Article

Comparative Political Studies
2018, Vol. 51(7) 938–974

DOI: 10.1177/0010414017720704
journals.sagepub.com/home/cps

Social Diversity Affects the Number of Parties Even Under First-Past-the-Post Rules

Caitlin Milazzo[1], Robert G. Moser[2], and Ethan Scheiner[3]

Abstract

Nearly all systematic empirical work on the relationship between social diversity and the number of parties asserts the "interactive hypothesis"—Social heterogeneity leads to party fragmentation under permissive electoral rules, but not under single-member district, first-past-the-post (FPTP) rules. In this article, we argue that previous work has been hindered by a reliance on national-level measures of variables and a linear model of the relationship between diversity and party fragmentation. This article provides the first analysis to test the interactive hypothesis appropriately by using district-level measures of both ethnic diversity and the effective number of parties in legislative FPTP elections and considering a curvilinear relationship between the variables. We find that there is a strong relationship between social diversity and the number of parties even under FPTP electoral rules, thus suggesting that restrictive rules are not as powerful a constraint on electoral behavior and outcomes as is usually supposed.

Keywords

representation, electoral systems, political parties, Duverger's Law

[1]University of Nottingham, UK
[2]University of Texas at Austin, TX, USA
[3]University of California, Davis, CA, USA

Corresponding Author:
Robert G. Moser, University of Texas at Austin, 1 University Station, Austin, TX 78712, USA.
Email: rmoser@austin.utexas.edu

The dominant view within the comparative politics literature on electoral systems and political parties can be termed the "interactive hypothesis"—Electoral rules interact with social diversity to determine the number of parties competing in elections (Amorim Neto & Cox, 1997; Clark & Golder, 2006; Cox, 1997; Duverger, 1954; Ordeshook & Shvetsova, 1994; Singer & Stephenson, 2009). As explained particularly explicitly and cogently by Duverger (1954) and Clark and Golder (2006), social diversity creates the foundation for the party system: Where there is little diversity, there tends to be few parties, but higher levels of social heterogeneity tend to be associated with greater party fragmentation. At the same time, this mapping of society onto the number of parties only occurs when electoral rules are "permissive," meaning that even small parties can win office. Put differently, it is generally agreed that, in permissive electoral systems, such as proportional representation (PR) with large numbers of seats per district and low legal thresholds of representation, greater social heterogeneity promotes party fragmentation.

Most well known, this rich literature also argues that "restrictive" rules constrain the number of parties, thus leading to few parties irrespective of the level of social heterogeneity. Framed most starkly by Clark and Golder (2006), but also noted in numerous empirical studies (see especially Amorim Neto & Cox, 1997; Ordeshook & Shvetsova, 1994; Singer & Stephenson, 2009), first-past-the-post (FPTP) systems—also known as single-member district (SMD) plurality systems—tend to limit party competition to no more than two candidates per district even if the district is socially diverse. Indeed, it is this very idea that underpins one of the most important and well-known theories of electoral rules: Duverger's Law, which holds that FPTP systems tend toward two parties.

To be sure, there have been challenges to Duverger's Law, but—and this is puzzling—even despite research that appears to undercut the microfoundations of the portion of the interactive hypothesis that relates to FPTP rules, these challenges have yet to provide evidence that seriously undercuts the hypothesis. The ideas underpinning Duverger's Law and the interactive hypothesis are founded in large part on the expectation of strategic voting under FPTP, but as Cox (1997) notes, there are multiple conditions under which voters will be unlikely to act strategically. Moreover, empirical work by scholars such as Alvarez, Boehmke, and Nagler (2006) indicates that in fact large numbers of voters do not behave strategically in FPTP elections even when they have a rational incentive to do so. This lack of strategic behavior ought to lead to a relationship between social diversity and party fragmentation even under these restrictive rules. Nevertheless, the literature offers only very limited evidence at best of a link between social heterogeneity and the number of parties under FPTP, as would be implied by the Cox and Alvarez et al. work.

In this article, we argue that the FPTP portion of the interactive hypothesis does not accurately represent real world outcomes. We argue that, in fact, *there is a relationship between social diversity and the number of parties under FPTP*. And we argue that the dearth of strong empirical evidence to counter the hypothesis is a result of shortcomings in the case selection, data used, and assumptions made to set up the tests conducted in the extant literature.

More specifically, we argue that an appropriate test of the relationship between social heterogeneity and the number of parties under FPTP will combine three features in its analysis: First, some may question the generalizability of the findings drawn from analyses of party fragmentation in plurality elections within mixed-member electoral systems, in new democracies, or presidential elections, so an appropriate test of the hypothesis will include analysis of legislative elections in pure plurality systems in established democracies, such as Great Britain and the United States. Second, the data in support of the analysis need to be drawn from the subnational (especially district) level. Duverger's Law is founded on a district-level logic, in which the district number of parties is a function of social diversity and electoral rules in the district, but with the exception of Potter (2014), all cross-national analyses of the interactive hypothesis are founded on national-level measures of at least one of the key variables. Third, recent evidence suggests that the effect of diversity on party fragmentation may not be linear (Dickson & Scheve, 2010; Moser & Scheiner, 2012; Raymond, 2015; Stoll, 2013), so an appropriate test of their relationship under FPTP will need to consider the possibility of a nonlinear relationship.

Following these principles, we create an original data set of the legislative *district-level* number of parties and social diversity—like most work on this topic (especially Amorim Neto & Cox, 1997; Clark & Golder, 2006; Singer, 2012; Singer & Stephenson, 2009), usually coded as ethnic diversity—under plurality rules in nine different cases, which we use to provide systematic evidence that social diversity does indeed shape the number of parties under FPTP rules. More specifically, we conduct district-level analyses of multiple plurality legislative elections in the most well-known established democracies that have used pure plurality electoral rules: Canada, Great Britain, India, (pre-reform) New Zealand, and the United States. In addition, to attempt to generalize further, we also include analyses within the FPTP rules in mixed electoral systems in (post-reform) New Zealand, Russia, Scotland, Ukraine, and Wales.

In all cases except Wales, we find that increases in social diversity tend to be associated with increases in the number of parties. The results for the United States are largely consistent with the interactive hypothesis, whereby the number of parties tends to be capped around two, but in all other cases (again, except for Wales), increases in social diversity are correlated with greater party fragmentation that goes beyond simply two principal candidates. Strikingly, this positive relationship between diversity and the number of parties halts at higher levels of diversity, and in fact we observe a statistically significant decline in party fragmentation at higher levels of diversity in all of the pure plurality systems in our study, as well as a number of the mixed systems we observe.

The Interactive Hypothesis: Social and Ethnic Diversity, Electoral Systems, and the Number of Parties

Most scholars, whether they are defenders or critics of the interactive hypothesis, have described the "conventional wisdom" regarding the relationship between social diversity, electoral systems, and party fragmentation as follows: Social heterogeneity provides the "raw material" for the number of parties found in a country (Clark & Golder, 2006; Duverger, 1954; Powell, 1982). Electoral rules provide a filter that allows or prevents such forces from being manifest as political parties. Consequently, societies with more social groups should have larger numbers of political parties as long as the electoral system is permissive enough to allow such groups to form the basis of their own parties. More specifically, electoral systems with higher district magnitudes (i.e., seats in the district), and thus lower effective electoral thresholds (i.e., the share of the vote needed to win a seat), allow more social cleavages to be represented by distinct parties. Electoral systems with lower district magnitudes restrict those opportunities because of their bias against smaller parties. Most notably, FPTP systems in single-seat districts create incentives for weak parties and candidates to exit the race and for voters to cast ballots strategically only for potentially competitive candidates. Following this logic, there is a winnowing process, whereby ultimately in FPTP systems there ought to be no more than two candidates per district, irrespective of the level of social diversity.

There is substantial empirical support for both parts of the interactive hypothesis. Most of the prominent cross-national work on the topic analyzes the relationship between social diversity and the number of parties using data at the national level—a single nationally aggregated measure of the effective number of ethnic groups for each country and the nationally aggregated

Laakso and Taagepera (1979) measure of the effective number of parties for each election in each country. Each of these studies finds strong empirical support for the same conclusion—There is a positive relationship between ethnic diversity and the number of parties at all but the smallest district magnitudes, but under SMDs there is no discernible relationship between the two variables (Amorim Neto & Cox, 1997; Clark & Golder, 2006; Cox, 1997; Ordeshook & Shvetsova, 1994). As the effects of electoral rules are expected to be most direct at the district level, more recent work by Singer and Stephenson (2009) and Singer (2012) takes the important step of analyzing the factors that shape the district-level number of parties. Still using a national measure of ethnic diversity, these studies also find no statistically significant relationship between social cleavages and the effective number of parties in plurality elections, thus supporting the interactive relationship between diversity and electoral rules.[1]

Based on this analysis, the leading empirical work on the topic has come to express a strong version of the interactive hypothesis in explicit terms: In perhaps the most well-known recent articulation of this interactive model, Clark and Golder (2006) argue,

> The central hypothesis Duverger's theory generates is that social heterogeneity should increase the number of parties *only once the electoral system is sufficiently permissive* . . . [and] [w]hen we actually test Duverger's theory with a fully specified model, we find that the results are remarkably consistent with Duverger's expectations. (p. 704, emphasis added)

Cox (1997) in his classic work, *Making Votes Count,* offers the same formulation: "A polity can tend toward bipartism either because it has a strong electoral system or because it has a few cleavages. Multipartism arises as the *joint* product of many exploitable cleavages and a permissive electoral system" (p. 221, emphasis added). Ordeshook and Shvetsova (1994) argue that "if district magnitude [the number of seats in the district] equals one, then the party system is relatively 'impervious' to ethnic and linguistic heterogeneity" (p. 122). Even Dickson and Scheve (2010), who offer a partial challenge to the conventional view, concede that "empirical studies demonstrate a positive relationship between ethnic fractionalization and the number of candidates or parties under so-called 'permissive' electoral systems, but little if any relationship when institutions are not 'permissive'" (p. 349).

Challenging the Interactive Hypothesis

Despite the long list of theoretical and empirical work in support of the interactive hypothesis, there are good reasons to challenge the thesis. Perhaps most striking, strategic defection by political actors in restrictive systems underpins the logic of the interactive hypothesis—and there is substantial evidence that large numbers of voters do cast strategic ballots (see, among many others, Cain, 1978; Cox, 1997)—but in reality only a subset of all voters actually does defect strategically when circumstances warrant (see, for example, Alvarez et al., 2006; Cox, 1997, pp. 83-84). For example, looking at the multicandidate context of elections in Britain, Alvarez et al. (2006) find that even among voters who prefer a totally uncompetitive candidate, only roughly half strategically defect from their top choice to a more competitive alternative.

There may be many reasons for such expressive voting. Significant theoretical work by Palfrey (1989), Myerson and Weber (1993), and Cox (1994, 1997) highlights how voters are apt to strategically abandon their first preference candidate and produce a two-candidate outcome in districts under FPTP, but Cox (1997), himself, notes explicitly the important—and potentially fragile—assumptions that underpin expectations of such behavior. Specifically, Cox articulates four preconditions for strategic behavior: (a) short-term instrumentally rational voters, (b) widespread knowledge of which candidates are "out of the running," (c) the absence of a widespread belief that one candidate is certain to win, and (d) relatively few voters who are so devoted to their first place candidate that they are indifferent between their second and lower choice candidates. There are, indeed, good reasons to believe that these conditions may often not hold, thus undercutting more strategic defection and promoting expressive voting.[2]

That being said, the focus of this article is on the empirical link between social diversity and the party system, and for our purposes what is most important is that, whatever the reason for expressive voting, a significant lack of strategic defection ought to undercut the logic of the interactive hypothesis: Even if a majority (i.e., >50%) of voters cast ballots for one of the top two candidates, if (a) not all small parties (or candidates) exit the race and (b) a relatively large proportion of voters continue to support its top preference irrespective of the chances of success, we are likely to see a relationship between social diversity and the number of parties under FPTP. That is, if a substantial number of voters and elites choose not to defect strategically in FPTP races, the restrictive rules will not eliminate the influence of social diversity on the number of parties.

Indeed, in related work that has implications for the interactive hypothesis, some scholars have pointed to the regular exceptions to Duvergerian equilibria in plurality elections as reason to question the generalizability of its underlying logic. In particular, scholars illustrate the presence of more than two candidates in scores of FPTP districts in Britain, Canada, and India (see, for example, Chhibber & Kollman, 2004; Gaines, 1999, 2009; Johnston & Cutler, 2009). Highlighting an increase over time in the vote totals of "hopeless" third place candidates who have no chance of winning in district FPTP races in Britain, Canada, and India, Chhibber and Kollman (2004) argue that "Duverger's Law works often enough to remain an important regularity in the social sciences, but not often enough to ascribe deviations as merely random disturbances from the norm" (p. 58). And noting this pattern, Grofman, Bowler, and Blais (2009) argue that Duverger's "prediction of two parties is not really so robust" (p. 6).

However, examples of these kinds do not address issues of social diversity and the interactive hypothesis directly, and to this point no work has been able to provide a thorough empirical challenge to the interactive hypothesis. Taagepera and Grofman (1985) offer perhaps the most strongly stated systematic challenge to the hypothesis, arguing that social diversity affects the number of parties under any electoral system, and that "the effective number of parties tends to be obtained by adding 'one' to the number of issue dimensions" (Taagepera & Grofman, 1985, p. 341). But their empirical analysis is more preliminary than definitive, as it is founded on fairly subjective (and probably endogenous) variable measurement (drawn from Lijphart, 1984; see Gaines, 1999) and only four plurality cases with little variation across them.

Nevertheless, in the face of significant evidence of large numbers of voters who do not cast strategic ballots under FPTP rules, the lack of evidence challenging the interactive hypothesis is somewhat puzzling—but two factors may be responsible for this lack of evidence. First, nearly all systematic cross-national analysis of the hypothesis uses nationally aggregated data, which typically will provide a poor measure of conditions at the district level, where Duverger's logic holds. Second, most empirical work on the relationship between social diversity and the number of parties under FPTP implicitly assumes a monotonically positive correlation, which may not be an accurate representation of the true relationship between the two variables.

The Importance of District-Level Analysis in Studying the Relationship Between Social Diversity and Party Fragmentation

The extant literature's use of *nationally aggregated* measures for at least one of the key variables it uses to test the interactive hypothesis creates at least some doubt about the accuracy of the analysis.[3] The direct effect of electoral systems is at the *district* level, and theories that seek to explain the effective number of electoral parties—that is, weighted by the share of votes each party/candidate wins—rely first on a district-level logic (Chhibber & Kollman, 2004; Cox, 1997; Duverger, 1954; Jones, 2004; Singer & Stephenson, 2009).

For example, Duverger's Law is founded on the logic that voters and elites *within a district* seek to affect the outcome of that district's race (Cox, 1997). Social diversity at the national level may be wildly different from diversity in any given district, and the number of parties in any given district may be very different from what we see once votes from all districts are aggregated (see Cox, 1997; Riker, 1986). Even where a group is in the minority across most of the country, it may be able to win a seat in a district in which it actually makes up a majority of the population. A study based purely on nationally aggregated data would miss the fact that this national minority is quite large in some districts, and such a study would misinterpret the reasons that a party representing such a group might win votes under FPTP rules. Moreover, the aggregate number of parties at the national level is the sum of the votes for parties across all districts in the country and could easily misrepresent what occurs at the district level. At the most extreme, the expected Duvergerian (two-candidate) outcome in each district could lead to hundreds of different parties winning votes and seats nationally—that is, a large number of parties according to the nationally aggregated measure—if different parties compete in each district. As Jones (2004) explains, "Any analysis of legislative elections must either be conducted at the district level or else make several, usually unrealistic, assumptions regarding the distribution of the vote and influence of electoral laws in aggregate at the national level" (p. 75).

Nevertheless, because it has long been difficult to acquire such data, it is unusual for comparative work on the number of parties in legislative elections to use district-level measures of ethnic diversity. Most cross-national studies (Amorim Neto & Cox, 1997; Clark & Golder, 2006; Ordeshook & Shvetsova, 1994) use nationally aggregated measures of both the number of parties and ethnic diversity. Singer and Stephenson (2009) examine the number of parties at the district level, but use nationally aggregated measures of ethnic diversity. Jones (1997) and Madrid (2005) offer insightful district-level work on diversity and the number of parties in legislative elections, but although both studies find a clear association between social diversity and the number of parties, neither study offers a direct analysis of the relationship in plurality elections.

It is in large part because of the difficulty of acquiring data on FPTP legislative districts that a number of clever studies look instead at the relationship between ethnic diversity and party fragmentation in national presidential elections, which are in fact SMD races (where the entire country is the district). These studies arrive at a number of conclusions but as a group suggest a strong link between social heterogeneity and party fragmentation.[4] The high-stakes nature of presidential elections provides a particularly unlikely context in which to find a link between diversity and party fragmentation under FPTP rules, because such contests will be likely to promote significant strategic behavior that should undercut the relationship. As such, findings of an effect in presidential elections therefore provide reason to expect a relationship in legislative races. At the same time, the applicability to legislative elections is merely an assumption, and the reduced stakes and regionally based competition may lead to very different kinds of behavior in legislative FPTP races.

Unfortunately, district-level analyses of legislative contests conducted under FPTP rules are very uncommon. Ferree, Gibson, and Hoffman (2017) offer a strong district-level empirical challenge to the interactive hypothesis. Examining local elections in South Africa, they use SMD-level measures of all of their key variables to show a positive relationship between racial diversity and party fragmentation. Ferree et al.'s analysis is compelling, but it is difficult to generalize from a study of behavior in local elections in a relatively new democracy.

Potter (2014) offers perhaps the most sophisticated and comprehensive district-level analysis of the interactive hypothesis. Potter cleverly uses probabilistic topic modeling of survey results to estimate district-level measures of latent social diversity across a wide array of dimensions (ethnicity, language, religion, income, the rural–urban divide, and support for the democratic regime) and countries. Interestingly, in his base model, Potter finds a positive correlation (with the 95% interval almost entirely positive) between district-level diversity and the number of parties, irrespective of the permissiveness of the electoral system, but the results are not robust. In his central analysis, Potter includes interactions (for both district magnitude and district-level diversity) with nationally aggregated measures of diversity and in turn finds no statistically significant relationship between any form of diversity and the number of parties under SMDs (see Potter, 2014, Table 5).

To be sure, Potter's analysis may be accurately capturing a lack of a correlation between social diversity and party fragmentation under restrictive rules, but it may also be that components of his analysis make it harder for him to capture the relationship. Potter's imputed measure of social diversity may not accurately represent true diversity within a given district and/or the multiple interactions in his model may inhibit our ability to draw inferences about the effects of district-level diversity, specifically. Another possibility, though, is that any study that treats the relationship between social diversity and party fragmentation as positive and linear may find it difficult to identify a clear link between the two variables.

The Relationship Between Social Diversity and Party Fragmentation May Not Be Linear

Indeed, most studies of the effect of social diversity on party fragmentation assume the relationship to be positive and linear, but there is good reason to suspect that this assumption is incorrect. To begin with, it is easy to imagine that the effect of social diversity could have diminishing returns, whereby at a certain level further increases in heterogeneity do not lead to further increases in the number of parties. Stoll (2013) takes this logic to an even greater extreme, explicitly arguing that there is an inverted U-shaped relationship between social diversity and the number of parties. Stoll's argument is straightforward at low to medium levels of diversity, where she posits that the number of salient cleavages ought to map onto the number of parties, but the expectation is more counterintuitive at higher levels of diversity: Where there are larger numbers of (especially small) groups, there is little likelihood that any of those groups will be an electoral or government player by itself. In such cases, political entrepreneurs have an incentive to create parties (or promote candidates) that can gain the support of multiple groups. In this way, the presence of a large number of groups may promote behavior by elites that actually leads to *fewer* parties.

Probably the most straightforward case of high social diversity leading to a small number of parties is where a single party becomes the principal target of most ethnic minorities' votes. We can see this pattern, for example, with the Labour Party in New Zealand and Great Britain. Figure 1 illustrates the relationship between ethnic diversity and the share of the vote won by the Labour Party in New Zealand in 2002 and Great Britain in 1997. As the figure shows, there is considerable variation in the share of Labour's vote at most levels of diversity, but the Labour Party tends to do exceptionally well in districts where the effective number of groups score is greater than two in both countries. In short, increases in diversity promote greater party fragmentation, but the presence of large numbers of ethnic minorities appears to promote greater Labour vote concentration and, thus, a smaller number of parties.

Indeed, Stoll provides compelling evidence of the existence of a nonlinear relationship between social diversity and the number of parties under permissive legislative electoral rules: Increases in diversity are associated with larger numbers of parties up to a point, after which greater diversity is correlated with fewer parties. The work of Raymond (2015) gives further reason to believe that there is no simple positive correlation between diversity and party fragmentation. In an argument that is similar to Stoll's, Raymond also posits and demonstrates empirically an upside-down U-shaped relationship

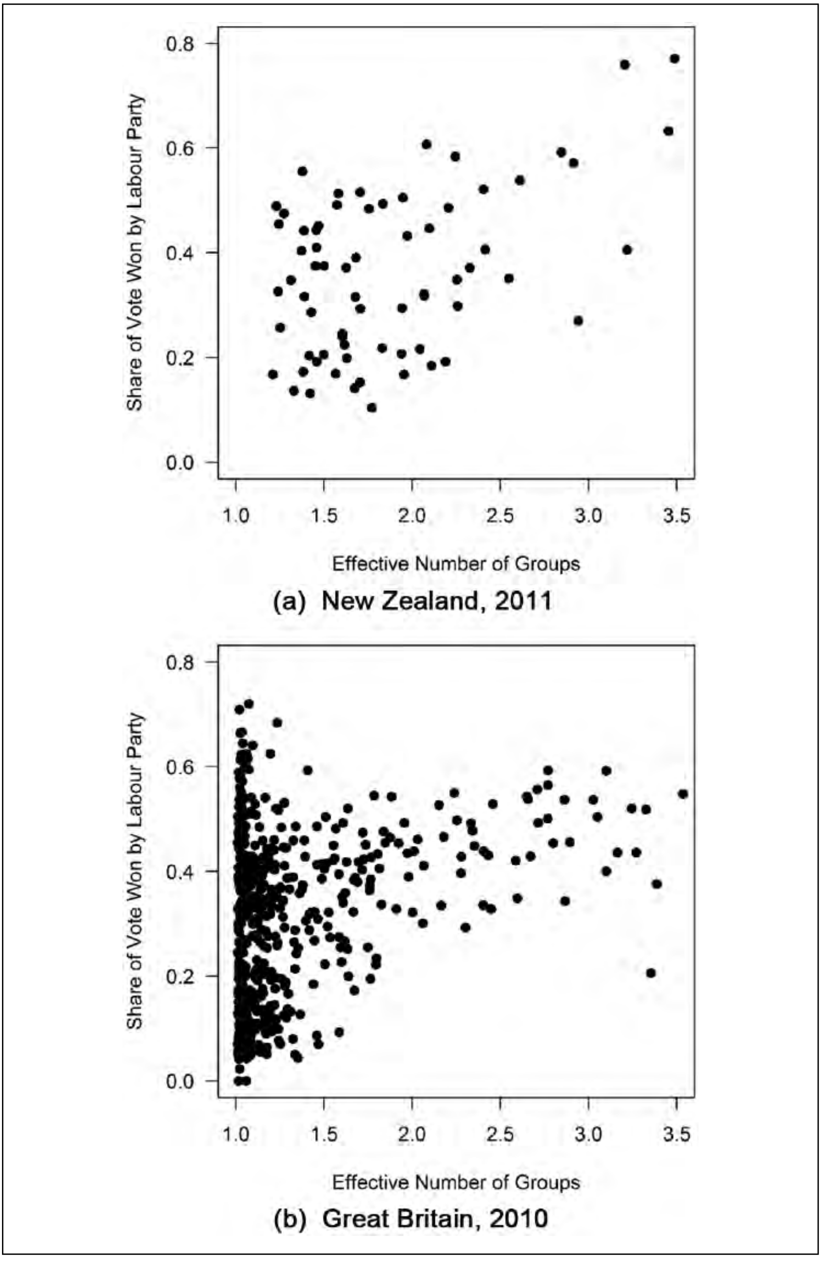

Figure 1. Ethnic diversity and the Labour Party's share of the district vote.

between ethnic diversity and party systems that is due to a dynamic in which extreme diversity produces groups that do not have the critical mass to elect their own representative. Consequently, in a context of many small ethnic groups, there are incentives to form multiethnic coalitions to win office. Such coalition building across ethnic groups tends to push down the number of parties competing (Raymond, 2015, pp. 109-110).[5]

Neither Stoll nor Raymond uses data from the district level in restrictive legislative elections—and largely avoid making definitive arguments about the likely effects of restrictive rules—but Moser and Scheiner (2012) examine the relationship between diversity and party fragmentation in four mixed-member electoral systems using SMD-level data for both social diversity and the number of parties. They find the same inverted U-shaped relationship demonstrated by Stoll and Raymond, and the curves representing the relationships are nearly identical under both PR and plurality balloting. However, it is not totally clear if these findings are generalizable, given they are based on cases with mixed-member electoral systems that are not necessarily comparable with pure plurality systems.

What Is Needed to Address the Issue of Social Diversity and Party Fragmentation Under Plurality Rules

In short, despite reasons to question its core theoretical assumptions, the interactive hypothesis remains the dominant view of the relationship between rules, society, and party systems. To be sure, there is good reason for this dominant view with respect to one half of the hypothesis: Empirical work demonstrates convincingly that there is a link between social diversity and party fragmentation under permissive electoral rules (although debate remains about the functional form of that relationship). However, evidence in support of the other half of the interactive hypothesis—that social diversity has no relationship to party fragmentation under restrictive FPTP rules—is on far less steady ground. Empirical challenges have emerged, but their empirics have been limited by case selection, key variable measurement, and, possibly, the specification of the form of the relationship between social diversity and party fragmentation.

To address more directly the core ideas underpinning this second part of the interactive hypothesis, we argue that empirical analysis of the relationship between social diversity and the number of parties under FPTP rules should take the following steps: First, to avoid criticisms of country case selection, any study should include multiple pure FPTP systems in established democracies. Second, studies should measure both party fragmentation and social diversity at the SMD level, where the direct effect presumably would exist. Third, studies should examine more than just the possibility of a linear arrangement, and specifically should run models that take into account the "curvilinear" relationship suggested by Stoll and Raymond. In addition, to test the conclusions of the extant literature most appropriately, it would be ideal to take on the literature's interactive hypothesis on its own terms by measuring social heterogeneity in the form that has been most used in extant analysis—ethnic diversity (see Amorim Neto & Cox, 1997; Clark & Golder, 2006; Singer, 2012; Singer & Stephenson, 2009).

A New *District-Level* Data Set to Study the Relationship Between Social Diversity and the Number of Parties Under FPTP Rules

We took these guidelines into account when we designed our study here. We use an original data set to analyze the relationship under FPTP between district-level social heterogeneity and the number of candidates in nearly 10,000 district-level races in lower house elections.[6] Our data set includes elections under pure FPTP electoral systems in the established democracies of Canada, Great Britain, India, (pre-reform) New Zealand, and the United States. We also extend the analysis by including the FPTP portion of mixed electoral systems in (post-reform) New Zealand, Russia, Ukraine, Scotland, and Wales.[7] For most of our cases, we include multiple years of elections (see Table 1). We measure the number of parties within each SMD in each election in our data set. In Canada, Great Britain, New Zealand, Russia, Scotland, the United States, Ukraine, and Wales, our measure of social diversity is the number of ethnic groups aggregated within the boundaries of each SMD.[8] Data on ethnicity were not available for India and, therefore, our measure of diversity is based on the number of religious groups.[9]

Table 1 presents the elections and measures of social heterogeneity used for each case. Our cases represent a range of political, electoral, and demographic contexts from a variety of geopolitical regions. There is variation in the level of democratic consolidation: Canada, Great Britain, India, New Zealand, and the United States have significant experience with democratic elections. Russia and Ukraine democratized much more recently—and Russia's democratization proved temporary. Our cases also display institutional variation. All of our data are drawn from district-level elections that use SMD plurality rules, but Russia, Scotland, Ukraine, and Wales have mixed electoral systems, where voters simultaneously cast a vote for a candidate in the FPTP balloting and one for a party in PR. New Zealand also has

Table 1. Description of Cases—Elections and Measures of Diversity.

Case	Elections used	Measure of diversity	Effective number of groups	Effective number of candidates
Canada	2004, 2006, 2008, 2011	Ethnicity	1.462 (1.006-4.861)	2.769 (1.397-4.132)
Great Britain	1987, 1992, 1997, 2001, 2005, 2010	Ethnicity	1.198 (1.003-6.269)	2.663 (1.434-4.269)
India	1998, 1999, 2004	Religion	1.406 (1.007-2.531)	2.718 (1.300-7.183)
New Zealand	1987, 1990, 1993, *1996*, *1999*, *2002*, *2005*, *2008*, *2011*	Ethnicity	1.601 (1.039-3.948)	2.702 (1.621-5.861)
Russia	*1995*	Ethnicity	1.373 (1.025-2.000)	5.924 (1.931-13.048)
Scotland	*1999*, *2003*, *2007*, *2011*	Ethnicity	1.309 (1.082-2.812)	3.112 (1.959-4.765)
Ukraine	*2002*	Ethnicity	1.526 (1.020-2.985)	6.286 (1.390-17.280)
United States	2002, 2004, 2006, 2008	Ethnicity	1.761 (1.047-3.788)	1.755 (1.000-2.690)
Wales	*1999*, *2003*, *2007*, *2011*	Ethnicity	1.096 (1.029-1.651)	3.125 (1.928-4.935)

Sources for all data are listed in the appendix. All measures of the effective number of parties and social diversity are at the single-member district level. Elections occurring under mixed electoral rules indicated in italics. Columns 4 and 5 present, for each case, the mean number of groups and the mean number of candidates, respectively. The range is given in parentheses. Our measure of the effective number of groups does not include individuals placed in a generic "Other" category, those who did not state a category membership, or those who stated that they belong to multiple groups.

used a mixed system since 1996[10] but used a pure SMD-plurality system before then. Our data set includes elections under both systems in New Zealand. Finally, there is considerable socioeconomic variation. Canada, Great Britain, and the United States have a GDP per capita of more than US$30,000, while India has a GDP per capita of less than US$2,000.[11]

To examine the effect of district-level social heterogeneity on the number of competitors, we calculate, first, the effective number of SMD *candidates* per district using the Laakso and Taagepera (1979) effective number of electoral parties (N) index, which weights all contestants by their share of the vote.[12] N ranges from 1 (there are a number of uncontested districts in the United States) to 17.28 (in the Kiev district in Ukraine).

For our principal explanatory variable, we follow the most common approach in the literature, which uses ethnic fragmentation as the measure of social diversity (see Clark & Golder, 2006).[13] To be sure, the variable is far from a perfect measure of the core divisions within society (see Ordeshook & Shvetsova, 1994, p. 109), and the focus on ethnic diversity is not due to a belief that more ethnic groups will lead to more ethnic parties (Amorim Neto & Cox, 1997, p. 166, fn. 21). Rather, for scholars, ethnicity has the advantage of being a more fixed voter trait than others such as religion and class, and, as such, can be seen as a more exogenous determinant of individuals' political preferences (Ordeshook & Shvetsova, 1994, pp. 107-108). The literature therefore uses ethnic diversity as a "proxy" measure of general social diversity (Amorim Neto & Cox, 1997, p. 166, fn. 21; Clark & Golder, 2006, p. 696): Higher levels of ethnic diversity ought to be associated with greater social heterogeneity in general. Following this assumption as well, we calculate the effective number of ethnic groups for each district in Canada, Great Britain, New Zealand, Russia, Scotland, Ukraine, the United States, and Wales, as well as the effective number of religious groups in India. The effective number of groups measure (G) ranges from 1 to 6.29—the latter being the East Ham district in Great Britain.

Testing Whether There Is a Relationship Between Diversity and Party Fragmentation Under FPTP

Using these data, we test whether there is a link between social diversity and the number of parties. As we suggested earlier—and contrary to the interactive hypothesis—there is good reason to expect a relationship even under FPTP. Our "naïve" hypothesis is that there will be a positive correlation between our core explanatory variable, G (the effective number of groups), and our outcome variable, N (the effective number of parties). We test this hypothesis by running a separate model for each case in our data set (with the district/election as the unit of analysis).[14] We report the results in Table 2.

The results in Table 2 provide mixed evidence for the naïve hypothesis of a positive relationship between social diversity and party fragmentation. In five of the cases (Great Britain, India, the United States, pre-reform New Zealand, and Scotland), the coefficient on G is positive and statistically significant, but in four of the cases (Canada, post-reform New Zealand, Ukraine, and Wales), the coefficient is statistically indistinguishable from zero (i.e., it is not significant). In one case (Russia), the coefficient is actually negative and significant.

Table 2. Linear Models of the District-Level Relationship Between Diversity and the Effective Number of Parties.

	Canada	Great Britain	India	United States	New Zealand		Russia	Scotland	Ukraine	Wales
	Plurality	Plurality	Plurality	Plurality	Plurality	Mixed	Mixed	Mixed	Mixed	Mixed
Effective number of groups (G)	−0.007	0.254**	0.179**	0.164**	0.213*	0.003	−1.404**	0.548**	0.897	0.027
	(0.022)	(0.056)	(0.067)	(0.058)	(0.085)	(0.072)	(0.487)	(0.181)	(0.458)	(0.615)
Constant	2.779**	2.358**	2.467**	1.467**	2.327**	2.760**	7.852**	2.394**	4.917**	3.095**
	(0.042)	(0.067)	(0.101)	(0.102)	(0.118)	(0.113)	(0.683)	(0.236)	(0.728)	(0.675)
N	1,232	3,804	1,563	1,741	281	371	222	292	223	160
R^2	.000	.022	.007	.009	.032	.000	.036	.053	.017	.000

Standard errors in parentheses. Dependent variable: SMD-level effective number of electoral parties (N). Effective number of groups (G) = SMD-level effective number of ethnic groups. Models used—Russia and Ukraine: OLS; Canada, India, New Zealand, and Scotland: OLS with standard errors clustered by district; Britain, the United States, and Wales: Cluster-specific fixed effects model with standard errors by clustered district. For Great Britain, the United States, and Wales, the R^2 corresponds to the within-cluster R^2 of the fixed effects model. SMD = single-member district; OLS = ordinary least squares.
$^*p < .05$. $^{**}p < .01$.

However, as we noted earlier, there is good reason to believe that the relationship between diversity and party fragmentation is not linear. Indeed, in the scatterplots in Figure 2, we can actually see the curvilinear relationship we had earlier hypothesized. Figure 2 plots the effective number of parties/candidates (on the vertical axis) against the effective number of groups (on the horizontal axis) for each case in our data (again, with the district/election as the unit of analysis). In Figure 2, the United States illustrates no obvious correlation between the number of groups and party fragmentation. However, other cases, in particular Canada, Britain, and New Zealand, seem to show that increases in social heterogeneity are accompanied by greater party fragmentation, but that at high levels of diversity, the number of parties actually declines.

In an effort to capture more systematically this seemingly nonlinear pattern, we rerun the models from Table 2, but we now include in our models G^2, the square of G (the effective number of groups). If, indeed, the number of parties increases at low levels of diversity, but then stops increasing or even declines at higher levels of social heterogeneity, our regression results should indicate a positive and significant coefficient on G, and a negative and significant coefficient on G^2, because larger values of G will actually lead to a decline (or a halt in the increase) in the number of candidates.

It is of course possible that other nonlinear patterns better describe the relationship between G and N, but we have two reasons for running our models with the quadratic term. First, and most important, Stoll's (and Raymond's) cogent analysis gives us good theoretical reason to believe that the relationship between G and N is in fact characterized by an inverted U-shaped pattern, with declines in party fragmentation at high levels of diversity. Including only G in the model would allow us to detect if there is a monotonically linear relationship between G and N but would be unlikely to capture other patterns. Including logged values of the key variables would capture a nonlinear relationship in which increases in G at high levels of diversity are associated with no further party fragmentation, but a model with logged values would miss an inverted U-shaped pattern. In contrast, including both G and G^2 allows us to test whether the relationship is characterized by the inverted u shape that Stoll suggests. Second, including G and G^2 gives our analysis flexibility: We can still test whether the relationship is purely linear (positive coefficient on G and nonsignificant coefficient on G^2) and whether the relationship between G and N is largely positive but also simply characterized by diminishing or nonexistent returns at high levels of diversity (which we would be able to see in predicted values drawn from the coefficients on G and G^2).

The results presented in Table 3 offer strong evidence that there is a relationship between social diversity and the number of parties under FPTP rules,

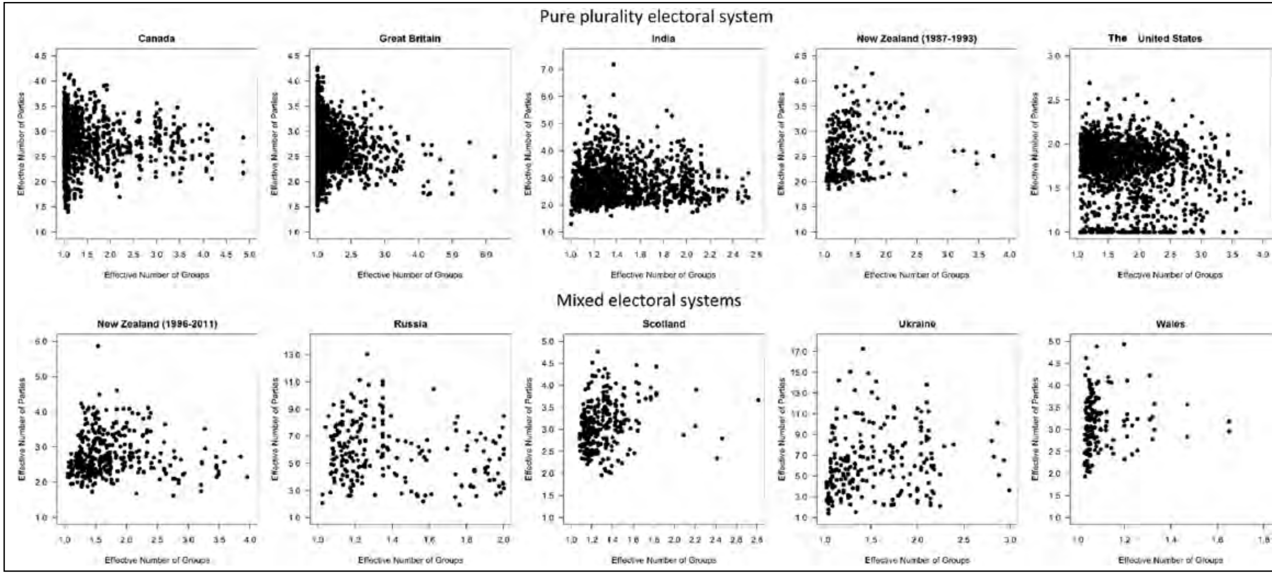

Figure 2. Scatterplots of district-level diversity and party fragmentation.

Table 3. Curvilinear Models of the District-Level Relationship Between Diversity and the Effective Number of Parties.

	Canada	Great Britain	India	United States	New Zealand		Russia	Scotland	Ukraine	Wales
	Plurality	Plurality	Plurality	Plurality	Plurality	Mixed	Mixed	Mixed	Mixed	Mixed
Effective number of groups (G)	0.256*	1.208**	2.765**	1.183**	1.550**	1.400**	12.612*	3.156**	7.728**	0.040
	(0.112)	(0.186)	(0.537)	(0.232)	(0.310)	(0.298)	(6.253)	(0.880)	(2.914)	(6.779)
G^2	−0.059**	−0.166**	−0.810**	−0.241**	−0.341**	−0.325**	−4.634*	−0.803**	−1.970*	−0.005
	(0.022)	(0.035)	(0.162)	(0.053)	(0.074)	(0.067)	(2.061)	(0.288)	(0.830)	(2.361)
Constant	2.551**	1.483**	0.516	0.493	1.170**	1.406**	−2.273	0.395	−0.539	3.088
	(0.109)	(0.168)	(0.418)	(0.235)	(0.280)	(0.292)	(4.554)	(0.662)	(2.409)	(4.573)
N	1,232	3,804	1,563	1,741	281	371	222	292	223	160
R^2	.008	.063	.027	.023	.087	.056	.058	.087	.042	.000

Standard errors in parentheses. Dependent variable: SMD-level effective number of electoral parties (N). Effective number of groups (G) = SMD-level effective number of ethnic groups. G^2 = The square of the effective number of groups. Models used—Russia and Ukraine: OLS; Canada, India, New Zealand, and Scotland: OLS with standard errors clustered by district; Britain, New Zealand, the United States, and Wales: Cluster-specific fixed effects model with standard errors by clustered district. For Great Britain, the United States, and Wales, the R^2 corresponds to the within-cluster R^2 of the fixed effects model. SMD = single-member district; OLS = ordinary least squares.

*$p < .05$. **$p < .01$.

and they support the idea that the association is not monotonically linear.[15] In every single case, the coefficients are in the expected direction (positive for G and negative for G^2), though coefficients do not reach conventional levels of statistical significance in the model for Wales.[16]

To illustrate the substantive relationship between district-level diversity and the number of parties, we use the estimates in Table 3 to calculate the expected number of parties across the range of G for each case.[17] Figure 3 presents these values, as well as their 95% confidence intervals. The visual pattern for every case is that at the lowest levels of diversity, increases in the effective number of groups are associated with greater party fragmentation, but at high levels of diversity, we see a decline in the number of parties.[18] In Table 4, we lay out numerically the expected effective number of parties associated with a variety of different levels of diversity in each case. As Table 4 indicates, the pattern is particularly stark in Great Britain, where increasing G from 1 to 2 is associated with a sizable jump of 0.712 in N (from 2.526 when G = 1 to 3.238 when G = 2), and a further increase in G from 2 to 3 sees a rise of 0.380 (from 3.238 to 3.618) in the effective number of parties. With the exception of Canada, the substantive effect of increases in diversity on party fragmentation appears to be quite large in the other cases as well: For every case except Canada and Russia, an increase from 1 to 2 in the effective number of groups is associated with an increase of at least 0.333 in the effective number of parties. Our measure of G in Russia actually has a maximum of 2 and has a peak effective number of parties when G is 1.361, so that an increase in G from 1 to that peak is associated with an increase in N of 0.604 (from 5.705 to 6.309). As Figure 3 and Table 4 indicate, Canada demonstrates a statistically significant relationship between G and N, but the substantive effect is small. Increasing G from 1 to 2 is only associated with an increase of 0.079 in N. Moreover, the expected effective number of parties in Canada peaks at 2.829 (when G is equal to 2.169), an increase of just 0.081 in N from when G is equal to 1.

We can also estimate the magnitude of the "drop" in the number of parties at the highest levels of diversity by comparing the difference between the expected number of parties at the inflection point—that is, the highest point of the curve—and the expected number of parties when G is set to its maximum value. Among the pure plurality systems, the decline is particularly sharp in Britain, where the drop in the expected N is 1.154 (from 3.688 to 2.534) from its highest point (when G is 3.639) to the maximum value of G (6.287). The drop in the expected effective number of parties from the inflection point to the maximum number of groups is sizable in every case, ranging from 0.429 in Canada to 2.062 in Ukraine (although this decline is not statistically significant in Ukraine).

The same pattern exists in the United States in terms of the relationship between diversity and party fragmentation, but the substantive meaning of our

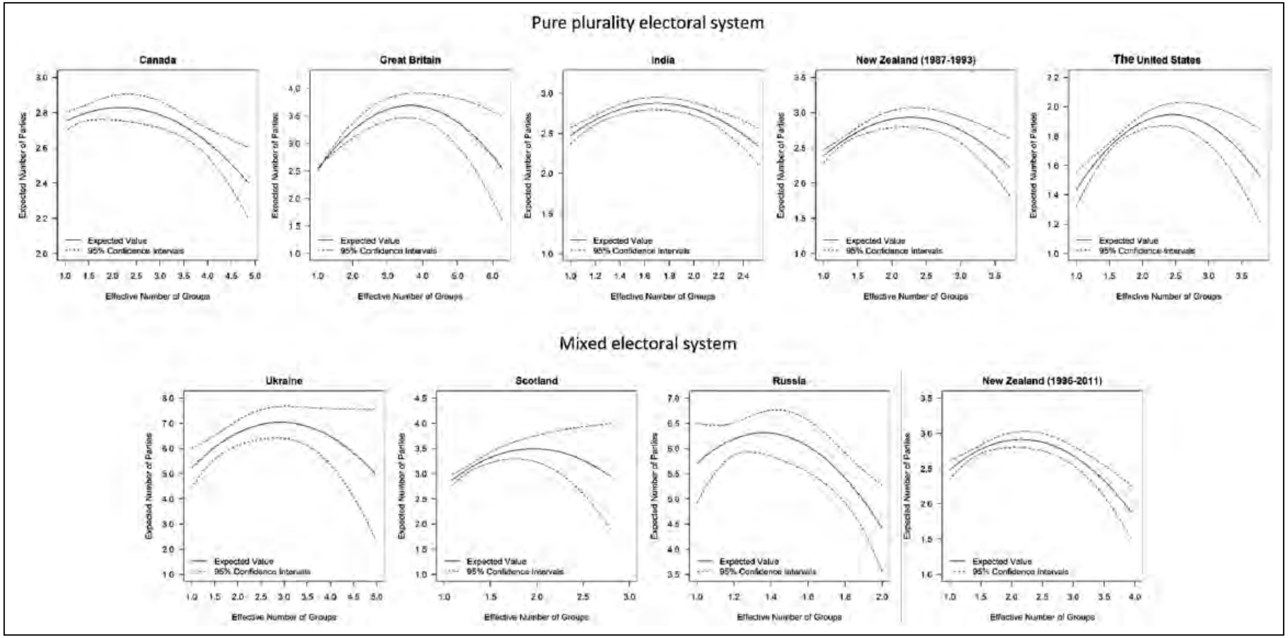

Figure 3. Predicted relationship between district-level diversity and party fragmentation.
The predicted effects are based on the estimates presented in Table 3.

Table 4. Expected Number of Parties (*N*) by Effective Number of Groups (*G*).

	$G = 1$	$G = 2$	$G = 3$	Inflection point	G = maximum	Districts (%) below inflection point
Plurality electoral system						
Canada	2.748	2.827	2.788	2.829 ($G = 2.169$)	2.400 ($G = 4.861$)	87.0
Great Britain	2.526	3.238	3.618	3.688 ($G = 3.639$)	2.534 ($G = 6.287$)	99.5
India	2.471	2.804	NA	2.875 ($G = 1.707$)	2.323 ($G = 2.531$)	82.0
New Zealand (1987-1993)	2.379	2.907	2.752	2.932 ($G = 2.273$)	2.203 ($G = 3.736$)	95.7
United States	1.435	1.896	1.875	1.946 ($G = 2.454$)	1.521 ($G = 3.788$)	88.1
Mixed electoral system						
New Zealand (1996-2011)	2.481	2.906	2.683	2.914 ($G = 2.154$)	1.870 ($G = 3.948$)	82.7
Russia	5.705	4.417	NA	6.309 ($G = 1.361$)	4.417 ($G = 2.000$)	67.1
Scotland	2.748	3.496	NA	3.496 ($G = 1.965$)	2.922 ($G = 2.812$)	97.4
Ukraine	5.220	7.039	NA	7.041 ($G = 1.961$)	4.979 ($G = 2.985$)	78.0

Based on coefficient estimates in Table 3. NA indicates cases where *G* never reaches 3. "Districts (%) below inflection point" refers to the percentage of districts in which effective number of groups is smaller than the value of *G* at the inflection point (where the effective number of parties peaks).

findings for the United States is quite different from the other cases we examine. Unlike the other cases, the effective number of parties in the United States is—as expected by Duverger and Cox (1997)—capped at two, with the highest expected *N* hitting 1.946. The curvilinear pattern we find in the United States makes great sense: Districts that are overwhelmingly White (homogeneous) tend to focus their votes on a single (Republican) party. Increases in diversity are for a time then associated with an increasingly two-party context, but highly diverse districts then see a large Democratic share of the vote, thus producing a much lower *N*. However, with only two parties (at most) seriously contesting elections and winning votes, there is no evidence that political actors are ignoring the strategic incentives of the restrictive FPTP rules in the United States. In contrast, all the other cases demonstrate a clear pattern where we observe a small number of parties in more homogeneous districts and significantly greater than two parties as the number of social groups increases. However, where there is substantial social heterogeneity, the number of parties declines again, which suggests that large numbers of actors are not acting strategically in the ways typically expected for FPTP districts.[19]

We are of course curious about the reasons for the inverted U-shaped pattern that we see across many of our cases, but ultimately the focus of this article is on demonstrating a relationship between diversity and the effective number of parties under FPTP. Indeed, our results show a strong positive relationship between *G* and *N* outside of the most diverse districts. Table 4 includes the "Districts (%) below inflection point," which indicates for each case the percentage of all districts whose effective number of groups is smaller than *G* at the inflection point (where the effective number of parties peaks). In every case, an overwhelming majority of districts are to the left of the inflection point in Figure 3. What this means is that in the expected values we plot in Figure 3, the effective number of parties continues to increase with the effective number of groups in all but a small number of (highly diverse) districts. Among the pure plurality cases, the smallest percentage of districts below the inflection point is in India at 82%: In India, the expected effective number of parties grows with increases in diversity until we reach the 18% most diverse districts. The largest percentage of districts below the inflection point is in Britain at over 99%, meaning the expected *N* continues to increase with *G* until we reach the 0.5% most diverse districts.

To address this point about the positive relationship between diversity and party fragmentation more systematically, we rerun our linear models from Table 2 but now drop the most diverse districts in each case. More specifically, we rerun each of the linear models from Table 2 a total of 7 times, first dropping the top 1% most diverse districts by case/election, and then dropping, respectively, the most diverse 5%, 10%, 15%, 20%, 25%, and 30% of the districts in each case/election (see Table 5).[20] For Great Britain, India,

Table 5. Linear Models of the District-Level Relationship Between Diversity and the Effective Number of Parties (Top 15% Most Diverse Districts Removed).

	Canada	Great Britain	India	United States	New Zealand		Russia	Scotland	Ukraine	Wales
	Plurality	Plurality	Plurality	Plurality	Plurality	Mixed	Mixed	Mixed	Mixed	Mixed
Effective number of groups (G)	0.204*	2.962**	0.492**	0.349**	0.452**	0.619**	−0.620	1.330**	1.491	−1.655
	(0.089)	(0.303)	(0.129)	(0.084)	(0.146)	(0.117)	(0.874)	(0.346)	(0.781)	(2.998)
Constant	2.534**	−0.478	2.069**	1.226**	2.002**	1.844**	6.897**	1.426**	4.123**	4.873
	(0.111)	(0.322)	(0.169)	(0.132)	(0.191)	(0.173)	(1.124)	(0.423)	(1.093)	(3.181)
N	1,044	3,233	1,327	1,477	259	295	188	248	183	136
R^2	.012	.105	.018	.022	.046	.066	.003	.068	0.020	0.004

Standard errors in parentheses. Dependent variable: SMD-level effective number of electoral parties (N). Effective number of groups (G) = SMD-level effective number of ethnic groups. Models used—Russia and Ukraine: OLS; Canada, India, New Zealand, and Scotland: OLS with standard errors clustered by district; Britain, the United States, and Wales: Cluster-specific fixed effects model with standard errors by clustered district. For Great Britain, the United States, and Wales, the R^2 corresponds to the within-cluster R^2 of the fixed effects model. SMD = single-member district; OLS = ordinary least squares.
*$p < .05$. **$p < .01$.

New Zealand (pre- and post-reform), the United States, and Scotland, the coefficient on G remains positive and significant every time. Earlier, we had noted that the substantive effect of G on N appeared especially small in Canada. In line with that finding, G is only statistically significant for Canada when we drop the 15% most diverse districts. That being said, for Canada the coefficient on G is always positive, and in the models in which we drop, respectively, the 20%, 25%, and 30% most diverse districts, the ratio of the coefficient to standard error is always quite large ($.050 < p < .10$). The coefficient on G in the Ukraine models is always positive and is always significant except when we drop only the 15% most diverse districts. For Russia, G has a positive coefficient once we drop the 30% most diverse districts, thus suggesting a positive relationship between G and N in the least diverse 70% of all districts.

We should note that our discussion here is not about the influence of outliers, but rather that the amount of diversity is likely to condition the relationship between G and N: Our analysis suggests that at high levels of diversity, there will usually be a negative relationship between G and N, but, otherwise, increases in diversity will promote greater party fragmentation.[21]

Conclusion

Scholars since Duverger have highlighted the overwhelming constraining effect of SMD, FPTP rules. Indeed, nearly all systematic empirical work on the effect of social diversity on the number of parties suggests that there is an interaction between electoral rules and diversity (see especially Amorim Neto & Cox, 1997; Clark & Golder, 2006; Cox, 1997; Duverger, 1954; Ordeshook & Shvetsova, 1994; Singer & Stephenson, 2009). That is, most analyses make a strong case that social heterogeneity leads to party fragmentation under permissive electoral rules, but that a psychological effect mitigates the power of social forces to promote party proliferation under FPTP rules. However, in this article, we provide substantial evidence that runs counter to most previous work: We demonstrate that even under FPTP electoral rules, there is a relationship between social diversity and the number of parties, thus suggesting that restrictive rules are not as powerful a constraint on electoral behavior and outcomes as is usually supposed.

Our findings are especially noteworthy in that they emerge from a new data set of *district-level measures of the central variables* that permits us to conduct cross-national district-level analysis of the relationship between social diversity and the number of parties in FPTP systems. Previous cross-national work on the topic founded its analysis principally on nationally aggregated measures, and therefore could not capture the direct relationship between diversity and the party system, which only exists at the district level.

Using the new data set, we find that increasing diversity usually leads to larger numbers of parties, but in districts with high ethnic diversity, we actually see, on average, a decline in the total number of electoral contestants. To be sure, the analysis of the downward slope of the curve that we find at higher levels of heterogeneity remains underdeveloped in this article, and future work will do well to consider in greater detail the causes of this pattern. However, whatever the pattern in the most diverse districts in each case, we find a consistently positive relationship between social diversity and party fragmentation at less high levels of diversity. Moreover, outside of Wales and Canada, the substantive effect of increasing social diversity is quite large, with increases of 1 in the effective number of groups seemingly producing increases of at least 0.333 (and often quite a bit more) in the effective number of parties. Indeed, with the exception of the United States, this analysis suggests that increasing social diversity helps drive the number of parties well beyond the two-candidate contests expected under FPTP rules by Duverger and most subsequent cross-national empirical studies.

This finding has several important implications for the study of party systems and the factors that shape them. First, this analysis contributes to a growing body of work that suggests that strategic defection by voters and elites is not a universal outcome under FPTP rules. For example, Cox (1997) highlights how particular conditions are necessary for strategic Duvergerian outcomes. And our analysis contributes to recent work that has found that plurality elections often do not constrain district-level electoral competition to two parties (Chhibber & Kollman, 2004; Diwakar, 2007; Gaines, 1999; Grofman et al., 2009; Moser & Scheiner, 2004, 2012). In many ways, these findings are not inconsistent with the words of Duverger (1954), who argues that FPTP "works in the direction of bipartism; it does not necessarily and absolutely lead to it in spite of all obstacles. This basic tendency combines with many others which attenuate it, check it, or arrest it" (p. 228). And, as our analysis suggests, in most cases social diversity is a major factor that "attenuates" the effect of plurality rules.

Our findings lead us to reject the commonly held strong version of the interactive hypothesis, but a weaker version of the hypothesis seems more plausible. That is, at the core of Duverger's theory about the effects of electoral rules is the idea that there ought to be more strategic behavior under FPTP than under other rules. Our analysis does not allow us to weigh in on this point, but future research would do well to consider it more systematically, especially in light of work (e.g., Abramson et al., 2010) that highlights the similarity of strategic voting patterns under both PR and FPTP rules.

Moreover, our findings open up new arenas for study. Earlier work suggested that social diversity had little or no effect on the number of parties in plurality elections. Thus, although it has been useful for observers to consider how district-level social cleavages become aggregated into two large parties (according to the predictions of scholars such as Duverger), there has been little reason to consider whether (and how) different minority (social) groups promote their own distinctive candidacies and how political elites might seek to attract them to options outside of the largest two parties. However, when we find evidence that social diversity affects the number of parties under FPTP, it becomes important to try to determine the ways that diversity actually becomes channeled into a variety of district-level candidacies in plurality elections.

Our dependence on ethnic diversity as our measure of social diversity is an important limitation of this article, and to decipher these mechanisms, future work will want to follow Potter's (2014) cue and compile district-level measures for a variety of different types of diversity. We believe that using ethnic diversity as a proxy for social diversity as a whole is a reasonable step to demonstrate that there *is* a relationship between social heterogeneity and the number of parties. Indeed, ethnic diversity has been the predominant proxy measure for social diversity in previous work that has established the interactive model as the conventional understanding of how social cleavages and electoral systems combine to shape party systems. Ethnic homogeneity tends to suggest low levels of social diversity, and ethnic heterogeneity suggests greater social diversity. In fact, using a blunt proxy measure for general social diversity creates an especially hard test, and our finding of a correlation between this proxy measure and party fragmentation provides especially strong evidence of a general relationship.

However, understanding the mechanisms that translate diversity into distinct parties requires more information on what types of cleavages are salient and the distinct ways that different salient cleavages affect party systems. For example, just because increases in ethnic diversity are associated with larger numbers of parties does not mean that parties use ethnic appeals to attract voters. Rather, ethnic diversity may become channeled into the party system by means of other types of cleavages, such as class: In many cases, districts that have a couple of different ethnic groups may be more economically diverse than wholly ethnically homogeneous districts—and thus may produce multiple parties that represent different income groups. But then districts with many ethnic groups may actually be more economically homogeneous if many voters within such districts have low incomes, and thus may promote the predominance in low-income districts of a single party (such as Labour Party in Great Britain) that appeals to lower classes. In this way, the curvilinear relationship, along with the success of the Labour Party in highly ethnically diverse districts, that we see in Great Britain and New

Zealand makes especially good sense if (a) low incomes predominate in highly ethnically diverse districts and (b) class is a salient cleavage.

Future work, therefore, will want to spend greater energy highlighting more thoroughly the mechanisms by which social diversity becomes channeled into the party system. Scholars will do well to ask, “To what extent does the number of parties track more closely to the amount of income (or other forms of) diversity in a district than to the degree of ethnic diversity?” Also, in a given polity, in what ways do parties (and candidates) try to appeal to voters and how do these appeals interface with the relationship between social cleavages and party systems? In highly ethnically diverse districts, do parties focus on ethnic appeals or, possibly, on class-based appeals, or both? Moreover, the menu of choices is likely to affect the links between social diversity and party fragmentation. The existence or absence of ethnic parties may have important implications for the relationship between diversity and the number of parties.

Finally, our fundamental explanation in this article for the correlation between social diversity and the number of parties even under FPTP rules is that many voters choose not to defect strategically even when electoral rules give them a strong incentive to do so. However, this lack of strategic voting—which Alvarez et al. (2006) demonstrate empirically—requires explanation. Future work will want to consider why there is so much less strategic voting than work by scholars such as Duverger (1954) and Cox (1997) argues should exist under FPTP rules.

Appendix

Sources of Subnational Census Data

Table A1 presents sources of the data for our district-level diversity measure, as well as the categories used to create the measure for each case. Where possible, we eliminate individuals placed in a generic “Other” category, those who did not state a category membership, and those who stated that they belong to multiple groups. In cases where we use data from multiple censuses, we take care to ensure that comparable categories are employed. In the case of Britain, we combine the subcategories of “White” groups in the 2001 and 2011 censuses to ensure that the categories used are comparable with the 1991 census.

In Canada, Great Britain, New Zealand, Scotland, the United States, and Wales, data were available at the level of electoral district directly from the source. In India, census data are collected within administrative units that do not correspond to national electoral districts. However, Jensenius (2016) uses

Table A1. Sources of Subnational Census Data and Categories Used.

Case	Year of census (elections)	Source	Categories used
Canada	2006 (All)	Statistics Canada (www.statcan.gc.ca)	Arab, Black, Chinese, Filipino, Japanese, Korean, Latin American, Not a visible minority, South Asian, Southeast Asian, West Asian
Great Britain	1991 (1987, 1992) 2001 (1997, 2001, 2005) 2011 (2010)	Office for National Statistics (https://www.ons.gov.uk/census) Scotland's Census Results Online (http://www.scrol.gov.uk)	Arab, Bangladeshi, African, Caribbean, Chinese, Indian, Other Asian, Other Black, Pakistani, White
India	2001 (All)	Jensenius (2016)	Christians, Buddhist, Hindu, Jain, Muslim, Sikh
New Zealand	1986 (1987, 1990) 1991 (1993) 1996 (1996, 1999) 2001 (2002, 2005) 2006 (2008, 2011)	Statistics New Zealand (www.stats.govt.nz)	Asian, European, Maori, Middle Eastern/Latin American/African, Pacific Islander
Russia	2002 (1995)	Collected by one of the authors from data provided by the Russian Census Bureau	Russian, Other minority
Scotland	2001 (1999, 2003, 2007) 2011 (2011)	Scotland's Census Results Online (http://www.scrol.gov.uk)	African, Arab, Bangladeshi, Caribbean, Chinese, Indian, Other Asian, Other Black, Other White, Pakistani, White British, White Irish, White Scottish
Ukraine	2001 (2002)	Collected by one of the authors from data provided by the Ukrainian Census Bureau	Hungarian, Romanian, Russian, Tatar, Ukrainian
United States	2000 (2002, 2004, 2006), 2010 (2008)	U.S. Census Bureau (http://www.census.gov)	American Indian/Alaska Native, Asian, Black/African American, Native Hawaiian/Other Pacific Islander, White
Wales	2001 (1999, 2003, 2007) 2011 (2011)	Office for National Statistics (https://www.ons.gov.uk/census)	African, Arab, Bangladeshi, Caribbean, Chinese, Indian, Other Asian, Other Black, Pakistani, White British, White Irish, Other White

Column 2 presents year of the census. The elections in the data set that correspond to the census are given in parentheses.

geographic information system (GIS) mapping techniques to identify the proportion of each administrative unit that overlaps with a given electoral district. These proportions are used to create area-weighted, district-level estimates of the variables in the 2001 Indian Census. We use Jensenius's estimates of religious groups to calculate the effective number of religious groups by electoral district. Finally, in the cases of Russia and Ukraine, our social diversity measures are mapped to electoral districts using data provided by the Census Bureaus in the two countries.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Supplemental Material

Supplementary material for this article is available online at the *CPS* website http://journals.sagepub.com/doi/suppl/10.1177/0010414017720704.

Notes

1. Singer (2012) introduces a different variable, the Third-First Loser (TF) ratio, which is the combined vote share received by all parties finishing fourth or worse divided by the vote share received by the second-place party. This variable is intended to capture electoral support for small parties. Using this variable, Singer (2012) finds some evidence that greater ethnic diversity increases support for minor parties in plurality elections (p. 215).
2. A number of studies indicate numerous ways in which the conditions may fail to exist. Ferree, Powell, and Scheiner (2014) highlight numerous electoral contexts in which Cox's conditions fail to obtain overwhelmingly. Chhibber and Kollman (2004) find no evidence that the number of votes for "hopeless" candidates increases in first-past-the-post (FPTP) districts with a near certain winner (p. 59), but they highlight how the other three preconditions that Cox highlights are apt not to hold. Chhibber and Kollman (2004) raise as likely the possibility that voters will use their votes, not to win the current election, but perhaps to keep a small party afloat for the future or to send a message to the leading parties (p. 59)—thus suggesting that many voters are not short-term instrumentally rational. Along with other scholars, Chhibber and Kollman express skepticism about the assumption of widely identical perceptions by voters of which candidates are out of the running. As Myatt (2007) shows formally, standard

formal theories may assume too much common knowledge when in fact voters may rely far more on private information, which can undercut their ability to coordinate strategically. Indeed, Blais (2002) shows convincingly that in the 1988 Canadian parliamentary election, large numbers of voters cast ballots for uncompetitive third place candidates because they overestimated the candidates' chances of victory. Interestingly, Potter (2016) highlights how in multimember districts increases in social diversity are associated with a lack of voter coordination around losing candidates. As such, there might be reason to expect more expressive behavior in more diverse districts even under FPTP. (At the same time, this logic might not hold where, as we suggest later, very high levels of social diversity can be associated with *fewer* parties.) Finally, there is evidence that in many cases, voters simply prefer their top choice by such a wide margin (with possible indifference about the alternatives) that they feel no compulsion to vote strategically. Chhibber and Kollman (2004) speculate about the presence in Britain of many moderate voters who cast their ballots for their uncompetitive but preferred centrist parties—and do so especially because of their distaste for those on the extremes (p. 58). This speculation is consistent with Blais's (2002) finding that large numbers of Canadian voters in 1988 strongly preferred their favorite candidate over all others and therefore did not strategically shift their vote away from their preferred choice even if the candidate was unlikely to be in the top two.

3. There are two especially noteworthy examples within the electoral system literature that run against the interactive hypothesis: In an analysis of legislative fractionalization, Powell (1982) finds evidence of an *additive* effect of diversity and electoral rules. Li and Shugart (2016) find evidence in support of a wholly institutional model, with district magnitude, assembly size, and the upper-tier seat share all shaping the number of seat and vote winning parties, and little effect of ethnic diversity. At the same time, both of these studies also focus on nationally aggregated measures of their variables.
4. Using different data sets, Jones (2004) finds no "noteworthy" effect of social heterogeneity on the effective number of presidential candidates, but Amorim Neto and Cox (1997) find that there is a relationship but only under runoff elections (and not plurality races). Using an expanded data set, Golder (2006) reports clearer evidence of a link between diversity and the effective number of presidential candidates, but again only under runoff rules. In models that control for presidential and central governmental power, Hicken and Stoll (2008) find a positive link between ethnic diversity and the number of presidential candidates under both runoff and plurality election rules, and Stoll (2013) demonstrates how the functional form of the relationship actually depends on the degree of presidential and/or central governmental power. Finally, as we discuss in greater detail below, Dickson and Scheve (2010) use presidential election outcomes to test an ingenious nonlinear model of the effect of social diversity.
5. Dickson and Scheve (2010) focus explicitly on how Duvergerian strategic behavior can promote a very different sort of nonlinear relationship between social

diversity and party fragmentation in plurality elections: More specifically, where a social group is sufficiently large that it can split its vote among more than one candidate (each of whom would represent the group) and still defeat a candidate of another group, then that (largest) group ought to be willing to run more than one candidate. In this scenario, the number of candidates peaks when the size of the largest group is quite large (and, therefore, the effective number of groups is small). However, once the largest group drops below 66% of the population (and the effective number of groups increases), the model posits that the total number of candidates will actually drop, as the largest group would risk losing the seat by running more than one candidate. In these scenarios, in contexts of relatively high social homogeneity, the relationship between social diversity and the number of parties should actually be negative. Dickson and Scheve find support for their model by testing it in the context of presidential elections, in which political actors have especially strong incentives to behave strategically. At the same time, Dickson and Scheve (2010), themselves, acknowledge that their model of highly strategic behavior may not hold in legislative elections (p. 365, fn. 20), which is the focus of the interactive hypothesis. Indeed, it is in the context of legislative elections, specifically, that work such as Alvarez, Boehmke, and Nagler's (2006) indicates that relatively few voters actually cast strategic ballots, thus further suggesting that Dickson and Scheve's model may be less likely to hold beyond presidential elections.

6. We limit our analysis to those cases for which we could obtain subnational-level data on both social diversity and electoral results within FPTP elections.
7. We use the term "Great Britain" rather than "The United Kingdom" because our data do not include districts located in Northern Ireland. For Scotland and Wales, we use the districts used to elect the members of the Scottish National Parliament and the National Assembly for Wales, respectively.
8. We present the sources of our subnational census data in the appendix, as well as additional information on the composition of our measure of social diversity for each case.
9. The Indian census has not collected data on race or ethnicity since 1941.
10. New Zealand, Scotland, and Wales use mixed-member proportional (MMP) systems, where parties' success in proportional representation (PR) voting is the principal factor shaping the total number of seats a party will win. In contrast, Russia and Ukraine use mixed-member majoritarian (MMM) systems, where each party's seat total is simply the sum of the seats it wins in the separate single-member district (SMD) and PR races.
11. Estimates based on the 2009 data compiled by the World Bank.
12. $N = 1/\sum(v_i^2)$, where v_i represents the proportion of the vote won by each candidate in the district.
13. For insightful exceptions, see Powell (1982), Potter (2014), and especially Stoll (2008, 2013).
14. The nature of our data varies across cases, and therefore we employ a number of different modeling strategies. For Russia and Ukraine, we use a simple ordinary

least squares (OLS) model, as we only have data from a single election for each of these cases. In Canada and India, we have multiple elections, but we measure the effective number of groups at a single point in time. For these cases, we estimate an OLS model and we cluster the standard errors by electoral district, as there is likely to be correlation between the errors of a given district over time. Finally, for Great Britain, New Zealand, Scotland, the United States, and Wales, we have data from multiple elections and the effective number of groups is measured at multiple points in time. In the cases of Scotland and New Zealand, the F test of the joint significance of the fixed effects intercepts indicates that we cannot reject the null hypothesis that all of the fixed effect intercepts are equal to zero. Here, we once again use an OLS model with clustered standard errors. Finally, for Great Britain, the United States, and Wales, we reject the null that all fixed effect intercepts are equal to zero. Therefore, for each of these cases, we estimate a cluster-specific fixed effects model that includes a separate intercept for each district to account for the within-cluster correlation in the data. However, given that the models do not generally account for all of the within-district correlation in the errors (Cameron & Miller, 2015), we also cluster the standard errors by the electoral district.

15. For each case, we performed a likelihood-ratio test to compare the fit of the linear and curvilinear models. With the exception of Wales, the inclusion of G^2 consistently improves the overall fit of the model compared with the linear specification.
16. In supplementary analysis, we also replicate Clark and Golder's (2006) analyses, which use the nationally aggregated measures of the key variables, precisely as they do, but we include the quadratic, G^2, that helps us recognize the existence of a curvilinear relationship between ethnic diversity and the number of parties. When we add G^2 (along with the interactions with G^2), we find a curvilinear relationship between diversity and the number of parties irrespective of the district magnitude for each replication model. In short, we analyze data used by previous work that appeared to show no relationship between diversity and the number of parties under nonpermissive rules—But when we consider the possibility that the relationship is nonlinear, we find strong evidence that social diversity affects the number of parties even under restrictive rules.
17. We omit Wales because neither G nor G^2 were statistically significant.
18. There are two cases in which this pattern is less definitive. For Scotland and Ukraine, the difference in N between the inflection point and the high value of G is not statistically significant, but the negative slope of the curve after the inflection point is statistically significant.
19. A number of factors make it likely that the United States will have a lower effective number of parties, typically hewing more closely to the two candidates (at most) per district expected by Duverger's Law. In the United States, the plurality rules used to elect the president, the concurrency of the presidential and legislative elections, and other campaign rules that privilege the Democratic and Republican Parties all give ambitious politicians strong incentive to affiliate with one of the

top two parties. Future work will do well to consider more systematically how such national strategic forces shape district-level behavior under FPTP. At the same time, it is striking that even despite the possibility of such national forces, there is a strong district-level relationship between diversity and the number of parties—even beyond two candidates per district—in all other cases.

20. For illustration sake, in Table 5 we present the coefficient estimates for models in which we drop the 15% most diverse districts for each case.
21. In addition, the impact and salience of diversity are unlikely to be the same everywhere, thus suggesting that the numerical cut-point between "high" and other levels of diversity will vary by context. Irrespective of what the precise cut-point is, we typically find a positive relationship between *G* and *N* past the most homogeneous 60% of districts in each case, and usually *well* past that line as well.

References

Abramson, P., Aldrich, J., Blais, A., Diamond, M., Diskin, A., Indridason, I., . . . Levine, R. (2010). Comparing strategic voting under FPTP and PR. *Comparative Political Studies, 43*, 61-90.

Alvarez, R. M., Boehmke, F. J., & Nagler, J. (2006). Strategic voting in British elections. *Election Studies, 25*, 1-19.

Amorim Neto, O., & Cox, G. (1997). Electoral institutions, cleavage structures, and the number of parties. *American Journal of Political Science, 41*, 149-174.

Blais, A. (2002). Why is there so little strategic voting in Canadian plurality rule elections? *Political Studies, 50*, 445-454.

Cain, B. (1978). Strategic voting in Britain. *American Journal of Political Science, 22*, 639-655.

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources, 50*, 317-373.

Chhibber, P., & Kollman, K. (2004). *The formation of national party systems: Federalism and party competition in Canada, Great Britain, India, and the United States*. Princeton, NJ: Princeton University Press.

Clark, W., & Golder, M. (2006). Rehabilitating Duverger's theory: Testing the mechanical and strategic modifying effects of electoral laws. *Comparative Political Studies, 39*, 679-708.

Cox, G. W. (1994). Strategic voting equilibria under the single nontransferable vote. *American Political Science Review, 88*, 608-621.

Cox, G. W. (1997). *Making votes count*. Cambridge, UK: Cambridge University Press.

Dickson, E., & Scheve, K. (2010). Social identity, electoral institutions, and the number of candidates. *British Journal of Political Science, 40*, 349-375.

Diwakar, R. (2007). Duverger's law and the size of Indian party system: A district level analysis. *Party Politics, 13*, 539-561.

Duverger, M. (1954). *Political parties: Their organization and activity in the modern state* (3rd ed.). London, England: Methuen.

Ferree, K. E., Gibson, C., & Hoffman, B. (2017). Why the salience of social divisions matters in party systems: Testing the interactive hypothesis in South Africa. *Party Politics*. Advance online publication. doi:10.1177/1354068817705124

Ferree, K. E., Powell, G. B., & Scheiner, E. (2014). Context, electoral rules, and party systems. *Annual Review of Political Science*, *17*, 421-439.

Gaines, B. J. (1999). Duverger's law and the meaning of Canadian exceptionalism. *Comparative Political Studies*, *32*, 835-861.

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 elections: The logic of party competition in Canada, India, the United Kingdom, and the United States* (pp. 115-134). New York, NY: Springer.

Golder, M. (2006). Presidential coattails and legislative fragmentation. *American Journal of Political Science*, *50*, 34-48.

Grofman, B., Bowler, S., & Blais, A. (2009). Introduction: Evidence for Duverger's law in four countries. In B. Grofman, A. Blais, & S. Bowler (Eds.), *Duverger's law of plurality elections: The logic of party competition in Canada, India, the United Kingdom, and the United States* (pp. 1-11). New York, NY: Springer.

Hicken, A., & Stoll, H. (2008). Electoral rules and the size of the prize: How political institutions shape presidential party systems. *Journal of Politics*, *70*, 1109-1127.

Jensenius, F. R. (2016). A sign of backwardness? Where dynastic leaders are elected in India. In K. Chandra (Ed.), *Democratic dynasties: State, party and political families in India* (pp. 77-96). Cambridge, UK: Cambridge University Press.

Johnston, R., & Cutler, F. (2009). Canada: The puzzle of local three-party competition. In B. Grofman, A. Blais, & S. Bowler (Eds.), *Duverger's law of plurality elections: The logic of party competition in Canada, India, the United Kingdom, and the United States* (pp. 83-96). New York, NY: Springer.

Jones, M. (1997). Racial heterogeneity and the effective number of candidates in majority runoff elections: Evidence from Louisiana. *Electoral Studies*, *16*, 349-358.

Jones, M. (2004). Electoral institutions, social cleavages, and candidate competition in presidential elections. *Electoral Studies*, *23*, 73-106.

Laakso, M., & Taagepera, R. (1979). "Effective" number of parties: A measure with application to West Europe. *Comparative Political Studies*, *12*, 3-27.

Li, Y., & Shugart, M. S. (2016). The seat product model of the effective number of parties: A case for applied political science. *Electoral Studies*, *41*, 23-34.

Lijphart, A. (1984). *Democracies: Patterns of majoritarian and consensus government in twenty-one countries*. New Haven, CT: Yale University Press.

Madrid, R. (2005). Indigenous voters and party system fragmentation. *Electoral Studies*, *25*, 689-707.

Moser, R. G., & Scheiner, E. (2004). Mixed electoral systems and electoral system effects: Controlled comparison and cross-national analysis. *Electoral Studies*, *23*, 575-599.

Moser, R. G., & Scheiner, E. (2012). *Electoral systems and political context: How the effects of rules vary across new and established democracies*. New York, NY: Cambridge University Press.

Myatt, D. P. (2007). On the theory of strategic voting. *Review of Economic Studies*, *74*, 255-281.

Myerson, R. B., & Weber, R. J. (1993). A theory of voting equilibria. *American Political Science Review*, *87*, 102-114.

Ordeshook, P., & Shvetsova, O. (1994). Ethnic heterogeneity, district magnitude, and the number of parties. *American Journal of Political Science*, *38*, 100-123.

Palfrey, T. (1989). A mathematical proof of Duverger's law. In P. C. Ordeshook (Ed.), *Models of strategic choice in politics* (pp. 69-92). Ann Arbor: University of Michigan Press.

Potter, J. D. (2014). Demographic diversity and district-level party systems. *Comparative Political Studies*, *7*, 1801-1829.

Potter, J. D. (2016). Constituency diversity, district magnitude, and voter coordination. *British Journal of Political Science*. Advance online publication. doi:10.1017/S0007123415000484

Powell, G. B. (1982). *Contemporary democracies: Participation, stability, and violence*. Cambridge, MA: Harvard University Press.

Raymond, C. D. (2015). The organizational ecology of ethnic cleavages: The nonlinear effects of ethnic diversity on party system fragmentation. *Electoral Studies*, *37*, 109-119.

Riker, W. (1986). Duverger's law revisited. In A. Lijphart & B. Grofman (Eds.), *Electoral laws and their political consequences* (pp. 19-42). New York, NY: Agathon Press.

Singer, M. (2012). Was Duverger correct? Single-member district election outcomes in 53 countries. *British Journal of Political Science*, *43*, 201-220.

Singer, M., & Stephenson, L. (2009). The political context and Duverger's theory: Evidence at the district level. *Electoral Studies*, *28*, 480-491.

Stoll, H. (2008). Social cleavages and the number of parties: How the measures you choose affect the answers you get. *Comparative Political Studies*, *41*, 1439-1465.

Stoll, H. (2013). *Changing societies, changing party systems*. New York, NY: Cambridge University Press.

Taagepera, R., & Grofman, B. (1985). Rethinking Duverger's law: Predicting the effective number of parties in plurality and PR systems—Parties minus issues equals one. *European Journal of Political Research*, *13*, 341-352.

Author Biographies

Caitlin Milazzo is an associate professor in the School of Politics and International Relations at the University of Nottingham. She is the author of *UKIP: Inside the Campaign to Redraw the Map of British Politics* (Oxford University Press, 2015), and her articles on electoral politics and political behavior have appeared in the *Journal of Politics*, *Electoral Studies*, *Political Behavior*, and the *Journal of Elections, Public Opinion and Parties.*

Robert G. Moser is a professor in the Department of Government at the University of Texas at Austin. He is the author of *Unexpected Outcomes: Electoral Systems, Political Parties, and Representation in Russia* (University of Pittsburgh Press, 2001) and *Electoral Systems and Political Context: How the Effects of Rules Vary Across New and Established Democracies* (Cambridge University Press, 2012). His articles on electoral systems, descriptive representation, and postcommunist states have appeared in *World Politics*, *Comparative Political Studies*, *Comparative Politics*, *Perspectives on Politics*, *Legislative Studies Quarterly*, *Electoral Studies*, and *Post-Soviet Affairs*.

Ethan Scheiner is a professor in the Department of Political Science at the University of California, Davis. He is the author of *Democracy Without Competition in Japan: Opposition Failure in a One-Party Dominant State* (Cambridge University Press, 2006) and *Electoral Systems and Political Context: How the Effects of Rules Vary Across New and Established Democracies* (Cambridge University Press, 2012). He has published articles on political parties, elections, and electoral systems across a range of political science and Asia studies journals.