

Does District Magnitude Matter? The Case of Taiwan

Carlisle Rainey*

Accepted for publication at *Electoral Studies*.

August 4, 2015

ABSTRACT

A sizable literature on electoral institutions argues that proportional electoral rules lead to higher voter turnout. However, recent work finds little evidence that the effect generalizes beyond western Europe and the theoretical arguments remain sparse, incomplete, and contradictory. I use a well-chosen data set to resolve problems of omitted variable bias and Bayesian model averaging to address model uncertainty and evaluate evidence both for and against the null hypotheses. I find evidence that the proportionality of electoral rules exerts no meaningful effect on turnout or the any of the theoretical mechanisms I test.

*Carlisle Rainey is Assistant Professor of Political Science, Texas A&M University, 2010 Allen Building, College Station, TX, 77843 (crainey@tamu.edu). I thank Jason Barabas, Bob Jackson, Dave Siegel, and Deb Sinha for helpful comments on previous versions of this manuscript. All data and code necessary to replicate these results can be found at github.com/carlislerainey/taiwan.

A literature extending forward from Powell (1986) suggests that majoritarian electoral rules generate chronically lower voter turnout. Since politicians have incentives to represent voters, who tend to have higher socioeconomic status, some political scientists have argued that countries with chronically low participation rates should switch to proportional rules in order to boost turnout (Lijphart 1997, 1999). However, recent work that extends empirical tests beyond western Europe casts doubt upon the general causal claim that proportional rules generate higher turnout. For example, using a large set of democracies inside and outside Europe, Blais and Dobrzynska (1998) demonstrate that, while electoral institutions might affect participation, the effect is most likely quite small. Further, Blais and Aarts (2006) criticize this literature for claiming that proportional rules cause higher turnout, while only observing a small, inconsistent correlation and having conflicting explanations for the effect. Theoretically, Jackman (1987) points out that voters have less incentive to participate in PR systems because elections are less decisive (Powell 2000). Further, Schram and Sonnemans (1996) and Herrera, Morelli, and Palfrey (2013) offer formal models and experimental results that suggest proportional rules *decrease* the incentives to mobilize voters, at least in some contexts. Karp, Banducci, and Bowler (2007) and Rainey (2015) use observational data to show that parties exert greater mobilization effort under *disproportional* rules. This research offers sufficient theoretical and empirical evidence to give political scientists pause, suggesting that until the effect is demonstrated in a wider range of cases and a more compelling theoretical argument emerges, skepticism is warranted and further study is required.

In this paper, I use survey data from the 2001 Taiwan legislative election to test the claim that more proportional electoral rules lead to higher turnout, as well as several of the causal mechanisms. These data are especially useful for testing this claim because Taiwan is one of few countries with substantial variation in the proportionality of electoral rules across electoral districts. In particular, the 2001 Taiwanese electoral system features several single-member districts and multimember districts with magnitudes up to 13. Using these data allows me to

hold the national political context constant as district magnitude varies, unlike earlier studies, making the inferences more compelling.

I use a Bayesian model averaging approach to combat problems of model uncertainty and assign probabilities to hypotheses (Montgomery and Nyhan 2010) and I find little support for the claim that proportional rules lead to higher turnout or for the theoretical mechanisms that some researchers have suggested to explain the purported relationship. In fact, using posterior probabilities, I argue that proportional rules have no meaningful effect on turnout and that none of the proposed methods operates.

Electoral Rules, Parties, and Turnout

As district magnitude increases, two important and related changes occur in the political system. First, the assignment of seats based on vote shares becomes more proportional as district magnitude increases (Cox 1997; Benoit 2000). Duverger (1954) first identified this as the “mechanical effect.” When district magnitude is small (e.g., single member districts), electoral rules punish small parties by assigning a smaller percentage of legislative seats to these parties than their vote share. In contrast, larger parties tend to receive a larger percentage of seats than their vote share. However, as district magnitude increases, the assignment of seats tends to be more proportional.

Second, a well-developed theoretical and empirical literature extending forward from Duverger suggests that district magnitude increases the number of political parties. In particular, as magnitude increases, the number of parties increases as well, but a larger increase occurs in systems with more social cleavages. Cox (1999*b*) shows formally that parties have a stronger incentive to coordinate and drop out of contests as the district magnitude shrinks. Sophisticated empirical work confirms many of these theoretical predictions (e.g. Clark and Golder 2006).

The Proportionality of Seat Assignment

All else constant, larger district magnitudes leads to more proportional outcomes, while smaller magnitudes lead to less proportional outcomes (Cox 1997; Benoit 2000). Banducci, Donovan, and Karp (1999) and Bowler and Lanoue (1992) argue that systems that disproportionately reward large parties leave supporters of small parties feeling under-represented, or not represented at all. In the extreme case of a single-member district with a plurality rule, the party that wins the most votes, even if it is less than a majority, represents the entire district. Often, 45% of voters or more find themselves with no candidate representing them in the legislature. While districts with magnitudes greater than one do not easily allow such severely biased outcomes, many small parties get no seats at all. In general, as district magnitude gets larger, smaller parties are able to win seats. Thus, as district magnitude increases, more voters receive representation because their preferred party receives enough votes to earn a seat. This leads to the first empirical hypothesis, which focuses on whether individuals feel represented or not.

REPRESENTATION HYPOTHESIS: As district magnitude increases, potential voters are more likely to feel represented in the political system.

Ideological Heterogeneity

Downs (1957) and Cox (1999a) show that when district magnitude is one, all but two parties have an incentive to exit the system and the two remaining parties have an incentive to converge to the median voter. However, as district magnitude increases, so does the number of political parties that can exist in the system in equilibrium (Cox 1997, 1999b; Clark and Golder 2006). More importantly, Cox (1990) formally shows that these rules also give parties an incentive to disperse across the ideological space. This has the effect of increasing the ideological heterogeneity of the parties in the district because as the district magnitude increases, parties have an incentive to appeal to more narrow constituencies and develop a political niche. As

the heterogeneity of parties' ideologies increases, voters should be more likely to find a party they agree with and feel close to (Bowler, Lanoue, and Savoie 1994). This leads to the second hypothesis, which focuses on whether or not voters feel close to a political party.

CLOSENESS HYPOTHESIS: As district magnitude increases, potential voters are more likely to feel close to a political party.

The Efforts of Political Parties

Previous work has often assumed that proportional districts create greater incentives for parties to mobilize voters, since single-member districts are plagued with the problem of noncompetitive districts. Proportional systems, on the other hand, create “nationally competitive districts” (Powell 1982, though see Rainey 2015). Cox (1999b) offers a formal extension of Powell’s suggestion, arguing that elites will “exert more mobilization effort when the probability of that effort being decisive is greater.” He argues that there is likely to be less variance in the effort exerted by parties in PR systems, and that the effort will be on average greater than in majoritarian systems. Selb (2009) offers a sophisticated empirical test of this argument, finding that turnout varies less in PR systems and is on average higher. This leads to the third hypothesis, which focuses on whether political parties contact citizens.

CONTACT HYPOTHESIS: As district magnitude increases, potential voters are more likely to be contacted by a political party.

Turnout

If citizens feel represented, feel close to a party, and are mobilized by a party, they are more likely to turn out. For example, Schattschneider (1960) argues that citizens whose views are not represented in the political system become chronically disengaged and Solt (2008, 2010) extends and tests this argument and finds strong support for the idea. Also, a large literature in

American politics stemming from Campbell et al. (1960) shows that citizens who feel closely attached to a political party are more likely to turn out. Using a rational choice model, Downs (1957) notes that campaigns serve the function of reducing information costs, helping voters overcome the costs of turning out to vote and Gerber and Green (2000) point out empirically that campaigns can get citizens to the polls, especially through canvassing. This leads to the fourth, and perhaps most important hypothesis, which focuses on whether citizens turn out to vote.

TURNOUT HYPOTHESIS: As district magnitude increases, potential voters are more likely to turn out to vote.

In summary, the literature has proposed that PR might lead to higher turnout and proposed several mechanisms. PR allows supporters of small parties to receive representation in the legislature, which gives citizens a greater stake in politics. PR also causes parties to disperse across the ideological space and make more effective, narrow appeals to voters, which allows voters to develop a close attachment to particular parties. Finally, PR increases parties' incentives to mobilize voters by ensuring that non-competitive districts do not emerge. Each of these consequences of PR should lead to an increase in voter participation.

Data and Measures

I use survey data from the 2001 Taiwanese legislative elections collected by Taiwan's Election Democracy Study and included in Module 2 of the Comparative Studies of Electoral Systems. These data offer a rich source of information district and individual-level variables to assess the effect of district magnitude on a variety of attitudinal and behavioral outcomes.¹

¹One might wonder about the impact of the choice to use individual-level data rather than aggregate-level data. A simple regression of district-level turnout on district magnitude for the 2001 election in Taiwan suggests about a 2.5 percentage point increase in turnout across the range of district magnitude (i.e., a small, insignificant effect). Unfortunately, a rich set of control variables is simply not available at the district level. However, using

Problems with Cross-National Research

One problem with current research is that PR countries differ in many ways from majoritarian countries. Indeed, Powell (2000) suggests the two are so different as to constitute different “visions” of democracy. Even more problematic, differences between countries lead to the adoption of different electoral rules (Blais, Dobrzynska, and Indridason 2005; Boix 1999). Therefore, making convincing causal inferences using cross-national observational data is difficult because it is impossible to control for all differences between PR and majoritarian countries. Any excluded variable that has a meaningful influence on the outcome will lead to biased estimates (Greene 2008). The problem lies in the fact that aspects of the national political environment likely differ between PR and majoritarian systems. Omitting variables (e.g., to preserve degrees of freedom or because they are unobserved) might lead the analyst to incorrectly conclude that the explanatory variable of interest has a meaningful effect on the outcome. In this setting, separating the effects of electoral rules and national political context is very difficult.

Unlike most countries, Taiwan features substantial variation in proportionality across districts. District magnitude in Taiwan ranges from one to 13, and includes several single-member districts. In this case, district magnitude serves as an excellent measure of the proportionality of the electoral rules (Cox 1990; Benoit 2000). That is, Taiwan offers a case in which the proportionality of the electoral rules varies *within* a constant national political context. Most countries have little variation in district magnitude, and those that do often have no single-member districts. For example, Portugal has district magnitudes ranging from three to twenty-seven, but for the purposes of this study, it is important that countries include both single- and multi-member districts, since a larger effect of district magnitude on proportionality occurs near one (Grofman and Selb 2011). Thus, Taiwan offers a rich opportunity to study the

individual-level demographic data to predict aggregate turnout suggests that about 80% of this 2.5 percentage point effect is due to variation in demographics across districts. Thus, the aggregate data seem consistent with my argument that the effect of district magnitude on turnout is negligible.

effect of district magnitude.

In trying to make a causal inference, an analyst must consider how the key explanatory variable is assigned. In the best case scenario, this variable is assigned randomly so that the variable of interest is not correlated with other variables that might influence the outcome. As discussed above, electoral rules are not randomly assigned across counties, which makes causal inference using observational data difficult. However, district magnitude in Taiwan is based on the size of the administrative regions, which also serve as electoral districts. This alleviates some concern over strategic politicians assigning district magnitude in a systematic and politically beneficial manner. However, the administrative regions may still differ in important ways. Economic and social conditions, ethnic makeup, education levels, and population density might vary systematically across the electoral districts. These factors must be controlled in the analysis. However, these Taiwanese data do eliminate the need to control for national-level variables. Because the national electoral environment is not changing in the single election I study, it is not necessary to control for its many features.

District Magnitude

The CSES data contain a host of national and district level data, including district magnitude, the variable of interest in the analysis.² The relationship between district magnitude and proportionality, however, is non-linear. Increasing district magnitude by one seat when district magnitude is small has a large effect on proportionality. When district magnitude is large, though, adding an additional seat does not increase the proportionality as much (Grofman and

²Unfortunately, respondents' electoral districts are not collected within the CSES data. As far as I am aware, no national-level surveys record this information. However, information is gathered about the respondents' county of residence, which closely corresponds to the electoral district. Only three counties contain multiple electoral districts (and all have similar magnitude) and 24 of the remaining 27 counties contain exactly one electoral district. Taipei County contains three electoral districts, while Taipei City and Kaohsiung City contain two electoral districts each. Although I do not know respondents' exact electoral district within these counties, the electoral districts are of similar magnitude and are geographically contiguous. I simply average the district magnitude across the three counties and use that average as the measure of district magnitude for that particular county.

Selb 2011). It seems appropriate to take this non-linearity into account by using the natural log of magnitude in a linear predictor. I follow this approach in the analyses presented here, but other methods of construction yield substantively similar conclusions.³

Survey Variables

Although there are many ways for citizens to participate in a democratic political system—including writing letters, donating time or money to political campaigns, or even discussing politics with a friend—I focus this analysis of participation on the decision to vote. Among the many forms of participation, turnout is perhaps the most important for democratic theory, as voting serves as the only formal mechanism for citizens to directly hold their representative accountable. If certain interests are under-represented at the polls, then the system has no formal mechanism to correct for the bias (Lijphart 1997, 1999). Thus, turnout is one of the most important modes of participation. Further, I am building theoretically and substantively on previous literature that focuses mainly on the decision to turn out. This allows me to show important results for a commonly analyzed form of political participation.⁴ To measure turnout, I use an indicator variable that equals one if the respondent self-reported turning out to vote, and zero otherwise.⁵

In addition to turnout, I attempt to explain whether the respondent feels represented. I measure this by a survey question that asks, “Would you say that any of the parties in Taiwan represents your views reasonably well?” The indicator variable equals one if respondents indicate

³I replicated the analysis simply using district magnitude. Non-linear marginal effects are still allowed in this model because of the non-linear logit link function. I also replicated the results using a dummy variable for single-member districts. Substantive conclusions using the alternative measures are quite similar. For simplicity of presentation, I present only models using the natural log of district magnitude.

⁴Substantive conclusions do not change for a wide range of participatory acts.

⁵As with most self-reported measures of turning out to vote, over-reporting is not trivial in these data, about 15 percentage points on average. If over-reporting were strongly related to district magnitude, such that respondents in smaller districts were much more likely to over-report, then this could obscure a positive effect of district magnitude. However, using the method suggested by Wright (1993) to model misreporting, I find that any effect of district magnitude on over-reporting is likely small and not nearly large enough to obscure a substantively meaningful effect of district magnitude.

they felt represented and zero if they did not. Respondents who did not know or refused to answer were coded as missing. This measure has the obvious problem that it asks whether any party *in Taiwan* represents the respondent's views. It might be the case that no party with a chance of winning a seat in the respondent's own district represents her views, but that some party in another, larger district does. This biases one toward a null effect, because all respondents would be looking to the same set of parties, leading to no differences across electoral districts. However, this survey question still offers a useful test because, all else equal, respondents almost certainly are less likely to feel represented by a party not seriously competing in their own district.

Also, I attempt to explain whether or not respondents feel close to a political party. I measure this through a survey question that asks, "Do you usually think of yourself as close to any particular political party?" The indicator variable equals one if the respondent indicated that he or she felt close to a political party, and zero if not. Again, respondents who did not know or refused to answer were coded as missing. This variable does not have the same small bias as the measure of whether the respondent feels represented.

Finally, I model whether a political party contacted the respondent. I measure this through a survey question that asks, "During the last campaign did a candidate or anyone from a political party contact you to persuade you to vote for them?" The indicator variable equals one if the respondent replied yes, and zero not. Those who refused or did not know are coded as missing.

Control Variables

Obviously, even to consider making tentative causal claims from these observational data, I must control for other potential causes of the four outcomes of interest, particularly other potential causes that are correlated with the respondents' region of residence. However, the literature on comparative political behavior offers little guidance, especially in the particular case of Taiwan. A large literature on turnout in American politics suggests "common suspects" such as age,

education, and income (Campbell et al. 1960; Rosenstone and Hansen 1993). However, these variables might be less important in Taiwan. A host of control variables are used in comparative studies of political behavior (e.g. Karp and Banducci 2008; Blais 2000). Beyond the “usual suspects,” variables such as union membership, marital status, gender, and population density might be important as well. Ethnicity is also an important variable to take into account in the analysis, since it is correlated with the respondents’ region of residence and might affect the respondents’ attitudes toward politics and participation.

I can exclude some variables *a priori*, especially those that do not occur causally prior to the “treatment” of district magnitude (see Gelman and Hill 2007, pp. 188-90, for a discussion of this point). For example, voter attitudes, such as feeling close to a political party, might predict turning out well, but do not belong in the model of turnout, since the goal is to estimate the effect of district magnitude, which theoretically causes higher turnout by making it more likely that citizens will feel close to a party. If the theory is correct, including an indicator for feeling close to a political party will lead to an underestimation of the effect of district magnitude. Therefore, political attitudes, which are unlikely to influence district magnitude, but are very likely to be influenced by magnitude, should be excluded from the model. In addition to attitudes, I should not include district-level variables, such as party ideologies or effective number of electoral parties. Again, these variables are very likely to be influenced by district magnitude, but unlikely to influence district magnitude. For the same reasons listed above, including these variables in the model will likely lead me to underestimate the effect of district magnitude.

The variables available in the CSES that meet the criteria for control variables include the following: age in years and indicators for whether the respondent is male, married, a union member, from a rural area, from a small or middle-sized town, from a suburb, white-collar, a worker, of Hakka ethnicity, of Min-Nan ethnicity, or a mainlander. To capture the potential non-linear effects of age, I also include the respondent’s squared age divided by 100.

The Empirical Approach

Most statistical analyses in political science proceed by choosing, *a priori*, a statistical model, often some linear regression or generalized linear model estimated with ordinary least squares or maximum likelihood. After estimating the coefficients, the researcher tests the hypothesis that some coefficient of interest is equal to zero. If the data are more extreme than would be expected if the null hypothesis were true, then the researcher rejects the null hypothesis in favor of an alternative hypothesis, concluding, for example, that the effect is greater than zero.

For the purposes of the analysis, this process has two important shortcomings. First, I do not have strong prior beliefs about which model is correct, particularly which control variables are important to include. The typical procedure offers little guidance in the presence of model uncertainty. Second, I would like to be able to evaluate evidence for and against the null hypothesis. The classic procedure only allows me to assess evidence against the null.

To deal with issues of model uncertainty, I adopt a Bayesian model averaging approach (Raftery 1995; Montgomery and Nyhan 2010). To assess evidence both for and against the null (see Gill 1999), I use model probabilities (often referred to as “posterior” probabilities) (Jackman 2004; Gill 2008; Jackman 2009). Below, I discuss each technique in terms of the theory and hypotheses. Because all of the outcomes of interest are dichotomous, I use logistic regression to model the outcome of interest.

Computing Probabilities

Posterior probabilities improve upon the common p -value in several important ways. Most importantly for my purposes, posterior probabilities allow me to assess the evidence in favor of a null hypothesis.⁶ Assume a large set of n models, one of which is “correct” in the sense that it

⁶The literature has typically used a ratio of two posterior probabilities known as “Bayes’ factor” to compare two models. However, for expository purposes, I only discuss the posterior probabilities of models. This quantity is more familiar and intuitive and comes in handy in dealing with model uncertainty.

accurately describes the process that generated a data set, \mathbf{D} . Each model, M_i , has a posterior probability that can be computing using Bayes' rule.

$$p(M_i|\mathbf{D}) = \frac{p(\mathbf{D}|M_i)p(M_i)}{\sum_{k=1}^n p(\mathbf{D}|M_k)p(M_k)} \quad (1)$$

Each term of the equation has an intuitive interpretation, but computation can be difficult. The first term in the numerator, $p(\mathbf{D}|M_i)$, is simply the probability that the observed data would occur under model i . The second term in the numerator, $p(M_i)$, is the prior probability of model i . When n models are under consideration, each model is typically assigned a prior probability of $1/n$. That is, most analyses consider all models equally likely before observing the data. Finally, the denominator simply serves as a normalizing constant, ensuring that the posterior probabilities sum to one.

In particular, the term $p(\mathbf{D}|M_i)$ is difficult to compute. On a technical level, this quantity is computed by integrating the likelihood function of the model across the parameter vector. This can be a high-dimensional (and thus computationally expensive) integration. To combat these difficulties, I use the Bayesian Information Criterion (BIC), which contains enough information to approximately calculate the posterior probabilities (Raftery 1995; Kass and Raftery 1995).⁷

Thus, the data combined with posterior probabilities provide a framework within which to evaluate evidence both for and against the current explanations for observed differences in turnout across countries, primarily that district magnitude increases the probability of turning out to vote. Further, I can use this approach to test the mechanism by which the theory works, namely that respondents in districts with larger magnitudes will be more likely to feel close

⁷The posterior probability of model M_i can be approximately computed from its BIC, using the fact that $p(M_i|\mathbf{D}) \propto \exp\left(-\frac{1}{2}\text{BIC}_i\right)$. By the law of total probability, one can approximately obtain the posterior probability

of a model by $p(M_i|\mathbf{D}) \approx \frac{\exp\left(-\frac{1}{2}\text{BIC}_i\right)}{\sum_{k=1}^n \exp\left(-\frac{1}{2}\text{BIC}_k\right)}$. For a more detailed discussion of posterior probabilities and their relationship to the BIC, see Raftery (1995) and Kass and Raftery (1995)

to a party, feel represented by a party, and be contacted by a party. However, in addition to evaluating evidence for and against the null, I must also overcome issues relating to model uncertainty.

Bayesian Model Averaging

As stated above, I have few prior beliefs about the correct model for each of the outcome variables. The large literature on political behavior in American politics points to the “usual suspects,” an emerging literature on comparative political behavior offers some insights, and I have some intuitions about how Taiwan might deviate from typical cases. I can further eliminate variables that do not occur causally prior to the key explanatory variable, district magnitude. Beyond these basic ideas, I have little insight into the “correct” models for the outcomes, particularly for an understudied case such as Taiwan. Many possible combinations of control variables exist. In the analysis, I identify 16 control variables that might be useful, meaning that I must adjudicate among $2^{16} = 65,536$ possible models.⁸

One solution to the problem of model uncertainty that seems especially useful for the problem at hand is Bayesian model averaging (BMA), which allows me to systematically compare and combine the inferences from a large number of models. Parameters are estimated by averaging across all models, weighting by the posterior probability of the model. For example, suppose that one particular model under consideration suggests a large positive effect of district magnitude. If this model has a high probability, it will be weighted heavily, pushing the BMA estimate upward. Conversely, if the model has a low posterior probability, then it will have very

⁸In addition to the many combinations of control variables, it is not clear a priori how to model district magnitude. A weak theoretical rationale suggests the effect of adding an additional seat in a district should be decreasing as the number of seats increases, but there are many ways to model this effect. I consider three potential models: the district magnitude, the natural log of district magnitude, and an indicator for single-member districts. Even with the patience to carefully evaluate and compare each of these models using tools such as Bayes factors or cross-validation, it is unclear how to test hypotheses and make inferences when conclusions about model specification are inherently uncertain. I perform the analysis using each of the three measures of district magnitude, and find similar substantive results for each. I limit the discussion to the most theoretically appealing, the natural log of district magnitude.

little influence on the BMA estimate.

Thus, the BMA procedure produces results not commonly found in social science research, in that the mean and standard error are of little use. This is because the uncertainty surrounding the estimated coefficient can be highly non-normal. Indeed, the uncertainty around almost all coefficients can be described as a mixture of several normal distributions (one for each model that does not set the coefficient equal to zero), and a mass of probability at zero.⁹ Each normal distribution is weighted by the posterior probability of its associated model, and the mass at zero is simply the sum of the posterior probabilities of models that set the coefficient to zero. If the most likely models exclude a variable, then I can be confident it has no effect. However, if the most likely models include a variable, then I can be confident it has an effect.

Because the uncertainty around each estimate is highly non-normal, the results from a BMA procedure are typically summarized using plots rather than tables. I use two types of plots to show the results: one which shows the uncertainty surrounding each coefficient, and another which ranks the models, summarizes the posterior probabilities, and shows the estimated effects of the variables for each model. The results section describes these figures in detail.

In summary, the BMA strategy allows additional leverage in two areas beyond traditional hypothesis testing in political science. First, a Bayesian estimation strategy allows one to test the hypothesis that the variable of interest, district magnitude, has no effect on each of the four outcome variables of interest: voting, feeling close to a party, feeling represented, and being contacted by a party. Second, I can take into account model uncertainty, especially in the inclusion and exclusion of control variables.

⁹It might seem strange to have a distribution that is partly continuous and partly discrete. This occurs because some models exclude certain variables (or equivalently set the coefficients equal to zero). These models have positive posterior probability and therefore there is some non-zero probability that the coefficients equal zero in the averaged result. Variables that are included in the model have distributions associated with them. These distributions are continuous and describe the uncertainty around the estimates. It is important to recall that the probability of getting a particular sample from a continuous probability distribution (often called a probability density function), is zero. The results from the BMA provide a weighted mixture of continuous and discrete distributions. In the case of BMA, these distributions are described with a density function over the real line with a mass of probability at zero, so that the probability of sampling a zero is positive.

Results

Results derived from a BMA approach have a slightly different interpretation than do those from more traditional methods, such as logistic regression estimated with maximum likelihood, which guarantee normal(ish) posterior distributions. Posterior distributions computed using BMA are almost guaranteed to be non-normal, with a discrete mass of probability at zero, and a continuous component consisting of a mixture of normal distributions that might be multimodal. For this reason, it makes sense to plot the posterior distributions and interpret them directly, rather than relying on a potentially misleading table of expected coefficients and standard deviations.

The Turnout Hypothesis

I begin by briefly summarizing the results from the models of the turnout decision. The best model, which has a posterior probability of 0.26, includes only the respondent's age and the indicator for whether the respondent is married. The best five models have a combined posterior probability of 0.47.

The key variable of interest—district magnitude—is equal to zero in each of the best five models. In fact, the posterior probability that district magnitude is non-zero is 0.03. That is, district magnitude almost certainly does not help to predict the turnout of respondents. There is considerable evidence for the null hypothesis that district magnitude does not affect turnout. I can be quite certain based on these models that district magnitude exerts no meaningful impact on turnout. Figure 1 provides a summary of the posterior distributions created by averaging across the models. This figure shows that there is a high probability that the coefficient equals zero. Figure 2 gives information about the model space, the most likely models, and the estimated coefficients in those models. Notice that district magnitude is included in very few of the most likely models. Also notice that every model including district magnitude predicts a

negative effect (is colored red/grey). This provides strong evidence for the null.

[Figure 1 about here.]

[Figure 2 about here.]

The Representation Hypothesis

A similar set of results emerges from the models of whether or not the respondent feels represented. The best model includes indicators for whether the respondent is male and married and has a posterior probability of 0.09. The cumulative probability of the best five models is 0.29.

As before, district magnitude is included in none of the best five models. The posterior probability that the coefficient is non-zero is 0.02. This provides considerable evidence for the null hypothesis. Based on these data, I can be quite confident that district magnitude has no meaningful effect on the probability that a respondent feels represented. Figure 3 provides the posterior distribution created by the BMA procedure in the models of feeling represented. This figure shows that district magnitude almost certainly has no effect. Figure 4 gives the information about the model space, the most likely models, and the estimated coefficients in those models. Notice that district magnitude is included in very few of the most likely models, even though the estimated effect is positive (is colored black) for some models including it. Again, these results provide strong evidence for the null.

[Figure 3 about here.]

[Figure 4 about here.]

The Closeness Hypothesis

The BMA results for whether the respondent feels close to a political party are similar to the previous results. The best model includes only household income, but the posterior probability

of this model is only about 0.07. The best five models have a cumulative probability of about 0.21.

As before, district magnitude has a positive effect in very few models, and those models receive very little weight. The posterior probability that the effect is non-zero is 0.03, which is, again, strong evidence against the closeness hypothesis. Figure 5 provides the posterior distribution created by the BMA procedure in the models of feeling represented. This figure shows that while the non-zero portion of the density is almost entirely positive, the vertical bar extends to about .97. This offers strong evidence of no effect. Figure 6 gives the information about the model space, the most likely models, and the estimated coefficients in those models.

[Figure 5 about here.]

[Figure 6 about here.]

The Contact Hypothesis

The BMA results for whether the respondent feels close to a political party are similar as well. The best model includes only household income, and the posterior probability of this model is about 0.22. The best five models have a cumulative probability of about 0.38.

Consistent with the trend in this analysis, district magnitude is excluded from most of the best fitting models, with the posterior probability of the effect being non-zero less than 0.03. This provides considerable evidence for the null hypothesis. Based on these data, I can be quite confident that district magnitude has no meaningful effect on the probability of a respondent being contacted by a political party. Figure 7 provides the posterior distribution created by the BMA procedure in the models of being contacted. This figure shows district magnitude almost certainly has no effect. Figure 8 gives the information about the model space, the most likely models, and the estimated coefficients in those models. Notice that district magnitude is included in very few of the most likely models, though estimated effect is positive (is colored

black) in some models including it. Again, this provides strong evidence for the null.

[Figure 7 about here.]

[Figure 8 about here.]

Robustness Check with Logistic Regression

Because the BMA procedure is unusual, I replicate these results using a standard logistic regression analysis. For each outcome variable, I estimate a logistic regression model including all the covariates as explanatory variables. I include all the variables to favor a high-variance (low-bias) estimator over a low-variance (high-bias) estimator. Thus, in theory, the confidence intervals from these models are conservative and, in practice, excluding subsets of controls tend to lend more support to my conclusion that district magnitude has no meaningful effect.¹⁰ After I estimate the model, I compute the effect (i.e., change in predicted probability) of moving from a single-member district to a ten-member district for each respondent in the data set and average across all respondents, as suggested by Hanmer and Kalkan (2012). I then use Clarify-like simulation to obtain a 90% confidence interval around this estimate (King, Tomz, and Wittenberg 2000). In order to provide strong support for the hypotheses of a negligible effect, the 90% confidence intervals should only contain only substantively negligible effects (Rainey 2014). Figure 9 provides the estimated effect of shifting district magnitude from one to ten and the confidence intervals.

[Figure 9 about here.]

But which effects are substantively meaningful and which effects are negligible? Blais and Aarts (2006) provide an excellent summary of prior estimates of the effect of district

¹⁰Indeed, uncertainty over model specification in a conventional setup partly motivated by choice to average across modeling. Particularly troubling, models that fit the data similarly well (according to fit criteria such as AIC or BIC) offered different substantive conclusions. BMA offers a principled way to combine these estimates

magnitude. They provide seven estimates from four studies that argue for a positive effect of district magnitude on turnout. The average of these estimates is about ten and the minimum is about five. In arguing that the United States switch to proportional representation, Lijphart (1997) suggests that effect would be a nine to twelve percentage point increase in turnout. Thus, any effect smaller than 0.10 casts some doubt on prior estimates and any estimate smaller than 0.05 is not substantively important.

In spite of emphasizing unbiasedness at the expense of variability, the data still strongly support the claim that district magnitude has little to no effect on turning out to vote. The upper bound of the confidence interval is about 0.04, suggesting that the effect of district magnitude is much smaller than suggested by previous research and substantively negligible. None of the other estimates are statistically significant, but the data offer less evidence against meaningful effects in these cases. However, it is worth pointing out that these data do not offer evidence *against* my claim that district magnitude has no effect on the mediating outcomes. Instead, the evidence is ambiguous, consistent with both no effect and small to moderate effects.

Discussion and Conclusion

The preceding results shed some light on the claim that proportional electoral rules lead to higher turnout and various explanations for why they may do so. In an earlier section, I describe four hypotheses suggested by previous research that purports to explain why turnout is higher under proportional rules. First, I examine the basic claim that proportional rules lead to higher turnout. Using a Bayesian approach to assess evidence both for and against the null, I find substantial evidence in favor of the null hypothesis, contrary to previous research. Second, I examine the theoretical mechanisms of feeling represented, feeling close to a party, and being contacted by a party. I find evidence that an increase in district magnitude does not affect any of these theoretical mechanisms.

These results cast doubt upon the claim that turnout is higher under proportional rules. I use the case of Taiwan to hold the national political context constant while varying the proportionality of the electoral rules substantially. If the claim that turnout is higher under proportional rules is correct, then turnout should be higher in districts with a large number of seats. This is not the case. Similarly, if the theory explaining the phenomenon is correct, then citizens in districts with large magnitudes should feel more represented, feel closer to political parties, and/or be contacted more often. The evidence in each case strongly favors the null.

The literature following Powell (1986) argues strongly that proportional rules lead to higher turnout, leading Lijphart (1997) to suggest that democratic political systems with low turnout consider switching to some form of proportional representation to boost citizen participation. Since this recommendation, scholars have suggested more skepticism (Blais and Aarts 2006). This analysis supports the accumulating evidence that the relationship between proportional rules and higher turnout is not as consistent and powerful as once thought.

References

- Banducci, Susan A., Todd Donovan, and Jeffrey A. Karp. 1999. "Proportional Representation and Attitudes about Politics: Evidence from New Zealand." *Electoral Studies* 18:533–555.
- Benoit, Kenneth. 2000. "Which Electoral Formula is Most Proportional?" *Political Analysis* 8:381–388.
- Blais, Andre. 2000. *To Vote or Not to Vote*. Pittsburg: University of Pittsburg Press.
- Blais, Andre, Agnieszka Dobrzynska, and Indridi Indridason. 2005. "To Adopt or Not to Adopt Proportional Representation: The Politics of Institutional Choice." *British Journal of Political Science* 36:182–190.
- Blais, Andre, and Agnieszka Dobrzynska. 1998. "Turnout in Electoral Democracies." *European Journal of Political Research* 33:239–261.
- Blais, Andre, and Kees Aarts. 2006. "Electoral Systems and Turnout." *Acta Politica* 41:180–196.
- Boix, Carles. 1999. "Setting the Rules of the Game: The Choice of Electoral System in Advanced Democracies." *American Political Science Review* 93:609–624.
- Bowler, Shaun, and David Lanoue. 1992. "Strategic and Protest Voting for Third Parties: The Case of the Canadian NDP." *Western Political Quarterly* 45:485–499.
- Bowler, Shaun, David Lanoue, and Paul Savoie. 1994. "Electoral Systems, Party Competition, and Strength of Partisan Attachment: Evidence from Three Countries." *Journal of Politics* 56:991–1007.
- Campbell, Angus, Philip E. Converse, Warren Miller, and Donald Stokes. 1960. *The American Voter*. New York: Wiley and Sons.
- Clark, Willam R., and Matt Golder. 2006. "Rehabilitating Duverger's Theory: Testing the Mechanical and Strategic Modifying Effects of Electoral Laws." *Comparative Political Studies* 39:679–708.
- Cox, Gary. 1990. "Centripetal and Centrifugal Incentives in Electoral Systems." *American Journal of Political Science* 34:903–935.
- Cox, Gary. 1997. *Making Votes Count: Strategic Coordination in World's Electoral Systems*. Cambridge: Cambridge University Press.
- Cox, Gary. 1999a. "Electoral Rules and Electoral Coordination." *American Political Science Review* 93(1):145–161.
- Cox, Gary W. 1999b. "Electoral Rules and the Calculus of Mobilization." *Legislative Studies Quarterly* 24(3):382–419.

- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper Collins.
- Duverger, Maurice. 1954. *Political Parties: Their Organization and Activity in the Modern State*. New York: John Wiley.
- Gelman, Andrew, and Jennifer Hill. 2007. *Data Analysis Using Regression and Multi-level/Hierarchical Models*. New York: Cambridge University Press.
- Gerber, Alan S., and Donald P. Green. 2000. "The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment." *American Political Science Review* 94(3):653–663.
- Gill, Jeff. 1999. "The Insignificance of Null Hypothesis Significance Testing." *Political Research Quarterly* 52(3):647–674.
- Gill, Jeff. 2008. *Bayesian Methods: A Social and Behavioral Science Approach*. 2nd ed. Boca Raton, FL: Chapman and Hall.
- Greene, William H. 2008. *Econometric Analysis*. 6th ed. Prentice Hall.
- Grofman, Bernard, and Peter Selb. 2011. "Turnout and the (Effective) Number of Parties at the National and District Levels: A Puzzle-Solving Approach." *Party Politics* 17:93–117.
- Hanmer, Michael J., and Kerem Ozan Kalkan. 2012. "Behind the Curve: Clarifying the Best Approach to Calculating Predicted Probabilities and Marginal Effects from Limited Dependent Variable Models." *American Journal of Political Science* 57(1):263–277.
- Herrera, Helios, Massimo Morelli, and Thomas Palfrey. 2013. "Turnout and Power Sharing." *The Economic Journal* 124(574):F131–F162.
- Jackman, Robert W. 1987. "Political Institutions and Voter Turnout in the Industrial Democracies." *American Political Science Review* 81:405–423.
- Jackman, Simon. 2004. "Bayesian Analysis for Political Research." *Annual Review of Political Science* 7:483–505.
- Jackman, Simon. 2009. *Bayesian Analysis for the Social Sciences*. United Kingdom: Wiley and Sons.
- Karp, Jeffrey A., and Susan A. Banducci. 2008. "Political Efficacy and Participation in Twenty-Seven Democracies: How Electoral Systems Shape Political Behavior." *British Journal of Political Science* 38:311–334.
- Karp, Jeffrey A., Susan A. Banducci, and Shaun Bowler. 2007. "Getting Out the Vote: Party Mobilization in a Comparative Perspective." *British Journal of Political Science* 38:91–112.
- Kass, Robert E., and Adrian E. Raftery. 1995. "Bayes Factors." *Journal of the American Statistical Association* 90(430):773–795.

- King, Gary, Michael Tomz, and Jason Wittenberg. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science* 44:341–355.
- Lijphart, Arend. 1997. "Unequal Participation: Democracy's Unresolved Dilemma." *American Political Science Review* 91(1):1–14.
- Lijphart, Arend. 1999. *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries*. New Haven: Yale University Press.
- Montgomery, Jacob M., and Brendan Nyhan. 2010. "Bayesian Model Averaging: Theoretical Developments and Preactical Applications." *Political Analysis* 18(2):245–270.
- Powell, G. Bingham. 1982. *Contemporary Democracies: Participation, Stability, and Violence*. Cambridge: Harvard University Press.
- Powell, G. Bingham. 1986. "American Voter Turnout in the Comparative Perspective." *American Political Science Review* 80:17–44.
- Powell, G. Bingham. 2000. *Elections as Instruments of Democracy*. New Haven: Yale University Press.
- Raftery, Adrian E. 1995. "Bayesian Model Selection in Social Research." *Sociological Methodology* 25:111–163.
- Rainey, Carlisle. 2014. "Arguing for a Negligible Effect." *American Journal of Political Science* 58(4):1083–1091.
- Rainey, Carlisle. 2015. "Strategic Mobilization: Why Proportional Representation Decreases Voter Mobilization." *Electoral Studies* 37(1):86–98.
- Rosenstone, Steven J., and John Mark Hansen. 1993. *Mobilization, Participation, and Democracy in America*. Chicago: University of Chicago Press.
- Schattschneider, E. E. 1960. *The Semi-sovereign People*. New York: Harcourt Brace.
- Schram, Arthur, and Joep Sonnemans. 1996. "Voter Turnout as a Participation Game: An Empirical Investiagtion." *International Journal of Game Theory* 25:385–406.
- Selb, Peter. 2009. "A Deeper Look at the Proportionality-Turnout Nexus." *Comparative Political Studies* 42(4):527–548.
- Solt, Frederick. 2008. "Economic Inequality and Democratic Political Engagement." *American Journal of Political Science* 52(1):48–60.
- Solt, Frederick. 2010. "Does Economic Inequality Depress Electoral Participation? Testing the Schattschneider Hypothesis." *Political Behavior* 32(2):285–301.
- Wright, Gerald C. 1993. "Errors in Measuring Vote Choice in the National Election Studies, 1952-88." *American Journal of Political Science* 37(1):291–316.

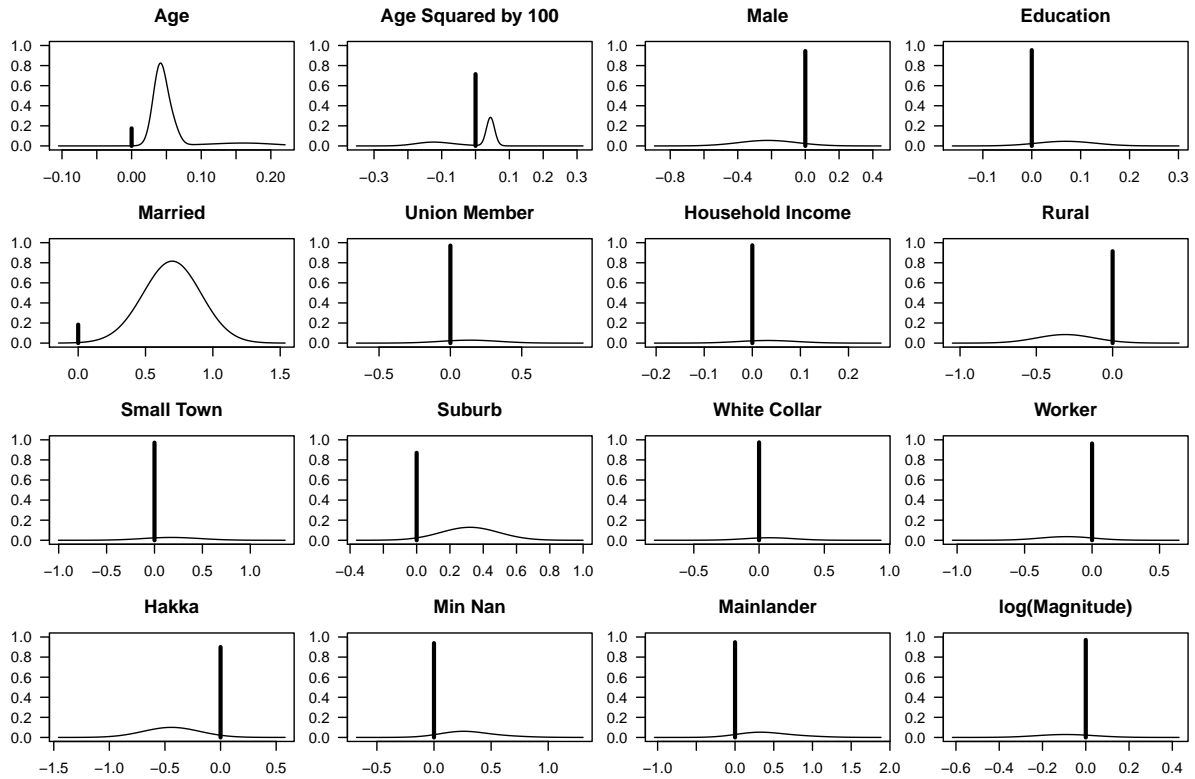


FIGURE 1: A figure showing the posterior densities of model coefficients for turnout based on the BMA procedure. Note that the height of the dark line indicates the posterior probability that the coefficient equals zero. The remaining density shows the distribution of coefficients in models that include the variable. Importantly, district magnitude almost certainly has no effect, as shown by the vertical line that nearly reaches one. Also, much of the posterior density is negative, indicating that while the effect is likely zero, it is almost certainly not positive.

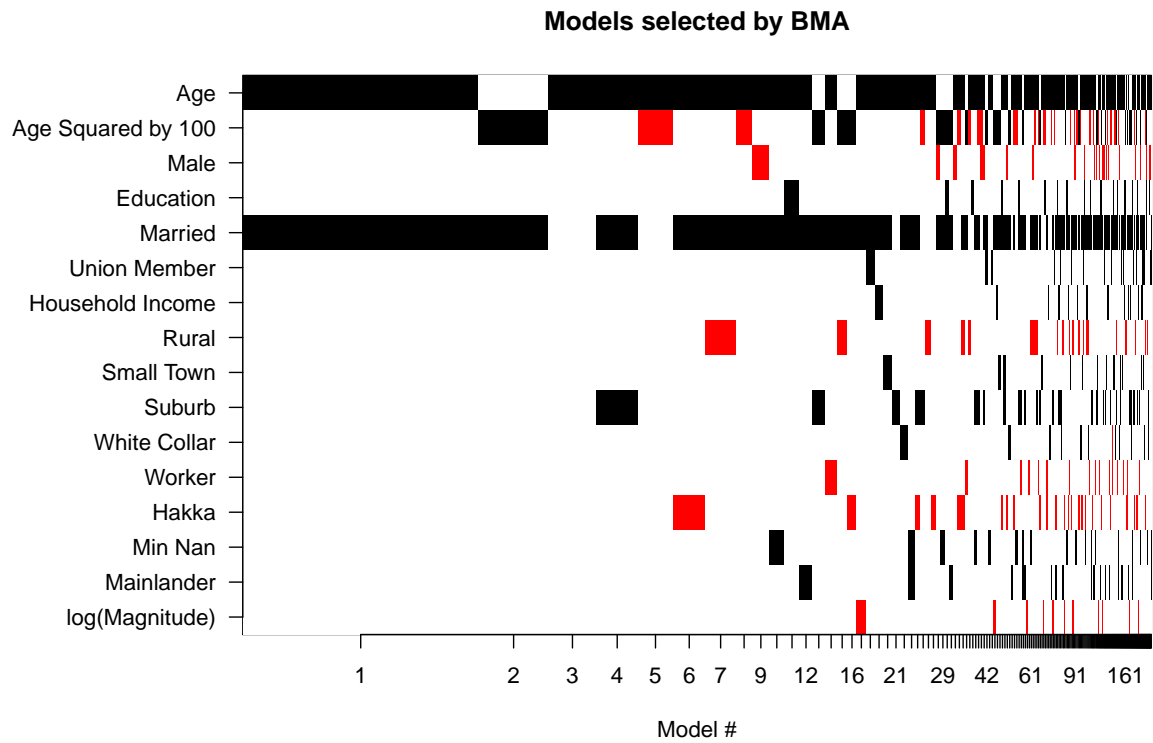


FIGURE 2: A figure showing the models of turnout considered by the BMA procedure. For each model, variables with positive effects are colored black, variables with negative effects are colored red/grey, and variables excluded from the model are white. The width of the area allotted to each model corresponds to the posterior probability of that model. For example the best model, Model #1, has a posterior probability of about 0.25 and includes the variables Age and Married, both of which have positive effects. This figure shows that district magnitude has little or no effect on the probability of turning out to vote, since it is included in very few of the most likely models.

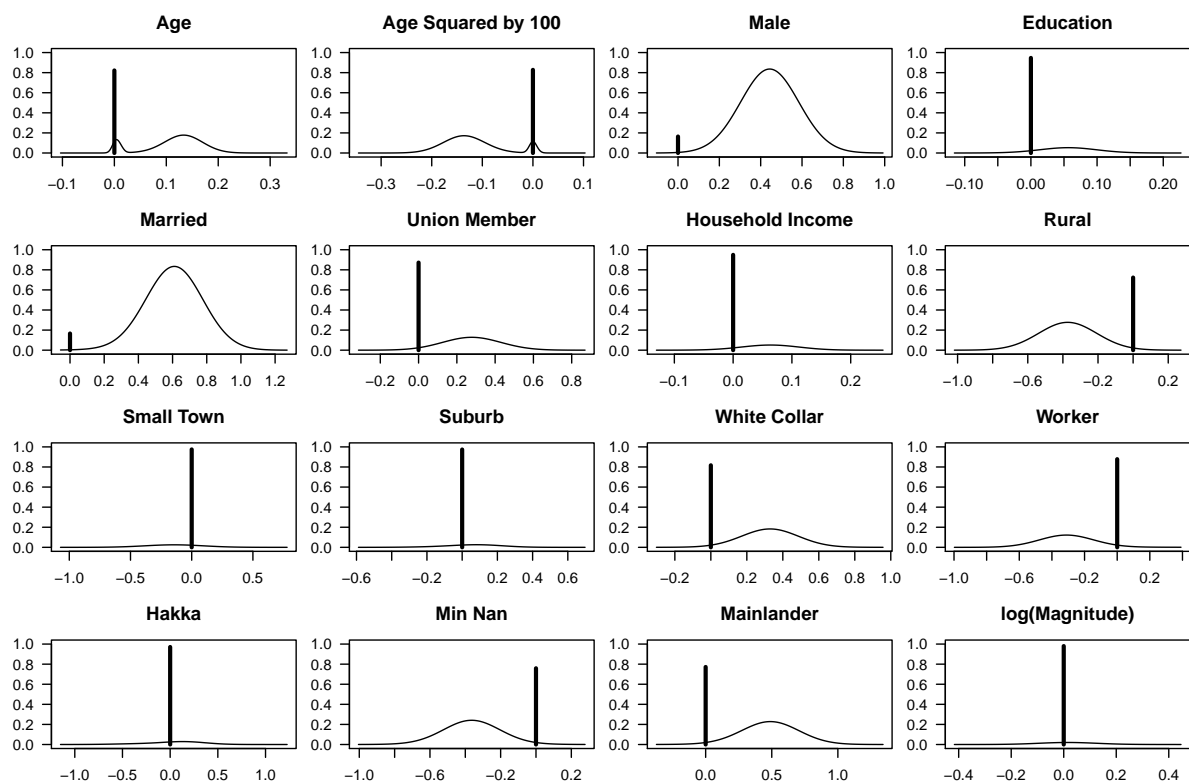


FIGURE 3: A figure showing the posterior densities of model coefficients for feeling represented based on the BMA procedure. Note that the height of the dark line indicates the posterior probability that the coefficient equals zero. The remaining density shows the distribution of coefficients in models that include the variable. Importantly, district magnitude almost certainly has no effect, as shown by the vertical line that nearly reaches one. This indicates that the effect is likely zero.

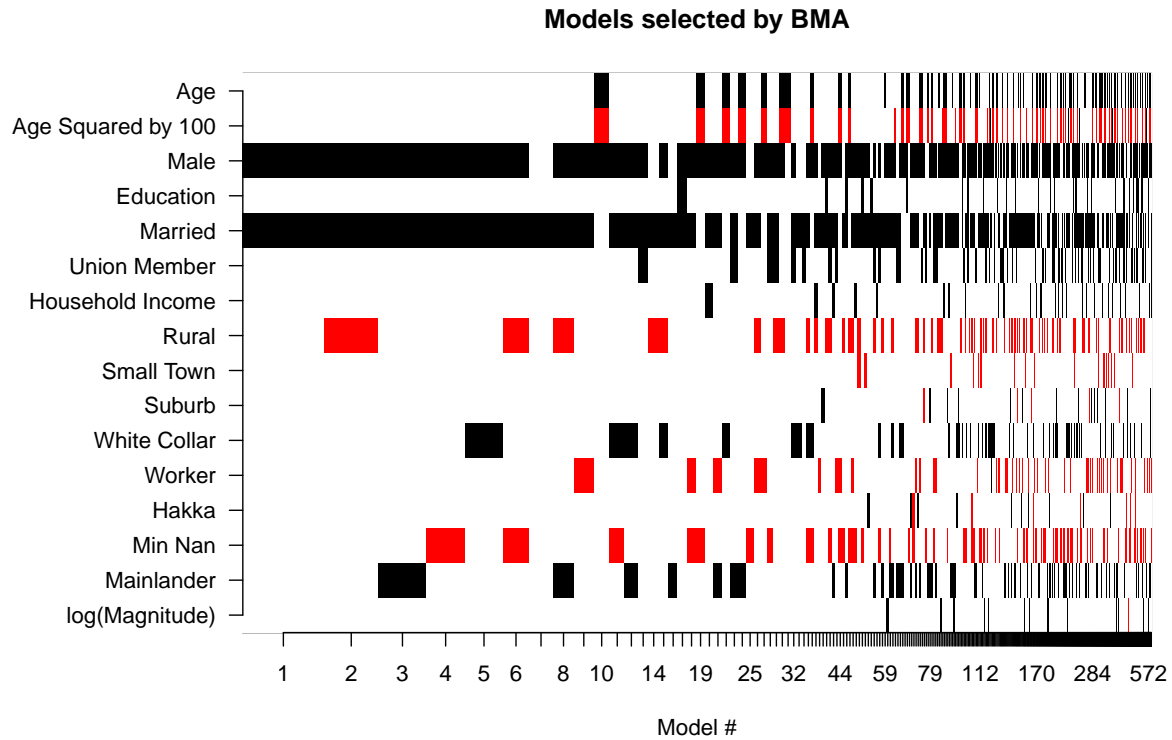


FIGURE 4: A figure showing the models of feeling represented considered by the BMA procedure. For each model, variables with positive effects are colored black, variables with negative effects are colored red/grey, and variables excluded from the model are white. The width of the area allotted to each model corresponds to the posterior probability of that model. This figure shows that district magnitude has little or no effect on feeling represented by a political party, since it is included in very few of the most likely models.

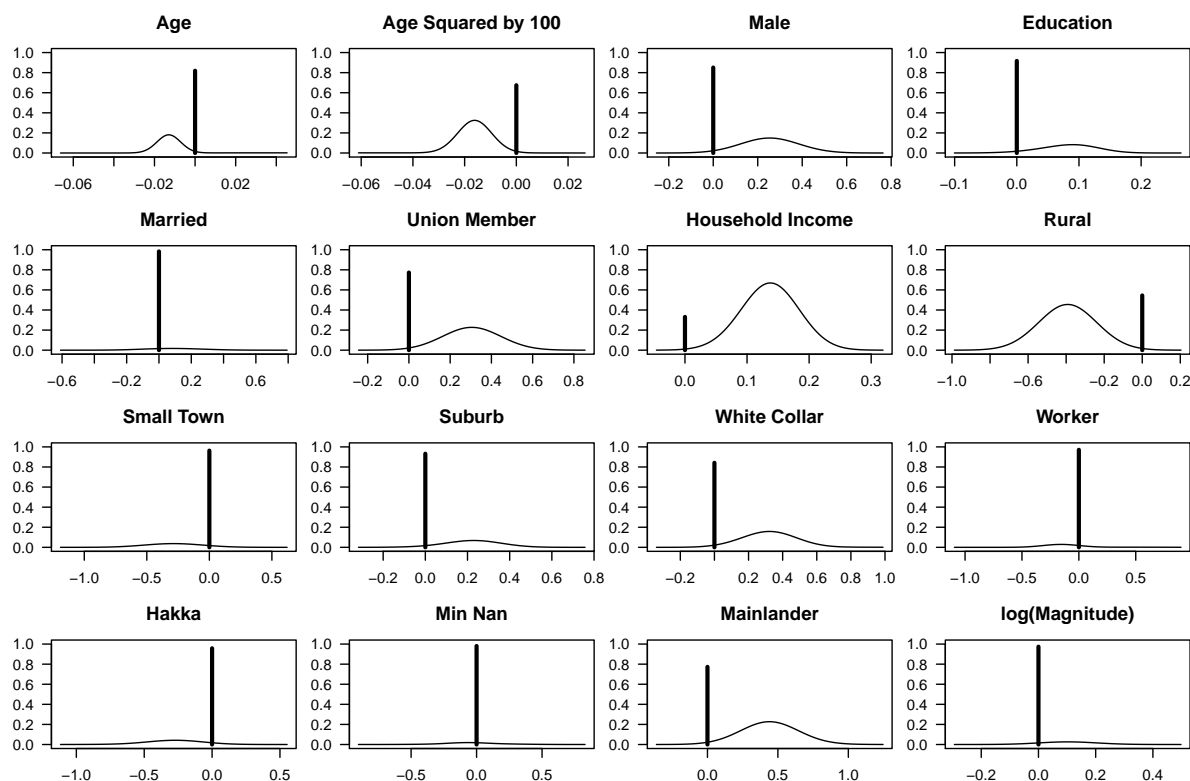


FIGURE 5: A figure showing the posterior densities of model coefficients for feeling close to a party based on the BMA procedure. Note that the height of the dark line indicates the posterior probability that the coefficient equals zero. The remaining density shows the distribution of coefficients in models that include the variable. Because the vertical line extends nearly to one, this figure shows that district magnitude has very little effect on the probability of feeling close to a political party.

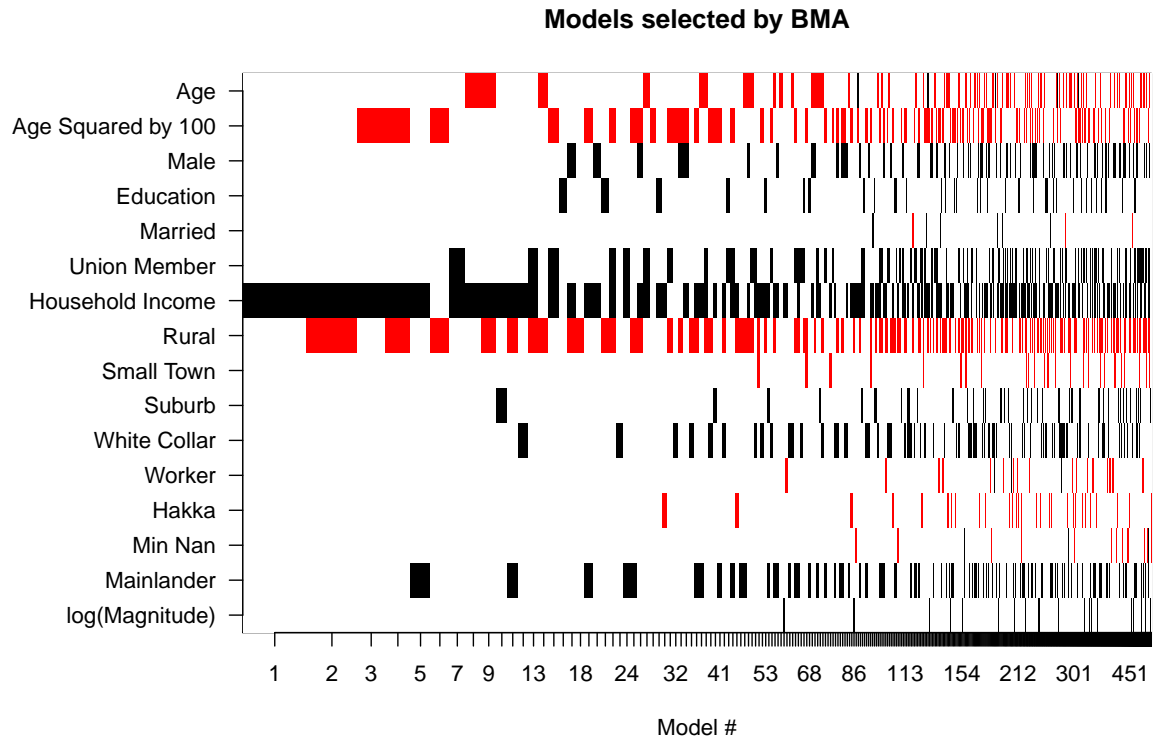


FIGURE 6: A figure showing the models of feeling close to a party considered by the BMA procedure. For each model, variables with positive effects are colored black, variables with negative effects are colored red/grey, and variables excluded from the model are white. The width of the area allotted to each model corresponds to the posterior probability of that model. This figure shows that the posterior probability of models including district magnitude is very low, even though its estimated effect when included tends to be positive.

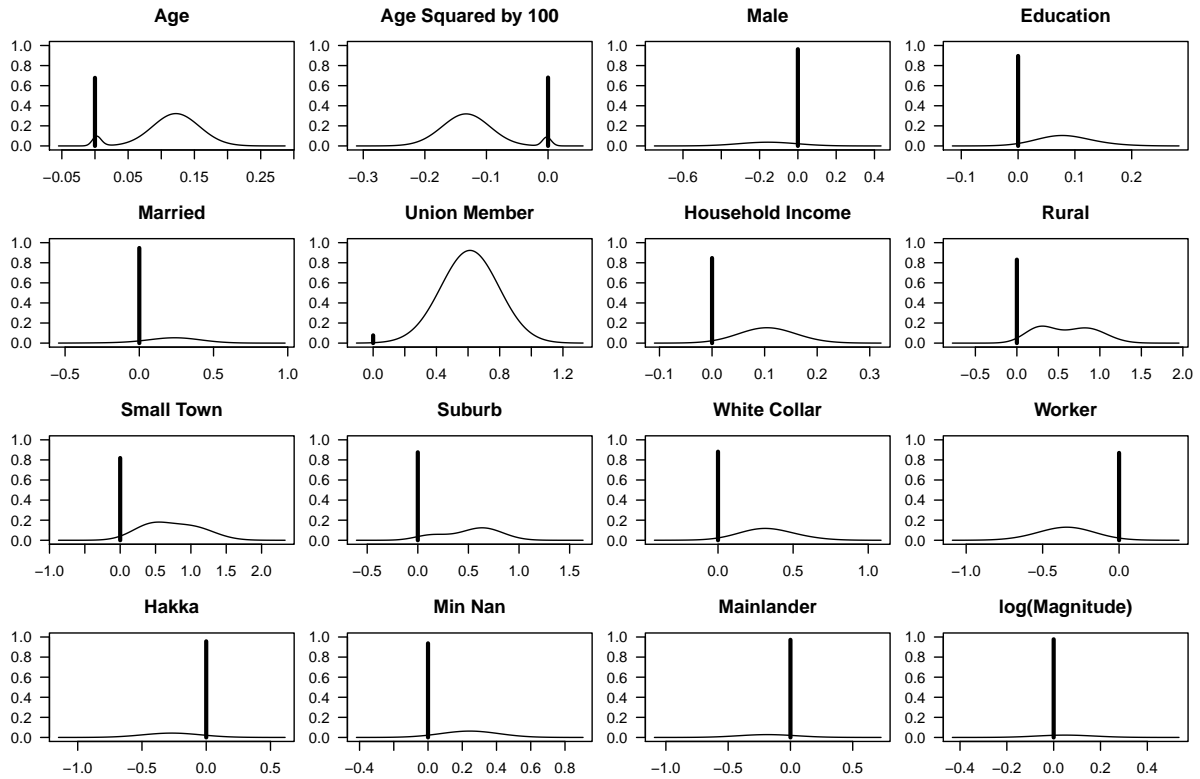


FIGURE 7: A figure showing the posterior densities of model coefficients for being contacted by a party based on the BMA procedure. Note that the height of the dark line indicates the posterior probability that the coefficient equals zero. The remaining density shows the distribution of coefficients in models that include the variable. Importantly, district magnitude almost certainly has no effect, as shown by the vertical line that nearly reaches one. Also, much of the posterior density is negative, indicating that while the effect is likely zero, it is almost certainly not positive.

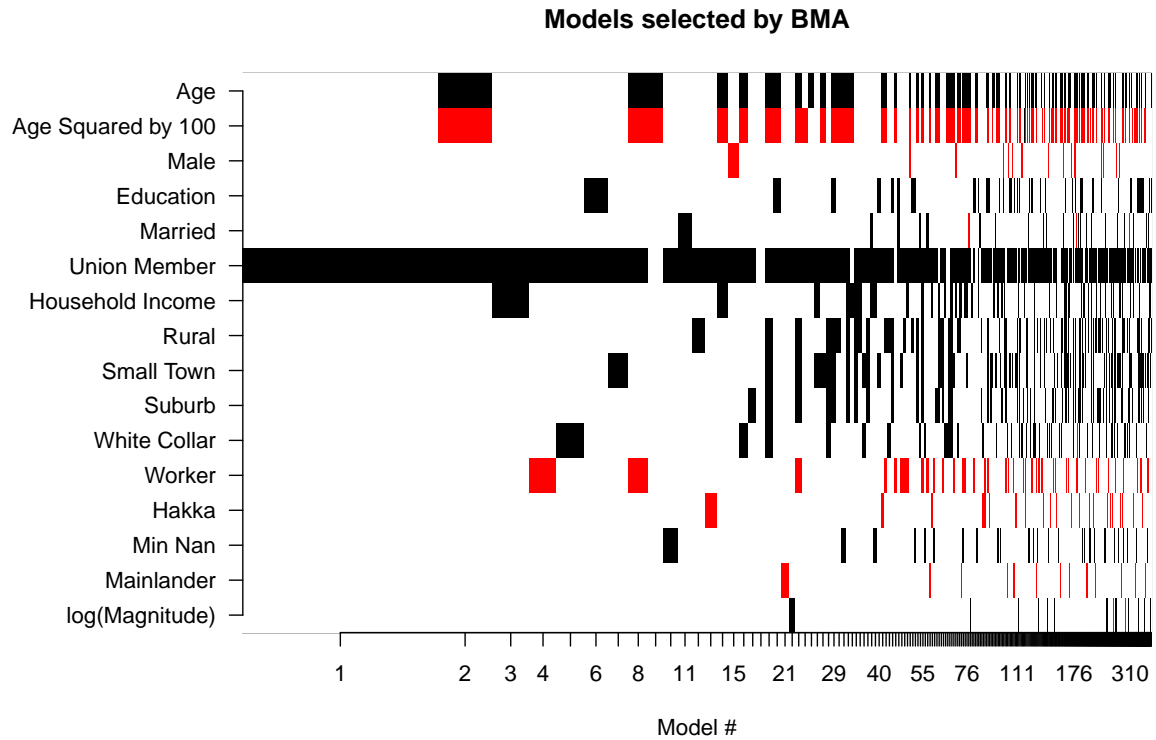


FIGURE 8: A figure showing the models of being contacted by a political party considered by the BMA procedure. For each model, variables with positive effects are colored black, variables with negative effects are colored red/grey, and variables excluded from the model are white. The width of the area allotted to each model corresponds to the posterior probability of that model. This figure shows that district magnitude has little effect on the probability of being contacted by a political party, since it is included in very few of the most likely models.

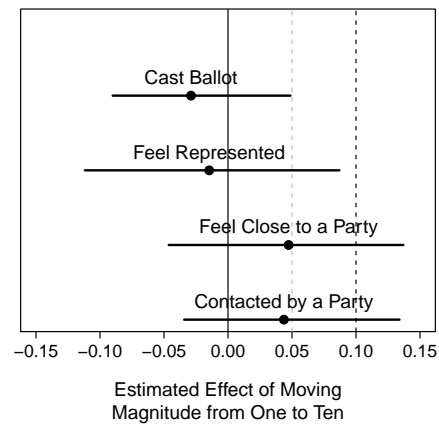


FIGURE 9: This figure shows the estimated effect of moving district magnitude from one to ten and 90% confidence intervals. Prior research suggests that this effect is about ten percentage points. Certainly an effect smaller than five percentage points is not substantively meaningful.