Legislator Replacement and Party Position Change in American Single Transferable Vote Elections

Jack M. Santucci*

January 19, 2017

Abstract

What explains parties' changing coalition alignments under proportional representation in American local government? I use dynamic ideal point estimation to determine their changing locations in hypothetical coordinate space. Dynamic ideal point results are sensitive to assumptions about how much legislators may "move," so I give results over a range of 100 plausible values. As in the United States Congress, legislator replacement is a significant predictor of party position change. Replacing an entire delegation implies an average, inter-term movement across seven percent of a policy space bounded by -1 and 1.

Keywords: ranked-choice voting, single transferable vote, party position change, replacement.

*The author thanks Georgetown University's Massive Data Institute for funding data collection.

1 Overview

Elsewhere I have shown that the enactment of proportional representation (PR) involved alliances between local minority parties and defecting ruling-party factions (see "Party splits, not Progressives"). I also have shown that minority parties and party factions, not pre-election coalitions thereof, were the salient units in an analysis of legislative discipline, then that these units were more often cohesive than not (see "Do vote transfers matter for party discipline").[1]

What about parties' coalition alignments? The core of PR is post-election coalition formation. That can be true even when parties try to form coalitions before elections. Witness change in Italian *apparentementi* from 2006 to present. We will see it again in American cities, with varying degree of drama. What accounts for this change in legislative parties' positions?

According to the literature on Congress, one possibility is that legislators change their minds. Another is that different-thinking (or acting) legislators replace old ones. The balance of evidence favors replacement (Fleisher and Bond 2004). By one estimate, two-thirds of post-1973 polarization has been from replacement (Theriault 2006). Bonica (2014) finds that replacement explains U.S. Senate polarization up to 1996. Poole (2007: 435) further claims that legislators "die in their ideological boots," which leaves little room for changing minds. Since the evidence for these claims is based on strategic action in government (i.e., roll-call voting), aggregate ideological change is more plausibly the result of change in parties' activist bases. Replacement, in other words, is the result of change in party coalitions. And Karol (2009: 19) finds that high-replacement periods accompany the most stable changes in party position. Whatever conversion does occur appears to happen by force. Newcomers tend to use rules and tactics to drag any laggards along (Jacobson 2007).

This chapter tests the replacement hypothesis. The empirical implication of "party position" is a coordinate in Euclidian space. I measure this with dynamic ideal point estimation to see how much absolute change is explained by the proportion of freshmen in a delegation. My big problem is that all dynamic models effectively seed a space with new ideal points. Smoothing techniques are one way to address this problem. Although I favor a level of smoothing with precedent in the literature, I also present results over a range

1. I will refer to both minority parties and party factions as "parties" in this paper.

of plausible levels.

The paper proceeds as follows. Section 2 offers some motivating examples from the three PR cities: Cincinnati; New York City; and Worcester, Massachusetts. Section 3 addresses data. Section 4 discusses methodological choices. Section 5 presents the aggregate data. Section 6 presents results from random-effects regression models of absolute party movement. Section 7 concludes.

2 Examples of party change in American PR

Cincinnati witnessed two periods of dramatic party change. One occurred in the regular Republican party in the late 1940s. The Progressive Democrats (PD) took shape in the mid-1930s as a New Deal alternative to the existing Democratic Party. They struck accommodations with both regular Republicans, then regular Democrats in the Charter Party coalition. The passage of PD leader Herbert S. Bigelow to Congress (D-OH) and then out of local politics left the party to Wiley Craig, who ran and won as a Republican in 1941, but continued to caucus with Democrats (Straetz 1958: 59). The Republican-New Deal working arrangement faltered completely by 1947. Republican leadership purged the party, installed two freshmen, and ensured that one was elected Mayor. Regular Republicans thereafter reflected increasingly conservative positions on race, taxation, and government finance.

A second change occurred in the Democratic wing of Charter in the mid-1950s. Prior to November 1953, the Democratic delegation included two labor leaders and Mayor Ed Waldvogel, a legacy New Dealer famous for brokering highway construction initiatives under the Republican administrations of the 1940s. Conventional middle-class liberals replaced all of but one these Democrats by 1954-5. Among the liberals was eventual Governor Jack Gilligan (D-OH), famous for bringing the income tax to Ohio. Another was Theodore Berry, eventual Mayor and president of the local National Association for the Advancement of Colored People. Berry led the charge for a city income tax over opposition by the local and national Congress of Industrial Organizations. The one labor leader remaining in 1954-5 lost his seat at the next election and, with the New Deal wing of the local Democratic Party, mobilized roughly 17,000 United Steelworkers behind the PR repeal campaign of 1957. This intra-party conflict came to an even earlier head in the spring of 1954, when, on Waldvogel's death, the freshmen liberal Democrats

chose a freshman liberal Charter Republican woman for Mayor.

A survey of Worcester history also reveals tension in the Democratic wing of that city's coalition party, the Citizens' Plan E Association (CEA). Well-off Democrats joined CEA in growing numbers after 1955. When several members of the local Young Democrats (YDs) landed positions on the CEA board, newspapers reported an "invasion of amateurs." At least one regular Democrat applauded them, and rumors circulated that the Republicans would leave CEA that November.[2] YDs in 1955 also attacked PR without success, challenging a CEA rule requiring candidates to support PR in public statements.[3] They gradually dominated its nominating process. In 1957, CEA nominated two regular Democrats from the pre-PR days.

A notable break in New York City came in November 1945, when a coalition of American Labor Party-endorsed (ALP) Tammany Democrats swept Mayor Fiorello La Guardia's Republican-Fusion coalition from city government. In 1944-5, Republicans controlled the Mayoralty, Comptrollership, the City Council Presidency, two borough presidencies, and eight of 11 votes on the Board of Estimate. By January 1946, Republicans controlled only one vote on the Board of Estimate, via the borough presidency of Staten Island. It is not clear from election returns alone whether the ALP endorsement of Democrats filtered to City Council. Only one of 28 members in 1946-7 carried both parties' endorsements. Yet 44 percent of Democrats who sat in 1946-7 were freshmen, the largest wave of replacement in Tammany's history under PR. We will see in the results whether this development coincides with change in the Democrats' aggregate position.

3 Data and measures

I use the same roll-call, electoral, and party-factional affiliation data from the previous chapter. Its sources are discussed at length in the introduction and next chapter, "Exit from proportional representation." We want to measure party position change. My dependent variable is therefore absolute change

2. Currier, Charles. "Many Democrats Here Supporters of PR." *Worcester Telegram*, March 15, 1955. And: Currier, Charles. "GOP to Run Council, School Board Ticket In Fall Plan E Election." *Worcester Telegram*, May 15, 1955. Both on file in the Plan E/CEA Collection, unsorted, at the Worcester Historical Museum.

3. Looney, Marguerite. Letter to the editor, *Worcester Telegram*, March 19, 1955. See also meeting minutes of the CEA Board for May 16 and June 20, 1955. All on file in the Plan E/CEA Collection, unsorted, at the Worcester Historical Museum.

in the median ideal point from each party from term t to term $t+1$. Because I am calculating change, I necessarily drop any party that sits for only one term. This means that six out of 25 parties are dropped from the analysis, all from New York: one American Labor-Democratic cross-endorsee (1946-7), one American Labor-Fusion cross-endorsee (1942-3), two Fusionists (1938-9), one Independent Democrat (1938-9), and two Liberal Party members (1946-7).

My data include a few party switchers: three in Cincinnati (out of 48 legislator-parties), two in Worcester (out of 23 legislator-parties), and three in New York City (out of 65 legislator-parties). I treat each legislator-party as a separate observation in the estimation of ideal points, such that John Doe (Party A) in term t and John Doe (Party B) in term $t+1$ are separate legislators.

I exclude no legislators from the estimation of ideal points. I follow common practice in treating abstentions and absences as missing data. As in the next chapter, I drop all unanimous votes. The only difference between the data here and in "Exit from proportional representation" is that, here, I drop any roll call for which only one legislator took the yea or nay position. I do this because model runs including these extremely lopsided votes place opposing parties implausibly close to each other in hypothetical space. This appears to result from the use of an expectation-maximization approach (see below).

The predictor of interest is the proportion of a party's council delegation that was not present in the previous term. Call this freshman replacement.

4 Issues in estimating dynamic ideal points

I want to measure parties' changing positions from the median of their respective legislators' ideal points. This implies projecting the legislators and every item on which they vote into a single space. The big problem in testing a replacement hypothesis against ideal point estimates is that new legislators necessarily introduce new ideal points. I have settled on the approximate-Bayesian variational expectation-maximization (EM) version (Imai et al. 2016) of the Bayesian dynamic estimator from Martin and Quinn (2002). This is a compromise solution.

The problem of new ideal points becomes clear from the most obvious way to generate common-space scores. That would be to combine all roll-

call matrices for a city into one big matrix, then estimate a single point for each legislator. Since we know the terms in which each legislator served, we could compute separate median party scores for each term. But since each legislator has a single ideal point, the only way for party change to manifest is for new legislators to enter the data. While the one-big-matrix strategy is appealing because it makes no assumptions about the structure of change (see below), it stacks the deck in favor of my finding positive results.[4]

At the other end of the strategy spectrum, one could estimate a hierarchical model that treats the degree to which legislator L changes as an estimated parameter (Bafumi et al. 2005). This is precisely what the DW-NOMINATE algorithm does (Poole and Rosenthal 1997). Every bill has an estimated location, every legislator-term has an estimated location, and change is a function of the legislator-term ideal point and the estimated inter-term rate of change for that legislator L. This is attractive because it lets change emerge from the data. The problem is that each legislator changes by a constant rate in each term. If we think change results from shocks (e.g., the arrival of freshman legislators), the DW-NOMINATE-style approach might swamp the effect of a shock.

A middle ground is to let the rate of change for each legislator vary across terms. Martin and Quinn (2002) accomplish this by estimating the rate from data with a random-walk prior. The prior variance must be fixed, and this raises some of the above issues. Arbitrarily small evolution-variance priors tend toward returning results consistent with the one-big-matrix approach: change effectively observed as the exclusive result of the introduction of new ideal points. Arbitrarily large priors impose more change than someone with a lot of historical knowledge would find plausible.[5] Martin and Quinn (2002: 147) name three evolution variance settings in their paper: 0.1, 0.01, and 0.001. I will present results for the entire range $\{0.001, 0.002, ..., 0.01, ..., 0.1\}$.

4. The one-big-matrix strategy makes sense if we are interested in the item parameters, not legislator estimates, as in Noel (2013). This is because items (e.g., roll-call votes on bills) do not bridge legislative terms.

5. Martin and Quinn (2002: 147) further specify different evolution-variance priors for different justices according to the justices' ideological extremity, necessarily perceived by the researcher, and the number of cases they heard over their tenures. Given the novelty of my data and the absence of any shared historiographic basis for such decisions – really, only I know the cases – I am uncomfortable with these decisions. So I instead present results from a series of one-size-fits-all solutions.

Since we are using EM, we have to tell the model where in the posterior distribution it should begin sampling values. It is well known that EM results vary with the choice of starting values. Given a multi-modal posterior distribution, the EM algorithm may "get stuck" in a local mode and fail to converge on the "true" posterior density, regardless of how many times it iterates. Different start values can change the point at which the model gets stuck (Karlis and Xekalaki 2003). One solution is to initialize the EM algorithm with the results of a simpler model. This can help sampling of each parameter's values begin in the approximate region of its "true" values (Roberts et al. 2016).

Multiple runs of my models with random starting values do produce very different results. The simpler model I use for initialization is the one-big-matrix approach. My starting values are legislator-party ideal points from a two-parameter Bayesian item-response model (Clinton et al. 2004).[6] Repeated model runs with these starting values and the same evolution variance prior give the same results every time.

The single-matrix starting values also help generate comparable coordinate spaces for each city, since the ideal points from the single-matrix Bayesian estimation are constrained to fall between -1 and 1. They also solve the rotational invariance problems that arise in estimating separate models for each term, which is what the Martin-Quinn approach does at its core.

Does using the static model for initialization wholly return us to the one-big-matrix approach, most biased of all in favor of finding results? That depends on how one views the evolution variance prior. As expected, smaller priors produce larger, more precisely estimated regression coefficients and vice-versa. Martin and Quinn (2002: 147) offer no rule of thumb. They simply observe that legislators (Supreme Court justices) move more or less with the choice of prior, which implies the differences we will see across the regression results. One piece of advice from the EM software authors, however, is to use a prior much smaller than the 0.1 used for all but one justice in the original Martin-Quinn paper.[7]

One other choice involves the prior mean for legislator ideal points. I use zero. This means every legislator is assumed to start in the same spatial

6. Item-parameter priors, the number of sampling iterations, and the thinning interval are the same as those used in the next chapter, "Exit from proportional representation."

7. Personal communication from James Lo, March 16, 2016, on the results of Monte Carlo trials using Martin and Quinn's original prior of 0.1.

location, letting differences unfold as they vote. The results I present below are similar to the result I get with more informative priors (i.e., assigning 1 to members of the pre-PR ruling party and -1 to members of the nominal opposition).

5 Descriptive plots

I begin by plotting the location of each party's median member over the course of the city's run with PR. Figures are based on an evolution variance prior of 0.01, the middle-ground value used by Martin and Quinn. Only parties that sat for more than one session are included. The figures make sense in light of what we know from case history.

In Cincinnati, Democrats gradually break with Charter Republicans as labor gains influence in the local Democratic party. Charter Republicans return to more firm coalition with the Democrats after the election of civil rights leader Ted Berry (as Charter Republican) in November 1949. Berry's re-election as a regular Democrat in 1951 takes the party even farther from the regular Republican position by 1953. Democrats lurch even further from the regular Republican Party from 1955 onward, reflecting the replacement of remaining labor figures with middle-class liberals.

For New York City, Republicans' progressive break with the anti-Tammany coalition is evident from 1943 onward. This was the year of city planner, Parks Commissioner, and ex-reformer Robert Moses' rapprochement with Tammany Hall (Prosterman 2013: 183). Democrats move closer to the American Labor Party (ALP) after 1945, consistent with the discussion of cross-endorsement and the Democratic sweep in section 2. The ALP and Communists also diverge in 1946-7, following a split in the ALP between pro- and anti-Soviet wings. Legislator-level ideal points (not shown) reflect a clear ALP split in 1944-5. Taking the median of ALP ideal points occludes this earlier split.

In Worcester, dramatic change in the median Democratic position is clear after the elections of November 1955. This likely reflects a split in the regular Democratic party, as Young Democrats colonized the CEA Board of Directors in this period to the chagrin of at least one regular Democrat. The median CEA Democratic position approaches the Republican position from the elections of November 1957 onward. Regular Democrats meanwhile vote less frequently with both CEA factions. This may reflect opposition to down-

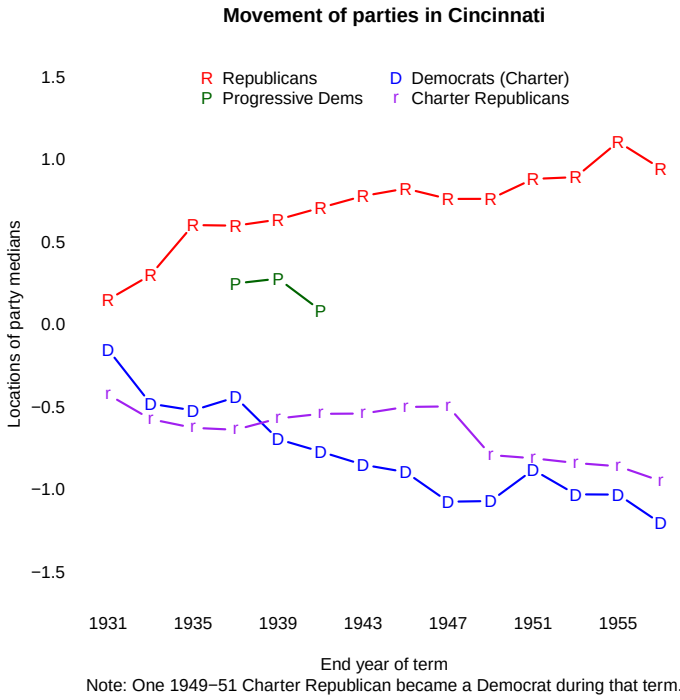

Figure 1: Evolution of the Cincinnati legislative party system.

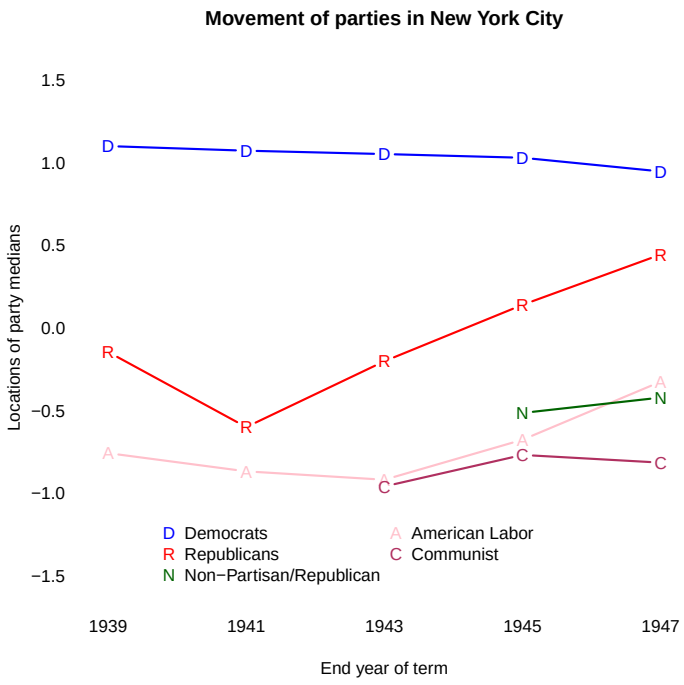

Figure 2: Evolution of the New York City Council party system. Not shown: independents.

town redevelopment and highway construction, which targeted working-class neighborhoods to the east of city center.[8]

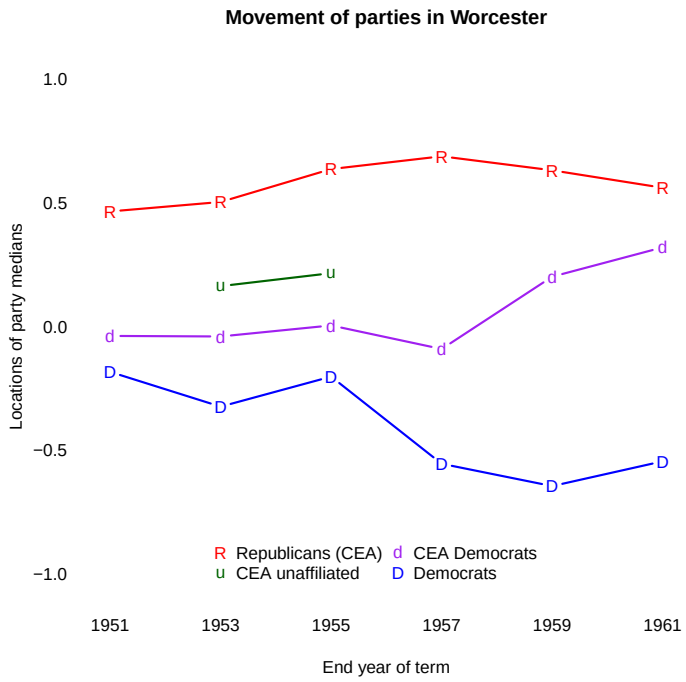

Figure 3: Evolution of the Worcester legislative party system.

6 Testing the replacement hypothesis

The overarching question is whether legislator replacement explains party movement. My dependent variable is the absolute value of change in a party's revealed position from time $t-1$ to t. The predictor of interest is the proportion of a party's delegation that are freshmen in time t. We want to know how strongly related are the proportion of freshmen elected in November and a summary measure of legislative party change over the following two years.

8. Apostola, Nicole. "Learning from our mistakes," *Nicole, Worcester*, May 2, 2016. `https://nicolecommawoo.wordpress.com/2016/05/02/learning-from-our-mistakes/`. Accessed January 6, 2017.

I begin with a bivariate model in which party movement is measured from the dynamic ideal point model with an 0.01 evolution variance (Martin and Quinn's middle ground). Figure 4 is a scatterplot with linear regression line and the bounds of a 95 percent confidence interval. The proportion of freshmen predicts absolute party movement with greater than 99 percent confidence ($\beta = 0.14$, $\sigma = 0.04$, $R^2 = 0.13$). I drop New York City independents from this model because they were not an organized party. (Table 1 includes results with and without this observation.)

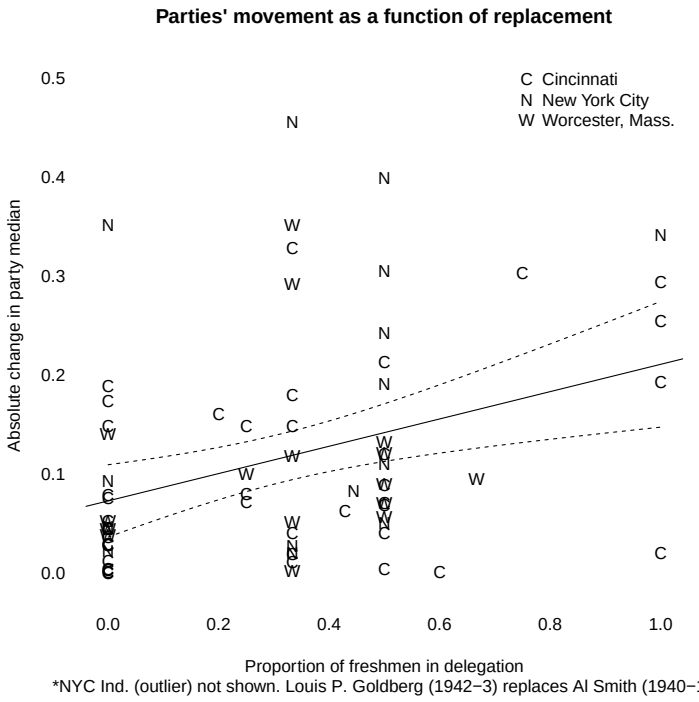

Figure 4: Underlying data and results of a bivariate regression of party movement on proportion of freshmen in delegation. Dotted lines indicate 95 percent confidence interval. The proportion of freshmen is significant with greater than 99 percent confidence ($\beta = 0.14$, $\sigma = 0.04$, $R^2 = 0.13$). Parties from each city are evenly distributed throughout the plot, which suggests the result does not depend on any one city.

Substantively speaking, replacing a party's entire delegation predicts movement across seven percent of a policy space bounded by -1 and 1. This makes

sense. Legislators are constrained. Their parties are not just collections of office-holders. Interest groups and other "policy demanders" outside of government ensure at least some continuity in the party's position, regardless of who their representatives are (Bawn et al. 2012).

Table 1 presents a series of results from models with random effects for city-terms. All models are based on the middle-ground evolution variance prior of 0.01. I control for the log of a party's delegation size because we might expect very large parties to manifest less aggregate change. The average delegation size in my data is 3.2, but Tammany Democrats enter with four observations of 12 or greater. Logging size pulls these outlier observations toward the rest of the distribution.

Two of the models include controls for concentration of the party's winners' vote shares, which I showed in the last chapter to predict party discipline. On the basis of case history and some example election returns, I interpreted this variable to capture leadership challenges within parties. Leadership challenges are an obviously possible source of party position change.

Vote-concentration models are based only on data from Cincinnati and Worcester. District structure in New York City makes it impossible to measure leadership challenges from election results. Each borough was its own multi-member district, and parties' delegations typically came from two or more districts. Contrast this with the other two cities, in which it is very clear when two same-party candidates effectively ran against each other.

Results from the random-effects regressions are consistent with results from the naive, bivariate model in figure 4. We recover the same coefficient and standard error in the outlier-free, three-city model.

How sensitive are these results to the choice of evolution variance prior? Figure 5 shows how coefficient estimates from the outlier-free, three-city, random-effects model change with the prior. I estimate 100 models over the range of omega implied by Martin and Quinn's original paper. As expected, the relationship is strongest when we allow ideal points to change the least (i.e., with an evolution variance prior of 0.001). In this case, we are effectively using the one-big-matrix approach, which imposes almost all observable aggregate change on new legislators. On the other hand, the predictive effect of freshmen replacement is statistically zero with an evolution variance prior of roughly 0.03 or greater. Note that the preceding models used a prior of 0.01.

	All cities	All cities (outlier removed)	CW	CW
Intercept	0.11^{**}	0.09^{***}	0.04	0.09
	(0.04)	(0.03)	(0.11)	(0.11)
Prop. freshmen	0.25^{***}	0.14^{**}	0.11^{*}	
	(0.06)	(0.04)	(0.04)	
Log of party size	-0.06^{*}	-0.02	0.02	0.02
	(0.03)	(0.02)	(0.06)	(0.06)
Concentration of			0.00	-0.02
winners' vote shares			(0.12)	(0.12)
AIC	-41.88	-96.96	-87.04	-87.35
BIC	-30.43	-85.58	-74.88	-77.22
Log Likelihood	25.94	53.48	49.52	48.67
Num. obs.	73	72	56	56
Num. groups: cityYear	22	22	18	18
Var: cityYear (Intercept)	0.00	0.00	0.00	0.00
Var: Residual	0.02	0.01	0.01	0.01

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 1: Results of random-effects regressions. Dependent variable is constructed with an evolution variance prior of 0.01, Martin and Quinn's middle-ground specification. "CW" means that only the Cincinnati and Worcester samples are included. This is because concentration of winners' vote shares cannot be calculated for New York City due to district structure.

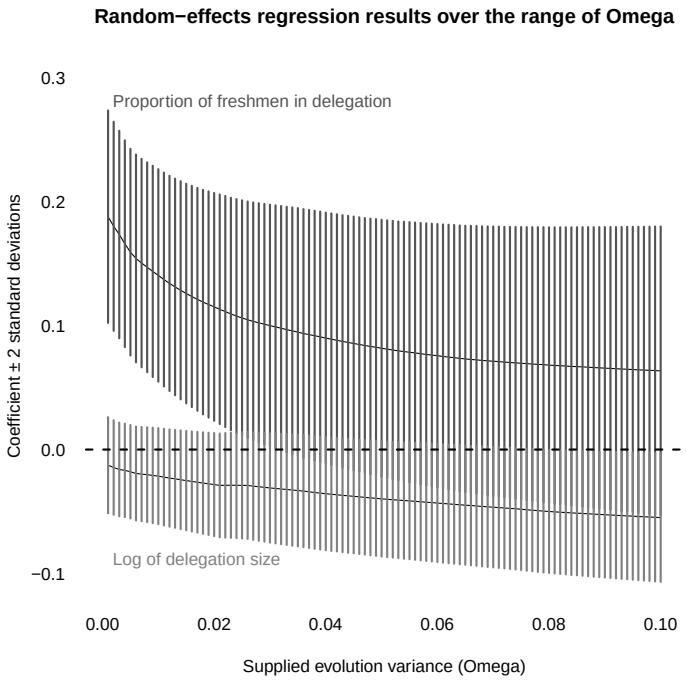

Figure 5: How the predictive effect of legislator replacement varies with the supplied level of tolerance for legislator movement.

7 Conclusion

I have given circumstantial evidence that parties' positions changed in response to change in their personnel. My quantitative evidence is the association of party movement with legislator replacement. I measure movement with dynamic ideal point estimation, which raises questions about how much legislators' should be allowed to move in coordinate space. Those questions do not have hard-and-fast answers. To address them, I presented results across 100 different values of that evolution variance prior. Together with a short description of the party system in each city, these trends suggest that replacement changed city parties under proportional representation, just as replacement changes parties in our national politics.

My next chapter addresses the joint consequences of within-party disloyalty and overall party position change. We will see that parties' relative placement and degree of legislator control predicts their behavior with respect to keeping proportional representation.

References

Bafumi, Joseph, Andrew Gelman, David K. Park, and Noah Kaplan. 2005. "Practical Issues in Implementing and Understanding Bayesian Ideal Point Estimation." *Political Analysis* 13 (2): 171–87.

Bawn, Kathleen, Marty Cohen, David Karol, Seth Masket, Hans Noel, and John Zaller. 2012. "A Theory of Political Parties: Groups, Policy Demands, and Nominations in American Politics." *Perspectives on Politics* 10, no. 3 (September): 571–597.

Bonica, Adam. 2014. "The Punctuated Origins of Senate Polarization." *Legislative Studies Quarterly* 39 (1): 5–26.

Clinton, Joshua, Simon Jackman, and Douglas Rivers. 2004. "The Statistical Analysis of Roll Call Data." *American Political Science Review* 98, no. 2 (May): 355–370.

Fleisher, Richard, and Jon R. Bond. 2004. "The Shrinking Middle in the U.S. Congress." *British Journal of Political Science* 34:529–51.

Imai, Kosuke, James Lo, and Jonathan Olmsted. 2016. "Fast Estimation of Ideal Points with Massive Data." *American Political Science Review* 110 (4): 631–56.

Jacobson, Gary C. 2007. "Explaining the Ideological Polarization of the Congressional Parties since the 1970s." Chap. 7 in *Party, Process, and Political Change in Congress, Volume 2: Further New Perspectives on the History of Congress,* edited by David W. Brady and Mathew D. McCubbins, 91–101.

Karlis, Dimitris, and Evodokia Xekalaki. 2003. "Choosing Initial Values for the EM Algorithm for Finite Mixtures." *Comptuational Statistics & Data Analysis* 41:577–90.

Karol, David. 2009. *Party Position Change in American Politics: Coalition Management.* Cambridge: Cambridge University Press.

Martin, Andrew D., and Kevin M. Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999." *Political Analysis* 10 (2): 134–53.

Noel, Hans. 2013. *Political Parties and Political Ideologies in America.* New York, NY: Cambridge University Press.

Poole, Keith T. 2007. "Changing Minds? Not in Congress!" *Public Choice* 131:435–51.

Poole, Keith T., and Howard Rosenthal. 1997. *Congress: A Political-Economic History of Roll Call Voting.* New York, NY: Oxford University Press.

Prosterman, Daniel O. 2013. *Defining Democracy: Electoral Reform and the Struggle for Power in New York City.* New York, NY: Oxford University Press.

Roberts, Margaret E., Brandon M. Stewart, and Dustin Tingley. 2016. "Navigating the Local Modes of Big Data: The Case of Topic Models." Chap. 2 in *Computational Social Science: Discovery and Prediction,* edited by R. Michael Alvarez, 51–98. New York: Cambridge University Press.

Straetz, Ralph A. 1958. *PR Politics in Cincinnati: Thirty-two Years of City Government through Proportional Representation.* New York, NY: New York University Press.

Theriault, Sean M. 2006. "Party Polarization in the U.S. Congress: Member Replacement and Member Adaptation." *Party Politics* 12 (4): 483–503.