Electoral Studies 40 (2015) 256–267

Contents lists available at ScienceDirect

Electoral Studies

journal homepage: www.elsevier.com/locate/electstud

The perils and pitfalls of ignoring disproportionality's behavioral components

Robin E. Best*, Andrei Zhirnov

Department of Political Science, Binghamton University (SUNY), 4400 Vestal Parkway East, Binghamton, NY 13902-6000, USA

ARTICLE INFO

Article history:
Received 28 January 2015
Received in revised form
25 September 2015
Accepted 25 September 2015
Available online 22 October 2015

Keywords:
Disproportionality
Electoral institutions
District magnitude
Measurement

ABSTRACT

Scholars commonly use a measure of the discrepancy between party vote shares and seat shares (observed disproportionality) as a proxy for the effects of electoral institutions. We illustrate the problems with doing so and, instead, recommend that scholars use more direct measures of institutional characteristics. Conceptually, we demonstrate that observed disproportionality cannot accurately capture institutional effects. Empirically, we show that (1) the variance of disproportionality is much higher when electoral rules are restrictive than when they are permissive, (2) the conclusions we draw about the effects of observed disproportionality differ substantially across samples of elections, and (3) replacing measures of observed disproportionality with more direct measures of electoral system characteristics such as district magnitude produces different and more reliable results.

1. Introduction

Electoral institutions vary in the ways in which they translate votes into legislative seats. It is common to think about these differences in terms of the proportionality of the vote-to-seat translation, which scholars then expect to affect a wide range of other outcomes of interest such as turnout (Jackman, 1987; Blais and Dobrzynska, 1998; Perez-Linan, 2001; Fornos et al., 2004; Henderson and McEwen, 2010), voting behavior (Carrubba and Timpone, 2005; Arzheimer, 2009), representational congruence (Blais and Bodet, 2006; Golder and Stramski, 2010), external efficacy (Davis, 2014) and policy outcomes (Bortolotti and Pinotti, 2008; Falcó-Gimeno and Jurado, 2011; Wigley and Akkoyunlu-Wigley, 2011). To examine the effects of an electoral system's (dis)proportionality on these other outcomes of interest, scholars commonly employ a measure of what we refer to as observed disproportionality: the discrepancy between party vote shares and party seat shares in any given election. Our aim is to demonstrate that such measures do a poor job of capturing the effects of electoral institutions and can lead scholars to draw faulty conclusions about the effects of electoral rules on other outcomes of interest. This is because any measure of observed disproportionality will invariably reflect the degree of strategic behavior of voters and political parties and will, therefore, capture the effects of the reactions to electoral rules. This is especially problematic under restrictive electoral rules, such as single-member-district plurality (SMD-p) systems, which can produce very low levels of disproportionality, very high levels of disproportionality, or anything in between, depending on the behavior of voters and parties.

In this paper we investigate the usefulness of observed disproportionality as a measure of institutional effects. We begin with a conceptual discussion of disproportionality as it relates to electoral institutions, where we demonstrate that observed disproportionality does a poor job of capturing the concept(s) of disproportionality. We then illustrate the empirical problems associated with employing measures of observed disproportionality as measures of institutional effects. Specifically, we demonstrate that: (1) measures of observed disproportionality have an inherently heteroskedastic relationship with electoral institutions, and that the degree to which they reflect institutional characteristics is typically quite weak; (2) measures of observed disproportionality are strongly determined by behavioral factors, which means that the conclusions we draw about the effect of disproportionality can differ across different samples of elections; and (3) substituting measures of electoral system characteristics (e.g., district magnitude) for measures of observed disproportionality can lead to very different conclusions about institutional effects.

We encourage scholars to think carefully about whether they are interested in the effects of electoral institutions or in the observed level of disproportionality in any given election. Our read of the

* Corresponding author.
E-mail addresses: rbest@binghamton.edu (R.E. Best), azhirno1@binghamton.edu (A. Zhirnov).

http://dx.doi.org/10.1016/j.electstud.2015.09.010
0261-3794/

literature suggests that scholars are often concerned with the effects of electoral institutions, and in those cases we generally advise against the use of measures of observed disproportionality as stand-ins for the effects of electoral institutions. Instead, we recommend the use of more direct measures of electoral system characteristics that affect the proportionality of the vote-to-seat translation.

2. Conceptualizing disproportionality

Our survey of the relevant literature suggests that it is common for scholars to include a measure of observed disproportionality as an independent variable used to capture some sort of institutional effect (e.g., Anderson and Beramendi, 2002; Arzheimer and Carter, 2006; Arzheimer, 2009; Blais and Bodet, 2006; Bortolotti and Pinotti, 2008; Carrubba and Timpone, 2005; Crepaz and Moser, 2004; Dahlberg, 2013; Ezrow, 2007; Jackman, 1987; Karp and Banducci, 2007; Kittilson, 2011; Kittilson and Schwindt-Bayer, 2010; Wigley and Akkoyunlu-Wigley, 2011).[1] It is usually difficult to discern the precise conceptual foundations of this measure because very few discuss the concept they seek to capture in any detail. In our read of the literature, when scholars ask about the relationship between disproportionality and other outcomes of interest they are often asking about the effects of the electoral institutions themselves: specifically, the mechanical effects of electoral institutions on the vote-to-seat translation. For instance, Bortolotti and Pinotti (2008) include a measure of observed disproportionality to capture the "barriers to entry imposed by different electoral systems (p. 339)." Carrubba and Timpone (2005) use disproportionality to measure "how disproportionately an electoral system translates votes into seats", which they explicitly refer to as an "institutional control (p. 268)." In his study of support for radical right wing parties, Arzheimer (2009) refers to disproportionality as an "institutional feature" and states that "disproportionality is of particular interest because the existing research seems to disprove the common wisdom that less proportional systems help to 'keep the rascals out' (pp. 263–264)."[2] Anderson and Beramendi (2002) use a measure of observed disproportionality to capture "institutional differences" in the "disproportionality of the electoral system (p. 723)." And Dahlberg (2013) appears to view disproportionality and district magnitude (an institutional feature) as interchangeable, stating "the degree of proportionality or the size of the electoral districts often has dramatic consequences for how parties organize and compete (p. 673)," which is then measured using an index of observed disproportionality. Furthermore, disproportionality is sometimes employed as a control variable to capture institutional effects (e.g., Falcó-Gimeno and Jurado, 2011; Pacek et al., 2009; Whiteley, 2011), and in these cases there is typically even less attention paid to the conceptual underpinnings of the measure.[3]

While there is good reason to expect the proportionality of the vote-to-seat translation under a set of electoral institutions to affect a wide range of political outcomes, there is little reason to believe that observed disproportionality is a good measure of these effects. This is because electoral institutions have both mechanical effects on the vote-to-seat translation and also psychological effects on the vote distribution. Duverger (1954) coined the use of these terms in discussing the tendency of single-member-district plurality (SMD-p) systems to reduce the size of the party system to equal two. The mechanical effects of electoral systems occur as a result of the way in which votes are translated into seats. The psychological effects of electoral institutions occur when voters, potential candidates, and other political actors observe or anticipate the mechanical effects of electoral systems and adjust their behavior to support only viable political parties.[4]

Under SMD-p rules, only the candidate with the largest share of district votes will win a legislative seat, which means that many candidates and parties contesting the election inevitably fail to gain representation due to mechanical effects.[5] As a consequence, voters who sincerely prefer a candidate with little chance of winning, but who are unwilling to waste their vote, will instead vote strategically for their most-preferred major party candidate. Candidates and small parties will either join with larger parties or avoid entering the contest if they are not electorally viable. According to Duvergerian logic, these strategic behaviors reduce the number of parties in a district to two: the number of viable contenders for one legislative seat (Cox, 1997).

Similar, but weaker, mechanical and psychological effects exist in proportional representation (PR) systems. Low district magnitudes, high electoral thresholds, and the electoral formula can all pose mechanical barriers to small parties entering the legislature and, as a consequence, create disproportionalities in the vote-to-seat translation (Taagepera and Shugart, 1989; Gallagher, 1992; Cox and Shugart, 1996; Cox, 1997; Benoit, 2000). Voters avoid voting for candidates and party lists that have little chance of entering the legislature, as well as for the obvious winners. Electorally nonviable candidates and parties – those that would not overcome the electoral threshold or do not expect to obtain any district seats – either refrain from contesting the election or coalesce with other parties (see, for example, Reed (1990) and Cox (1994) for the analysis of single non-transferable vote systems, Cox and Shugart (1996) for the analysis of strategic voting under proportional representation). Thus, one can think of the mechanical and psychological effects of electoral systems as a continuum, ranging from the highly disproportional and restrictive vote-to-seat translation of SMD-p systems to an almost perfectly proportional representation system, with most systems falling in between these two extremes.

When voters and candidates behave strategically, their observed behavior differs from the behavior that we would see if they acted sincerely. This means that the psychological effects of electoral systems can produce levels of observed disproportionality that deviate substantially from the mechanical disproportionality of the electoral rules. To aid our conceptual discussion we draw on two idealized electoral scenarios, which we present in Fig. 1. In the first

[1] In addition, observed disproportionality is often included as an institutional component in Lijphart's index of majoritarian-consensus democracy (Lijphart, 1994; see also Anderson and Guillory, 1997; Tavits, 2004; Vatter et al., 2014).

[2] See also Arzheimer and Carter (2006).

[3] In other instances, it is less clear whether scholars are referring to institutional effects or observed disproportionality. For example, in Jackman's (1987) analysis of turnout he includes a measure of observed disproportionality to capture "the degree of proportionality in the translation of votes into seats in the lower legislative house (p. 407)", which could be interpreted as both the actual, observed disproportionality of the system or the disproportional nature of the electoral institutions themselves. Without further clarification of the concept at issue, it is unclear whether there is an institutional component to this definition of disproportionality. This lack of clarity is found throughout the literature on voter turnout (for a relevant discussion, see Blais and Aarts, 2006). In our view, both observed disproportionality and the institutional effects of electoral systems are likely to impact voter turnout.

[4] Although originally Duverger talked about the national party systems, recent studies analyzed the mechanical effects and psychological effects at the level of electoral districts (see Reed, 1990; Cox, 1997, 1999; Singer and Stephenson, 2009; Singer, 2013).

[5] The winner-take-all nature of SMD-p also means that a candidate can win a seat even with a small lead in terms of votes, which creates some nonlinearity in the translation of major parties' vote shares into their seat shares at the national level. This effect is especially pronounced if there are only two major parties in the system and these parties are competitive throughout the country (e.g., Theil, 1970).