters dislike uncertainty. Bartels (1986) and Alvarez (1997) present evidence that voters dislike candidates with uncertain policy positions.[8] Finally, easy access to information may increase political participation.

We will analyze five variables related to voter information. The variables are based on survey data from the American National Election Studies (ANES) from 1982 to 2004. The ANES is biannual, coinciding with congressional elections, and contains an average of about 1,800 respondents per year. The first two variables are based on direct knowledge questions. The ANES asked respondents if they could name the candidates in the U.S. House races in their district. We code the dummy variable $NameRecall$ as one if a respondent could do this, and zero otherwise. We code the variable $NameRecognition$ as one if the respondent could identify the incumbent from a list of the major party candidates for the U.S. House in their district, and zero if they provided an incorrect answer. In our sample, 31% of the respondents could correctly name at least one of the candidates, while a consdirably larger share, 92%, could recognize the incumbents name on a list; see Table 3.

The final three variables describe respondents willingness to describe or rate their House representative. These types of variables have been used to proxy voter information, since voters who know little or nothing about a representative are presumably less willing to describe or rate this representative.[9] These variables are also directly related to accountability. In order to hold representatives accountable, it helps to have an opinion on where they stand ideologically and what is to like or dislike about them. The ANES asked respondents to place their House representative on a seven point ideological scale. We code the variable $IdeologicalRatingProvided$ as one if the respondent provided a rating, and zero if the respondent answered "don't know" or "don't recognize" the incumbent. The survey also asked respondents to rate their feelings towards the incumbent on a scale from 0 to 100. We code

[8] Although, see Berinsky and Lewis (2004) for a critique of some of this literature.

[9] See, e.g., Delli Carpini and Keeter (1989).

the variable *FeelingThermometerProvided* as one if such a rating was provided, and zero if the respondent answered "don't know where to rate," "can't judge" or "doesn't recognize the incumbent's name." Finally, we code the variable *LikesOrDislikesProvided* as one if the respondent mentioned at least one thing that he or she liked or disliked about the incumbent House representative in their district, and zero if the respondent said that they did not know anything about this candidate. In our sample, 70% were willing to describe the incumbent's ideological position, 82% provided a feeling thermometer rating, and 53% mentioned something that they liked or disliked about the incumbent; see summary statistics in Table 3. Our five variables are positively correlated, but far from perfectly so. For example, of respondents who can name a candidate, 80% mention something that they like or dislike about the incumbent, compared to 40% among those who cannot name any candidate. The simple correlations are all larger than 0.2 and less than 0.4. Not all survey questions were asked all years, which accounts for the differences in the number of observations in Table 3.

Our key independent variable is *Congruence*, and we use the following specification:

$$y_{it} = \gamma\, Congruence_{cdt} + \mathbf{x}_{icdt}\delta + \alpha_t + \alpha_r + \varepsilon_i,$$

where y_i are our voter information variables, x_{icdt} is a vector of control variables and α_t a year-specific fixed effect. Finally, α_r is a fixed effect for each representative and three consecutive terms, for example, a fixed effect for Linda Smith's first three terms as representative for Washington's third congressional district, another for her next three terms, etc. Since the dependent variable is dichotomous, the specification assumes a linear probability model. We employ this because the linear probability model is consistent under weak assumptions, it works well with fixed effects, and its coefficient estimates are simple to interpret. All reported standard errors are heteroscedastic consistent.

Table 4 contains regression results. The columns contain different specifications, and the rows contain results for our five dependent variables. The first column of Table 4 shows the result of a regression of our information variables on *Congruence*, controlling only for the representative- and year-specific fixed effects. *Congruence* is significantly and positively correlated with all of these except *NameRecognition*. For this variable, the estimated coefficient on *Congruence* is of the same size as the standad errors in these regressions. Note that the additional effect of *Congrunce* increasing *NameRecognition* is likely to be small, since recognizing the representative's name is fairly simple and over 90% do this. The other estimated coefficients of *Congruence* are large. For example, a change from the lowest to the highest values of *Congruence* is associated with a 28 percent increase in the probability of correctly recalling a candidate's name. This is about as large as the effect of changing a respondent's education from grade school to some college. We now try to establish that the coefficient estimates in this column are good measures of the causal effect of *Congruence*

on voter knowledge. We extensively discuss identification concerns here. As many concerns will arise again in the coming sections, we will frequently refer back to this discussion, for example, for definitions of sets of control variables and specifications.

Adding controls Political knowledge is likely to be correlated with respondent characteristics such as education, income, and age. To distinguish the effect of the news market, we include a large set of *individual-specific controls* in the vector x_{icdt}. We include dummy variables for whether the respondent's party identification matches that of the incumbent, for the respondent's education (4 categories), income (5 categories), age (7 categories), gender, race, the number of years the respondent has lived in the community. These are 15 variables in total, after dropping one category of education, income and age to avoid perfect collinearity.

We include seven *county controls* characterizing the county in which the respondent lives: population (logged), average education levels (share with 1-11 years, share with 12 years, and share with more than 12 years), per capita income (logged), share aged below 20, share aged above 65, share black, and share female. We include eight *representative controls*: the representative's age, tenure, and dummy variables for whether the representative is a freshman, belongs to the majority party, died or retired, serves on a powerful committee, or is party leader, chair, or ranking member. There are three *electoral-race controls*: the closeness of the race, measured by the negative absolute difference between the democratic share of the vote and 0.5; a dummy variable for whether there was an incumbent running; and a dummy variable for whether the incumbent ran unopposed. We include the following *urbanism controls*: the degree of urbanism in the community where the respondent lives (3 dummy variables), the share urban population, logged population density. There are four variables in total, after dropping one category to avoid perfect collinearity. Finally, we include state-by-year fixed-effects, α_{sy}, because growing states receive additional congressional districts after redistricting, which typically changes $Congruence$, and state growth may be correlated with our outcome variables.

The second column of Table 4 shows the result of regressions of our information variables on $Congruence$, adding all above mentioned controls. The estimated coefficients and standard errors do not change considerably. For example, $Congruence$ remains significantly and positively correlated with knowledge about the candidates' names. Most demographic variables correlate with voter information in the expected direction. Older, better educated, and higher-income respondents are more likely to be able to name a candidate, as are male, white respondents. Moreover, respondents with the same party identification as the incumbent are significantly better at naming a candidate. More respondents can correctly name a candidate when the incumbent representative is running opposed, and when the incum-

bent is young. Finally, voters are better able to name their representative when the county population is larger. This is the only county-specific variable which is significant.

Within-race and redistricting specifications Our main identification concerns fall into two categories: those related to representative characteristics and those related to population or location characteristics. Regarding the first, more interesting and well known congressmen attract more media coverage. If this additional coverage impacts newspaper sales, then it will induce a spurious correlation between *Congruence* and voter knowledge. However, the impact of coverage of the House representative on newspaper sales is likely to be very modest. A typical newspaper prints about 100 stories per congressman and year; this can be compared to the 1000 stories that are printed about the governor. Another apparent representative-related concern is that certain types of congressmen self-select into highly congruent districts. However, candidate selection is part of the media effects we wish to study. A third representative-related concern is that congressmen who become increasingly powerful manage to increase *Congruence* in their district during redistricting. We find no correlation between being on powerful committees and increases in *Congruence*. Still, this could be a potential concern.

We address these concerns by a *within-congressional-race* specification. By including a district-by-year fixed-effect, we control for all factors that are constant within congressional races, such as the characteristics of the incumbent and challenger, and the type of race. This specification asks the question, are voters who live in counties with higher *Congruence* than average in a certain district in a particular year better able to name their incumbent than the average in that district in that year? We can ask this question because many districts contain or overlap more than one county, so some voters (and survey respondents) in a given district live in parts of the district with a high degree of *Congruence*, and others live in parts with a low level of *Congruence*.

The results from the *within-race* specifications are show in Table 4, Columns III to IV. The estimated coefficients on *Congruence* do not change to any considerable extent, nor do the estimated standard errors.

The other category of concerns relates to population or location characteristics. The first of these is that *Congruence* is lower in big cities. For example, New York is carved up into a large number of congressional districts. Newspapers in the New York media market sell in many districts and will not have a large share of their readers in any district. Other concerns in this category are that *Congruence* may, for some reason, be correlated with unobserved features that are also related to voter knowledge, such as unobserved income, education, and age structure.

We address these concerns by a *redistricting* specification. This uses county fixed effects

and identifies *Congruence* effects through changes due to redistricting. We apply the 1982-2004 average sales of each newspaper in each county to all years. Consequently, *Congruence* in a county-district only changes when the congressional district boundaries change due to redistricting. For example, a county becomes more congruent if moved to a district which their newspapers cover more (have larger sales in). In 1992, 1376 counties became more congruent, 1532 became less congruent, and in 220 there was no change. The average absolute change in Congruence was 0.1.

The results from the *redistricting* specifications are show in Table 4, Columns V and VI. The estimated effects of *Congruence* retain their significance levels and do not change considerably in magnitude. Finally, removing big cities from the sample and estimating the specification in Column VI yields very similar results. For example, the estimated coefficient on *Congruence* on *NameRecall* is 0.31 with a standard error of 0.07 (not shown in table).

Auxiliary regressions Another piece of evidence that *Congruence* is unlikely to be correlated with unobserved correlates of knowledge is that it is not significantly correlated with those we do observe. As mentioned earlier, the estimated coefficients on *Congruence* in Table 4 are not very much affected by including the large set of control variables, even though these variables are very significant. The reason is that our key variable, *Congruence*, is only weakly correlated with the variables that affect individual respondents' information levels. We now investigate this directly by running auxiliary regressions of *Congruence* on all our control variables. It would be worrying if, for example, demographic variables were correlated with *Congruence*. Note, however, that some control variables may be correlated with *Congruence*. For example, in section 7 we will show that *Congruence* increases the support for the incumbent. This means that more respondents may prefer the incumbent's party in highly congruent areas, and that incumbents in these areas may have longer tenure.

Table 5 shows the number of variables that are individually significant at the five-percent level, with the total number of variables in that category in parenthesis, and the p-value of a F-test for excluding all variables in that category.

Of the individual controls, the variable indicating that the respondent's party matches that of the incumbent is significant both in the baseline and the within-race specifications. This correlation may be explained by a positive incumbency advantage effect of *Congruence*; see below. As expected, *Congruence* is correlated with county characteristics, but not in the Redistricting regression which includes county dummy variables. Similarly, *Congruence* is correlated with some race and representative characteristics, but not in the within-race specification where these partial correlations are not identified. In the redistricting specification, only 2 variables out of 40 are individually significant at the five percent level: the dummy variables on the incumbent being unopposed (consistent with a positive incumbency

advantage effect) and living in a big city. To further guard against non-linear effects of urbanism from now on, we add 5 dummy variables for share urban and log population density to the set of *urbanism controls*.

General knowledge effects As a final robustness check, we see whether *Congruence* is related to general measures of voter knowledge. The congruence between a newspaper's market and a congressional district influences how much the newspaper writes about a particular congressman, but it should not influence how much the newspaper writes about many other things, such as U.S. senators or the party that controls the U.S. House or the U.S. Senate. Therefore, we replaced the dependent variable with knowledge questions not related to a particular representative or district. Table 6 shows results from regressions using our baseline specification. Columns I and II shows the results for the questions corresponding to *NameRecall* and *FeelingThermometer* but for senators. Column I shows a regression on whether the respondent can name at least one Senator correctly (46 percent of the respondents were able to do this). In the second column, the dependent variable is a dummy for whether the respondent was willing to place the incumbent senator on a feeling thermometer (91% where). We also check two other knowledge questions: whether the respondent knows which party had a majority in the House before the election (56% correct), and which party had a majority in the Senate (50% correct). The results show that *Congruence* is not correlated with any of these measures of "non-district-specific" political knowledge. This increases our confidence that the positive correlations shown in Table 4, between *Congruence* and district-specific information, are not spurious.

TV and radio Next, we investigate whether the effect mainly goes through newspapers or other media. Newspaper markets and radio and TV markets are both centered around large cities and towns. The two measures of *Congruence* and *TVCongruence* are fairly highly correlated, with a correlation coefficient of .64.

We first evaluate how these two congruence measures explain differential media exposure. In the surveys conducted between 1982 and 1994, the ANES asks respondents whether they have read about the incumbent House representative in a newspaper or magazine (50% say they have), or heard about him or her on TV (50%) or radio (25%). We constructed dummy variables for positive responses to each of these three questions. Table 7 shows regressions of these dummy variables on both *Congruence* and *TVCongruence*. The first column shows that *Congruence*, but not *TVCongruence*, is significantly positively correlated with respondents saying that they have read about the incumbent in a newspaper or magazine. The next two columns show that the opposite is true for radio and TV exposure. The estimated effects are large. *Congruence* increasing from zero to one is associated with a 45% increase in the

share of respondents who say that they have read about the incumbent. This is consistent with our newspaper content analysis, finding that an increase in *ReaderShare* from zero to one is associated with a 50% increase in articles about the representative. A caveat regarding these dependent variables is in place. We know from other studies that people have a hard time remembering exactly where they read or heard something. For example, people report seeing advertisements in areas where these advertisements have not been shown.

Interestingly, we find that *TVCongruence* is not significantly related to *NameRecall*, while *Congruence* is; see the last column of Table 7. This is consistent with the view that voters get most of their information about their local congressmen from newspapers rather than television. This is plausible, since content analyses of television news suggest that local television stations devote much less news coverage to local congressmen than do local newspapers (Hess, 1991; Vinson, 2003; see also, the discussion in Arnold, 2004).

5 Representatives' behavior in Congress

The results from the previous section indicate that voters in highly congruent districts are better informed about their representatives in congress. These voters are better able to name their representatives, more willing to place them ideologically, more willing to mention something that they like or dislike about the representative, and more willing to rate them on a feeling thermometer. This is reasonable since newspaper coverage of the representatives varies greatly with *Congruence* and the most commonly covered policy-related activities of the congressmen are roll-call voting, advance position taking, and acting as local agents (Arnold, 2004). We now expore whether news coverage influences the positions that representatives take, especially on roll calls, and the effort they devote to acting as local agents.

To provide a more concrete picture of what news coverage looks like, consider the case of Representative Jack Brooks of Texas ninth district. Brooks represented a highly congruent district, and was frequently in the press. In the 1994 election, his unpopular stance on a crime bill was extensively covered. For example, the *Houston Chronicle* reports "Now, Stockman [the challenger] is trying to portray Brooks as being soft on the issue because the congressman supported the recently enacted crime bill, which bans assault weapons."[10] and speculates about other provisons in the bill: "Was this a payoff by the Clinton administration to buy Jack Brooks' cooperation as he ignores his pro-gun constituents?"[11]. The paper also discusses Brooks performance as a local agent: "If Brooks is so powerful, where's the pork?

[10] The *Houston Chronicle*, October 9, 1994, "Campaign '94; Brooks feels right at home in the House; Challengers saying incumbent's 42 years in Congress are enough."

[11] The *Houston Chronicle*, August 21, 1994, "Did pork buy gun bill?"

... Brooks can tick off a long list of public works projects he takes credit for bringing to the district..."[12] Brooks was serving on the Judiciary Committee, with limited direct influence over grants. He was, however, actively lobbying for federal money to the district in frequent witness testimonies before congressional hearings, as we discuss below. He lost the election of 1994, and the defeat was partly attributed to his position on a crime bill.

Because they receive more press coverage, representatives from highly congruent areas might have a greater incentive to engage in constituency service and pork-barrel politics. If successful, they can look forward to local newspaper stories reporting on their success – free advertising, from their point of view – and if unsuccessful they might find that reported as well. Members representing districts with low values of *Congruence* will not receive much free reporting of their activities, and even though the franking privilege reduces the costs of promoting their records, self-promotion is probably viewed more skeptically than reporting by independent journalists. Similarly, representatives from highly congruent districts might have greater incentives not to cast their roll-call votes against the interest of their constituency. Doing so is more likely to generate publicity and electoral loss.

In this section, we investigate whether congressmen respond to these incentives. We study whether representatives from highly congruent areas appear more as witnesses before congressional hearings, are over-represented on "constituency-oriented" committees and under-represented on broad "policy-oriented" committees, and more frequently vote against their party line.

5.1 Specification and identification

The level of analysis is now at the congressional district, and *Congruence* is now measured at this level. Suppose that M newspapers, indexed by m, sell in district d, and let $MarketShare_{md}$ be the market share of newspaper m in district d. Then,

$$Congruence_d = \sum_{m=1}^{M} MarketShare_{md} ReaderShare_{md}.$$

At the district level, $Congruence_d$ has a mean of 0.45 and a standard deviation of 0.24, see Table 8. Our unit of analysis will be congressional district by congress. We match witness appearances, roll-call voting, and committee assignments each congress with $Congruence_d$ in the coming election. For example, we match witness appearances in 2001 and 2002 with $Congruence_d$ computed using the district lines of the 2002 election.

[12] The *Houston Chronicle*, October 9, 1994, "Campaign '94; Brooks feels right at home in the House; Challengers saying incumbent's 42 years in Congress are enough."

We include various *district controls*: the percent urban and five dummy variables for urbanness, log population density, and five dummy variables for log population density intervals, the log median income, the percent of, respectively, people aged 65 or older, military population, people employed in farming, foreign born, and blue collar workers. We also include the *representative*, and congressional *race controls* discussed in Section 4.

The identification concerns are similar to those discussed earlier, which is why we carefully control for urban-rural differences. Since the analysis is now at the district level, we cannot apply the within-race specification. We study the effects of redistricting, using district-fixed effects. However, note that since the districts change with redistricting, the fixed effects will not absorb all time-invariant heterogeneity. In sum, the identification is not as strong in this section as elsewhere.

5.2 Witness appearances

One way for House representatives to advance their constituencies' interests is by appearing as witnesses at congressional hearings. This is a costly procedure and, on average, representatives only do it a few times per congress. To build the case that a certain project should receive (or keep its) funding, the representative may have to gather data on and hire experts discussing impacts on their district and the nation. Jack Brooks was a frequent witness, appearing 14 times in 1994, of which 8 were before the appropriations committee. Consider, for example, his testimony before the House Appropriations Committee, March 23, 1994.

> "Mr. Chairman, thank you for this opportunity to appear today to urge the subcommittee's favorable consideration of the public works projects for the ninth congressional district of Texas included in the fiscal year 1995 budget.
>
> Mr. Chairman, I support funding for all of the projects included in the budget proposal... All of these federal public works projects are critical to the economic development of my district, which includes six major ports, and many oil and chemical plant operations which compose one of the largest industrial complexes in the United States. Ports along the Gulf coast are an integral part of the economy of the nation... I have with me today a number of persons vitally concerned with these projects. They are here to support these statements and to answer any questions the subcommittee may have. I would like to introduce them at this time..."

We have collected data on witness appearances by House representatives from 1982 to 2004. The variable $Witness$ contains the number of appearances per congressional session by each representative. Representatives stand witness before a congressional hearing on average 3.8 times per congress. However, there is considerable variation, as the standard deviation is

3.7 (see Table 8). The Appropriations and Ways and Means committees have a considerable direct influence over spending and taxes, and witness appearances before these committees may be more directly linked to influencing the allocation of funds. We therefore define the variable $WitnessBudget$, containing the number of appearances before the Appropriations or the Ways and Means Committees. Representatives stand witness before hearings in these committees 1.7 times per congress, on average.

We want to investigate whether representatives from highly congruent districts stand witness more often. First, we run a Poisson regression, controlling for our *district controls* and state and year fixed effects. The results are shown in the first column of Table 9. The coefficient on $Congruence$ in the district is significant and positive. The point estimate implies that an increase in $Congruence$ from zero to one is associated with an average increase of 46 percent in representative appearances. This implies 1.5 additional appearances per congress. The next column shows the results from estimating the same specification using a negative binomial regression (testing for goodness-of-fit rejects the Poisson specification). The following Column (III) adds the *representative* and *race controls*. To that specification, Column (IV) adds district fixed effects, and Column (V) replaces the district- and year- by state-by-year-fixed effects.

The final two columns show the results from regressions on witness appearances before the Appropriations and Ways and Means Committees only. The proportional effects are of about the same size as witnessing before all committees. However, since the representatives only witness before these committees an average 1.5 times per congress, the absolute effect is smaller.

Representatives who are more closely monitored by the newspapers more often appear as witnesses before committee hearings. The evidence presented in this section shows that this correlation is significantly positive and robust across a wide range of specifications. Surprisingly, even within the same congressional district, becoming more congruent over time is associated with more witness appearances. The effects are also significant in appearances before committees with a more direct say over the budget.

5.3 Committee choice

We now investigate whether representatives from congruent districts more frequently serve on committees which are effective in promoting constituency interests. Congressional scholars often divide the set of committees into four categories: constituency-oriented, policy-oriented, prestige, and undesirable. The presumption is that representatives serving on constituency-oriented committees are better able to engage in constituency service and pork-barrel politics. The opposite is true for representatives serving on policy-oriented committees, who instead

are better positioned to influence national policies.[13]

The committee assignments are influenced both by the individual representatives and the party leadership. In the beginning weeks of the first session of every congress, new congressmen submit requests, and returning congressmen may submit requests for second assignments or transfers. Based on these requests, each party's committee-on-committees determines the committee seat allocation. The single factor regarded as most important in distributing assignments to all but the prestige committees is whether a particular place will help insure the reelection of the member in question (Masters, 1961). Other goals are policy influence; see Fenno (1973) for a longer list of personal goals.

We now discuss our data and present the evidence. We have the committee assignments of each representative from 1982 to 2004. While these are around 5,000 observations, the number of independent observations is considerably smaller since many representatives remain on the same committee between congresses. In our sample period, there are about 1,000 representatives.

Operationally, we use Deering and Smith's (1997) classification of constituency committees.[14] We define the dummy variable $DistribComm$ to equal one if more than 50 percent of the representative's committee assignments are on constituency-oriented committees, and zero otherwise. This is true for about 28 percent of the representative-year observations. We also constructing the variable $PolicyComm$ which equals one if the representative serves on the House Judiciary or the International Relations/Foreign Affairs Committees. This variable is one for about 15 percent of the representative-year observations.

The committee assignments have been found to correlate with district demographics. For example, Rhode and Shepsle (1973) find they are correlated with the population share residing in urban areas, and the share of the labor force employed in the armed forces or in agriculture. These variables are included in our *district controls*.

Our results from regressions of committee assignments on $Congruence$ are presented in Table 10. The first column shows that $Congruence_d$ is positively and significantly correlated with $DistribComm$, including only fixed effects for each year-state combination. In all specifications, the standard errors are clustered by representative. However, when we add

[13] Appropriations, Budget, Rules, and Ways and Means are designated as "prestige" committiees. See Fenno (1973) for a detailed discussion and early classification. See, e.g., Baumgartner, Jones and MacLeod (1997), Thorson, Glieden, and Lina (1999), Frisch and Kelly (2002) for recent papers using similar classifications.

[14] Before the 104th Congress: Agriculture; Armed Services; Interior and Insular Affairs (Natural Resources in the 103rd Congress); Merchant Marine and Fisheries; Public Works and Transportation; Science, Space, and Technology; Small Business; and Veterans Affairs. In and after the 104th Congress, we designate the following as constituency committees: Agriculture; National Security; Resources; Science; Small Business; Transportation and Infrastructure; and Veterans Affairs.

the *district, representative*, and *race controls*, $Congruence_d$ loses significance; see Column II. Columns III and IV use district fixed effects, with and without controls. Congruence is insignificant in both these specifications. Columns V-VIII replace $DistribComm$ with $PolicyComm$ as the dependent variable. Policy committee assignments are negatively correlated with Congruence, and significantly so in all specifications but one.

To sum up, we find weak evidence that $Congruence$ affects committee assignments. $Congruence$ is mostly significantly negatively correlated with being on a policy committee, but not consistently so across specifications. The effect of being on a distributive committee is positive, but insignificant except in the most basic specification. This might reflect a low statistical power because of the high persistence in assignments. The variable $DistribComm$ changes between congresses for only 7 percent of the representatives. A larger fraction, 22 percent, of representatives change their $PolicyComm$ value between congresses.

5.4 Roll-call voting

Jack Brooks' vote in favor of gun-control was in step with his party's leadership, but not with his constituency. In general, in determining whether to vote for or against a particular bill, elected officials must balance the wishes of the overall electorate, specific constituencies within the electorate, pressure from party leaders within the Congress, and their own ideology. We hypothesize that representatives in more congruent districts vote more in line with constituency interests, because their constituencies are better able to monitor them.

We restrict the analysis to the ideological dimension of roll-call voting. There is some empirical evidence that representatives' voting is typically more ideologically extreme than what their constituencies would prefer. A number of studies find a significant negative correlation between vote support and representatives' "ideological extremism" in roll-call voting (Democrats voting more to the left and Republicans voting more to the right); see e.g. Canes-Wrone et al. (2002). In a related vein, Ansolabehere et al. (2001) find that the ideological gap between Republican and Democratic candidates is smaller in races that are expected to be close. This presumably occurs because the competition for votes makes representatives more prone to consider constituency interest over party and personal interests.

We test whether the ideological gap between Republican and Democratic representatives' voting is smaller in highly congruent districts. To measure the ideological position of the representatives, we use NOMINATE scores, constructed by Poole and Rosenthal (1997). NOMINATE is a nonlinear latent variable model, somewhat akin to a factor analytic model, used to estimate the underlying ideology that drives observed roll call behavior. Roughly speaking, it assigns an ideological position to each legislator in order to correctly predict as many roll-call votes as possible. The range of scores is $[-1, 1]$, and scores are increasing in conservativeness. The first dimension NOMINATE scores correlate well with other mea-

sures of ideology, such as ADA scores. We define the variable *NominateScores* as the first dimension ideal point estimates from DW-NOMINATE.[15]

We have data on NOMINATE scores for House representatives for 1982-2002. Figure 4 plots these scores against the share of voters that supported the Republican candidate for the House election in the congressional district. Scores from districts where *Congruence* is above (below) the median are in grey (black). The solid lines show the predicted values from separate linear regressions of Nominate scores on Democratic vote shares and a Democrat-wins dummy variable within the two sub-groups. The gap in voting is smaller in the highly congruent districts. In other words, representatives are more moderate in their voting behavior in these districts. The same relationship is apparent in the "party loyalty" scores, which is not surprising since ideologically extreme representatives tend to be those who vote consistently with the party line. Recall from Figure 1(f) above that party loyalty (measured as the percentage of times a member votes with the majority of the party leadership) is inversely related to *Congruence*.

We now analyze these differences more carefully by regressing the *NominateScores* on a dummy variable for the Democratic candidate winning the election, *Congruence*, and the interaction of the two. The interaction allows the jump in *NominateScores* at 50% Democratic votes to vary with *Congruence*. A positive interaction coefficient implies that the negative jump at 50% is lower for more congruent districts. We also include fifth degree polynomials in Democratic vote shares in the House and presidential elections to account for any continuous influence on Nominate scores; using a specification similar to that in Lee, Moretti and Butler (2004). The presidential vote share is interpolated for years in which there was no presidential election. We limit the sample to elections where both parties had at least a one percent vote share.

The results are presented in Table 11. The first column only contains only the above mentioned variables. There is an estimated drop of -0.69 in *NominateScores* when a Democrat is elected instead of a Republican (at zero *Congruence*). The drop is -0.50 (-0.69+0.19) when *Congruence* is one. The next column adds the *district, race* and *representative characteristics* discussed above. This causes the estimated coefficient on the interaction term to drop to 0.17.

Because the drop may depend on district urbanism, we interact the dummy variable for the Democratic candidate winning, not only with *Congruence*, but with the percent of urban and five dummy variables for urbaneness, the log population density, and five dummy variables for log population density intervals. The estimated coefficient on the

[15] See Poole and Rosenthal (1997) for an exact definition of the statistical model underlying NOMINATE and DW-NOMINATE, as well as discussion of the advantages of NOMINATE over other techniques and indices.

interaction term is not much affected, see Column III. Because we interact with our urbanism controls, the direct effect of the Democratic candidate winning does not have the same interpretation as in Columns I and II. Column IV includes congressional-district and year-fixed effects. This specification identifies the coefficient on Democratic victory only through districts where the party in power has changed. The final column adds the controls and Democratic win dummy variable interacted with district urbanness. The interaction term remains significantly positive.

We find that the ideological distance between Democratic and Republican representatives' roll-call voting is about 25 percent lower (.17/.68=.25) in perfectly congruent than in perfectly non-congruent areas. This is consistent with the hypothesis that representatives moderate their behavior in highly congruent districts because the better monitoring of the newspapers increases the electoral costs of ideologically extreme or highly partisan roll-call voting.

6 Policy

We have argued that representatives that are closely monitored by local newspapers have a greater incentive to provide services to their constituency. We have also provided some evidence that representatives react to these incentives, they appear more as witnesses before congressional hearings, perhaps sit less on policy committees, and display less extreme roll-call voting patterns. If the representatives' efforts in Congress are successful, then we should observe more federal funds flowing into highly congruent districts, and to more congruent counties within districts.

To test this, we assembled data on federal expenditure allocations across counties from the Consolidated Federal Funds Report. The expenditures we analyze include grants, procurement contracts, salaries and wages, and direct payments for retirement and disability, and other direct payments. We exclude loans, insurance and social security payments.[16] The expenditures are in constant 2000 dollar values. The total value of the expenditures we study is about $ 2,700 per capita, or about 10 percent of U.S. GDP. We have this data yearly for the period 1983-2004. However, we collapse the data by Congressional session, and study mean annual expenditures per Congress. We define the variable *Spending* as the (log of) this mean annual spending per Congress.

Since the unit of analysis is now counties and some counties cross district boundaries, we average the variables containing district controls. We compute the population weighted

[16] The loans data does not report the value of the loan subsidy but only the value of the loan, the insurance data reports obligations and not payments, and the social security data is excluded because representatives are unlikely to affect its distribution.

average $Congruence_{cd}$ in a county

$$Congruence_c = \sum \frac{n_{cd}}{n_c} Congruence_{cd},$$

where n_{cd}/n_c is the share of the population of county c that lives in district d. Similarly, we compute the population weighted averages of district level variables such as representative seniority, committee assignments, etc., and use these in the regressions.

Table 12 shows the results of a set of regressions of *Spending* on $Congruence_c$. We apply the *baseline*, *within-race* and *redistricting* specification with and without of our sets of *county, urbanism, representative*, and *race controls*. We also include a dummy variable for the county containing the state capital. Once more, it is important to control for the urban dimension since it correlates with both federal spending and *Congruence*. In the *within-race* specification, we only look at counties which lie entirely in one district. We also add a "within neighbor-counties" specification where each county that lies along a district border has been matched with the closest county on the other side of the border. This specification includes dummy variables for each such neighbor pair.

The coefficient on *Congruence* is significantly positive in all specifications that use cross-sectional variation (which is all but the last two). Comparing counties within the same district and year, significantly more funds were spent in more *Congruent* counties. Comparing neighboring counties on different sides of district borders, significantly more funds were spent in the more congruent neighbor. The estimated effect is consistently about 0.1 which implies that a shift in Congruence from zero to one is associated with a 10-percent increase in spending.

In the redistricting regression, the coefficient on $Congruence_c$ is only significant at the 10-percent level after adding controls. The size of the coefficient also drops to 3-5 percent. The reason, we believe, is that there is a very high inertia in spending. Much of the funding that the government spends in a given year was authorized in previous sessions of Congress as the total money allocated is the sum of continuing awards still in existence as well as new funds being awarded that year. For this reason, the stock of spending is less related to congressional action in any given year. It may be possible to address this issue by using data that identifies new awards by year, such the U.S. Domestic Assistance Programs Database used by Bickers and Stein. We leave this for future work.

In sum, we find strong evidence that media coverage substantially influences federal spending. This is the last link in the accountability chain. Our results on news coverage, voter information, politicians' actions and policy consistently support the idea that an active media coverage of politics increases electoral accountability.

7 Extensions: Voter turnout and incumbency advantage

We next turn to two important additional issues: the effects of press coverage on voter turnout and the incumbency advantage.

7.1 Voter turnout

For most Americans, voting is the main form of political participation. This participation is, in the view of many observers, beneficial per se. A high turnout in fair elections gives legitimacy to public officials and their decisions. Voting is also an effective means of holding elected officials accountable for their decisions and behavior in office. Voter abstention may signal that citizens are dissatisfied with their political institutions, or lack a sense of efficacy and confidence in government, and habitual abstention may further erode citizens' satisfaction and confidence.

It is clear that *Congruence* could influence political participation. We have seen that it affects voters' information, and numerous surveys find a strong and positive correlation between citizens' information levels and political participation.[17] While causation is more difficult to establish, several studies are suggestive. The strong correlation between education and turnout may also be interpreted as evidence that decreasing the cost of information increases participation.[18]

The analysis of voter turnout in House elections is complicated by the fact that people vote for multiple offices (and sometimes also controversial and important referenda). The decision to go to the voting booth depends on all these elections. One way of dealing with this is to control for turnout for the most important office. This specification analyzes "roll-off." That is, given that a voter has come to vote for, say, the office of president, how likely is it that he or she will not vote for Congress?[19]

To analyze voter turnout, we use data on votes in congressional elections, by county and congressional district , for 1984-1992 and 1998-2004.[20] Voter turnout in the presidential

[17] See, e.g., Verba and Nie (1972).

[18] The underlying assumption is that education reduces the costs of acquiring information. See, e.g., Wolfinger and Rosenstone (1980) for evidence of the correlation.

[19] Another possibility is to study elections when there was no presidential, gubernatorial or senatorial contest on the ballot, so the House race was arguably the "top of the ticket." Unfortunately, we cannot do this because there are two few such cases, from a small number of states.

[20] We could also use the self-reported turnout data in the ANES. However, it is well known that respondents grossly overstate their propensity to vote, and also overstate their propensity to vote for winning candidates and incumbents (at least in U.S. House races). For this reason, we prefer the actual voting data. When we

election will be used as a control, and we have data on presidential election returns by county and district for the 1984, 1988 and 2000 elections. We drop uncontested elections and observations where the reported election turnout is larger than the total population.

In our sample, the average share of the population who voted in the House election was 36 percent. The turnout rate is substantially higher in presidential election years (about 40 percent) than in other years (about 30 percent). When we have data on the number of votes cast both in congressional and presidential elections, the average turnout is 1.2 percent lower in congressional in presidential elections.

We regress the turnout in House elections on *Congruence*, controlling for turnout in the presidential election. The results are shown in Table 14. The coefficient on *Congruence* is highly significant in the baseline and within-race specifications. In the redistricting specification, the standard errors are large and the estimated effect is not significant. Apparently, there is not enough time-series variation to identify effects.

In the baseline specification, an increase in *Congruence* from 0 to 1 is associated with a 0.8 percent increase in House election turnout (as share of the population). This is two-thirds of the average 1.2 percent roll-off if presidential election years. While this is a large percentage decline in the "roll off", the absolute increase in turnout is not large. It is possible that the effect is larger in elections when the House election is at the "top of the ticket." However, running the same regression as in Table 14 for this sub-sample yields an estimated effect in the baseline regression of 0.8 with a standard error of 0.6. So it is unlikely that the impact is larger than, say, 2 percent.

Given earlier large estimated effects of media access on turnout, these moderate effects are somewhat surprising. Strömberg (2004) estimates that the increasing use of radio in 1920-1940 led to an increase in votes per capita of about 5.5 percent in gubernatorial elections. More closely related to our study, Gentzkow (2006) estimates that the increasing use of television in 1950-1970 *reduced* House turnout in the non-South by 5.6 percent in off-year elections and 3.1 percent in presidential elections. The reason is, he argues, that people stopped reading newspapers and listening to the radio and instead turned to the uninformative TV. Our estimates suggest that if local newspapers completely stopped covering House politics, then turnout would drop by less than one percent in presidential-election years. One possible explanation for the differences is that the introduction of radio and TV affected voter turnout for all offices, while *Congruence* only affects House turnout.

analyze the ANES we find significant and even larger effects of *Congruence* both on voter turnout and the incumbency advantage.

7.2 Incumbency advantage

The electoral advantage of incumbents is one of the most studied issues in U.S. congressional politics. The dramatic increase in the incumbency advantage since the 1950s has sparked a discussion of its origins, and electoral, democratic and policy consequences. A central question in this research is why this advantage has been growing so rapidly since the 1950s. Researchers have proposed a variety of potential explanations including: the electorate has become more ideologically polarized, incumbents have been able to create safe districts by redistricting, incumbents control increasingly valuable resources for constituency service, and incumbents have received increasing media exposure. We investigate a question relating to this last point. Specifically, we study whether people in areas with more media coverage of congressional politics are more likely to vote for the incumbent.

It is not obvious that voters with more information will be more supportive of incumbents. However, if pure name recall acts as a signal of candidate quality, then media coverage may increase electoral support.[21] The incumbency advantage may also increase if media coverage of incumbents is biased towards positive news, or if media provides more precise information about incumbents and voters are risk averse. Finally, changes in representatives' behavior induced by media coverage may influence voter support.

Using the ANES survey data, we first analyze whether newspapers differentially inform voters about challengers and incumbents. This can be done since the ANES records separately whether respondents correctly recalled the names of both the incumbent and the challenger running in their district. The ANES also records separately whether respondents report having read about the incumbent or the challenger. We created dummy variables for each of these four items. Incumbents are better known than challengers, and more people read about them. In contested elections with an incumbent running for re-election, about 50% of the respondents say that they have read about the incumbent, compared to 15% for the challenger; and about 30% can correctly name the incumbent, compared to 10% for the challenger.

What are the differential impacts of *Congruence*? Table 15 shows the results from regressions of the "read about" and name recall dummy variables on *Congruence*. Higher *Congruence* significantly increases both incumbent and challenger name recall. However, the effects are larger for incumbents than for challengers (0.33 as compared to 0.12). Moreover, while higher *Congruence* also significantly increases the probability of reading about the incumbent, it appears to have little effect on reading about challengers.

We now turn to voting data to see whether incumbents receive more vote support in highly

[21] For example, good representatives may be more active in Congress and hence receive more media coverage. See Nelson (1970) or Kihlstrom and Riordan (1984) for a similar argument for consumer product advertisements.

congruent areas. Following the previous literature, we measure the incumbency advantage by regressing the Democratic vote share on an indicator variable for whether the incumbent House representative is a Democrat, controlling for other factors that determine the election outcome. To test whether incumbents are more advantaged in highly congruent areas, we add the interaction between the dummy variable for Democratic incumbent and *Congruence*. We demean the incumbency advantage and *Congruence* variables before interacting them, so that the coefficients on the main effects show the impact of these variables measured at their sample means. The specification is of the form:

$$\begin{aligned} vh_{cdt} &= \beta_1 I_{dt} + \beta_2 I'_{dt} Congruence'_{cdt} + \beta_3 Congruence_{cdt} + normalvote_{cdt} \\ &\quad + \mathbf{x}_{cdt}\delta + I'_{dt}\mathbf{x}'_{cdt}\delta + \alpha_{st} + \varepsilon_{cdt}, \end{aligned}$$

where vh_{cdt} is the Democratic share of the two party vote in county c in district d in the House election at time t, and I_{dt} is a dummy variable for whether the incumbent in district d was a Democrat. The prime superscript denotes demeaned variables, for example, I'_{dt} equals I_{dt} minus the sample mean of I_{dt}. As controls, $\mathbf{x}_{icdt}$, we include the *race*, *representative*, *county,* and *urbanism controls* discussed in section 4. We also interact the controls with the incumbency dummy variable to allow the incumbency effect to vary with the degree of urbanness, etc.

We use two sets of variables proxying $normalvote_{cdt}$, the constituencies' underlying partisan preferences (e.g., Levitt and Wolfram, 1997, and Gelman and King 1990). The Levitt and Wolfram (1997) specification uses a district-by-county fixed effect for each districting period. The other specifications include the lagged Democratic vote in district d and a dummy variable for the Democrats being the incumbent party. Both specifications include the Democratic share of the two-party vote in the *presidential* election in the county-district part. Since our unit of observation is county by district by election, we can both use redistricting and within-district specifications. The results are show in Table 16.

We estimate an average personal incumbency effect of 8-13 percent, which is similar to previous estimates in the literature. In areas where we expect the press to cover congressional politics more, the incumbency advantage is about 1 percentage point larger. This is true whether or not we include controls, and whether we use the baseline, within-race, or redistricting specifications.

The incumbency advantage in House elections has risen from less than two percent in the 1940s to about eight percent today. Our estimates indicate that a relatively small share of this can be explained by changes in the media market. Given the large estimated effects on voter information, it also seems unlikely that increased voter information is a major determinant of the rise in the incumbency advantage.

8 Discussion and Conclusion

All is not safe, even where the press is free and every person is able to read. The press must also have economic incentives to cover politics. Those incentives might be muted by a poor fit between media markets and political districts. Stated from the positive side, we find that a good fit (high *Congruence*) between newspaper markets and congressional districts leads to more press coverage of the local congressmen. In areas where coverage is high for this reason, voters are better informed about their congressmen and more likely to participate in elections. Politicians respond to the increased media coverage by more actively pursuing their constituencies' interests. They moderate their partisan voting, more frequently stand witness before committee hearings, and (perhaps) more frequently work on constituency committees. As a result, we see more federal money flowing into more congruent districts.

Quantitatively, the effects of news coverage on voter information, politicians' actions and policy are large. We estimate that increasing *Congruence* from zero to one increases the number of articles written about the representatives by 50 percent, it increases voters' ability to correctly name the candidates in the House race by 30 percent and their willingness to describe and rate the representative by 20-30 percent, it increases the representatives' witness appearances by 46 percent, and reduces the gap between Republican and Democratic ideological differences in voting by 25 percent. Finally, it increases federal spending per capita by 10 percent. A one standard deviation increase in *Congruence* (0.45) induces effects that are slightly less than half this size. So the variation in *Congruence* induces a significant variation in all these outcomes. Effects on voter participation and the incumbency advantage are significant, but not large.

Our findings support the idea that press coverage is important for electoral accountability. Voters need information to keep politicians accountable and the press delivers this information. This was the belief of the Framers, and, that information improves accountability is also a standard feature of political agency models; see Persson and Tabellini (2000) or Grossman and Helpman (2001). The U.S. House districts constitute a suitable testing ground for this idea. While the districts are of similar size and political competence, press coverage varies (we argue) exogenously with the *Congruence* between media markets and congressional districts. Moreover, the importance of the U.S. House makes the welfare consequences of public action or inaction of high order.

Our results are helpful in understanding the political consequences of trends in media markets. At the start of the twentieth century, every large and medium-sized U.S. city and even many small towns had at least one newspaper, geared to local political needs. The number of newspapers has shrunk so that many communities no longer have a paper of their own. Newspapers have merged and been replaced by broadcast media, serving even larger regions. This trend is likely to decrease the congruence at local levels. This might not only

affect the congruence of congressional districts, but also that of municipalities, counties, and other local government units. Our results suggest that this is likely to reduce voter information, political participation, and political accountability.

The internet may influence this trend. Already at this stage, more than 1,300 domestic daily newspapers have a web presence, and 15 percent of all American adults say that the internet was the primary source of campaign news during the 2006 election (compared to 34% for newspapers).[22] The effect on information about local politics is unclear. On the one hand, there is a low fixed cost of starting an internet outlet, basically no space constraint, and good search technologies. This should increase the number of viewable relevant articles. On the other hand, most people do probably not search intensively for information about local politics, rather they get this information as a by-product. In addition, the cost of gathering, editing and writing of news is the same for print and online editions, and these costs make media write more about issues that concern larger shares of their readership. Moreover, since the potential internet audience is global, local reader shares may decrease.[23]

Because of the large effects, *Congruence* should be an important factor to consider in contemplating the regulation of the press to improve voter information and political accountability. It may be even more important than, say, the degree of competition between local newspapers. There exists an externality argument for why congruence may be too low. Better informed voters induce politicians to work more for their constituencies, which benefits all voters. However, individual voters do not fully internalize these benefits and, consequently, their willingness to pay for political information is low (see Downs, 1957). This causes media markets to conform too little to political districts.

Note that our results may generalize to media coverage of issues. In the same way that a representative works more for geographical constituencies that are better informed, they may also work more for better informed issue-defined constituencies (such as environmental groups, or people who care strongly about taxes). If our results generalize in this way, they have wide-ranging implications. The media coverage of issues varies greatly across issues in a systematic way. This could create a systematic policy bias disfavoring minority groups and groups not valuable for advertisers, see Strömberg (1999).

Finally, as an application, the strong "personal vote" in U.S. elections might in part be due to the local nature of much newspaper coverage, and could therefore be effected by these

[22] PEW Research Center (2007), "Election 2006 Online."

[23] The UK's *The Guardian* newspaper receives 78 per cent of its web readers from overseas (Mayes, 2004), compared to 30 per cent for *The New York Times* and 17 per cent for *The Washington Pos*t (New York Times, 2005; Tedeschi, 2004). The concentration in online media use is also very high. Although thousands of sites have contain information on any given topic, only a few sites account for most of the traffic. For example, three old media-related sites—*MSNBC, CNN*, and *The New York Times*—account for three-quarters of traffic online for U.S. news (Barnhurst, 2001).

trends. Today, relatively few voters in the U.S. read a "national" newspaper; rather, almost everyone who reads a newspaper reads a local paper. In 2003, the combined circulation of the three largest newspapers – *The New York Times, The Wall Street Journal,* and *USA Today* – was 6 million out of a total circulation in the U.S. of 50 million, or just 12 percent. In many other countries, the opposite is true. Citizens in these countries are therefore unlikely to encounter much news about their local politicians. This is even the case in countries with geographically defined constituencies, such as the U.K.

9 References

Ahrend, Rudiger (2002) "Press Freedom, Human Capital, and Corruption" , Working Paper no. 2002-11, DELTA.

Alvarez, R. Michael (1997) "Information and Elections" , Ann Arbor, MI: University of Michigan Press.

Ansolabehere, Stephen, James M. Snyder, Jr., and Charles Stewart, III (2001), "Candidate Positioning in U.S. House Elections" , American Journal of Political Science 45 (January), 136–59.

Ansolabehere, Stephen, Eric C. Snowberg, James M. Snyder, Jr. (2006) "Television and the Incumbency Advantage in U.S. Elections", Legislative Studies Quarterly 31(4), 469-490.

Arnold, R. Douglas (2004), "Congress, the Press, and Political Accountability", Princeton, NJ: Princeton University Press.

Barnhurst, Kevin G., (2002), "News Geography & Monopoly: the form of reports on U.S. newspaper internet sites", Journalism Studies, 4(3), 477–489.

Bartels, Larry M., (1986), "Issue Voting Under Uncertainty: An Empirical Test", American Journal of Political Science 30, 709-728.

Baumgartner, Frank R., Bryan D. Jones, and Michael C. MacLeod, (1997), "New Issues and Old Committees Jurisdictional Change in Congress, 1947–93."Unpublished manuscript, Pennsylvania State University.

Berinsky, Adam J., and Jeffrey B. Lewis, (2004), "Testing the Assumption of Quadratic Preferences: Voter Choice and Candidate Uncertainty in Spatial Models of Issue Voting." Unpublished manuscript, Massachusetts Institute of Technology.

Berkowitz, Dan, and David Pritchard, (1989), "Political Knowledge and Communication Resources", Journalism Quarterly, 66, 697-701.

Besley, Timothy, and Robin Burgess, (2002), "The Political Economy of Government Responsiveness: Theory and Evidence from India", Quarterly Journal of Economics 117, 1415-1452.

Besley, Timothy, and Andrea Prat, (2006), "Handcuffs for the Grabbing Hand? Media Capture and Government Accountability", The American Economic Review 96(3), 720-736.

Brunetti, Aymo, and Beatrice Weder (2003), "A Free Press Is Bad News for Corruption", Journal of Public Economics 87, 1801-1824.

Campbell, James E., John R. Alford and Keith Henry (1984) "Television Markets and Congressional Elections", Legislative Studies Quarterly 9, 665-678.

Canes-Wrone, Brandice, David W. Brady and John F. Cogan, (2002) "Out of Step, Out of Office: Electoral Accountability and House Members' Voting", American Political Science Review 96(1), 127-140.

Cohen, Marty, Hans Noel, and John Zaller (n.d.) "Local News and the Quality Of Congressional Representation", Unpublished manuscript, UCLA.

Delli Carpini, Michael X., and Scott Keeter (1989) "What Americans Know about Politics and Why It Matters", New Haven, CT, Yale University Press.

Deering, Christopher J., and Steven S. Smith (1997) "Committees in Congress", Washington, DC, Congressional Quarterly Press.

Djankov, Simeon, Caralee McLiesh, Tatiana Nenova, and Andrei Shleifer (2003), "Who Owns the Media?", Journal of Law and Economics 46, 341-383.

Downs, Anthony (1957) "An Economic Theory of Democracy", New York, Harper Collins.

Fenno, Richard F., Jr. (1973) "Congressmen in Committees", Boston, Little, Brown.

Frisch, Scott A., and Sean Q. Kelly (2002) "House Committee Assignments and Public Policy: Committee Theories Reconsidered", Unpublished manuscript, California State University at Bakersfield.

Gelman, Andrew and Gary King, (1990) "Estimating Incumbency Advantage without Bias", American Journal of Political Science 34(4), 1142-1164.

Gentzkow, Matthew (2006), "Television and Voter Turnout", The Quarterly Journal of Economics 121(3), 931-972.

Grossman, Gene, and Elhanan Helpman (2001) "Special Interest Politics", Cambridge, MA, MIT Press.

Hess, Stephen (1991) "Live From Capitol Hill!", Washington, DC, Brookings Institution.

Kihlstrom, Richard E., and Michael H. Riordan (1984) "Advertising as a Signal", Journal of Political Economy, 92(3), 427-450.

Lee, David S., Enrico Moretti and Matthew Butler (2004) "Do Voters Affect or Elect Policies? Evidence from the U.S. House", Quarterly Journal of Economics 119(3), 807-859.

Levitt, Steven D. and Catherine D. Wolfram, (1997) "Decomposing the Sources of Incumbency Advantage in the U. S. House", Legislative Studies Quarterly 22(1), 45-60.

Levy, Dena, and Peverill Squire (2000) "Television Markets and the Competitiveness of U.S. House Elections", Legislative Studies Quarterly 25, 313-325.

Masters, Nicholas A. (1961) "Committee Assignments in the House of Representatives", The American Political Science Review 55(2), 345-357.

Mayer, William G. (1993) "The Polls: Trends in Media Usage", Public Opinion Quarterly 57, 593-611.

McCarty, Nolan M., Keith T. Poole, and Howard Rosenthal (1997), "Income Redistribution and the Realignment of American Politics", Washington, AEI Press.

Mondak, Jeffrey J. (1995) "Newspapers and Political Awareness", American Journal of Political Science 39, 513-527.

Nelson, Phillip (1974) "Advertising as Information", Journal of Political Economy 82, 729-754.

Niemi, Richard G., Lynda W. Powell, and Patricia L. Bicknell (1986), "The Effects of Congruity between Community and District on Salience of U.S. House Candidates", Legislative Studies Quarterly 11, 187-201.

Poole, Keith T., and Howard Rosenthal (1997) "Congress: A political-economic history of roll call voting.", New York: Oxford University Press, 1997

Persson, Torsten and Guido Tabellini (2000) "Political Economics: Explaining Economic Policy", MIT Press.

Rohde, David W., and Kenneth A. Shepsle (1973) "Democratic Committee Assignments in the House of Representatives: Strategic Aspects of a Social Choice Process", The American Political Science Review 67(3), 889-905.

Robinson, John P., and Mark Dennis K. Davis (1990) "Television News and the Informed Public: An Information-Processing Approach."Journal of Communication 40, 106-119.

Robinson, John P., and Mark K. Levy (1986) "The Main Source: Learning from Television News", Beverly Hills, CA, Sage.

Stewart, Charles, III and Mark Reynolds. 1990. "Television Markets and U.S. Senate Elections."Legislative Studies Quarterly 15:495-523.

Strömberg, David (1999) "The Political Economy of Public Spending", Ph. D. Dissertation, Princeton University.

Strömberg, David (2004) "Radio's Impact on Public Spending", Quarterly Journal of Economics 119(1), 189-221.

Thorson, Greg, Lynnae Glieden, and Jason Lina (1999) "Swinging Left: An Examination of the Change in Committee Compositions in the U.S. House during the Reform Period", Unpublished manuscript, University of Minnesota at Morris.

Verba, Sidney, and Norman H. Nie (1972) "Participation in America", Chicago, University of Chicago Press.

Vinson, C. Danielle (2003) "Local Media Coverage of Congress and Its Members: Through Local Eyes", Cresskill NJ, Hampton Press.

Weaver, David, and Dan Drew (1993) "Voter Learning in the 1990 Off-Year Election: Did the Media Matter?"Journalism Quarterly 70, 356-368.

Welch, Susan (1978) "Election to the Legislature: Competition and Turnout", In John C. Comber and James B. Johnson, eds., Nonpartisanship in the Legislative Process: Essays on the Nebraska Legislature, Washington, DC, University Press of America.

Wolfinger, Raymond E., and Steven J. Rosenstone (1980) "Who Votes?", New Haven, CT, Yale University Press.

Appendix

Computation of Congruence

Let q_{md} denote the number of articles that newspaper m writes about House Representative from district d. Assume that

$$q_{md} = \alpha ReaderShare_{md}.$$

Next, assume that the share who reads any article is proportional to the share of households who buy the paper (that we call $ReaderPenetration$). This assumption is standard in the advertising literature. Let p_{mc} be the probability that a random person in county c reads any particular article in newspaper m. We assume that

$$p_{mc} = hReaderPenetration_{mc}.$$

Typically, the share of buyers who read an article, h, will depend on the article's size, location and attractiveness as well as features of the newspaper. Since we do not have data on these features, we simply use the average probability h.

Now, if newspaper m writes q_{md} articles about a congressman, and each time the probability that a person reads the article is p_{mc}, then the expected number of articles that this person will read about the congressman in this newspaper is $p_{mc}q_{md}$. Suppose that M newspapers sell in county c and each newspaper m prints q_{md} stories about the congressman. Then the total expected number of articles that a random person in county c will read about his or her congressman will be[24]

$$\sum_{m=1}^{M} p_{mc}q_{md} = \alpha h \sum_{m=1}^{M} ReaderPenetration_{mc}ReaderShare_{md} = \alpha h Exposure_{cd}.$$

To address a potential endogeneity problem, we need to rewrite the equations before estimation. The $Exposure$ variables are mixtures of two concepts: (i) "congruence," i.e., the extent to which a district's boundaries match the boundaries of newspaper markets; and (ii) readership, i.e., overall newspaper penetration in a district or county. It is more than likely that newspaper penetration is correlated with unobserved factors which are correlated voter behavior, since people who buy newspapers are richer, better-educated, more interested in politics, etc.

[24] Not only the expected number, but also the distribution of articles read could matter for average learning. The MEDIAC model of advertisement impact discusses how to account for this. The whole distribution matters if learning is not linear in the number of articles read. Marginal effects may increase initially since the first articles generate interest, so called "wearin". At higher levels, the effect on learning is decreasing in the number of articles read (Pechmann and Stewart, 1989).

To avoid this problem, we define

$$Congruence_{cd} = \frac{Exposure_{cd}}{\sum_m ReaderPenetration_{mc}},$$

or equivalently,

$$Congruence_{cd} = \sum_{m=1}^{M} MarketShare_{mc} ReaderShare_{md}. \tag{3}$$

$Congruence_{cd}$ is the average $ReaderShare_{md}$, weighted by market-share. It is a proxy for how actively the average newspaper read in county c covers congressional district d, since each newspaper's incentive to cover district d is increasing in $ReaderShare_{md}$. Intuitively, $Congruence_{cd}$ is the value $Exposure_{cd}$ would take if everyone in county c reads exactly one newspaper – this would be the average share of newspapers' markets which lie in district d. *Congruence* varies between zero and one. Note that since *Congruence* is defined using market shares, it is not dependent on the total newspaper penetration in the county.

Empirically, we compute *Congruence* as follows. We use the newspaper sales data to estimate the number of people in each district that read each newspaper. (Although we estimate this separately for each year in our sample, we suppress time subscripts in what follows.) Let n_{cd} be the number of people who live in county c and congressional district d, let $n_c = \sum_d n_{cd}$ be the total number of people who live in county c, and let $n_d = \sum_c n_{cd}$ be the total number of people who live in district d. Let x_{mc} be the number of copies of newspaper m sold in county c, and let $x_m = \sum_c x_{mc}$ be the total number of copies sold of newspaper m. Assume that the number of copies per capita of newspaper m is the same in the part of county c that live in district d as for the county as a whole (we do not have data on newspaper sales for sub-county units). Then the number of copies of newspaper m sold in the part of county c that lies in district d is approximately

$$x_{mcd} = \frac{x_{mc}}{n_c} n_{cd}. \tag{4}$$

Also, the number of copies of newspaper m sold in district d is approximately $x_{md} = \sum_c x_{mcd}$. These variables can then be used to compute $MarketShare_{md}$, $ReaderShare_{md}$ and $Congruence_{cd}$ using the definitions

$$MarketShare_{mc} = \frac{x_{mc}}{\sum_{m\prime} x_{m\prime c}},$$

and

$$ReaderShare_{md} = \frac{x_{md}}{x_m}.$$

Combining these, we simply use equation (2) to compute $Congruence_{cd}$.

Table 1. Newspaper data summary statistics

Variable	Obs	Mean	Std. Dev	.	Min	Max
CoverShare$_{md}$	4213	0.15	0.15		0.00	1.00
ReaderShare$_{md}$	4213	0.15	0.20		0.00	0.99
Scandal	4213	0.01	0.14		0.00	2.00
Party leader	4213	0.02	0.13		0.00	1.00
Higher office	4213	0.03	0.17		0.00	1.00
Close race	4213	0.26	0.16		0.00	0.50
Freshman	4213	0.14	0.35		0.00	1.00
Retired	4213	0.09	0.28		0.00	1.00
Out of state	4213	0.13	0.34		0.00	1.00
Pct. urban	4213	0.68	0.30		0.00	1.00

Table 2: Newspaper Coverage of U.S. House Members, 1991-2000

Dependent Variable: *CoverShare$_{md}$*	I	II
ReaderShare$_{md}$	0.64	0.62
	(0.02)***	(0.02)***
Party leader		0.10
		(0.02)***
Scandal		0.04
		(0.01)***
Higher_office (ran or appointed)		0.08
		(0.01)***
Out_of_state		-0.03
		(0.01)***
Pct. urban		-0.09
		(0.01)***
Close_race		0.04
		(0.01)***
Freshman		0.00
		(0.00)
Retired		0.02
		(0.01)***
Additional controls	no	yes
N	4213	4213
R2	0.69	0.74

Results from OLS regressions. The unit of observation is a newspaper by year. All regressions include year-fixed effects. Robust standard errors, clustered by newspaper, in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 3. NES data summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Congruence	15014	0.46	0.29	0.01	0.97
NameRecall	14139	0.31	0.46	0	1
NameRecallIncumbent	11734	0.29	0.45	0	1
NameRecallChallenger	11734	0.11	0.32	0	1
NameRecognition	9624	0.92	0.28	0	1
FeelingThermometerProvided	12459	0.82	0.38	0	1
IdeologicalRatingProvided	7441	0.70	0.46	0	1
LikesOrDislikesProvided	10775	0.53	0.50	0	1
NameRecallSenator	5337	0.46	0.50	0	1
FeelingThermometerSenator	6441	0.91	0.29	0	1
KnowsHouseMajority	14153	0.56	0.50	0	1
KnowsSenateMajority	14146	0.50	0.50	0	1
ReadAboutIncumbent	8985	0.50	0.50	0	1
ReadAboutChallenger	8957	0.17	0.38	0	1
Hear about incumbent on TV	8985	0.50	0.50	0	1
Hear about incumbent on radio	8985	0.26	0.44	0	1

Table 4. Voter knowledge of House Representative

	Baseline		*Within-race*		*Redistricting*	
	I	II	III	IV	V	VI
Controls	no	yes	no	yes	no	Yes
Fixed effects	year	state*year	district*year	district*year	year, county	state*year, county
Dependent variable: *NameRecall*						
Congruence	0.28	0.35	0.33	0.42	0.28	0.27
	(0.07)***	(0.07)***	(0.08)***	(0.07)***	(0.07)***	(0.06)***
Observations	14139	14139	14139	14139	14139	14139
R-squared	0.16	0.27	0.21	0.30	0.10	0.24
Dependent variable: *NameRecognition*						
Congruence	0.04	0.08	0.04	0.10	0.15	0.07
	(0.05)	(0.05)	(0.06)	(0.06)*	(0.06)**	(0.06)
Observations	9624	9624	9624	9624	9624	9624
R-squared	0.27	0.39	0.41	0.42	0.08	0.31
Dependent variable: *FeelingThermometerProvided*						
Congruence	0.20	0.19	0.19	0.19	0.35	0.29
	(0.05)***	(0.06)***	(0.07)***	(0.07)***	(0.09)***	(0.09)***
Observations	12459	12459	12459	12459	12459	12459
R-squared	0.18	0.25	0.23	0.28	0.10	0.19
Dependent variable: *IdeologicalRatingProvided*						
Congruence	0.23	0.25	0.25	0.30	0.32	0.20
	(0.09)***	(0.09)***	(0.10)**	(0.10)***	(0.09)***	(0.09)**
Observations	7441	7441	7441	7441	7441	7441
R-squared	0.18	0.25	0.21	0.27	0.08	0.19
Dependent variable: *LikesOrDislikesProvided*						
Congruence	0.28	0.26	0.24	0.21	0.30	0.30
	(0.08)***	(0.09)***	(0.10)**	(0.09)**	(0.07)***	(0.07)***
Observations	10775	10775	10775	10775	10775	10775
R-squared	0.17	0.29	0.22	0.32	0.1	0.24

Results from OLS regressions. Robust standard errors, clustered by county, in parentheses: * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5. Auxiliary regressions. Dependent variable: *Congruence*

	I *Baseline*	II *Within-race*	III *Redistricting*
Explanatory variables:			
Individual controls	2(17), 0.13	1(17), 0.22	0(17), 0.24
County controls	5(8), 0.01	4(8), 0.08	0(8), 0.34
Urbanism controls	1(4), 0.08	2(4), 0.02	1(4), 0.08
Race controls	0(3), 0.60		1(3), 0.04
Rep. controls	0(8), 0.18		0(8), 0.57
Fixed effects	State*year, incumbent	District*year	State*year, county
Observations	14139	14139	14139

Results from OLS regressions. Each entry of format x(y), z, where x is the number of variables significant at the 5% level, y is the total number of variables in that category, and z is the p-value of a test of joint significant of all y variables. Robust standard errors, clustered by county, were used.

Table 6. Placebo: General political knowledge

Dependent variable	*NameRecall Senator*	*Feel.Therm. Senator*	*KnowsHouse Majority*	*KnowsSenate Majority*
Congruence	0.04	-0.02	0.00	0.02
	(0.14)	(0.09)	(0.05)	(0.06)
Observations	5337	6441	14153	14146
R-squared	0.30	0.22	0.31	0.27

Results from OLS regressions. All regressions include state*year- and incumbent-fixed effects, and controls ("baseline specification"). Robust standard errors, clustered by county, in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7. TV and radio

	Read/hear about Rep. from			
Dependent variable	Newspaper	Radio	TV	*Name Recall*
Congruence	0.45	0.10	0.19	0.33
	(0.10)***	(0.07)	(0.11)*	(0.09)***
TVCongruence	-0.29	0.62	0.68	0.17
	(0.22)	(0.19)***	(0.22)***	(0.19)
Observations	8985	8985	8985	14139
R-squared	0.22	0.16	0.24	0.27

Results from OLS regressions. All regressions include state*year- and incumbent-fixed effects, and controls ("baseline specification"). Robust standard errors, clustered by county, in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 8. Representatives in Congress summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Congruence	4768	0.45	0.24	0.00	0.99
NominateScores	3317	0.01	0.36	-0.72	0.89
DistribComm	4508	0.27	0.44	0	1
PolicyComm	4768	0.16	0.36	0	1
Witness	4553	3.79	3.72	0	28
WitnessBudget	4553	1.67	2.20	0	21

Table 9. Dependent variable: Number of witness appearances before congressional hearings

	I	II	III	IV	V	VI	VII
Congruence	0.45	0.47	0.46	0.45	0.43	0.45	0.43
	(0.16)***	(0.15)***	(0.15)***	(0.21)**	(0.15)***	(0.22)**	(0.21)**
District controls	x	x	x	x	x	x	x
Race and Representative controls			x	x	x	x	x
Fixed-effects	state,year	state,year	state,year	district,year	state*year	state,year	state,year
Estimation procedure	Poisson	NB	NB	NB	NB	Poisson	NB
Appearance before committee	All	All	All	All	All	Appr. W&M	Appr. W&M
Observations	4553	4553	4553	4553	4553	4553	4553

Estimation procedure: NB = negative binomial regression; Poisson=Poisson regression. The unit of observation is House Representative by congressional session. Appearance before committee: Appr.=Appropriations; W&M=Ways and Means. Standard errors, clustered by congressional district in parenthesis: * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 10. Committee assignments

Dependent variable	Distributive committee assignment				Policy committee assignment			
	I	II	III	IV	V	VI	VII	VIII
Congruence	0.41	0.11	0.15	0.03	-0.18	-0.10	-0.38	-0.24
	(0.07)***	(0.09)	(0.09)*	(0.08)	(0.06)***	(0.08)	(0.09)***	(0.08)***
Controls	no	yes	no	yes	no	yes	no	yes
Fixed effects	state*year	state*year	year, district	year, district	state*year	state*year	year, district	year, district
Observations	4508	4508	4508	4508	4768	4768	4768	4768
R-squared	0.17	0.36	0.45	0.56	0.12	0.22	0.48	0.53

Results from OLS regressions. The unit of observation is House Representative by congressional session. Standard errors clustered by House representative: * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 11. Dependent variable: NOMINATE scores first dimension

	I	II	III	IV	V
*(Democrats win)*Congruence*	0.19	0.17	0.18	0.17	0.14
	(0.04)***	(0.04)***	(0.05)***	(0.04)***	(0.05)***
Democrats win	-0.68	-0.67	-0.79	-0.67	-0.80
	(0.03)***	(0.03)***	(0.08)***	(0.03)***	(0.08)***
Congruence	-0.04	-0.10	-0.09	-0.05	-0.02
	(0.03)	(0.04)***	(0.04)**	(0.03)	(0.03)
Controls	no	main	main, urb. interacted	no	main, urb. interacted
Fixed effects	no	no	no	year, district	year, district
Observations	3317	3317	3317	3317	3317
R-squared	0.88	0.90	0.90	0.96	0.96

Results from OLS regressions. The unit of observation is House Representative by congressional session. All specifications control for 5th order polynomials in Democratic vote share in the House and presidential elections. The “main” controls denote the direct effect of all our control variables. The “urb. interacted” controls denote interaction between the Democrats win dummy variable and share urban, density, and the 10 dummy variables for urban and density. Standard errors, clustered by congressional district in parenthesis: * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 12: Distribution of federal funds across counties, 1984-2004
Dependent variable, log spending per capita

	Baseline		Within-race		Within neighbor-counties		Redistricting	
	I	II	III	IV	V	VI	VII	VIII
Congruence	0.092	0.137	0.064	0.094	0.106	0.096	0.051	0.035
	(0.030)***	(0.027)***	(0.030)**	(0.030)***	(0.039)***	(0.038)**	(0.021)**	(0.020)*
Controls	no	yes	no	yes	no	yes	no	yes
Fixed effects	state*year	state*year	district* year	district* year	year, neighbor	year, neighbor	state*year, county	state*year , county
Observations	33085	33085	28787	28787	16698	16698	33085	33085
R-squared	0.259	0.393	0.441	0.516	0.638	0.677	0.882	0.890

Results from OLS regressions. The unit of observation is county by congressional session. Standard errors, clustered by county, in parentheses: * significant at 10%; ** significant at 5%; *** significant at 1%

Table 13 Voting data summary statistics.

Variable	Obs	Mean	Std. Dev	Min	Max
Democratic vote share	17221	48	19	0	100
Incumbent vote share	17221	66	12	0	100
Congruence	17221	0.49	0.30	0	1
House election turnout	7106	40.5	9.3	3.0	99.9
Presidential election turnout	7106	41.6	8.8	4.1	99.5
House election turnout (top of ticket)	999	31	9	5	71

Table 14 Dependent variable:
Percent of population voting in House election, 1984, 1988, 2000

	Baseline		Within-race		Redistricting	
	I	II	III	IV	V	VI
Congruence	0.83	0.73	0.83	0.81	0.44	0.40
	(0.14)***	(0.14)***	(0.14)***	(0.14)***	(0.61)	(0.60)
Presidential elec. turnout	0.92	0.89	0.94	0.92	0.84	0.83
	(0.01)***	(0.01)***	(0.01)***	(0.01)***	(0.03)***	(0.03)***
Controls	no	yes	no	yes	no	yes
Fixed effects	state*year	state*year	district*year	district*year	state*year, county	state*year, county
Observations	7106	7106	7106	7106	7106	7106
R-squared	0.92	0.92	0.94	0.94	0.95	0.95

Results from OLS regressions. The unit of observation is county by district by election. Standard errors, clustered by county in parentheses: * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 15 Incumbent and challenger effects.

	Read about		Name recall	
	Incumbent	Challenger	Incumbent	Challenger
Congruence	0.38	0.05	0.33	0.12
	(0.12)***	(0.10)	(0.09)***	(0.05)**
Observations	5945	5930	10424	10424
R-squared	0.24	0.29	0.29	0.30

Results from OLS regressions. Sub-sample of contested elections. All regressions include state*year- and incumbent-fixed effects, and controls ("baseline specification"). Robust standard errors, clustered by county, in parentheses: * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 16 Incumbency advantage:
Dependent variable: Democratic percent of two-party vote

	Levitt and Wolfram		Baseline		Within-race		Redistricting	
			V	VI	VII	VIII	IX	X
Incumbent	9.16	9.71	9.33	8.27			10.82	12.77
	(0.23)***	(0.26)***	(1.25)***	(1.38)***			(1.93)***	(1.78)***
*Incumbent*Congruence*	0.68	0.68	0.74	1.13	0.64	0.77	1.43	1.21
	(0.27)**	(0.26)***	(0.29)**	(0.30)***	(0.26)**	(0.26)***	(0.46)***	(0.44)***
Congruence	0.86	0.82	1.50	1.74	0.60	0.63	2.57	2.55
	(0.28)***	(0.27)***	(0.30)***	(0.31)***	(0.27)**	(0.26)**	(0.74)***	(0.73)***
Presidential vote share	0.74	0.74	0.67	0.66	0.76	0.75	0.6	0.59
	(0.01)***	(0.01)***	(0.01)***	(0.01)***	(0.01)***	(0.01)***	(0.02)***	(0.02)***
Democratic Incumbent			1.22	2.58			-2.57	-5.5
			(2.48)	(2.75)			(3.84)	(3.54)
Lagged Dem vote			20.72	23.39			19.51	17.47
			(1.13)***	(1.11)***			(1.42)***	(1.41)***
Controls	no	yes	no	yes	no	yes	no	yes
Fixed effects	state*year, district*plan	state*year, district*plan	state*year	state*year	district*year	district*year	state*year, county	state*year, county
Observations	17221	17221	14147	14147	14147	14147	14147	14147
R-squared	0.89	0.9	0.87	0.86	0.94	0.95	0.91	0.92

Results from OLS regressions. The unit of observation is county by district by election. Standard errors, clustered by county in parentheses: * significant at 10%; ** significant at 5%; *** significant at 1%.

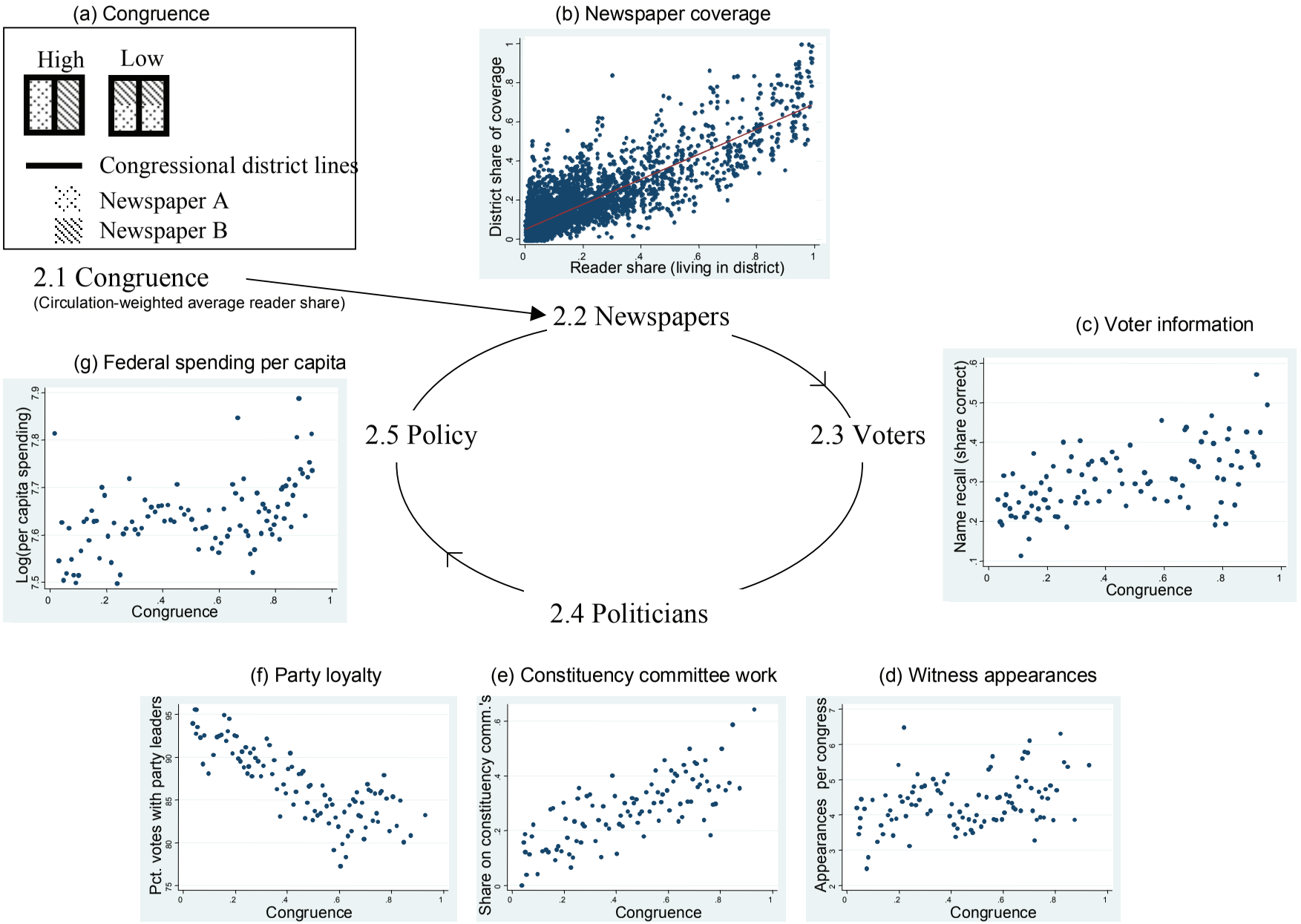

Figure 1: Structure of empirical investigation

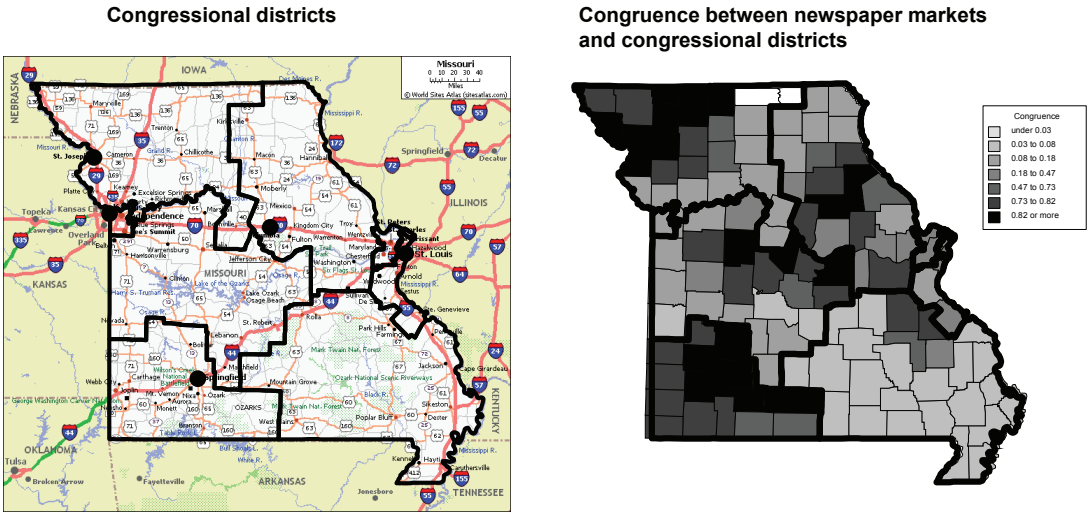

Figure 2: *Congruence* in Missouri

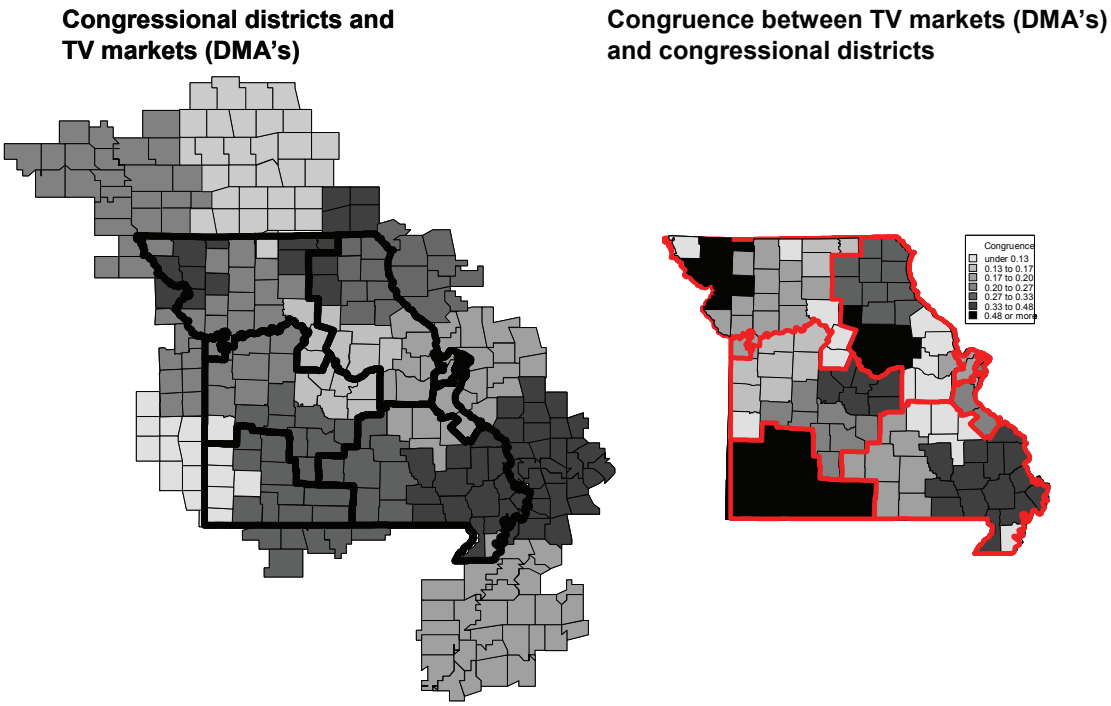

Figure 3: *TVCongruence* in Missouri

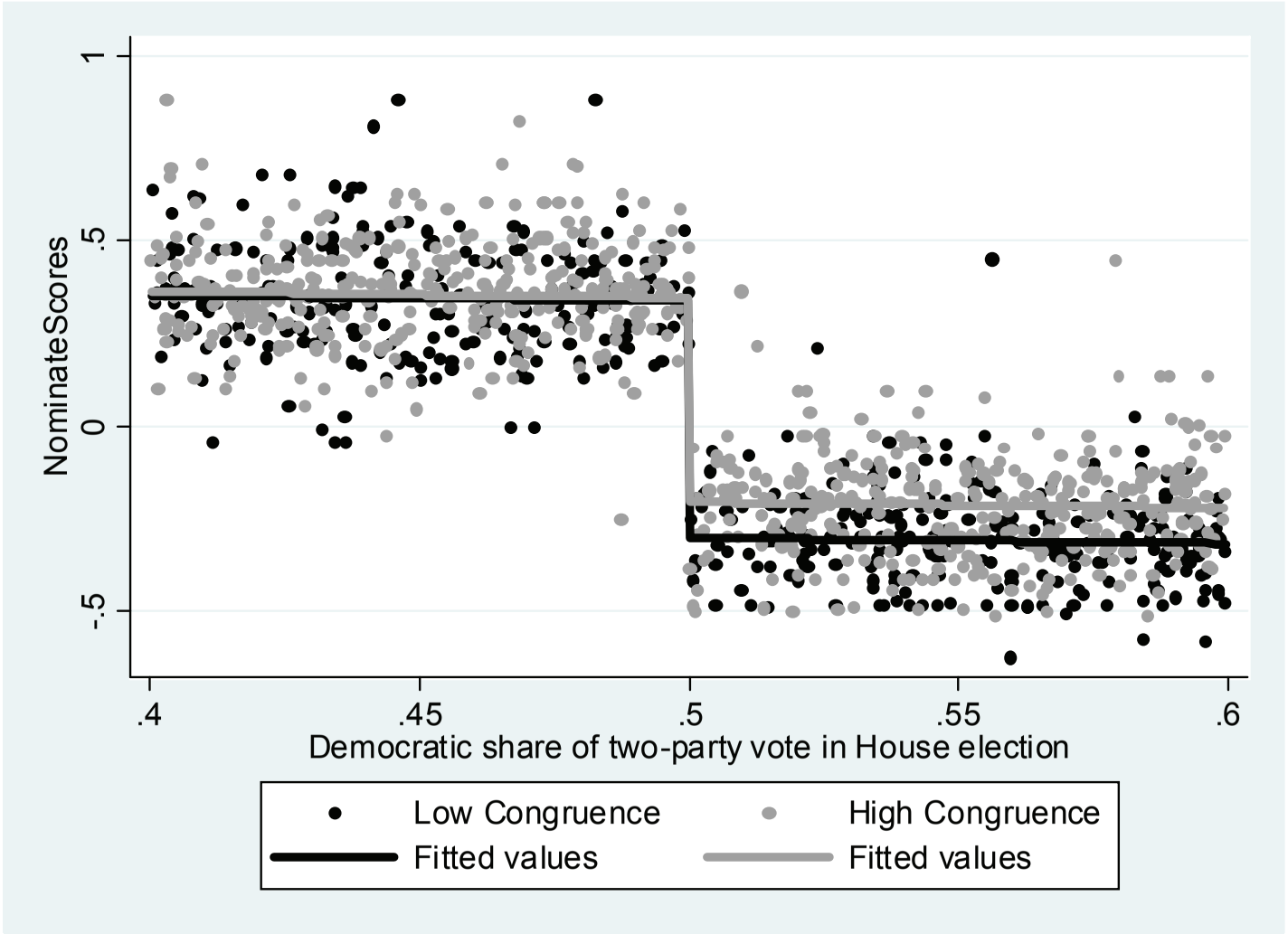

Figure 4. Nominate Scores by Congruence