Party system effects on electoral rules

Josep M. Colomer
Abstract and Keywords

A long tradition of empirical studies has focused on the consequences of electoral systems on party systems. A number of contributions have turned this relationship upside down by postulating that it is the parties that choose electoral systems and manipulate the rules of elections. The most remote shaping of innovative electoral rules, the choice of electoral systems in new democracies, and further electoral reforms in well-established regimes can be explained on the basis of political parties’ relative strength, expectations, and strategic decisions. In a broader institutional context, political parties can also trade off electoral systems with choices and changes of other institutional rules.

Keywords: political parties, electoral system choice, Duverger’s propositions, micromega rule, plurality rule, multimember districts, proportional representation

Introduction

Political parties and electoral systems have been analyzed both as a cause and as a consequence of each other. First, a long tradition of empirical studies focused on the consequences of electoral systems on party systems. Most prominently, Maurice Duverger postulated that “old” political parties were created internally in elected (and also in nonelected) assemblies and parliamentary groups (Duverger 1951). Empirical analyses of party systems, which typically focus on democratic regimes from the mid-twentieth century on, usually assume that political parties derive from given elections and electoral systems that can be taken as the independent variable in the explanatory framework.

A number of contributions have turned this relationship upside down by postulating that it is the parties that choose electoral systems and manipulate the rules of elections. The origins of the invention and adoption of different electoral rules and procedures,
especially during the nineteenth and early twentieth centuries, can be found in the incentives created by political party competition. Further electoral reforms, as well as the choice of electoral systems in new democracies, can also be explained on the basis of political parties’ relative strength, expectations, and strategic decisions. With this approach, it is the political parties that can be taken as given and work as the independent variable to explain the emergence of different electoral rules. It can also be hypothesized, nevertheless, that changes in the party system—which is the focus of this chapter—tend to have significant effects on the electoral systems in the long term, while changes in electoral rules may have short-term effects on political party competition.

This chapter reviews, first, the most remote origins of political parties and how they began to shape innovative electoral rules; second, how different political party configurations influence the choice, permanence, or change of electoral rules; and third, the tradeoffs that political parties consider between electoral systems and other institutional rules.

Origins of Parties and Electoral Systems

To understand the origins of political parties and how they shaped innovative electoral rules, we must pay attention to the role of traditional elections in local settings. A very simple type of electoral system was used widely in local and national assemblies in predemocratic or early democratic periods before and during the nineteenth century. The essential elements of this type of electoral system are multiseat districts, open ballot (permitting the voter to vote for a number of individual candidates equal or lower than the seats to be filled), and plurality or majority rule.

This combination of rules has been widely used, especially in relatively simple elections with rather homogeneous electorates, and particularly at the beginning of modern suffrage regulations and for small-size local governments. It is still a common type of procedure for local or municipal elections in many countries, as well as in many meetings and assemblies of modern housing condominiums, neighborhood associations, school and university boards and delegates, professional organizations, corporation boards, and students’ and workers’ unions. In these, as well as in many of the traditional communities just mentioned, individual representation is well suited to contexts of high economic and cultural homogeneity in which it is relatively easy to identify common interests and priority public goods to be provided by the elect.

This set of electoral rules appears indeed as almost “natural” and “spontaneous” to many communities when they have to choose a procedure for collective decision making based
on votes, especially because it permits a consensual representation of the community. Such a simple electoral system is able to produce satisfactory and acceptable citizens’ representation. But it also creates incentives for self-interested, would-be political leaders to coordinate on “factional” candidacies or voting coalitions—in a word, “parties.” Under that system, forming or joining a “party” may increase the prospects of winning additional votes and seats. “Party” is thus defined here in a minimalist way that is not substantially different from traditional meanings of “faction” in early periods of voting and elections.

Under the previously identified set of rules, factions or parties tend to induce “voting in bloc” for a list of candidates, which may change election results radically. Once party candidacies are presented, it is not necessary that all or most voters follow the advice of factional leaders to vote for all of and only the members of a list of candidates to attain a party sweep. It may be sufficient that a few people do it, since, even if they are few, they can make a difference, especially under simple plurality rule where no specific threshold of votes is required to win. Note that, in historical terms, voting “in bloc” was not an institutionally induced behavior, but a party-strategy-induced behavior.

In some crucial cases, it was largely as a consequence of this type of experience that different political leaders, candidates, activists, and politically motivated scholars began to search for alternative, less intuitive, or “spontaneous” electoral rules able to reduce single-party sweeps and exclusionary victories. This new period began to develop by the mid-nineteenth century. It can be held that from that moment on, it was the previously existing political parties that chose, manipulated, and promoted the invention of new electoral rules, including the Australian ballot, single-seat districts, limited ballot, and proportional representation rules, rather than the other way around.

Virtually all the new electoral rules and procedures that were created since the nineteenth century can be understood as innovative variations of the previous, simple, “originating” system, which can be called “originating” precisely for this reason. The new electoral systems can be classified in three groups, depending on whether they changed the district magnitude, the ballot, or the rule.

The first group implied a change of the district magnitude from multiseat to single-seat districts, keeping both individual-candidate voting and majoritarian rules. With smaller single-seat districts, a candidate that would have been defeated by a party sweep in a multiseat district with plurality rule may be elected. Thus, this system tends to produce more varied representation than multiseat districts with voting in bloc, although less varied than multiseat districts with open ballot.

The second group of new electoral rules implied new forms of ballot favoring individual-candidate voting despite the existence of party candidacies, while maintaining the other
two essential elements of the traditional system: multiseat districts and majoritarian rules. In particular, by limited vote, one party can sweep as many seats as the voter has votes, but it is likely that the rest of the seats will be won by candidates of different political affiliation.

Finally, the third group of new electoral rules implied the introduction of proportional representation formulas, which permit the maintenance of multimember districts and in some variants also open or individual-candidate ballot. Single-transferable vote, double vote, preferential voting, and open ballots make individual-candidate voting compatible with proportional representation (Colomer 2007).

This discussion can clarify the conundrum of the relationship between party systems and electoral systems, as if they were related like the chicken and the egg. Actually, biologists have an answer for the latter inquiry: in the beginning was an egg, but a very simple one, not a chicken’s egg. Analogously, we can say that first, there were elections, but with very simple rules, not party elections. Once the parties were formed, they delivered more complex and varied “eggs,” or electoral systems.

The Parties’ Strategic Choices of Electoral Systems

This early history can be expanded for the further period since the late nineteenth century. In contrast to the basic assumption in the old tradition of electoral studies referred to earlier, it is the number of parties that can explain the choice of new electoral systems, in particular regarding the introduction of proportional representation rules, rather than the other way around.

In fact, Maurice Duverger himself briefly noted that “the first effect of proportionality is to maintain an already existing multiplicity” (Duverger, 1950, 1951, 344). In a review of Duverger, John G. Grumm (1958, 375) held that “the generally-held conclusions regarding the causal relationships between electoral systems and party systems might well be revised…. It may be more accurate to conclude that proportional representation is a result rather than a cause of the party system in a given country.” Leslie Lipson (1964, 343) developed a historical narrative from the premise that “chronologically, as well as logically, the party system is prior to the electoral system.” Stein Rokkan noted:

In most cases it makes little sense to treat electoral systems as independent variables and party systems as dependent. The party strategists will generally
have decisive influence on electoral legislation and opt for the systems of aggregation most likely to consolidate their position, whether through increases in their representation, through the strengthening of the preferred alliances, or through safeguards against splinter movements.”

(Rokkan 1968; see also Lipset and Rokkan 1967)

Arend Lijphart and Bernard Grofman (1984) explored the factors for the “not highly probable, but possible” changes and choices in electoral systems. Rein Taagepera (2003, 5) more recently also suggested a “causality following in the reverse direction, from the number of parties towards electoral rules.”

Some recent contributions in this approach compare not only different countries using different rules, as is customary in the previously cited works, but also every single country before and after the introduction of new electoral rules. By putting “the Duverger’s laws upside down,” Josep Colomer (2004, 2005) showed how previously existing political party configurations dominated by a few parties tend to establish majority-rule electoral systems, while multiparty systems already existed before the introduction of proportional representation.

The emphasis on this line of causality does not deny that existing electoral systems offer different positive and negative incentives for the creation and endurance of political parties. Precisely because electoral systems can have important consequences on shaping the party system, it can be supposed that they are chosen by already existing political actors in their own interest. Accordingly, it can be expected that, in general, electoral systems will crystallize, consolidate, or reinforce previously existing political party configurations, rather than (by themselves) generate new party systems.

We know that electoral systems based on the majority principle, which tend to produce a single, absolute winner and subsequent absolute losers, are riskier for nondominant actors than those using rules of proportional representation—a principle that was forged to create multiple partial winners and far fewer total losers than majority rules. In general, the “Micromega rule” can be postulated: the large will prefer the small and the small will prefer the large. A few large parties will prefer small assemblies, small district magnitudes (the smallest being one), and small quotas of votes for allocating seats (the smallest being simple plurality, which does not require any specific threshold), to exclude others from competition. Likewise, multiple small parties will prefer large assemblies, large district magnitudes, and large quotas (like those of proportional representation), which are able to include them within the legislature.

Changing electoral rules can be a rational strategy for likely losers or threatened winners if the expected advantages of alternative rules surpass those of playing by the existing
rules minus the costs of change. In particular, an alteration of the electoral system can be
more successfully promoted by parties with high decision, negotiation, or pressure power
under the existing institutional framework. This makes incumbent rulers submitted to
credible threats by new or growing opposition parties likely candidates to undertake
processes of institutional change.

Thus, it can be expected that in situations in which a single party or two parties
alternating or sharing power are institutionally dominant and expect to obtain or
maintain most voters’ support, restrictive rules based on majority requirements will be
chosen or maintained. Since this type of electoral rule tends to produce a single absolute
winner, it can give the larger parties more opportunities to remain as winners and retain
control—as can happen, in particular, during long processes of gradually broadening
suffrage rights and democratization, giving the incumbent rulers significant opportunities
to define the rules of the game.

In contrast, it may be that no single group of voters and leaders, including the incumbent
ruling party, is sufficiently sure about its support and the corresponding electoral
prospects in future contests. In other words, there can be uncertainty regarding the
different groups’ relative strength or it can be clear that electoral support is going to be
widely distributed among several small parties. Here changes in favor of less risky, more
inclusive electoral rules, such as mixed or proportional representation electoral systems,
are more likely to be promoted and established by the currently powerful actors in their
own interest. This tends to be the development preferred by new or newly growing
parties in opposition to traditional rulers, including, in particular, multiparty opposition
movements against an authoritarian regime. But it can also be favored by threatened
incumbent rulers to minimize their possible losses (Colomer 2004, 2005).

These analyses have been developed for the introduction of proportional formulas for the
elections of parliaments. Analogous strategies have been observed more recently for the
choice and change of rules to elect the president in separation-of-power regimes. In Latin
America, it has been observed that while dominant and large parties are likely to choose
plurality rule and concurrent elections, small parties are likely to choose nonconcurrent
elections with majority rule with a second-round runoff, which permit multiparty
competition. It has also argued that military rulers and military–civilian coalitions in
processes of redemocratization, as they may feel threatened by newly emerging party
configurations, tend to follow the logic of electoral choice of small parties (Negretto
2006).

When dealing with presidential elections, shifting partisan political fortunes have also
been placed at the center of analysis. Changes in the rules of the electoral game tend to
reflect the political self-interest of dominant political parties as defined in relation to
mounting electoral uncertainty. The impact of electoral reforms on party system change, in contrast, appears to be less consistent with the expectations derived from the more traditional literature. In particular, party system change in Latin America has generated institutional change more predictably than vice versa (Remmer 2008).

Additional caveats can be introduced. If the incumbent rulers are still sufficiently powerful, they may prefer mixed-member systems, rather than openly proportional rules, in the aim of reducing political fragmentation and limiting the potential for new entrants to the party system (Shugart and Wattenberg 2003; Reilly 2007). Also, the incumbents’ willingness to introduce electoral system change may derive not only from a serious threat to their dominance but also from their loss of control of the situation or some internal division of interests among its members. The rules for changing the rules also matter. If the incumbent government can change the electoral law without seeking a broader agreement, there is greater likelihood that partisan self-interest will dominate. On the other hand, change may be deterred by high barriers to change, such as the constitutionalization of the electoral system or the requirement of a parliamentary supermajority or a popular referendum for its change (Gallagher 2005; Katz 2005).

**Electoral Behavior and Institutional Choice**

By reassuming the two approaches referred to previously, it can be postulated that self-interested parties competing in elections can develop two strategies at the same time: behavioral and institutional. In the behavioral field, the basic decision is to create or not an electoral partisan candidacy. More specifically, to create a new partisan candidacy might imply either a new effort of collective action or splitting from a previously existing party, while not to create a new candidacy may mean entering a previously existing party, forming an electoral coalition, or merging with a rival party. In the institutional field, the decision is to promote or not a change in the electoral system, the two basic polar alternatives being either majoritarian or proportional representation rules.

Given the electoral system, the actors’ relevant strategy lies mainly in the behavioral field. Under the “originating” system, there were incentives for self-interested actors to create new partisan candidacies. If there is a majoritarian electoral system, the rational strategy is not to create a party. Instead, it makes sense to coordinate efforts with other would-be leaders and groups to form only a few large parties or coalitions in the system (typically two), each of them able to compete for offices with a reasonable expectation of success. By contrast, if the existing electoral system is based on the principle of proportional representation and is inclusive enough to permit representation of small parties, rational actors may choose either coordination or running on their own
candidacies, in the expectation that in both cases they will obtain the corresponding office rewards.

However, emerging parties may produce unwelcomed sweeps, coordination may fail, and, especially under majoritarian systems, lack of coordination may produce defeats and no representation for candidates, groups, and parties with some significant, real or potential, support among voters. In these cases, the alternative strategic field—the choice of electoral institutions—becomes relevant. Parties unable to coordinate themselves into a small number of large candidacies will tend to prefer electoral systems able to reduce the risks of competing by giving all participants higher opportunities to obtain or share power. Two-party configurations are likely to establish or maintain majoritarian electoral systems, while multiple-party configurations will tend to mean that political actors choose systems with proportional representation rules. Also, majoritarian electoral systems tend to restrict effective competition to two large parties, while proportional representation permits multiple parties to succeed.

This discussion brings together “institutional theories,” which include those about the political consequences of electoral systems, and “theories of institutions,” in this case regarding the choice and change of electoral systems. A “behavioral-institutional equilibrium” can be produced by actors with the ability both to choose behavioral strategies (such as a party, candidacy, or coalition formation deciding on electoral platforms or policy positions) and to choose institutions regulating and rewarding those behaviors.

In the long term, two polar behavioral-institutional equilibria can be conceived. In one, political actors coordinate into two electoral parties or candidacies under majoritarian electoral rules. In another, multiple parties compete separately under proportional representation electoral rules. Either of the two results can be relatively stable and durable, but there is no deterministic relationship able to predict which one is going to prevail (there are also other intermediate pairs of consistent behavioral and institutional alternatives, including imperfect two-party systems, mixed electoral systems, and so on).

In this approach the presumed line of causality is double. Two-party configurations tend to maintain or choose majoritarian rules, while multiparty systems tend to establish or confirm proportional representation rules. In turn, majoritarian rules induce the consolidation of two large parties, while proportional representation confirms the potential for the development of multiple parties.

A crucial point is that coordination failures can be relatively more frequent under majoritarian electoral systems than under proportional rules, especially for the costs of information transmission, bargaining, and implementation of agreements among previously separate organizations, as well as the cost of inducing strategic votes in favor
of the larger candidacies. With coordination failures, people will waste significant amounts of votes, and voters’ dissatisfaction with the real working of the electoral system may increase. Large numbers of losing politicians are also likely to use voters’ dissatisfaction and their own exclusion, defeat, or underrepresentation to develop political pressures in favor of changing to more proportional electoral rules.

In contrast, coordination failures, properly speaking, should not exist under conditions of flawless proportional representation. Even if the number of candidacies increases, each of them can expect to obtain about the same proportion of seats that they would have obtained by forming part of more encompassing candidacies. In reality, coordination failures are relevant under proportional systems to the extent that they are not properly proportional, particularly when small assembly sizes and small district magnitudes are used.

An important implication is that, in the long term, we should expect that most electoral system changes should move away from majoritarian formulas and in favor of systems using rules of proportional representation. Reverse changes, from proportional toward more majoritarian rules, may be the bet of some potentially dominant, growing, or daring party. But they can imply high risks for a partial winner to be transformed into a total loser, if its optimistic electoral expectations are not confirmed. This occurrence may be more frequent when actors are risk prone or badly informed about electoral systems. In a historical perspective, there have been increasing numbers and proportions of electoral systems using proportional representation formulas rather than majoritarian rules.

### Party System Configurations

Several discussions of specific political party configurations in which we should expect different types of electoral system change have been proposed. Stein Rokkan (1968) analyzed the origins of proportional representation by taking inspiration from Karl Braunias (1932), who distinguished two phases in the spread of proportional representation electoral rules: the “minority protection” phase, before World War I, and the “antisocialist” phase, in the years immediately after the armistice. This approach has been further developed with a focus on the turn of the twentieth century by Boix (1999); Blais, Dobrzynska, and Indridason (2005); and Cusack, Iversen, and Soskice (2007) regarding Europe, and by Calvo (2009), Wills-Otero (2009), and Gamboa and Morales (2015) for Latin America. Boix, in particular, analyzes the shift from the plurality/majority rule to proportional representation as a result of the entry of new voters (assumed to be left-wing voters) and a new party (socialist) at the turn of the twentieth century in the
Western world. He analyzes the shift from the plurality/majority rule to proportional representation as a result of the entry of new voters.

Kenneth Benoit (2004, 2007) sketches a more general model of electoral system change at political parties’ initiative. He assumes that the parties’ objective is to maximize their share of seats. Benoit gives some real-life examples of electoral system change and discusses some empirical implications of his model. He predicts that the electoral rule will be changed when a coalition of parties, which have sufficient power to change the rule, exists such that each party in the coalition would gain more seats under the new rule.

In Colomer’s (2005) analysis, parties will be interested in replacing majoritarian rules with more inclusive systems, typically proportional representation rules, if none of them can be sure of winning by a majority. In other words, electoral system change from majority rule to a more inclusive electoral system permitting representation of minorities can be a rational choice if no party has 50 percent of popular votes, which is the threshold guaranteeing representation under the former system. Configurations in which one party has more than 50 percent of votes have values for the effective number of parties of between 1 and 4, with an expected average at 2.5. Changes in favor of proportional representation will not take place with values of the effective number of parties below 2, for lack of powerful actors with an interest in such a change. Above four effective parties, maintaining or establishing a majority-rule electoral system would be highly risky for the incumbent largest party, and possibly not feasible either, due to pressures for an alternative system supported by a majority of votes.

We mentioned that the elections held immediately after the adoption of proportional rules tend to confirm, rather than increase, the previously existing multiparty configuration. However, the new rules transform the share of each party in votes into fairer party seat shares, thus making it more attractive for voters to give their support to new or emerging political parties. As a consequence, the effective number of parties tends to increase in the long term, thus creating further pressures in favor of maintaining proportional rules.

In further elaboration, Selim Ergun (2010) holds that for a change to occur, the government should be formed by a coalition. He finds that a change is more likely to occur when there is a larger number of parties and also when the spoils of office are shared equally among the members in the governing coalition. These results are extended to analyze partial reforms from a less proportional rule to a more proportional one.

Ergun finds that a change can also occur when the effective number of parties is between two and three. The crucial point is not necessarily the size of the largest party but the difference between the parties. In the case of three parties, the largest party is always
against a change and the smallest party always in favor, and thus the electoral system will be changed depending on the size of the median-sized party. In a three-party configuration, the second party in size will successfully promote proportional representation if its size is sufficiently distant from the largest party size and, as a consequence, it cannot guarantee leading a majority coalition with the third party under majority rule.

Looking to the question from the other side, some authors have discussed the political party conditions for establishing majority rule. Eunju Chi (2014) examines how party competition led to electoral reforms in Taiwan. The two larger parties, driven by the goal of maximizing the number of seats, formed a coalition and passed reform bills to change the electoral system from a single nontransferable vote and multiseat district system to a first-past-the-post mixed system. Konstantinos Matakos and Dimitrios Xefteris (2015) identify more general, favorable conditions for introducing majority rule, such as the expected vote share of the smaller parties, the high rents from a single-party government, and sufficient uncertainty over the electoral outcome.

Further contributions include the examination of situations in which the party system is weakly institutionalized and high levels of electoral volatility can be observed. For relatively recent democracies, such as in Eastern Europe, these situations deter the stabilization of electoral systems and make relevant actors prone to engage in reforms (Bielasiak and Hulsey 2013). The role of actors other than parties, such as voters, academics, and reform activists, has also been highlighted in specific cases (e.g., Renwick 2010; Leyenaar and Hazan 2011).

### Institutional Tradeoffs

Political parties usually address the maintenance or change of electoral rules in the broader context of the institutional system, thus developing tradeoffs with changes of other institutional rules. Most significant rules to consider are the extension of suffrage rights, the size or number of seats of the assembly, and the number of territorial governments in a federal-like structure.

The extension of suffrage rights was usually perceived as a potential source of emergence of new political parties than can challenge the status quo. The dominant parties in the existing system were wary of expanding suffrage if no parallel institutional rules were introduced to reduce the risk of being defeated and becoming noncompetitive parties in the new system. That is why many successful processes of expansion of suffrage rights were accompanied by reforms in the electoral rules.
A case for reference is the United Kingdom, the oldest democracy still using plurality rule in single-seat districts. Major change of electoral rules was prevented by avoiding the sudden introduction of universal (male) suffrage and following a long process of gradual enlargements of suffrage rights in the nineteenth and early twentieth century. Requirements of wealth, property, income, or literacy for voting were safeguards against a possible turnabout of the political and party system by the sudden irruption of new mass voters. While the political system was dominated by alternations in government between the Conservative and the Liberal Parties, the electoral system (at the time still including a number of multimember districts by plurality rule) was not challenged.

It was the emergence of the new Labour Party, broadly supported by recently enfranchised workers, at the beginning of the twentieth century that introduced demands for proportional representation. Yet, when in the late 1920s the Labour leaders forecasted a possible party victory under the existing electoral rules they turned against their own former proposals for proportional representation. The new alternation between Conservatives and Labourites since the end of the World War II had the effect of converting the Liberals, which were excluded then from government, into the heralds of electoral reform.

A different type of process could be observed in Germany and, even more clearly, in northern European countries such as Sweden, Norway, and Finland. In these countries, the sudden enfranchisement of a very large electorate was made compatible with appreciable degrees of political stability by the introduction “from above” by the incumbent rulers of new electoral systems favoring political pluralism. By the early twentieth century, proportional representation or similar institutional “safeguards” promoting multiparty politics were adopted. Governments were then able to rely on parliamentary coalitions in which centrist and moderate parties were expected to play a decisive role. In this way, the risk of instability and the threat of turnabouts were limited, and incumbent voters and leaders enjoyed continued opportunities of being included in government and maintaining a significant influence on the political process.

In these countries, suffrage rights were traded off with electoral rules. Proportional representation was adopted as an institutional safeguard in place of the traditional qualifications for voting rights. New electoral systems permitting multipartism were conceived to be protective devices. The Conservatives, the Liberals, or the Agrarians would become a minority, but they would not be expelled from the system as might be risked with a majoritarian rule. The incumbent rulers took the initiative of introducing universal suffrage “with guarantees” rather than witnessing their own defeat (Lewin 1989; Acemoglu and Robinson 2000; Colomer 2001; Ahmed 2013).
Another institutional rule that can be traded off with changes of electoral rules is the number of seats of the assembly. The total number of seats to be elected in a country may enlarge or restrict the number of viable parties and, more generally, the degree of inclusiveness of the political system. According to Rein Taagepera (2001, 2007), the number of parties in parliament, \( P \), is related to the number of seats in the average electoral district, \( M \), and the total number of seats in the assembly, \( S \). In his notation

\[ P = (MS)^{1/3}. \]

Taagepera’s focus on predicting the number of parties may suggest that the number of parties is always a dependent variable of the basic elements of the electoral systems. But his formula accepts two-direction lines of causality. It can indeed be turned the other way around to present the electoral system as derived from the number of parties. Specifically, \( M = P^4/S \).

This formula shows that the number of previously existing parties (which is raised to the fourth power) is more important than the size of the assembly to explain the choice of electoral system (as operationalized by \( M \)). As long as the size of the assembly is not manipulated, for a small country with a small assembly, just a few parties can be sufficient to produce a change of electoral system in favor of proportional representation. In contrast, for a large country and a large assembly, many parties would be necessary to produce such a result—as reform activists in the United Kingdom, for instance, know very well.

But we can also deduct from the previous formula that, for a similar number of parties, \( P \), the larger the country, and hence the larger the assembly, \( S \), the smaller the expected district magnitude, \( M \), can be. Very large countries, precisely because they have large assemblies, should be associated to small (single-seat) districts. For example, the institutional designers in large India are likely to choose single-seat districts, while the institutional designers in small Estonia are likely to choose multimember districts, typically associated with proportional representation rules. Thus, we should usually see large assemblies with small districts, and small assemblies with large districts.

The interest of this finding is that it is counterintuitive, since apparently small countries should have more “simple” party configurations, so that they could work with simple electoral systems with single-seat districts and majority rule in acceptable ways (actually this tends to happen in very small and micro-countries with only a few dozen thousand inhabitants in which only one or two significant parties emerge). But now we have an answer to the very intriguing question of why some very large countries, including the United States, in spite of the fact that large size is typically associated with high heterogeneity, keep small single-seat districts and have not adopted proportional representation. The answer may be that in large countries, such as Australia, Canada,
France, India, the United Kingdom, and the United States, a large assembly can be sufficiently inclusive, even if it is elected in small, single-seat districts. By contrast, in small countries, including Belgium, Denmark, Estonia, Finland, the Netherlands, Norway, Switzerland, and so many others, the size of the assembly is small and, as a consequence, the development of multiple parties has favored more strongly the adoption of more inclusive, large multimember districts with proportional representation rules.

The large size of a country and its assembly is also usually associated with a federal structure. In fact, the electoral system can also be traded off with federalism, as a high number of decentralized territorial governments can play an inclusive role of the variety of the population in large countries and make proportional representation unnecessary, as discussed in Colomer (2010). As always, the tradeoff is operated through human collective action. If in a large country multiple territorial governments are established, much political action will focus on those local institutions and it will be less likely that multiple political parties will be formed at the countrywide level. As a consequence, there will be less pressure to adopt a federal large assembly and a federal electoral system of proportional representation. The United States, with a very high number of fifty states and extremely decentralized nationwide political parties, is a case in point.

In contrast, in a medium-sized country with a unitary territorial structure but with a variety of economic interests or cultural allegiances among the intertwined population, the formation of multiple political parties may push for a sufficiently inclusive assembly elected by proportional electoral rules, rather than for territorial governments (like, say, in the Netherlands).

This tradeoff between electoral system and federalism can also involve the size of the assembly. If the electoral system implies high levels of political party pluralism, the general assembly must be sufficiently large to capture that pluralism. If, in contrast, constitution makers privileged the representation of varied territories but were wary of the perils of multipartism when they chose the electoral institutions (or just imported them from the colonial metropolis before new rules of proportional representation had been invented), then the size of the federal assembly can remain relatively small. This is, in particular, the case of the United States, which has the largest number of territorial units and both the smallest single-seat electoral district magnitude and the smallest assembly in proportion to the population. For each country size there can be multiple equilibrium sets of institutions, but each of the sets involving different combinations of institutional alternatives will be in equilibrium if it is consistent with certain quantifiable tradeoffs between institutions.

A potentially fruitful exploration to investigate is the population density of the country (i.e., the quotient between population and area). Low density usually implies territorial
dispersion of the population in several distant or separate groups. Even if the country’s population as a whole is relatively homogeneous in economic and cultural terms, the costs of governance related to physical distance might make federalism an advisable formula for a durable democracy. Australia, for one, would be an example of this. If, in addition to being dispersed, the country’s population taken as a whole is highly heterogeneous, it is likely that it will be concentrated into separated groups with relatively high degrees of internal homogeneity, as largely happens, for instance, in Canada. In any of the cases, nonideological territorial governments could be based on relatively homogeneous communities, which somehow would replace the representative and aggregative role of multiple political parties.

On the other side, high population density is likely to imply local heterogeneity, whether in economic terms, as high density is usually related to high degrees of urbanization and diversification of economic activity, or in cultural terms, which may be produced by recent migrations. With this type of structure, federalism may not be a suitable solution since the creation of small territorial governments would not reduce much the complexity of the communities, while multiple, nonterritorially based political parties able to represent different interests and values within a mixed population may require proportional representation electoral rules (related discussion about Africa appears in Mozaffar, Scarritt, and Galaich 2003 and Brambor, Clark, and Golder 2007). In fact, multiple parties successfully pressured for the adoption of proportional representation in the early twentieth century in a few medium-sized European countries, such as Belgium, Finland, Norway, and Sweden, soon followed by Austria, Denmark, Ireland, and Switzerland, and this has spread widely among new democracies in medium-sized countries across the world in recent decades.

In contrast, in large countries, federalism can be more effective for good governance and durable democracy than any variant of party systems or electoral rules. Large federal countries include, for instance, multiparty proportional Argentina, Brazil, Germany, and South Africa, as well as two-party majoritarian Australia, Canada, and the United States, while the extremely large size of India has forced both federalism and a multiparty system in spite of a single-seat electoral system (partly favored by significant territorial concentration of different ethnic groups in different states and electoral districts).

All in all, political parties tend to be the main actors in the choice of electoral rules, which they tend to develop in their own self-interest, although under the incentives and limits imposed by existing rules and procedures for institutional change. Broader sets of other political institutions, such as basic rules for democracy, voting rights, the representative assembly, or the territorial structure of the country, can also constrain the processes and outcomes of electoral system parties’ choice.
References


Josep M. Colomer

Josep M. Colomer is Research Professor in Political Science in the Higher Council of Scientific Research, Barcelona.