WORKING PAPER

Where's the Post in First-Past-the-Post (and Beyond)?

A Logical Model of the Effective District-Wide Threshold

Jack Bailey

Department of Politics, University of Manchester, Oxford Road, Manchester, UK

This version: August 22, 2025
Correspondence: jack.bailey@manchester.ac.uk

ABSTRACT I set out to answer a deceptively simple question: where does the "post" lie in first-past-the-post and other electoral systems? To this end, I present a new logical model of the *effective district-wide threshold*: the vote share at which any given party has a 50/50 chance of winning its first seat in a district. To derive my model, I rely only on simple combinatorics. The resulting expression suggests both that, all else being equal, we should expect parties to be more likely than not to win at least one seat in a district, and that the effective district-wide threshold depends only on the district magnitude and the number of vote-winning parties. Using district-level data from 198 elections in 60 countries, I then estimate a nonlinear Bayesian model that strongly corroborates my theoretical logic. As such, my model serves both to advance the broader Taageperan research agenda of interlocking electoral models and to unlock new quasi-experimental research designs that exploit as good as random variation around electoral thresholds.

1 INTRODUCTION

Democracies use elections to convert *the will of the people* into *political power*. Yet, different electoral systems can lead to different results, especially when it comes to electing smaller and less popular parties (Duverger 1964). One notable obstacle that many small parties face is the *electoral threshold*: the share of the vote that parties must receive in order to win a seat in the legislature. In some cases, these thresholds are explicit and written into the law. Spain, for instance, mandates that parties gain at least 3% of all votes in a district before they can win a seat. In other cases, however, thresholds are implicit instead and arise due to the mechanics of the electoral system itself (Rae, Hanby, and Loosemore 1971). Whatever the case, the conclusion remains much the same: electoral thresholds constrain party systems.

https://doi.org/10.33774/apsa-2025-l4d92-v3 **ORCID:** https://orcid.org/0000-0001-8517-5018 Content not peer-reviewed by APSA.

In this note, I present a new quantitative logical model (Taagepera 2008) of the *effective district-wide threshold*—the point at which any given party has a 50/50 chance of winning its first seat in a district. In effect, this provides an answer to a deceptively simple question: where does the post lie in first-past-the-post and other electoral systems? When compared to past attempts to model this quantity (Taagepera 2002, 1998; Lijphart 1994), my model has two clear advantages: it uses fewer assumptions or approximations, and out-competes possible alternatives by a wide margin. Consequently, we can be confident that it produces reliable predictions and that it generalises well to a range of electoral systems. As well as advancing the broader Taageperan research agenda of parsimonious and interlocking electoral models (for recent examples, see Hanretty 2022; Shugart and Taagepera 2017), the model that I present here also has broader implications for causal inference in electoral studies. In particular, it should allow future research to compute the propensity that any party will be assigned to a seat and to exploit the quasi-random and fuzzy assignment to seat-winning parties that occurs around effective electoral thresholds.

2 REPRESENTATION AND ELECTORAL THRESHOLDS

The term "first-past-the-post", though now synonymous with elections in single-member districts, derives from a wager common in horse-racing. Though now widespread and well-understood, it obscures as much as it reveals. Unlike in horse-racing, the electoral "posts" are not fixed: parties can win seats even if they have very different shares of the vote. What's more, "posts" are not exclusive to single-member districts: *all* electoral systems enforce this type of threshold. One obvious question, then, is where they tend to occur.

In any district-level election, parties face two clear electoral cut-offs (Rae, Hanby, and Loosemore 1971). The first—the *threshold of representation*, t_R—reflects the *smallest* vote share that a party could win while still having a chance of winning a seat in a district. For example, in a single-member district with four vote-winning parties, the threshold of representation occurs where all parties receive near-equal support, save for one that receives an additional vote (i.e. t_R = 25% + 1). The second—the *threshold of exclusion*, t_X—reflects the *largest* vote share that a party could win while still having the chance *not* to win a seat in a district. Using the previous example, the threshold of exclusion would occur where one of the four parties won just under 50% of the vote, since another party could potentially beat it (i.e. t_X = 50% – 1).

For any electoral district with a district magnitude of m and n_{V0} vote-winning parties, Taagepera and Shugart (1989, 276) show that, in all "simple" electoral systems,[1] the threshold of representation, t_R, must be greater than or equal to the lower threshold of the Hare formula:

$$t_R = \frac{1}{mn_{V0}} \tag{1}$$

[1] "Simple" electoral systems allocate all seats within districts, use categorical ballots, and take place in a single round (see Shugart and Taagepera 2017). So, for example, first-past-the-post and any districted system of proportional representation are simple, but mixed-member systems or ranked-choice voting are not.

https://doi.org/10.33774/apsa-2025-l4d92-v3 ORCID: https://orcid.org/0000-0001-8517-5018

Likewise, they also show that the threshold of exclusion, t_X, again in any district with a magnitude of m in a simple electoral system, must be less than or equal to the Droop quota:

$$t_X = \frac{1}{m+1} \tag{2}$$

We can think of these quantities as hard bounds on the *effective district-wide threshold*, t: the point at which any party has a 50/50 chance of winning their first seat in a district. Since the thresholds of representation and exclusion represent the lower and upper bounds on the effective district-wide threshold, respectively, we know that $t_R \leq t \leq t_X$. However, we still *do not* know exactly where t lies between them. At present, our best guess comes from Taagepera (2002), who models the effective district-wide threshold as follows:

$$t = \frac{0.75}{m+1} \tag{3}$$

Taagepera's model draws on earlier work by Lijphart (1994), who estimates the effective district-wide threshold by taking the arithmetic mean of the thresholds of representation and exclusion subject to two key assumptions: that the number of vote-winning parties, n_{V0}, equals the district magnitude, m,[2] and that the effective district-wide threshold lies exactly halfway between each of the bounds. This gives the following equation:

$$t = \frac{0.5}{m+1} + \frac{0.5}{2m} \tag{4}$$

Lijphart then notes that we can rewrite this expression as:

$$t = \frac{0.25\left(3 + \frac{1}{m}\right)}{m+1} \tag{5}$$

Which, if we omit $\frac{1}{m}$ in the numerator on the basis that it matters only for small values of m (see Lijphart 1994, 183–84), reduces to Taagepera's (2002) model as shown in Equation 3.

While Taagepera's (2002) work may represent our best guess of the effective district-wide threshold at present, his approach is not without criticism (see, for example, Bischoff 2009). One clear problem stems from the simplifying assumption that the number of vote-winning parties, n_{V0}, equals the district magnitude, m. To see why, consider a simple contest between two parties in a single-member district (i.e. $n_{V0} = 2$, $m = 1$). Here, both the threshold of representation and the threshold of exclusion equal exactly 50%. The effective district-wide threshold *must*, therefore, be 50%. Taagepera's equation, however, suggests that it is 37.5%. Indeed, wherever n_{V0} or m are both small, Taagepera's model's predictions will most likely be wrong. Furthermore, the assumption that we should weight the thresholds of representation, t_R, and exclusion, t_X, *equally* when computing the effective threshold, t, still remains to be tested and, as we will see in the section that follows, there is reason to believe that it is false.

[2]Though I prefer to preserve both the number of vote-winning parties and the district magnitude rather than assume they are equal, this is not an unreasonable assumption. As Shugart and Taagepera (2017) argue, we might expect around $n_{S0} + 1 = m^{1/2} + 1$ parties to contest a district with a magnitude of m.

3 MODELLING THE EFFECTIVE DISTRICT-WIDE THRESHOLD

The thresholds of representation, t_R, and exclusion, t_X, are nonlinear functions of the district magnitude, m, and the number of vote-winning parties, n_{V0}. Thus, any model of the effective district-wide threshold, t, must be nonlinear too. With this in mind, I follow a two-step approach inspired by Shugart and Taagepera (2017), who advocate in favour of theoretically informed, parsimonious, and interlocking nonlinear electoral models. In the first step, I use simple combinatorics to derive a model that predicts the effective district-wide threshold *in the absence of all other data*. In the second step, I then specify a corresponding nonlinear Bayesian model using the `brms` package (Bürkner 2017) in `R` to put my logic to the test.

3.1 THEORETICAL MODEL

To see how I derive my logical model, imagine a single electoral district. Yet, assume that you know *nothing* about the area that it covers, the people who live there, or their customs. Instead, assume that the *only* facts that you know are that it has a magnitude of m, n_{V0} vote-winning parties, and n_{V2} *effective* vote-winning parties. Given such scant information, what is the probability that any party won *at least one* seat in the district?

Simple combinatorics provide a potential solution. The key is to recognise that the probability that *at least one trial is a success* is exactly the same as the probability that *not all trials are failures*. In the absence of any other information, let us assume that all n_{V0} vote-winning parties had an equal share of the vote, such that each party had a $\frac{1}{n_{V0}}$ chance of winning any given seat. Thus, if there is a $\frac{1}{n_{V0}}$ chance that any party *did* win a seat, then there is a $1 - \frac{1}{n_{V0}}$ chance that it *did not*. The chance that a party failed to win *any* of the district's m seats must, therefore, be $\left(1 - \frac{1}{n_{V0}}\right)^m$ and, so, the chance that this *did not* happen, and that any given party won *at least one* seat in the district, must be:[3]

$$\Pr(m_i \geq 1) = 1 - \left(1 - \frac{1}{n_{V0}}\right)^m \tag{6}$$

The problem, of course, is that assuming all parties had an equal chance of winning a seat is totally unrealistic: at any election, some parties will garner very little support and others will be vanity projects. After all, the cost of participating in elections can often be surprisingly low. Interestingly, however, though this assumption may not hold true for the *absolute* number of parties, it does for the *effective* number instead. As Laakso and Taagepera (1979) note: "The effective number of parties is the number of hypothetical *equal*-size parties that would have the same total *effect* on fractionalization of the system as have the actual parties of *unequal* size" (4, emphasis in original). In other words, effective parties are equal in size *by design*. As such, we can simply substitute the effective number of parties, n_{V2}, into Equation 6 to get the probability that any party will win at least one seat in a district:

[3] One interesting implication of this equation is that, as the district magnitude, m, and the effective number of vote-winning parties, n_{V0}, each tend to ∞, the probability that any party will win at least one seat in a district will converge on $1 - \frac{1}{e} \approx 0.63$ (where $e \approx 2.71$ is Euler's number, a common mathematical constant).

$$\Pr(m_i \geq 1) = 1 - \left(1 - \frac{1}{n_{V2}}\right)^m \tag{7}$$

Now that we know the probability that any party gained representation, next consider the two bounding thresholds: the threshold of representation, t_R, and the threshold of exclusion, t_X. The threshold of exclusion, t_X, depends only on the district magnitude, m, so is not sensitive to party numbers. The threshold of representation, t_R, however, is a function of *both* the district magnitude, m, and the number of vote-winning parties, n_{V0}. Recall that it occurs where all vote-winning parties have an *equal* share of the vote, save for one that receives one additional vote. Typically, we ignore this single vote on the assumption that it contributes next to nothing at almost any realistic election. Thus, since all parties receive *exactly* the same share of the vote in this case, the *actual* and *effective* number of vote-winning parties are equal, suggesting that we can again substitute effective for absolute parties as follows:

$$t_R = \frac{1}{mn_{V2}} \tag{8}$$

Where only a single party contests an election, thresholds must remain undefined. After all, if no other party competes in a district the sole party running will win all of the seats without having to win any votes. Consequently, when we compute the threshold of representation, t_R, using *actual* parties, the largest value that it can take occurs where $n_{V0} = 2$ and $m = 1$, giving $t_R = \frac{1}{2}$. Note, however, that this is *not* the case when we use *effective* parties instead: since the effective number of vote-winning parties is not limited to only integer values, we can define the threshold of representation for *any* case where $n_{V2} > 1$.

The left-hand panel of Figure 1 shows the thresholds of representation and exclusion where the district magnitude is fixed such that $m = 1$ and where we allow the effective number of vote-winning parties, n_{V2}, to take all possible values between one and ten. By definition, parties are *guaranteed* to win a seat where they exceed the threshold of exclusion. However, note that, in this case, the threshold of representation, t_R, *exceeds* the threshold of exclusion, t_X, where the effective number of vote-winning parties takes any value lower than 2, implying that the point at which any given party first secures a chance to win a seat occurs at a *larger* vote share than the point at which it *must* win a seat.

Clearly, this situation makes little sense: the thresholds of representation, t_R, and exclusion, t_X, represent the lower and upper bounds on the chances of winning a seat, respectively. Accordingly, the former should *never* come to exceed the latter. Instead, the threshold of representation should *increase* as the effective number of vote-winning parties *decreases* up until it reaches the threshold of exclusion, at which point the two should remain equal in value. Thankfully, computing this new lower bound is straightforward: as the middle panel of Figure 1 shows, we need only take the minimum of the two thresholds, $t_{\min} = \min(t_R, t_X)$.

So far, we have determined the probability that any party will win at least one seat in a district (Equation 7) and have defined appropriate lower ($t_{\min}$, Equation 8) and upper (t_X, Equation 2) bounds on the effective district-wide threshold, t. We can now turn to the quantity

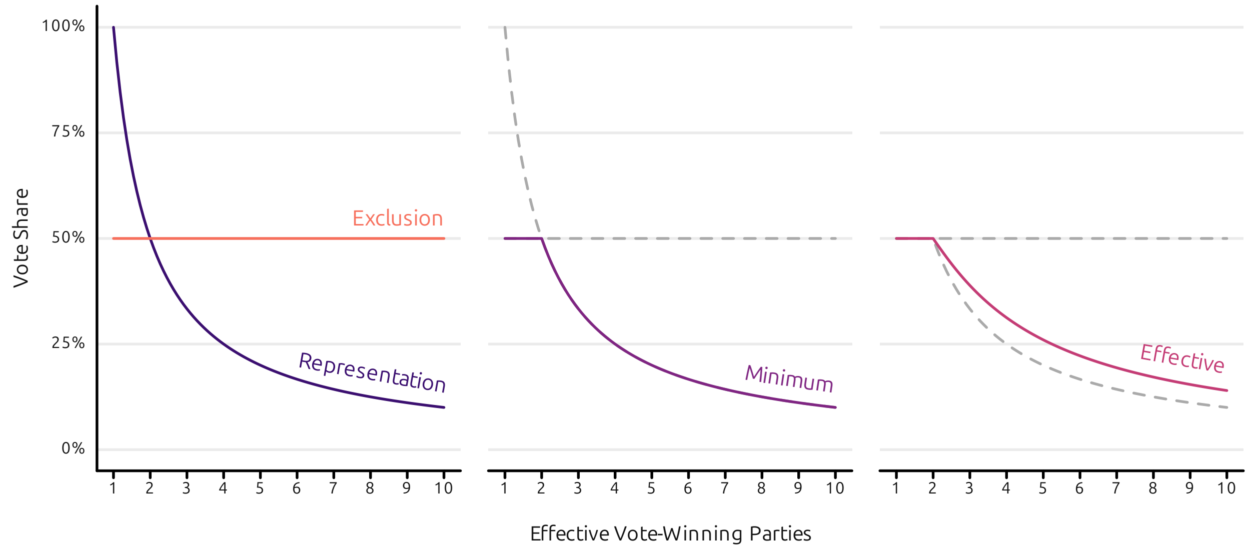

Figure 1: The thresholds of representation and exclusion (left panel) represent hard bounds on the effective district-wide threshold. However, the threshold of representation cannot exceed the threshold of exclusion, so we must take their minimum as the true lower bound (middle panel). Weighting these bounds according to the probability that any party will win at least one seat in the district, we get the effective district-wide threshold (right panel). Note that, in this case, the district magnitude is held constant at one.

itself. Recall that Lijphart (1994) and Taagepera (2002) assume that we must *estimate* this quantity by assuming that the bounding thresholds each have equal weight. However, so long as we know the district magnitude, m, and the effective number of vote-winning parties, n_{v0}, we can compute the *exact* probability that any party will win a seat in a district and use this to weight the thresholds instead. To avoid any visual clutter, let θ equal the probability that any party will win at least one seat in a district, as shown in Equation 7:

$$\theta = 1 - \left(1 - \frac{1}{n_{v2}}\right)^{m} \tag{9}$$

We can use this probability, θ, to weight the threshold of exclusion, t_x, since it represents the point at which a party is *guaranteed* to win its first seat in a district. Likewise, we can assign a weight of $1 - \theta$ to our lower bound, the minimum of the thresholds of representation and exclusion, $t_{\min}$, which represent the lowest possible vote share on which a party might win a seat. Importantly, since we know θ *exactly*, this implies that we can also know the effective district-wide threshold, t, *exactly* as well. Thus, we arrive at my logical model:[4]

$$t = (1 - \theta)\, t_{\min} + \theta t_x \tag{10}$$

$$\theta = 1 - \left(1 - \frac{1}{n_{v2}}\right)^{m} \tag{11}$$

[4]To make it as easy as possible for others to implement these models in their own research, I define a series of convenience functions in the R programming language in Section A in the appendix to this paper.

The right-hand panel of Figure 1 plots this quantity for the case where $m = 1$. As we can see, the effective district-wide threshold follows a similar path to the lower bounding threshold, $t_{\min}$, though does not have as steep a slope. Further, this slope suggests that, for any given value of n_{v2}, relatively little additional success is required to move from contention (the lower bound, $t_{\min}$) to an even chance of winning a seat (the effective district-wide threshold, t). After this point, however, the task then becomes much more difficult: to make up the remaining difference between an even chance (t) and total certainty (the upper bound, t_x) requires a substantial increase in support.

3.2 EMPIRICAL CORROBORATION

To test my logical model, I draw on district-level data from the Constituency-Level Elections Archive (Kollman et al. 2024). After removing errant cases—for example, where districts have a magnitude of zero—and subsetting to elections that used "simple" electoral systems (Shugart and Taagepera 2017)—i.e. first-past-the-post or districted proportional representation—I am left with 81,832 unique observations, each representing some party in some district at one of 198 elections in one of 68 different countries between 1946 and 2022.

To test my logic, we must model the probability that any party does ($w_i = 1$) or does not ($w_i = 0$) win at least one seat in a district. Since this variable is binary by nature, we can treat it as distributed according to a Bernoulli distribution with some probability π:

$$w_i \sim \text{Bernoulli}(\pi_i) \tag{12}$$

Next, we can model the log odds of π according to some slope parameter, β, that determines how fiercely to discriminate between cases that did and did not win a seat, and a threshold parameter, t, that estimates the effective threshold accordingly to each party's vote share, v_i:

$$w_i \sim \text{Bernoulli}(\pi_i) \tag{13}$$

$$\text{logit}(\pi_i) = \beta\,(v_i - t) \tag{14}$$

This model resembles a simple two-parameter logistic model common in item-response theoretic research. However, it can make impossible predictions. For example, where the threshold parameter, t, is low, which might occur in districts with large magnitudes and many vote-winning parties, the model would be able to predict that parties still have a chance of winning a seat even where their vote share is 0%. To account for this, we can put each party's vote share, v_i, and the model's estimate of the effective district-wide threshold, t, on the logit scale by taking their log odds. In effect, this forces the model to interpret vote shares of 0% and 100% as implying probabilities of winning at least one seat of 0% and 100%, respectively:

$$w_i \sim \text{Bernoulli}(\pi_i) \tag{15}$$

$$\text{logit}(\pi_i) = \beta\,[\text{logit}\,(v_i) - \text{logit}\,(t)] \tag{16}$$

Another limitation that the model faces is that it has only a single effective district-wide threshold. However, we are in the fortunate situation of knowing exactly what value this threshold should take given Equations 10 and 11. As such, we do not need to *estimate* the effective district-wide threshold at all. Instead, we can substitute in the equations:

$$w_i \sim \text{Bernoulli}(\pi_i) \tag{17}$$

$$\text{logit}(\pi_i) = \beta\,[\text{logit}\,(v_i) - \text{logit}\,(t_i)] \tag{18}$$

$$t_i = (1 - \theta)\,t_{\min} + \theta t_X \tag{19}$$

$$\theta = 1 - \left(1 - \frac{1}{n_{v2}}\right)^m \tag{20}$$

The one final change that we must make concerns the discrimination parameter, β. We know that this parameter must vary depending on the district magnitude, m, and the effective number of vote-winning parties, n_{v2}. After all, the bounding thresholds, t_R and t_X, differ from case to case as a function of these variables, so the value that β takes must change to accommodate the distance between them. The *smallest* distance between the bounding thresholds occurs where $n_{v2} \leq t_X$ and the two thresholds converge on a single value (e.g. where $m = 1$ and $n_{v2} = 2$, $t_{\min} = t_X = 0.5$). The *largest* distance, instead, occurs where $m = 1$ and n_{v2} tends to ∞, since the threshold of exclusion equals its maximum value of 50%, but the threshold of representation converges on 0%, producing a 50 percentage point distance.

This suggests that we should let the discrimination parameter, β, vary as a function of the distance between lower, $t_{\min}$, and upper, t_X, bounding thresholds. To do so, recall that, where $t_R \leq t_X$, both bounding thresholds are equal. Thus, the slope here *must* be infinite to allow for a discontinuous jump in the probability of winning a seat at the electoral threshold. Imagine some new parameter, λ, that represents this 100 percentage point change, though measured in units of logits. If we set $v_i = t_X$ and $t_i = t_{\min}$ to reflect the limiting cases where $n_{v2} \leq t_X$, and express λ in terms of β and the two bounding thresholds such that $\lambda = \beta\,[\text{logit}\,(t_X) - \text{logit}\,(t_{\min})]$, we can rearrange to solve for β such that:[5]

$$\beta = \frac{\lambda}{\text{logit}\,(t_X) - \text{logit}\,(t_{\min})} \tag{21}$$

[5]In all cases where $t_R \leq t_X$, a division by zero occurs. This prevents the model from fitting, since it evaluates to a value of infinity (`Inf`), thereby producing a singularity. Due to the model's functional form, these cases are all perfectly identified. In effect, they have an infinitely large amount of statistical power since we know the exact value that the effective district-wide threshold must take: $t = t_X = t_{\min}$. As such, I drop any cases where $t_R \leq t_X$ from my data to allow the model to sample reliably.

Table 1: My combinatorial model outperforms similar models that rely on Taagepera's (2002) and Lijphart's (1994) models of the effective district-wide threshold. Judging by the relative discrimination parameters, k-fold information criteria, and R^2 values, it better discriminates between parties that do and do not gain representation, makes more precise out-of-sample predictions, and explains more variation in data from the Constituency-Level Elections Archive.

	Discrimination, λ	K-Fold IC	R^2
Combinatorial	8.10 [7.87, 8.34]	10320.16	0.88
Taagepera (2002)	1.80 [1.77, 1.82]	21685.64	0.74
Lijphart (1994)	2.07 [2.04, 2.10]	20820.78	0.78
Observations			81,832
Elections			198
Countries			68

Finally, we can substitute this into Equations 17 to 20 above to get our empirical model:[6]

$$w_i \sim \text{Bernoulli}(\pi_i) \tag{22}$$

$$\text{logit}(\pi_i) = \beta\,[\text{logit}\,(v_i) - \text{logit}\,(t_i)] \tag{23}$$

$$\beta = \frac{\lambda}{\text{logit}\,(t_x) - \text{logit}\,(t_{\min})} \tag{24}$$

$$t_i = (1 - \theta)\,t_{\min} + \theta t_x \tag{25}$$

$$\theta = 1 - \left(1 - \frac{1}{n_{v2}}\right)^m \tag{26}$$

To determine this model's performance, I compare it to two possible competing models. These resemble Equations 22 to 26 exactly, though use Taagepera's (2002) and Lijphart's (1994) models to predict the effective district-wide threshold instead (shown in Equations 3 and 4, respectively). To evaluate how well each model performs, I test the models in two ways. First, I compute simple R^2 statistics to see how much of the variation they explain in the probability that a party will win at least one seat in a district. Second, I conduct k-fold cross-validation by reserving one tenth of all cases at random, refitting the model to the remaining data, testing the model's out-of-sample predictions, and then repeating this step ten times. This produces a 'k-fold information criterion' which we can interpret in the same way as we would the Akaike or Bayesian information criteria (i.e. lower values imply a better fit to the data).

Table 1 presents the results. First, note that my combinatorial model has a much larger discrimination parameter, λ, than the models that estimated the effective district-wide threshold using the equations set out in Taagepera (2002) or Lijphart (1994), implying that it is more capable of discriminating between parties that did or did not win a seat in the district. Next, consider the models' respective k-fold information criteria (further, note that the values themselves are not meaningful, only the relative difference between them). While the models based

[6]Note that I also include a weakly-informative Normal(0, 1) prior on β. For more information on the model's implementation, including all relevant code, see the replication package that accompanies this paper.

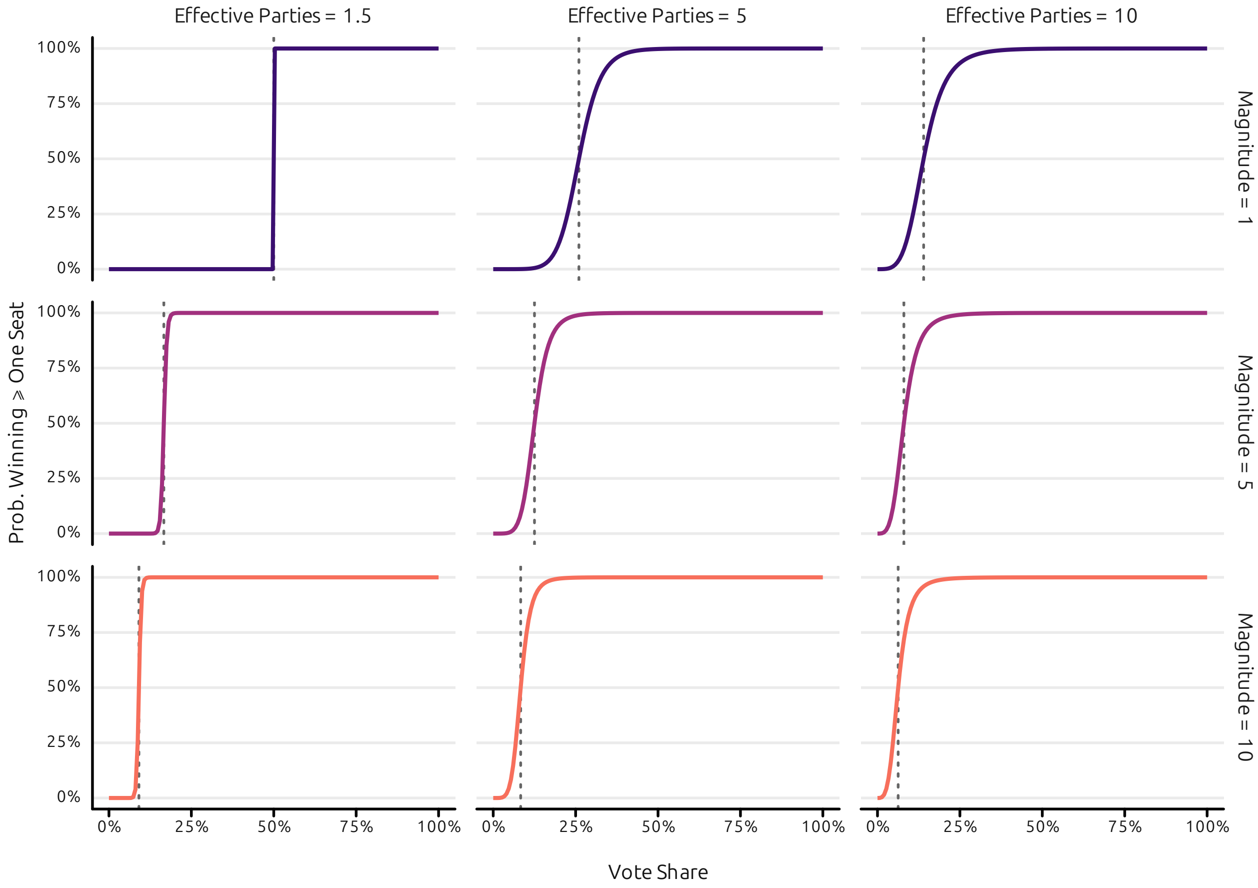

Figure 2: As the district magnitude and the effective number of vote-winning parties *increases*, the effective district-wide threshold (here shown as a dotted vertical line) *decreases* in turn. Likewise, where the district magnitude and the effective number of vote-winning parties are small, we are better able to discriminate between parties that did and did not win a seat. Though hard to perceive, predictions include a point estimate and a 95% credible interval.

on Taagepera (2002) or Lijphart (1994) have values of 21685.64 and 20820.78, respectively. By contrast, my combinatorial model has a score half the size at just 10320.16, implying that it has a much greater out-of-sample predictive accuracy. The same is true when we compare the models based on the amount of variation that they explain in the probability of winning a seat. The models based on Taagepera (2002) or Lijphart (1994) have R^2 values of 0.74 and 0.78, respectively, whereas my combinatorial model has an R^2 of 0.88 instead.

We now have compelling evidence that my combinatorial model out-competes both alternatives: it better discriminates between parties that do and do not gain representation, makes more precise predictions, and explains much more variation. Next, we must consider how it performs. Figure 2 shows predictions from the model where we let $m = 1, 5, 10$ and $n_{v2} = 1.5, 5, 10$. The model behaves much as we would expect. First, as the effective number of vote-winning parties *increases*, the effective district-wide threshold (shown here as a dotted vertical line) *decreases* to account for the growing distance between the two bounding thresholds. Second, the effective district-wide threshold also *decreases* as the district magnitude *grows*, though this change is smaller moving from $m = 1$ to $m = 5$ than from $m = 5$

to $m = 10$. Finally, the discrimination between parties that did and did not win a seat is more pronounced where there are fewer effective parties and lower magnitudes. In particular, where the bounding thresholds converge (top-left panel), the curve becomes a step function, allowing for perfect discrimination between parties that do and do not gain representation. So, the effective threshold is not a sharp discontinuity, but neither is it overly fuzzy.

4 CONCLUSION

I set out to answer a seemingly simple question: where does the "post" lie in first-past-the-post and other "simple" electoral systems? To this end, I derived a logical model of the relationship between a district's magnitude, m; the number of vote-winning parties that it contains, n_{v0}; and its effective district-wide threshold, t: the point at which any given party has a 50/50 chance of winning its first seat in the district. I then specified several competing nonlinear Bayesian models that reflected my theory and competing alternatives and showed that, in all cases, my model was better able to discriminate between cases where parties do and do not win a seat, makes more precise out-of-sample predictions, and explains a greater proportion of the variation present in the data.

Equation 7, which presents a logical model of the probability that any party will gain representation, and Equations 10 to 11, which present a logical model of the effective district-wide threshold, each serve to advance the broader Taageperan research agenda by providing more parsimonious and interlocking predictive models of parameters inherent to electoral systems. Logically, future research should proceed in two directions. First, it should seek to establish equivalent quantities at the national level. Second, it should use the interlocking nature of these models to explore how factors that predict district magnitudes and the number of vote-winning parties affect party systems by altering the probability of success and the placement of electoral thresholds.

These models and the model that I describe in Equations 22 to 26, which describes the probability of winning representation conditional on the district magnitude and effective number of vote-winning parties for any given vote share, also have the potential to unlock new causal designs. For example, future research might use the latter model to compute propensity scores that they can use to recover causal effects. And, though these methods are somewhat maligned since we almost never know the true propensity score given a pattern of covariates (King and Nielsen 2019), that is not true in this case, since the true effective threshold is determined exactly by only n_{v2} and m. Similarly, much like recent research by Matsunaga and Winzen (2024), Valentim and Dinas (2024), and others who use Taagepera's (2002) model to exploit the quasi-random assignment to seat-winning parties that occurs around electoral thresholds, future research might use the thresholds that I estimate to specify new fuzzy regression discontinuity designs that explain the impact that assigning new parties to a district has on other electoral factors.

5 REFERENCES

Bischoff, Carina S. 2009. "National Level Electoral Thresholds: Problems and Solutions." *Electoral Studies* 28 (2): 232–39. https://doi.org/10.1016/j.electstud.2008.09.003.

Bürkner, Paul-Christian. 2017. "Brms: An R Package for Bayesian Multilevel Models Using Stan." *Journal of Statistical Software* 80 (1): 1–28. https://doi.org/10.18637/jss.v080.i01.

Duverger, Maurice, ed. 1964. *Political Parties: Their Organization and Activity in the Modern State*. University Paperbacks 82. London, UK: Methuen.

Hanretty, Chris. 2022. "Party System Polarization and the Effective Number of Parties." *Electoral Studies* 76 (April): 102459. https://doi.org/10.1016/j.electstud.2022.102459.

King, Gary, and Richard Nielsen. 2019. "Why Propensity Scores Should Not Be Used for Matching." *Political Analysis* 27 (4): 435–54. https://doi.org/10.1017/pan.2019.11.

Kollman, Ken, Allen Hicken, Daniele Caramani, David Backer, and David Lublin. 2024. "Constituency-Level Elections Archive." Ann Arbor, MI.

Laakso, Markku, and Rein Taagepera. 1979. "'Effective' Number of Parties: A Measure with Application to West Europe." *Comparative Political Studies* 12 (1): 3–27. https://doi.org/10.1177/001041407901200101.

Lijphart, Arend. 1994. *Electoral Systems and Party Systems: A Study of Twenty-Seven Democracies, 1945-1990*. Oxford ; New York: Oxford University Press.

Matsunaga, Miku, and Thomas Winzen. 2024. "Strengthening Mainstream Consensus? The Effect of Radical Right Populist Parties on the Defense Policies of Left Parties." *Political Science Research and Methods*, October, 1–16. https://doi.org/10.1017/psrm.2024.40.

Rae, Douglas, Victor Hanby, and John Loosemore. 1971. "Thresholds of Representation and Thresholds of Exclusion: An Analytic Note on Electoral Systems." *Comparative Political Studies* 3 (4): 479–88. https://doi.org/10.1177/001041407100300406.

Shugart, Matthew S., and Rein Taagepera. 2017. *Votes from Seats: Logical Models of Electoral Systems*. Cambridge, UK: Cambridge University Press.

Taagepera, Rein. 1998. "Effective Magnitude and Effective Threshold." *Electoral Studies* 17 (4): 393–404. https://doi.org/10.1016/S0261-3794(97)00053-X.

———. 2002. "Nationwide Threshold of Representation." *Electoral Studies* 21 (3): 383–401. https://doi.org/10.1016/S0261-3794(00)00045-7.

———. 2008. *Making Social Sciences More Scientific: The Need for Predictive Models*. Oxford, UK: Oxford University Press.

Taagepera, Rein, and Matthew Soberg Shugart. 1989. *Seats and Votes: The Effects and Determinants of Electoral Systems*. New Haven, Conn.: Yale University Press.

Valentim, Vicente, and Elias Dinas. 2024. "Does Party-System Fragmentation Affect the Quality of Democracy?" *British Journal of Political Science* 54 (1): 152–78. https://doi.org/10.1017/s0007123423000157.

A R FUNCTIONS

```r
# Convert probabilities to logits

logit <- function(p){ log(p) - log1p(-p) }


# Convert logits to probabilities

inv_logit <- function(x){ 1 / ( 1 + exp(-x) ) }


# Compute the threshold of representation

tR <- function(m, n){ 1 / ( m * n ) }


# Compute the threshold of exclusion

tX <- function(m){ 1 / ( m + 1 ) }


# Compute the probability of winning at least one seat

prob_rep <- function(m, n){ 1 - ( 1 - 1 / n )^ m }


# Compute the effective district-wide threshold

effective_threshold <-
  function(m, n){

    ( 1 - prob_rep(m, n) ) * min( tR(m, n), tX(m) ) + prob_rep(m, n) * tX(m)

  }
```


https://doi.org/10.33774/apsa-2025-l4d92-v3 **ORCID:** https://orcid.org/0000-0001-8517-5018

```r
# Compute propensity of winning at least one seat in a district

seat_propensity <-
  function(v, m, n){

    # Compute beta

    b <- 8 / ( logit( tX(m) ) - logit( min( tR(m, n), tX(m) ) ) )


    # Compute effective threshold

    t <- effective_threshold(m, n)


    # Return propensity score

    inv_logit( b * ( logit(v) - logit(t) ) )

  }
```


https://doi.org/10.33774/apsa-2025-l4d92-v3 **ORCID:** https://orcid.org/0000-0001-8517-5018