UC Irvine

UC Irvine Previously Published Works

Title

How to Operationalize Two-Partyness

Permalink

<https://escholarship.org/uc/item/9jc9800w>

Journal

Journal of Elections, Public Opinion and Parties, 23(4)

ISSN

1745-7289

Authors

Gaines, BJ

Taagepera, R

Publication Date

2013-11-01

DOI

10.1080/17457289.2013.770398

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at

<https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

How to Operationalize Two-Partyiness

BRIAN J. GAINES* & REIN TAAGEPERA**

*University of Illinois, USA; **UC Irvine, USA

ABSTRACT *Because of the prominence of Duverger's law, a great deal of work has focused on assessing whether or not given election outcomes can be regarded as exhibiting two-party competition. The most common metric for assessment is whether the effective number of parties is close to two. We introduce two statistics that better measure conformity to two-partyiness, briefly survey their logical properties, and then demonstrate their utility with data on parliamentary elections in England.*

Introduction

“The simple majority single ballot system favors the two-party system” according to a claim that, almost uniquely in political science, is dubbed a “law” (Duverger, 1954: 217). Indeed, it is “perhaps the most famous theoretical generalization in political science” (Grofman et al., 2009b). Causality may work in the opposite direction too, if a founding body with only two groupings picks the single member plurality (First-Past-The-Post, FPTP) rule, so as to keep out new competitors (Benoit, 2002, 2004; Boix, 1999; Colomer, 2005). Hence, Duverger's law can be restated without causal direction: “Seat allocation by FPTP tends to go with two major parties” (Taagepera, 2007: 103).¹

But what, exactly, is a two-party system? That component of Duverger's famous claim is usually taken to mean two parties having roughly equal vote shares, with any other parties having only negligible support. The ideal case is thus 50-50(-0).² But which is closer to the ideal, 65-35-0 or 45-40-15? Is either sufficiently close to be regarded as a two-party constellation, in the event that we treat two-party to be a dichotomous standard? Does the term “system” connote national rather than district-level results?

Hundreds of studies have been published on Duverger's law, including many that attempt to judge whether or to what degree given outcomes agree with the law. The conclusions have been qualitative, impressionistic, and often hedged, because there has never been a fully quantitative operationalization of “two-party system.” Frequently, the Laakso-Taagepera (1979) effective number of parties (N) is invoked,

Correspondence Address: Brian J. Gaines, University of Illinois, Political Science & IGPA, 1007 W Nevada, Urbana, 61801 United States. Email: bjgaines@illinois.edu

with assessment resting on whether N is sufficiently close to 2. We argue below that N , the workhorse of this literature, is often misleading when one's goal is to assess conformity to two-way competition. We introduce two alternative statistics that seem preferable on logical grounds, and, incidentally, seem better to match intuition of experts. We do not provide a precise statistical test for two-partyity, and retain the practice of making impressionistic judgments based on somewhat arbitrary thresholds. But given the popularity of trying to distinguish two- from one- and more-than-two-party competition, statistics designed expressly to measure that criterion represent a useful advance just the same.

The Effective Number of Parties: Popularity and Reservations

The Laakso-Taagepera (1979) effective number of parties (N) is defined as the inverse of the sum of squared fractional shares of components, such as votes or seats. For k shares $p_1, p_2, \dots, p_k$ that sum to 1, $N = (p_1^2 + \dots + p_k^2)^{-1}$. It is the inverse of the Herfindahl-Hirschman index, and is by far the most common entropy-related measure employed in electoral studies. It can also be expressed as a simple function of the mean and variance of these components, $N = (\mu + k\sigma^2)^{-1}$ where k is the number of components, $\mu = 1/k$, and σ^2 is the population variance (Feld & Grofman, 2007). N might be the most frequently used index in electoral and party studies – possibly in all political science – because scholars often wish to summarize compositional data, de-emphasizing the smaller components.³ The effective number is useful for many purposes, but not for all. Its strengths as well as limitations have been reviewed by one of its co-inventors in Taagepera (2007: 47–64), where it is compared to various alternatives. As a measure of two-partyity, $N = 2.0$ can mislead.

Indeed, while N is ubiquitous in work on Duverger's law, so too are expressions of reservations about its use. In the recent volume *Duverger's Law of Plurality Voting* (Grofman et al., 2009a), six chapters out of nine invoke this effective number. But the very first chapter points out “an important issue of measurement: An effective number of parties near 2 need not tell us that the fundamental mechanical and psychological logic underpinning Duverger's law is working well over time” (Grofman et al., 2009b).

Moreover, quite apart from its silence on mechanisms, an approach based on N falls short of capturing two-partyity. First, $N \approx 2.0$ can result not only from 50-50 but also from patterns suggesting one-party dominance, such as 66-17-17 ($N = 2.03$). In the other direction, N sometimes seems to exaggerate the degree of departure, as with 45-40-15 ($N = 2.60$). In the aforementioned volume, Gaines (2009: 125) points out that two cases seeming very different in strategic terms, 49-38-13 and 60-10-10-10-10, map to N values of 2.49 and 2.50, respectively, and concludes, “Notwithstanding the popularity of counting effective parties, information is unavoidably discarded in collapsing the vote-share vector, and it remains unclear exactly what ‘two-party competition’ should mean once one makes the move to weighting candidates.” Myriad existing studies have made similar points, before proceeding with inspection of deviation from the neighborhood of $N = 2$ just the same.

Apart from worries about the suitability of the Laakso-Taagepera measure, disagreement reigns about where it should be applied: the stages in an election (votes or seats), the level considered (district or nationwide), and the temporal extent (one election or long-term averages). This point is again exemplified in Grofman et al., who first say that Duverger introduced his law in terms of “the number of political parties whose representatives contest election *in the constituency*” (2009b: 2, emphasis added), but then assert, “In Duverger’s original formulation the ‘number’ of political parties is simply n_s , the number of parties whose representatives are elected” (2009b: 6). So is the right focus district votes or nationwide seats?

In the absence of consensus on which or how many cells of this $2 \times 2 \times 2$ matrix are appropriate venues for testing Duverger’s law, we focus hereafter on votes at the district level, one election at a time. The measures we propose might be applied differently (e.g. to long-term averages of seat shares), but we think the two-party standard is clearest and most defensible when applied to votes won in individual races. We now turn to operationalization.

Two Novel Indices

Consider first the vote shares (proportions) in one single-member-plurality electoral district wherein k parties compete. We have $p_1 + p_2 + p_3 + \dots + p_k = 1$, and we allocate the party labels so that $p_1 \geq p_2 \geq \dots \geq p_k$. Because the shares sum to 1, there are only $k-1$ independent variables. The effective number is bounded between 1 (Party 1 wins all votes) and k (a k -way tie).

We assume absence of any further knowledge about the parties, such as their possible affinities on ideological or historical grounds. This decision is the basis for all general indices; a party is a party is a party. If more information about the parties is available, it can be incorporated into analysis in some manner, of course, but a general metric for counting should not impose unnecessary assumptions or conflate variables.

We also note that not all candidates carry party affiliations, and some systems permit single individuals to stand as the candidate for more than one party. In a single district contest, whether candidates represent no party, one party, or more than one party is unimportant for counting purposes. As soon as one aggregates across districts or considers sets of districts, however, independent, unaffiliated and multiply affiliated candidates complicate assessment of an election’s fit to “two-party” competition. We return to this distinction in the empirical section below, but proceed now by treating “party” and “candidate” as synonyms, mindful that this assumption is sometimes incorrect.

When one asserts that FPTP electoral systems tend to go with two-party systems, what does “two-party” mean? First, the third, fourth, etc. parties should be of negligible size, compared to the two largest: $p_i > > p_j$ for $i \in \{1, 2\}$, $j \in \{3, \dots, k\}$. Second, the two largest shares should be rather evenly balanced—not $p_1 > > p_2$ but $p_1 \approx p_2$. Finally, there should not be so many parties winning votes that the top two can be large compared to any given other, but not large as compared to the aggregate vote

won by all other opponents, collectively. We take these relationships to be the basic definition of two-party system in Duverger's law:

$$p_1 \approx p_2 > p_j, \text{ and } p_1 + p_2 > \sum_j p_j \text{ for } j \in \{3, \dots, k\} \quad (1)$$

These relationships are also what Gary Cox (1997) terms a Duvergerian equilibrium. The ideal two-party constellation is 50-50, which precisely satisfies them. The outcome 42-42-16 is weaker on both $p_1 > p_3$ and $p_2 > p_3$, whereas 40-40-5-5-5-5 is a worse fit to $p_1 + p_2 > \sum_j p_j$.

We now wish to translate these desiderata into an index of two-partyness. It should satisfy the following requirements:

- When there are only two parties, and they are equal in size ($p_1 = p_2 = 0.5$), the index must take its maximum value, 1.
- When second and third parties are equal ($p_2 = p_3$), it should take the minimum value, 0.
- The index should decrease as the vote share captured by third through k th parties rises.

The second claim might require elaboration. We certainly should assign a score indicating a complete lack of two-partyness (0) for $p_1 = p_2 = p_3 = 1/3$, because it is a precise three-party constellation. Likewise, we want 0 when $p_1 = 1$ and $p_2 = p_3 = 0$, a clear one-party constellation. What about 50-25-25? It has both three-party and one-party aspects, as it has clearly more than two parties, one of which is dominant. On account of both aspects, 50-25-25 should score 0. There is no two-partyness: to assume similarity to 50-50, which of the two smaller vote shares should we ignore, and which one should we double? It is an instance of what Cox (1997) calls a non-Duvergerian equilibrium. But if 0 is apt for 50-25-25, then it should also apply to all $p_2 = p_3$ constellations, because they are intermediaries between 50-25-25 and either the pure one-party case of 100-0-0 or the pure three-party case of 33-33-33. Parallel reasoning applies as we consider cases with additional parties, such as 40-25-25-10, which is intermediate between 70-10-10-10 and 30-30-30-10.

As for the third claim, there are two issues to weigh: how much of the vote is captured by parties other than the top two, and how that vote is distributed. Comparing 50-50, 40-40-20 and 36-36-28, we have already noted that fit with two-partyness worsens as p_3 rises. But what about 40-40-10-10, 40-40-1, ..., 1 (denote this result 40-40-1 $\times$ 20) and 36-36-18-10 or 36-36-1 $\times$ 28? For any given two-way tie at the top, as the residual vote rises, two-partyness diminishes, but as it scatters more widely, the top-two finishers' shares rise relative to those of the other competitors, taken one at a time, offsetting the sense of worsening fit to bipartism somewhat.

Return to $p_1 \approx p_2 > p_3$, momentarily restricting attention to $k=3$. One measure of deviation from the requirement $p_2 > p_3$ is the ratio p_3/p_2 , which is 0 when $p_3=0$,

and is 1 when $p_2=p_3$. The converse measure of *agreement* with $p_2 > p_3$ is $1-p_3/p_2$, which can be expressed as the ratio $(p_2-p_3)/p_2$, equal to 1 when $p_3=0$ and to 0 when $p_2=p_3$. A measure of agreement with $p_1 \approx p_2$ is the ratio p_2/p_1 , which is 1 when $p_2=p_1$, and is 0 when $p_1=1$ and $p_2=p_3=0$. To reach its maximum value of 1, our index must maximize both ratios, hence one possible two-partyness index is simply the product of $(p_2-p_3)/p_2$ and p_2/p_1 . With p_2 cancelling out, the result is:

$$I = (p_2 - p_3) / p_1 \quad (2)$$

We have $I=1$ only for 50-50-0. When $p_2=p_3$, $I=0$, regardless of p_1 . Whenever $p_3=0$, the expression simplifies into $I=p_2/p_1$. Whenever $p_2=p_1$, the expression simplifies into $I=(p_2-p_3)/p_2$.

However, I assigns identical scores of $2/3$ to 42-42-14-2, 39-39-13-9, 30-30-10 $\times$ 4 and 30-30-10-1 $\times$ 30, defying the criteria that p_1 and p_2 ought to be large relative not only to p_3 but also as compared to the total vote won by all others. A simple way to deflate I accordingly is to multiply by the sum p_1+p_2 . While the resulting index makes no explicit use of p_4 through p_k , the incorporation of both $p_2 > p_3$ and $p_1+p_2 > \sum_{j \neq 1} p_j$ ensures that any outcome with large surplus vote is assigned an appropriately reduced two-party score.

$$T = (p_2 - p_3)(p_1 + p_2) / p_1 \quad (3)$$

In the examples offered previously, 45-40-15 leads to $N \approx 2.60$ and $T \approx 0.47$, while 65-35 has $N \approx 1.83$ and $T \approx 0.54$. The T metric thus judges them as fairly similar intermediate cases, whereas N more strongly favors the latter for two-partyness. The cases of 60-10-10-10-10 and 49-38-13 nearly match in N , but their T scores are 0 and 0.44. Clearly N and T can differ greatly. While this T is obviously not the only index that can be constructed on the basis of the requirements specified by equation (1), it is a reasonably simple way to cover all the required relationships.⁴

The index T takes into account the vote won by losers other than the top two, in deflating I , but it discards all information about the distribution of those votes. Hence, 30-30-10 $\times$ 4 and 30-30-10-1 $\times$ 30 are scored identically. An alternative, that more closely resembles the effective number of parties in using all components, but, like T , is formulated expressly as a measure of conformity with two-party competition is Euclidean distance in k -space from an exact two-way tie:

$$d_2 = \sqrt{(0.5 - p_1)^2 + (0.5 - p_2)^2 + (0 - p_3)^2 + \dots + (0 - p_k)^2} \quad (4)$$

Since the maximum distance from (0.5,0.5,0,...) in a vote-share space of any dimension is that to (1,0,...), it can be convenient to compute a slight variation, $D_2 = 1 - d_2 / \max(d_2) = 1 - \sqrt{2}d_2$, which necessarily falls between 0 and 1, with higher values indicating better conformity, so that it is aligned with T .

The observation that, no matter how many parties are competing, the furthest point from two-way competition is one-party monopoly might worry some about face validity of D_2 . Should not (perfect) four-party competition be scored as being twice as different from two-party competition as are either one- or three-party competition? Interval and ratio properties have intuitive appeal, and effective-party metric obviously retains them in precisely that sense. However, distances also have ratio properties, and the fact that effective-party scoring is non-linear in geometric distance could just as easily be seen as a problem with N , rather than a problem with relying on geometry. Most people have poor intuition for space of more than three dimensions, but the asymmetry in how D_2 penalizes departure from two-party competition by way of concentration of votes in first place (i.e. movement towards 100-0-...) and departures by way of dispersion to extra parties (e.g. movement towards a k -way even split) is not a mistake or quirk, but a true reflection of geometry.

Subjective Assessments, N , T and D_2

Even before the effective number of parties was devised, Duverger's law was accepted without explicit operationalizing of "two-party system," because of a tacit assumption that experts recognize a two-party constellation when they see one. As a rough test of this assumption (and of its agreement with use of the effective number), we asked a few colleagues who have dealt with Duverger's law to assess various single-member-district three-way vote constellations for two-partyness. The set included all 44 possible combinations of three vote shares, at 5% intervals, from 50-50-0 down to 35-35-30 and up to 100-0-0, and the response options were "yes, two-party" (scored as 1), "no" (0) and "hard to tell" (scored as 0.5) (see Appendix 1 for details).

The eight participants gave ratings that had mean scores across the 44 scenarios varying from 8 to 24.5, with an overall mean of about 16. The top panel of Figure 1 shows the mean ratings of constellations graphed against their effective numbers. The horizontal line at 0.5 marks the value we assigned to the response "hard to tell." While no constellations with $N < 1.7$ or $N > 2.6$ won majority support as two-party, those in between generated disparate reactions. Some were universally accepted (e.g. 50-50), and others overwhelmingly rejected (e.g. 65-25-10, with $N=2.02$, produced six no votes, one hard-to-tell, and only one yes). There is no range of T values that identifies the cases our subjects collectively regarded as most clearly two-party.

The middle panel of Figure 1 plots the average expert ratings against the index T . While some scatter remains, the association is much clearer. Constellations with T above 0.50 all earned large majorities of yes votes, while those below $T=0.3$ were rejected by most evaluators. Most of the "hard to tell" responses went to vote combinations in between, those having T values in the range $0.3 < T < 0.5$. Finally, the bottom panel reveals that D_2 also partitions the cases fairly well: only three cases scored 0.5 or higher while $D_2 \geq 0.7$.

It would certainly take a larger sample to make more precise and systematic claims about the mental images of experienced researchers and political practitioners

Figure 1. Subjective assessment of three-way vote splits for two-partyness, by N , T and D_2 . *Note:* Vertical-axis values are mean scores for eight scholars' agreement that a three-way vote-share vector having the given N , T , or D_2 score represents two-party competition, from 0=no, $\frac{1}{2}$ =hard to say, 1=yes. See Appendix 1.

regarding the “two-party” standard. Our coarse inquiry, however, suffices to show that N does not appear to operationalize intuition well, and that there are appreciable differences in the mental images of at least some researchers in the field. Because we do not all converge when classifying cases for “two-party” competition, better operational measures of two-partyness should improve clarity of communication.

Direct Comparison of N , T and D_2

To illustrate more thoroughly how these three measures compare, Figure 2 displays N versus T (top), N versus D_2 (middle) and T versus D_2 (bottom) for competition among up to five parties. We superimpose thresholds of 0.5 for T and 0.7 for D_2 without thereby implying that these particular arbitrary values are the only reasonable choices for cutoffs. Likewise, we mark the interval between $N=1.5$ and $N=2.5$ as a (generous) interpretation of N being roughly equal to 2. Hence, the distinct regions created by the dashed lines correspond to passing or failing the relevant two-party test for the particular index, and regions where these test disagree (e.g. not-two-party by N , but two-party by T) are of particular interest.

Figure 2. N , T and D_2 compared.

The top panel shows the logically possible range for T and N given at most five competitors. Comparing $T \geq 0.5$ and $1.5 \leq N \leq 2.5$, there are five non-empty regions. Both approaches reject designating as two-party all outcomes in the bottom left, where one party dominates, and bottom right, where support is widely dispersed. Both accept the top middle region, where outcomes most closely match 50-50-0-0-0. The remaining differences are instructive. On the right frontier, one finds all ties between the first- and second-placed parties where the residual vote divides evenly. While 41-41-6-6-6 yields $N \approx 2.88$, its T score of 0.77 suggests a reasonably close approximation of two-way competition, appropriately we think. As the race at the top becomes less evenly balanced, T and N both fall, until (58,24,6,6,6), which lies just inside the $1.5 \leq N \leq 2.5$ zone, but falls very short of $T=0.5$. If we instead fix the first two shares, but allow the residual vote to concentrate, we get (41,41,18,0,0), with $T=0.46$ and $N=2.71$. The shrinkage in T follows the intuition that p_2 ought to be substantially larger than p_3 . With a threshold of $T=0.5$, the highest N for an outcome (with integer vote shares) that T treats as two-party is 3.63, for (35,35,10,10,10). As votes shift from p_2 to p_1 , we eventually reach (60,10,10,10,10), with $T=0$ but a surprising N score of 2.5. Surely, however, the former result is a better (albeit far from perfect) example of two-way competition.

The middle panel shows N plotted against D_2 for all five-part vote tuples consisting of integers (i.e. the 46,262 five-component partitions of the integer 100). It loosely resembles the top panel, and, again, the most important contrasts are those cases on which the two statistics (for the given thresholds) disagree. As illustrated by (81,5,5,5,4) and (61,10,10,10,9), there are cases with one dominant party that appear to be very far from the two-party ideal but that nonetheless fall into the $1.5 \leq N \leq 2.5$ region. Meanwhile, some cases with N values above 3 meet the standard of $D_2 \geq 0.7$, as (39,38,8,8,7) demonstrates.

Finally, the bottom panel of Figure 2 plots T versus D_2 , again for integer-vote-share five-tuples. While the two measures will generally agree, for any given choice of thresholds, there will be regions with cases rejected by one standard and accepted by the other. The two points labeled with vote shares provide some sense of which cases are ambiguous or borderline in this sense of causing a discrepancy between T and D_2 . Comparatively large third-place vote shares are more penalized by T , while D_2 is comparatively more sensitive to large third-through-fifth totals (more generally, third through k th). Given these definitions, there is no way to choose thresholds for both indices that agree perfectly: if one chooses values to prevent T -based assessment suggesting two-partyness while D_2 does not (i.e. no data in the bottom right region), there will necessarily be logically possible cases in the top left, where D_2 values pass the threshold while T do not. Empirically, logically possible cases of each kind might or might not exist, and the UK data examined below show that it is sometimes not difficult to choose (sensible) thresholds that leave one of the four regions mostly empty.

In sum, Figure 2 shows that the effective number of parties, useful for many other purposes, is a questionable metric for gauging two-partyness as posited by Duverger's law. Even close agreement with $N=2.0$ does not necessarily signify

two-party competition.⁵ One might prefer either T or D_2 as an alternative measure. We think two-party is a flexible enough concept that neither index is unambiguously superior, and, indeed, examination of both can be useful. But to test to what degree competition in a given FPTP district fits the two-party image requires a good measure of two-partyness, and the effective number of electoral parties (or candidates), though popular, is not well suited to the task. The list of studies that have made use of the effective-number criterion in exactly this manner is too long to be presented, and it includes our own work. Old habits die hard, but we can all do better.

The vote-share tuples labeling the points in Figure 2 cannot be easily graphed, given their high dimensionality, but their at-most-three-competitor analogues can be plotted in a triangular portion of the p_1 versus p_2 plane or in a ternary plot, showing the district votes on triangular graph paper that treats parties 1, 2 and 3 on an equal basis. Appendix 2 uses such figures to provide further exposition on how N and T are related mathematically, clarifying why they cannot both be good measures of the same concept, given their fundamentally different behavior.

Aggregation from Districts to National Totals

Thus far, we have assigned party labels on a district-by-district basis, ignoring the existence of multiple districts and the point that the indexing would not, generally, match across districts. If the labels instead reflect national totals, so that party 1 is the national top finisher, equations (1) to (3) would no longer be sensible, and would need to be re-written so that p_1 is replaced by $\max(p_i)$ and so on. Duverger's law suggests that we should see rough parity between two parties, but not necessarily that they be the same two parties in all districts.

Arguably, however, Duverger's law is usually understood to suggest a weaker form of that same condition, that top-two finishes by the two largest parties should at least be more common than ties between either of the nationally largest parties and some other party. The nationwide means of district indices T or D_2 provide a meaningful measure of the degree to which Duverger's law works *at the district level*, irrespective of which two parties compete in a given district, but it says little about the nationwide balance. Indeed, if all districts have $p_1=p_2=0.5$ or $p_1=p_3=0.5$ or $p_2=p_3=0.5$, with equal frequency, then $T=1$ for each district, while the nationwide mean T is 0 (assuming equal sized districts). On the other hand, if half the districts support only party 1 ($p_1=1, p_2=p_3=0$) and half are purely 2 ($p_2=1, p_1=p_3=0$), then $T=0$ for each district, while the nationwide mean T is 1. Hence, one might wish to keep track not only of how two-way competition is, district by district, but also how often the two top rivals are the nationally dominant pair.

Party Competition in England, 1950–2010

We now apply the indices of two-partyness broached above to some district-level data, using the well-studied case of the UK. For simplicity, we restrict attention to competition in England.

Assume, again, that $T \geq 0.5$ and $D_2 \geq 0.7$ are taken as two alternative cutoffs for a single district to be counted as having two-party competition. Then the degree to which Duverger's law applies to all districts in a given election might be measured as the share of districts where either or both of these thresholds is/are met. A secondary measure would incorporate both the two-partytess indices and information about whether the two-way competitions do or do not take place mainly between the two parties with highest national vote shares. After all, for nationwide $p_1 > p_2 > p_3$, all districts can have T near 1 if parties 1 and 3 roughly tie and win most of the vote in half the districts while 2 and 3 tie do so in the other half. While such competition would be Duvergerian at the district level, knowing that the share of seats with close races

Figure 3. N , T and D_2 compared for English parliamentary constituencies in three general elections.

between parties 1 and 2 is near zero would correct the mistaken impression of purely two-partisan competition between p_1 and p_2 . There is a clear difference between “any two-partytiness” and “major-party two-partytiness.”

We computed these indices for districts in England only, not the whole of the UK, to sidestep the effect of regional parties that do not mount candidacies nationwide. Figure 3 mimics Figure 2 in format, displaying data from English parliamentary constituencies in the general elections of 1950, October 1974 and 2005.

At a glance, the data for October 1974 and 2005 look rather similar, and resemble those for 1950 little. The distinct curve on the left of the 1950 panel in the top row consists of races with exactly two competitors, and these vanish in the later years. By either T or D_2 standards, the middle election seems least Duvergerian. For all three columns, but especially the later two, there is very little overlap in the cases that qualify according to $1.5 \leq N \leq 2.5$, as opposed to either $T \geq 0.5$ or $D_2 \geq 0.7$. For these years, races generating large values for T but small ones for D_2 are rare, so the choice in defining the share that qualify as two-party is essentially between using only D_2 or using both T and D_2 . Note that cases that fail both the T and D_2 tests can be one-party dominant (with low N s) or three- or more way splits (with high N s).

The figures do not reflect which parties are finishing in the top two spots, but Table 1 shows what proportion of all English seats met various standards of two-party competition for the elections from 1950 to 2010, and the final column includes the criterion

Table 1. Two-party measures for English constituencies, 1950–2010

Year	(1)	(2)	(3)	(4)	(3) (4) &			<i>n</i>
	$1.5 \leq N \leq 2.5$	$1.75 \leq N \leq 2.25$	$T \geq 0.5$	$D_2 \geq 0.7$	(3)&(4)	(3) (4)	Top2-Con-Lab	
1950	0.80	0.43	0.40	0.59	0.39	0.60	0.39	506
1951	0.96	0.90	0.75	0.75	0.71	0.79	0.70	506
1955	0.96	0.87	0.70	0.71	0.65	0.76	0.65	511
1959	0.86	0.68	0.56	0.62	0.53	0.64	0.52	511
1964	0.56	0.39	0.40	0.50	0.38	0.51	0.38	511
1966	0.73	0.50	0.48	0.59	0.46	0.61	0.45	511
1970	0.91	0.55	0.53	0.65	0.52	0.66	0.51	511
F 74	0.26	0.14	0.10	0.16	0.09	0.16	0.07	516
O 74	0.31	0.12	0.07	0.20	0.07	0.20	0.05	516
1979	0.76	0.19	0.27	0.50	0.27	0.50	0.25	516
1983	0.37	0.15	0.06	0.12	0.06	0.12	0.01	523
1987	0.48	0.16	0.10	0.24	0.10	0.24	0.06	523
1992	0.63	0.15	0.20	0.40	0.20	0.40	0.16	524
1997	0.44	0.21	0.11	0.28	0.11	0.28	0.09	528
2001	0.37	0.14	0.11	0.33	0.11	0.33	0.07	528
2005	0.13	0.03	0.08	0.19	0.08	0.19	0.03	528
2010	0.11	0.02	0.19	0.16	0.13	0.22	0.08	533

that the candidates from the Conservative and Labour parties – which won the two largest vote shares in England and the UK in every one of these elections – finished one-two or two-one. In answer to the question, “Does England satisfy Duverger’s law?” the columns of Table 1 offer distinct responses. Probably the modal view would match column 2, based on N values within 0.25 of 2: roughly speaking, two-way competition was the norm from 1950 to 1970, thereafter, it was the exception, and has become vanishingly rare in the last two elections. Column 1, with a greater tolerance of deviation from $N=2$, tells a broadly similar, but more complicated story, with such anomalies as 1992 seeming more Duvergerian than 1964. The better measures of two-partyness, however, resist a very simple description. If one accepts cases based on passing either a T or D_2 test, the majority of English seats were two-party until February 1974, but the period from October 1974 to 2010 remained somewhat Duvergerian, with only two elections having seen less than one-fifth of the seats meeting a two-party standard (1983 and 2005, barely). The second column from the right clarifies that most of the two-way fights are not Tory–Labour battles, so it is safer to conclude that England has not been *strictly* Duvergerian, in the sense of having the same two-party competition both nationally and in many/most of the districts, since 1970.

Conclusions

What do quantitative measures of two-partyness add to perceptive qualitative and semi-quantitative analysis? Having an explicit measuring stick not only permits more precise claims, but also clarifies how, where, and when the standard is met or not. For Duverger’s law to be testable, “two-partyness” needs to be operationalized. As long as no specific measure of two-partyness existed, using the effective number was better than nothing, despite occasionally puzzling outcomes. But if we take the topic seriously, we should not be wed to clumsy assessment, and it is possible to devise measures better suited to the concept. Alternatives to the Laakso-Taagepera effective number of components have been proposed (e.g. Golosov, 2009; Molinar, 1991), but have seen limited use, and have sometimes been sharply criticized. Instead of tweaking a party-count index yet again, we have operationalized “two-party” at the district level. Superficially similar quantities may be unrelated. In retrospect, one may wonder why the effective number and two-partyness have been confused to the point of so many taking the former as a proxy for the latter. An effective number very coarsely around 2.0 is neither a sufficient nor a necessary condition for two-partyness. For the same reason, neither T nor D_2 should be mistaken for a replacement for N – they are expressly designed for only one purpose, and are narrowly tailored.

Cox’s SF ratio, p_3/p_2 for the SMP races we consider here, is a much simpler measure of strategic coordination, an ingredient in Duverger’s law to be sure. But in ignoring p_1 , it rates very different scenarios as identical, and is thus also not a viable gauge of two-partyness (Gaines, 1997). T and D_2 are less parsimonious because we see no way adequately to capture conformity to bipartite competition simply. It might seem messy to propose two indices, not one, but we submit that there is some fuzziness in how “two-party” is understood, and suspect that examination of multiple measures

can often be useful. We propose more tools, not “the tool” because we doubt that there is one plainly optimal measure. For those who understand Duverger’s prediction to have been “no more than two parties competing,” for instance, D_2 is likely to be unsatisfactory alone, given how sharply it declines as the first-place share approaches 1. Comparison to N , poor though N is as a sole measure of two-partyness, would be one way to recover the roughly one-way cases.

On the other hand, D_2 , also has the potential virtue that it is more easily generalized to a family of statistics measuring fit to three-, four-, etc. way competition, should these be required. We do not necessarily expect these two indices to stand as the last word on how to operationalize two-party competition, but we do hope that readers will concur that the most popular gauge, N , is unsatisfactory.

Acknowledgment

We wish to thank Jake Bowers, Josep Colomer, Bernard Grofman, Arend Lijphart, Anthony McGann, Matthew Shugart, Allan Sikk, Matt Winters, and the editors and anonymous reviewers for helpful assistance and advice.

Notes

1. Sometimes the “law” is understood also to encompass claims about proportional representation and multi-party systems (aka “Duverger’s hypothesis”). Our aim hereafter is not to review Duverger’s works in search of the best representation of his many arguments. Rather, we focus solely on the meaning and measurement of one ingredient, the two-party system.
2. Throughout the article, we use labels expressed in (more familiar) percentages, but do calculations with (more convenient) proportions. So a “50-50” race means a vote-share tuple of (0.5,0.5).
3. As of November 2012, Web of Knowledge and Google Scholar identify 504 and 1,802 citations to the 1979 Laakso-Taagepera article, respectively.
4. Other possible indices sometimes used in evaluating Duverger’s law, beyond N , include Cox’s SF ratio, p_3/p_2 , and the simple margin of victory, p_1-p_2 , both of which incorporate less information than T . Since I was devised to measure p_3 being small, an alternative T' would deflate less, multiplying I by $(p_1+p_2+p_3)$. Yet another option is $T'=(p_2/p_1) \prod_j(p_2-p_j)/p_2$.
5. We would resist applying this index to seat shares in a legislative chamber. Votes do not form coalitions so as to achieve majority, while holders of seats do. In particular, if a majority is not obtained, the implications are quite different for district votes and seats. For district votes, 49-45-6 is clearly close to a pure two-party contest, as intimated by $T=0.75$. But if these were nationwide seats, then a coalition would have to be formed (formally or informally, in case of a minority cabinet). In terms of bargaining power, the outcome is thus rather tripartite in nature.

References

- Benoit, K. (2002) The endogeneity problem in electoral studies: a critical re-examination of Duverger’s mechanical effect. *Electoral Studies*, 21, pp. 35–46.
- Benoit, K. (2004) Models of electoral system change. *Electoral Studies*, 23, pp. 363–389.
- Boix, C. (1999) Setting the rules of the game: the choice of electoral systems in advanced democracies. *American Political Science Review*, 93, pp. 609–624.
- Colomer, J.M. (2005) It’s parties that choose electoral systems (or, Duverger’s laws upside down). *Political Studies*, 53, pp. 1–21.

- Cox, G.W. (1997) *Making Votes Count: Strategic Coordination in the World's Electoral Systems* (Cambridge: Cambridge University Press).
- Duverger, M. (1954) *Political Parties: Their Organization and Activity in the Modern State* (London: Methuen).
- Feld, S. L. & Grofman, B. (2007) The Laakso-Taagepera index in a mean and variance framework. *Journal of Theoretical Politics*, 19, pp. 101–106.
- Gaines, B. J. (1997) Where to count parties. *Electoral Studies*, 16, pp. 49–58.
- Gaines, B. J. (2009) Does the United Kingdom obey Duverger's law?, in: B. Grofman, A. Blais & S. Bowler (eds) *Duverger's Law of Plurality Voting: The Logic of Party Competition in Canada, India, the United Kingdom and the United States* New York: Springer), pp. 115–134.
- Golosov, G. V. (2009) The effective number of parties: a new approach. *Party Politics*, 16, pp. 171–192.
- Grofman, B., Blais, A. & Bowler, S. (eds) (2009a) *Duverger's Law of Plurality Voting: The Logic of Party Competition in Canada, India, the United Kingdom and the United States* (New York: Springer).
- Grofman, B., Blais, A. & Bowler, S. (2009b) Introduction: evidence for Duverger's law from four countries, in: B. Grofman, A. Blais & S. Bowler (eds) *Duverger's Law of Plurality Voting: The Logic of Party Competition in Canada, India, the United Kingdom and the United States* New York: Springer), pp. 1–11.
- Laakso, M. & Taagepera, R. (1979) Effective number of parties: a measure with application to West Europe. *Comparative Political Studies*, 12, pp. 3–27.
- Molinar, J. (1991) Counting the number of parties: an alternative index. *American Political Science Review*, 85, pp. 1383–1391.
- Taagepera, R. (2007) *Predicting Party Sizes: The Logic of Simple Electoral Systems* (Oxford: Oxford University Press).

Appendix 1: Mental Images of Two-Partyness

To see if practiced eyes agree on what constitutes a two-party constellation, we asked a few colleagues who have dealt with Duverger's law to rate the two-partyness of various three-way-vote constellations in a single-member district, using Yes (1), No (0) or "Hard to tell" (0.5). The set included all 44 possible combinations of three parties, at 5% intervals, from 50-50-0 down to 35-35-30 and up to 100-0-0. The precise wording of the questionnaire circulated to the evaluators is shown below.

Duverger's law says that FPTP electoral systems tend to go with two-party systems. But if we want to test the law's degree of validity, we must define what we mean by a two-party system. One starting point would be: We recognize a two-party constellation when we see one. In terms of votes, we probably do not feel Duverger's "two-party systems" includes vote constellations 34-33-33 or 90-10. But where do we draw a line (or rather a grey zone)?

Let us construct a mapping of what almost everyone considers a two-party constellation, what almost everyone considers NOT to be a two-party constellation, and the grey area in-between. Once we have such a mapping, we can start figuring out what the underlying criteria are.

It matters whether we think in terms of district vote shares or nationwide seat shares of parties. Consider here the **vote** shares in a single district. Votes for candidates in a single FPTP district do not form coalitions. Hence we do not have to consider the ideological distances of parties. In the absence of any other knowledge, consider all parties equidistant ideologically.

For the purposes of testing whether Duverger's law holds for *votes* in a single FPTP district, which of the following are two-party constellations?

Please indicate Y (yes), N (no) or H (hard to tell).

50-50	60-40	65-35	70-30	75-25	80-20	85-15
55-45	55-40-5	60-35-5	65-30-5	70-25-5	75-20-5	80-15-5
50-45-5	50-40-10	55-35-10	60-30-10	65-25-10	70-20-10	75-15-10
45-45-10	45-40-15	50-35-15	55-30-15	60-25-15	65-20-15	70-15-15
	40-40-20	45-35-20	50-30-20	55-25-20	60-20-20	90-10
		40-35-25	45-30-25	50-25-25		85-10-5
		35-35-30	40-30-30			80-10-10
				95-5	90-5-5	100-0

Appendix 2: The Mathematical Relationship Between T and N

Since T ignores the distribution of vote across parties 4 through k , we designate $\sum_j p_j = p_r$, and note $p_r = 1 - p_1 - p_2 - p_3$. Then, from the definition $T = (p_1 + p_2)(p_2 - p_3)/p_1$ we get:

$$p_1 T = (p_1 + p_2)(p_2 - 1 + p_1 + p_2 + p_r) = (p_1 + p_2)(p_1 + 2p_2 + p_r - 1) = p_1^2 + 3p_1 p_2 + p_1 p_r - p_1 + 2p_2^2 + p_2 p_r - p_2$$

or

$$p_1^2 + p_1(3p_2 + p_r - T - 1) + 2p_2^2 + p_2 p_r - p_2 = 0$$

Solving the quadratic equation by completing the square yields:

$$p_1 = (1/2)((p_2^2 + 2p_2 p_r - 6p_2 T - 2p_2 + p_r^2 - 2p_r T - 2p_r + T^2 + 2T + 1)^{1/2} - 3p_2 - p_r + T + 1)$$

This identity, along with the limits $p_1 \geq p_2$, $1 \geq p_1 + p_2$, $1 \geq p_r$, and $0 \geq p_i \geq 1$ imply bounds on p_1 , p_2 , p_3 and p_r at any given T , as reflected by the top panel of Figure 2. The expressions, however, are algebraically cumbersome, in a way that hampers intuition. Hereafter, then, we focus on the special case with at most three competitors. Then, the p_r terms fall out and we get:

$$1/(1 + T) \geq p_1 \geq 1/(3 - T)$$

$$1/(3 - T) \geq p_2 \geq T/(1 + T)$$

$$(1 - T)/(3 - T) \geq p_3 \geq 0$$

The lower limit on p_1 and upper limits on p_2 and p_3 correspond to $p_1 = p_2$ and, in turn, the largest possible N at any given T . The upper limit on p_1 and lower limits on p_2 and p_3 follow from $p_1 + p_2 = 1$ and $p_3 = 0$, and correspond to the smallest possible N at a given T .

Figure A.1. Triangular graphs of N and T for three-way vote splits, two options.

Substituting these bounds into $1/N=(p_1^2+p_2^2+p_3^2)$ leads to:

$$N_{\max}|T = (3 - T)^2/(3 - 2T + T^2).$$

$$N_{\min}|T = (1 + T)^2/(1 + T^2).$$

Note that for pure two-party constellations ($p_1+p_2=1$), which lie along the curve marking the left side of the top panel of Figure 2, the $N_{\min}$ equation holds, and T increases with increasing N ($dT/dN>0$). But when $p_1=p_2$ (Duvergerian equilibria), on the right side of the figure, T decreases with increasing N ($dT/dN < 0$), as the two are connected by the first equation above, for $N_{\max}$.

Figure A.1 illustrates both options, in separate panels. The bisectors of the triangles divide the field into six symmetric sections, only one of which has $p_1 \geq p_2 \geq p_3$, and both panels also mark this subregion.

These triangular graphs makes highly visible the non-relationship between two-partytiness and the effective number of parties, due to the following remarkable properties of N and T . In the lower, equilateral triangular graph for three parties:

- *equal- N contours are concentric circles around the center of the triangle;*
- *$T=0$ contours are the spines extending from triangle center to its tips; i.e. these contours are those parts of bisectors which involve a dominant party;*
- *the equal- N circles and the $T=0$ lines are orthogonal;*
- *the lines $T=0.25$, $T=0.50$ (pictured) and $T=0.75$ also cut equal- N circles at steep angles.*

At a given value of T , p_1 and p_2 are related as $p_2 = [1 - p_1(1 - T)]/2$, given that $p_1 + p_2 + p_3 = 1$. Thus, equal- T contours in a the plots are straight lines.

The circle for $N=2.0$ is inscribed into the triangle, barely touching the centers of its sides (50-50-0, 50-0-50, 0-50-50), where $T=1$. In large part, however, the circle $N=2$ passes through low- T areas, including 66.6-16.7-16.7. Strikingly, the $T=0$ lines are precisely orthogonal to the equal- N circles, meaning that, at this level of T , the correlation of T and N is nil. As T increases, the equal- T lines intersect the equal- N circles at decreasing angles, so that some correlation of T and N appears, but their relationship goes in opposite directions inside the circle $N=2$ (where T decreases with increasing N) and outside (where T increases with increasing N). At the circle $N=2$, $dN/dT=0$. More districts tend to lie inside this circle, and hence the prevalent pattern is “ T up, N down,” but it cannot be taken for granted.