

# Using Geographic Information Systems to Measure District Change, 2000–2002

Michael H. Crespin

*Department of Political Science, Michigan State University,  
303 S. Kedzie Hall, East Lansing, MI 48824  
e-mail: crespim@msu.edu*

In this article, I use geographic information systems to develop a continuous measure of district continuity and change following the 2000–02 congressional redistricting cycle. The new measure provides details of where the new population in a district came from and how the old population was distributed within new districts. This measure is then used to demonstrate the independent and interactive influence of district change on competition for congressional elections.

## 1 Introduction

How can we improve on our current understanding of the redistricting process? One way is by using the same tools that states use to draw district boundaries to create better measures of district change. In this brief research note, I outline how to use geographic information systems (GIS) to create a continuous measure of district continuity and change.<sup>1</sup> The primary goal of this analysis is to track the changes in the populations of congressional districts in the pre- (2000) and post-redistricting (2002) years using basic GIS techniques. The measure allows us to precisely state the fraction of constituents in district  $d$  in election year  $y$  that has candidate  $c$  as the incumbent.

## 2 Previous Measures of Redistricting

As mandated by the U.S. Supreme Court following *Baker v. Carr* (1962), *Wesberry v. Sanders* (1964), and *Reynolds v. Simms* (1964), states must redraw district boundaries every 10 years to ensure equality of representation. Soon after the initial rounds of redistricting, political scientists have tried to determine the influence of district changes on competition (Erikson 1972; Tufté 1973), patterns of turnover (Bullock 1975), and more recently, challenger entry and incumbent exit decisions (Cox and Katz 2002; Hetherington et al. 2003).

---

*Author's note:* This article was previously presented at the 2004 Summer Political Methodology Meeting, Palo Alto, California. I would like to thank David Rohde and the Political Institutions and Public Choice program for supporting this research, Gary Jacobson for sharing his congressional elections data, Jamie Carson, Ko Maeda, Nathan Monroe, Ashton Shortridge, Carl Snook, and Pariwate Varnakovida for comments and assistance, and the three anonymous reviewers. District change data are available on the *Political Analysis* Web site.

<sup>1</sup>In particular, I use Arc/Info 9.0 workstation for spatial analysis and ArcMap 8.3 to display the resulting shape files.

Previous research has measured redistricting in different ways. The recent accounts of redistricting by Cox and Katz (2002),<sup>2</sup> Hetherington et al. (2003), and others measure redistricting with a simple dichotomous variable when the boundaries are changed. In contrast, some are concerned with “major” redistricting as defined by some professional source such as *Congressional Quarterly* (see, e.g., Jacobson and Dimock 1994). Finally, Carson et al. (2004) go one step further and rely on county-level maps to classify districts as “continuous” if 50% or more of the population is shared between the old and the redrawn district and “new” otherwise.

While each of these measures are valid and have provided us with many useful insights into the effects of redistricting, it is now possible to move beyond simple dichotomous measures of redistricting and adapt the same techniques used by the states to draw district boundaries. This is important since, as scholars have pointed out, there are advantages in paying attention to measurement when trying to capture abstract concepts (Zeller and Carmines 1980; Jacoby 1991). In this case, the concept is changing district boundaries, but the measures to date have not indicated to what extent the districts have actually changed. It is now feasible to use geographic information systems (GIS) programs such as ARC/Info in combination with boundary shape files and census data to produce continuous measures of district continuity and change.<sup>3</sup>

### 3 Data and Methods

Creating a continuous measure of district change requires only simple spatial techniques and readily available census data.<sup>4</sup> The data include district and census tract boundary files that are then combined with population tables for the tracts within each state.<sup>5</sup> Along with the coordinates in each file is a polygon attribute table that provides data for each shape (i.e., district or tract), including the area and perimeter of each polygon as well as an identification code that matches up polygons with congressional district and census tracts.

The steps to create the measure are demonstrated in Fig. 1, using Iowa as an example.<sup>6</sup> The first step in the process is to spatially intersect the district boundary files for the 107th (labeled A in Fig. 1) and 108th (B) Congresses. This new shape file is then overlaid on top of the census tract file (C).<sup>7</sup>

The result (D) includes three types of census tracts.<sup>8</sup> The first types of tracts deemed “continuous” appear in both the new and the old districts. I define “new” tracts as those

<sup>2</sup>Cox and Katz (2002, pp. 15–18) discuss a similar continuous measure but only use it sparingly.

<sup>3</sup>ARC/Info is one of several programs that can be used to manipulate, explore, and display data that are spatial in nature.

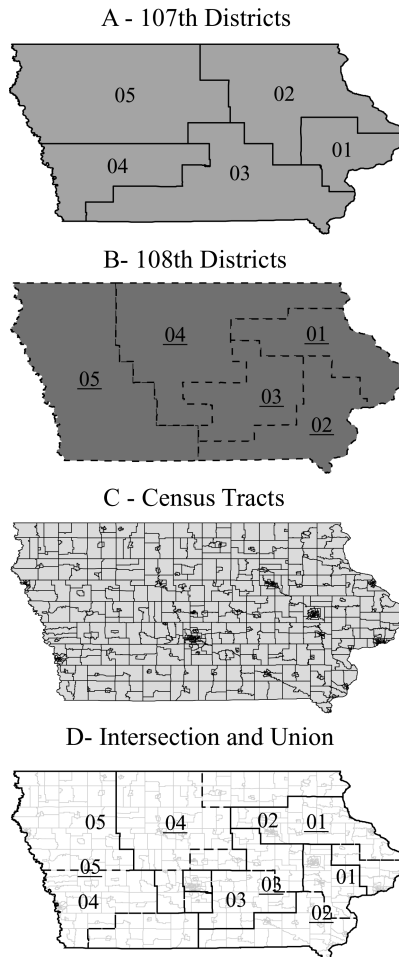
<sup>4</sup>Creating a continuous measure is so simple, in fact, that the only ARC/Info commands used to create the measure are intersect and union. See Bolstad (2002) for an introduction to GIS and spatial analysis.

<sup>5</sup>The 2000 census data used in this paper are available at [www.census.gov](http://www.census.gov). A cartographic boundary file is a series of coordinates that ARC/Info can read to produce a visual depiction of the shape.

<sup>6</sup>This state provides an interesting case study because of the degree of district change produced by the nonpartisan redistricting commission.

<sup>7</sup>According to the census, tracts are small, relatively permanent statistical subdivisions of a county with populations ranging from 1500 to 8000. The tracts grow smaller in dense urban areas, making it possible to measure change in all districts.

<sup>8</sup>Since some tracts do cross district boundaries, some post-processing of the data is necessary. Further, spatial analyses such as intersect and overlay frequently create small spurious polygons called “slivers,” which are an artifact of imprecise overlay (Bolstad 2002, p. 398). The results in the attribute tables are the repeated listing of tracts with the same tract ID (and population), but with different areas. To erase these slivers and place tracts in the appropriate district if there was an overlap, I dropped all of the repeated tracts except for the one with the largest area. This technique produced a final attribute table that had only one unique entry for each of the original census tracts.



**Fig. 1** Example of spatial analysis—Iowa.<sup>10</sup>

that have been drawn into the 108th congressional district and “old” tracts that have been drawn out of the 107th districts after redistricting.<sup>9</sup> The new attribute table, which is automatically updated by Arc/Info, makes it relatively easy to identify continuous, new, and old tracts since each tract has an identifier that labels where it resided in both the 107th and 108th congressional districts. To produce a measure of change (continuity), simply divide the population for the new (continuous) tracts by the total population of the district. The final results are actually two measures of change or continuity. The first tells us what happened to the population of the 107th districts while the second reports where the population came from for the new 108th districts.

<sup>9</sup>This means that a tract can be simultaneously new for one district and old for another. Terms like new, old, and continuous are really a mental shortcut for ways to think about tracts. An alternative way to think about tracts would be to identify each tract with a double  $(i, j)$  with  $i$  identifying the district number in the old scheme and  $j$  identifying the district number in the new scheme. In fact, this is essentially how they are treated in the data set.

<sup>10</sup>This map of Iowa may appear to be “squished” to some readers. However, it is only due to the choice of projection system from the round earth to the flat paper. GCS North American 1983 is the default system applied here.

**Table 1A** Displacement of Iowa's 107th congressional districts into the 108th

		<i>108th Congressional Districts</i>					
		<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>Total</i>
107th Congressional Districts	1	37.93	62.07				100
	2	62.63		12.56	24.81		100
	3		36.68	24.27	27.89	11.16	100
	4			60.28	8.81	30.90	100
	5				41.07	58.93	100

#### 4 Uses of New Measure

In order to better understand the different measures of district change, refer to Tables 1A, 1B, 2A, and 2B, which display the results of the spatial analysis for Iowa and New Hampshire. Each of the A tables reads from left to right across the rows and shows the first measure of district change, the distribution of population from the 107th congressional districts to the new districts in the 108th Congress.<sup>11</sup> The B tables read from the top down in the columns and display where the population came from for each of the new districts. For example, Table 1A shows that 37.93% of the population in Iowa's first district in the 107th Congress remained in the first district for the 108th Congress. Meanwhile, the same district lost 62.07% of its population to the second district in the 108th Congress. The same population movements can be tracked for each of the other four districts.

Table 1B shows that new district one is made up of 39.13% of old district one and 60.87% of old district two. This measure also makes it relatively easy to match up new districts with their "parent" districts (see Cox and Katz 2002, p. 107, n. 1).<sup>12</sup> Clearly, new district one's parent district is not old district one but district two, since old two and new one share the largest percentage of the population.

In some states such as New Hampshire, redistricting produces only minor changes, as Table 2A and 2B demonstrate. In fact, the 100% in the row for district two in Table 2A means that all of the population in that district remained there during redistricting. Meanwhile, 98.73% remained in district one while it gave 1.27% away to district two. This

**Table 1B** Continuity of Iowa's 108th congressional districts from the 107th

		<i>108th Congressional Districts</i>					
		<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	
107th Congressional Districts	1	39.13	64.04				
	2	60.87		12.21	24.11		
	3		35.96	23.79	27.33	10.94	
	4			64.00	9.36	32.81	
	5				39.20	56.25	
Total		100	100	100	100	100	

<sup>11</sup>While it may rarely be the case that we would want to know where voters have gone, this technique can be used to examine where contributors reside in two different election cycles (see, e.g., Crespin 2005).

<sup>12</sup>Some districts will not have parents as a result of reapportionment. Depending on the research question, it is also possible for a district to have one or several "children" if districts are added to a state.

**Table 2A** Displacement of New Hampshire’s 107th congressional districts into the 108th

		<i>108th Congressional Districts</i>		
		<i>1</i>	<i>2</i>	<i>Total</i>
107th	1	98.73	1.27	100
Congressional Districts	2		100	100

demonstrates that district two lost population relative to district one between the 1990 and 2000 censuses.

As a final descriptive example, the measure can also be used to determine which districts are added to a state as a result of reapportionment. Table 3, read from top to bottom, describes where the population came from to make up the 108th congressional districts for Arizona, a state that gained two seats prior to the 2002 elections. It is possible to identify districts one and four as “new” since they share the smallest population with parent districts from the 107th Congress. Although these examples only make use of basic population counts, they hopefully show the worth of the new measure.<sup>13</sup>

**5 Does Change Matter?**

The above examples are interesting but they do not mean much if change does not matter. To show the impact of change on election outcomes, I add the new measure of district changes to a basic model of congressional elections (Jacobson 2004). The dependent variable is the change in the margin of victory measured as the difference between the incumbents vote in the 2002 congressional elections to the 108th Congress and the 2000 election, for only contested races in both years.<sup>14</sup> The main independent variable of interest is the percent new (*% new 108th*) for each district before the 2002 election.<sup>15</sup> To be clear, this variable is measured using the same incumbent’s district for both elections. Although this measure does not take into account the types of voters who are being shuffled in and out of districts, we do know one thing about them for certain: *they have not*

**Table 2B** Continuity of New Hampshire’s 108th congressional districts from the 107th

		<i>108th Congressional Districts</i>	
		<i>1</i>	<i>2</i>
107th	1	100	1.29
Congressional Districts	2		98.71
	Total	100	100

<sup>13</sup>The district change data are available on the author’s Web site, [www.msu.edu/~crespinm](http://www.msu.edu/~crespinm).

<sup>14</sup>To be clear, this means that for a district to be included in the regression there must be an incumbent facing a major party challenger in *both* the 2000 and 2002 elections. Conversely, open seat races, or races without a major party challenger in either year, were excluded.

<sup>15</sup>For example, referring back to Table 1B, the value of the variable would be  $100 - 60.87 = 39.13$  for district 1,  $100 - 64.04 = 35.96$  for district 2, and  $100 - 23.79 = 76.21$  for district 3 since Leonard Boswell decided to remain in district three, even though the bulk of his old district is in new district 4.

**Table 3** Finding “new” districts in Arizona

		<i>108th Congressional Districts</i>							
		<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>
107th Congressional Districts	1			5.77	17.89	72.65	33.72		
	2		0.77		48.25			72.24	2.40
	3	31.41	90.51	8.63	13.80			9.57	
	4		8.57	74.54	20.06	8.72			
	5	8.14						15.03	97.60
	6	60.45	0.14	11.07		18.62	66.28	3.16	
	Total	100	100	100	100	100	100	100	100

*been represented by the incumbent.*<sup>16</sup> If part of the incumbency advantage lies in the fact that all an incumbent has to do to win reelection is get all of the voters who voted for her the last time to do it again, then some of the advantage may be taken away after redistricting, some of those voters no longer reside in the district. In other words, the incumbency advantage is not present for incumbents if the voters are new. For this reason, it is expected that the greater the percent new for a district, the greater the decline in the incumbents' vote share, all else equal. This is similar to the “personal vote” measure developed by Ansolabehere et al. (2000).

It is also necessary to control for changes in the underlying political makeup of the district. This is measured by the change in the district presidential vote for the incumbent party's presidential vote in the district ( $\Delta$  *Presidential Vote*). This variable is operationalized, for Democratic incumbents, by subtracting Al Gore's vote in the new 108th congressional districts from his vote share in the old 107th districts as reported by *Politics in America*. For Republican incumbents, the measure is based on the change in George W. Bush's vote share. Including this control variable allows us to test for the independent influence of district change while controlling for the political nature of the change. As a district becomes more Democratic (Republican), Democratic (Republican) incumbents should see an increase in their margin of victory. I also include an interactive variable between the percent new in the district and the change in district partisanship. This variable shows not only how much the district changed in terms of population, but also the ideological direction of the change.

The other control variables are standard measures taken from the congressional election's literature (see Jacobson 2004 for an example). First, I include a dummy variable for *quality challenger* coded one in the presence of a quality challenger (where, as Jacobson 1980 suggests, previous electoral experience is used as a proxy for “quality”), zero otherwise. There are also controls for the challenger's and incumbent's spending. In each instance, the natural log is used (*ln Incumbent spending* and *ln Challenger spending*).<sup>17</sup> There is also a dichotomous variable that records whether the incumbent is up for reelection for the first time. This variable, *Freshman incumbent*, will control for the presence of the sophomore surge and *party* (coded 1 for Democrats) will control for changes in national trends. Finally, I include a dummy variable controlling for *South*,

<sup>16</sup>See Fenno 1977 for a discussion of how members represent different parts of their geographic constituency.

<sup>17</sup>I employ the convention adopted by Jacobson (1980) in assuming a minimum of \$5000 spent by each candidate.

**Table 4** Influences on change in incumbent victory margins, 2000–2002

<i>Variable</i>	<i>Coefficient (Robust Standard Error)</i>
% New 108th	−.037* (.019)
Δ Presidential vote	.060 (.170)
% New 108th × Δ presidential vote	.019* (.005)
Quality challenger	1.40 (1.04)
Ln Incumbent spending	.160 (.527)
Ln Challenger spending	−.532* (.185)
Party	−4.03* (.631)
Freshman incumbent	5.22* (1.40)
South	.503 (.840)
<i>N</i>	274
Adj. R <sup>2</sup>	.39
Root MSE	5.08
F-stat	20.10*

\* $p < .05$ .

*Note.* Dependent variable = incumbent 2002 vote minus incumbent 2000 vote.

defined as the 11 states of the Confederacy. Since the dependent variable is theoretically continuous, ordinary least squares (OLS) is used and Huber-White robust standard errors are used to correct for heteroskedasticity.

Table 4 displays the results of the regression. The negative and significant coefficient on *percent new 108th* is evidence that the more a district changes, the incumbent's vote share declines. Although the coefficient on the change in presidential vote is not significant, the interaction between *percent new* and *presidential vote* is significant and in the expected direction. If the district changes in terms of population and that change is in the direction of the incumbent's party, then the incumbent's vote share increases.<sup>18</sup>

In terms of the control variables, the more a challenger spends, the greater the decline in the incumbent's vote share (Jacobson 2004). The party variable suggests that the Republicans did better in 2002 compared to 2000. Freshman incumbents also did better in their second election, as expected.

## 6 Conclusion

This project provides evidence that not all redistricting should be treated equally and offers an example to better measure district change. As illustrated above, the results suggest that this change can have an influence on the margins of victory in congressional races. If members wish to remain in office, they will need to adjust their homestyle accordingly (Fenno 1977).

More important, this note has provided a new way to measure redistricting that can be used to answer other interesting questions. For example, the measure can be used in future research to track changes in member behavior before and after redistricting. As members pick up different types of voters, we would expect them to change their voting behavior accordingly. This measure allows us to learn where the new voters came from and, with the addition of other demographic census data, to make predictions about how members

<sup>18</sup>A joint hypotheses test between the three variables is significant at  $p < .001$ . I also examined scatter plots and added-variable plots to check for nonlinearity and did not uncover any.

will react to their new constituents. We can also move beyond voting and learn how potential campaign contributors will behave in response to their new representatives. Finally, we can also determine if different types of plans, i.e. Democratic plans or judicial plans, differ in their propensity for change. Using this new measure, these questions may now be easier to answer.

## References

- Ansolabehere, Stephen, James M. Snyder Jr., and Charles Stewart III. 2000. "Old Voters, New Voters, and the Personal Vote: Using Redistricting to Measure the Incumbency Advantage" *American Journal of Political Science* 44:17–34.
- Baker v. Carr*. 1962. 369 U.S. 186.
- Boldstad, Paul. 2002. *GIS Fundamentals*. White Bear Lake, Minnesota: Elder Press.
- Bullock, Charles S. 1975. "Redistricting and Congressional Stability, 1962–1972." *Journal of Politics* 37: 569–575.
- Carson, Jamie L., Michael H. Crespín, Charles J. Finocchiaro, and David W. Rohde. 2004. "Linking Congressional Districts across Time: Redistricting and Party Polarization in Congress." Paper presented at the Annual Meeting of the American Political Science Association, Chicago, IL.
- Congressional Quarterly. 2003. *Politics In America: 2004. The 108th Congress*. Washington, DC: Congressional Quarterly Press.
- Cox, Gary W., and Jonathan N. Katz. 2002. *Elbridge Gerry's Salamander: The Electoral Consequences of the Reapportionment Revolution*. New York: Cambridge University Press.
- Crespín, Michael H. 2005. "Out with the Old, In with the New: Redistricting and Individual Campaign Contributions to Congressional Candidates." Paper presented at the Annual Meeting of the Southern Political Science Association, New Orleans, LA.
- Erikson, Robert S. 1972. "Malapportionment, Gerrymandering, and Party Fortunes in Congressional Elections." *American Political Science Review* 66:1234–1245.
- Fenno, Richard F. Jr. 1977. "U.S. House Members in Their Constituencies: An Exploration." *American Political Science Review* 71:883–917.
- Green, Donald Philip, and Jonathan S. Krasno. 1988. "Salvation for the Spendthrift Incumbent: Reestimating the Effects of Campaign Spending in House Elections." *American Journal of Political Science* 32:884–907.
- Green, Donald Philip, and Jonathan S. Krasno. 1990. "Rebuttal to Jacobson's New Evidence for Old Arguments." *American Journal of Political Science* 34:363–372.
- Hetherington, Marc J., Bruce A. Larson, and Suzanne Globetti. 2003. "The Redistricting Cycle and Strategic Candidate Decisions in U.S. House Races." *Journal of Politics* 65:1221–1234.
- Jacobson, Gary C. 1980. *Money in Congressional Elections*. New Haven: Yale University Press.
- Jacobson, Gary C. 1990. "The Effects of Campaign Spending in House Elections: New Evidence for Old Arguments." *American Journal of Political Science* 34:334–362.
- Jacobson, Gary C. 2004. *The Politics of Congressional Elections*. New York: Addison Wesley Longman.
- Jacobson, Gary C., and Michael A. Dimock. 1994. "Checking Out: The Effects of Bank Overdrafts on the 1992 House Elections" *American Journal of Political Science* 38:601–624.
- Jacoby, William G. 1991. *Data Theory and Dimensional Analysis*. Newbury Park, CA: Sage.
- Reynolds v. Sims*. 1964. 377 U.S. 533.
- Tufte, Edward R. 1973. "The Relationship between Seats and Votes in Two-Party Systems." *American Political Science Review* 67:540–554.
- Wesberry v. Sanders*. 1964. 376 U.S. 1.
- Zeller, Richard A., and Edward G. Carmines. 1980. *Measurement in the Social Sciences: The Link between Theory and Data*. New York: Cambridge University Press.