2

Components of Simple Electoral Systems

Elections determine in a democracy who the voters' representative agents will be. Before the voters can choose, however, someone has to have chosen an electoral system. This book will not focus on the reasons why a specific set of rules is found in a country. Rather, we will take them as given. Nonetheless, one of our goals with this book is to offer those who might adopt rules for elections where none exist (as in a transition from dictatorship to democracy), or reform those already in place in an established democracy, a deeper understanding of the main choices and what their consequences tend to be. In this chapter and Chapter 3, we offer a broad overview of the basic components of rules that make up the "electoral system."[1]

This chapter focuses principally on those electoral systems that can be considered "simple." We will offer a more precise definition of "simple system" (drawing on Taagepera 2007) later. The general idea is that among the many choices and components that can go into the design of an electoral system, as we review in this chapter and the next, simplicity means sticking to the basics and avoiding unusual combinations of distinct components. At times there may be valid reasons to stray from simplicity in design, but it is generally best to avoid overcomplicating an electoral system. When systems are made more complex, their effects on political parties and voters may be unpredictable.

Basic Choices

One of the biggest choices one would make in setting the rules of the electoral game is: Do we want the electoral system to enhance the seats of the largest vote-earning parties, or do we want the seat share of nearly every party to be very close to its vote share? This question lies squarely on the interparty dimension, and primarily concerns the tradeoff between "majoritarian" and

[1] Some portions of this chapter and Chapter 3 – particularly some of the tables – build directly off Taagepera (2007).

"proportional" systems (Lijphart 1999). Another question, which belongs mainly to the intraparty dimension, is: Do we want the primary representative agents to be parties as collective actors, or should voters be selecting candidates to represent them as individual representatives?[2]

The world of electoral systems has many ways of answering these basic questions, and the ways in which the electoral system answers them go well beyond a straightforward "this" or "that"; an almost infinite range of intermediate solutions exist. A good starting point, however, is to think of the tradeoff between plurality, also known as First-Past-The-Post (FPTP), and proportional representation (PR). If the electoral system is FPTP, only the candidate who gets the most votes in a district is chosen as a representative agent of that district. If the same party happens to have nominated the candidates who obtain these pluralities in more than half the districts, then the system is "majoritarian" in the sense of bolstering the position of a large party.

The more parties there are that run but fail to get very many district-level pluralities, the greater the tendency of the FPTP system to "manufacture" a majority, because three or more parties divide the votes, but two big ones get most of the seats. These are the dynamics that make up a tendency often known as "Duverger's law"[3]: big parties are favored in turning votes into seats under FPTP; seeing this, opponents of the leading party have an incentive to "coordinate" around one other party that can replace it.[4] Hence we might wind up with a two-party system, with not very many votes – and even fewer seats – remaining for parties other than the big two. Despite the term "law" it is not a given that such a party system will result, and in fact, several chapters of this book (especially Chapters 7–10) will cast doubt both on the empirical validity of the tendencies to which Duverger referred and on their suitability for being called "law."

On the other hand, if the system is PR, normally a party will not win a majority of seats unless it has a majority of votes. As we saw in our example of Poland, 2015, in Chapter 1, it is nonetheless possible for a party under PR to get a majority of seats despite less than half the votes. However, as reviewed there, other features aside from the PR method were involved in producing that result (notably the thresholds). Normally, PR rules reduce the chances of a majority party because even small parties can get seat shares that are quite close to their vote shares.

Additionally, the opportunity for quite small parties to win seats may encourage more of them to form, and voters will be more inclined to stick with them. Districts under PR systems need to elect more than a single legislator,

[2] See Carey (2009) for extensive treatment of the tradeoffs involved in collective versus individual accountability of legislators.
[3] Originally derived from the propositions laid out by Duverger (1954).
[4] For detailed analysis of electoral coordination, see Cox (1997, 1999).

in order to be able to divvy up the district's representation according to parties' vote shares. As more seats are available in a district, more opportunities exist for smaller parties to get at least one of them (as shown already in Figure 1.2 in the Introduction). Thus PR increases our odds of having a multiparty system, although the relationship is not deterministic.

On the intraparty dimension, do we prefer that voters select candidates or whole parties? In the FPTP system, every nominee of a party runs in a unique district, and hence receives votes as an individual. While parties can be cohesive and policy-focused even under FPTP, it is also very likely that the candidates and legislators elected in this way have a strong incentive to appeal to the parochial concerns of voters who reside in the district (Cain, Ferejohn, and Fiorina 1987). Thus FPTP may produce contradictory results – a majority party (interparty dimension) and local focus (intraparty). Where a given system ends up may be depend on many things other than the electoral system, such as the executive type (parliamentary or presidential), legislative rules, party rules, and so on.[5]

What about proportional systems? Here the intraparty rules make a big difference. In most PR systems, parties present *lists* of candidates. With the exception of very small parties under certain rules, parties want to nominate more than one candidate per district for the very basic reason that, by definition, PR means districts with more than one seat. A party – at least a larger one – may be able to win two or more seats in a district, but only if it nominated two or more before the election! Thus under PR rules, there is a further question that must be encountered in the design of the system: Should the voter select only a party, or should the voter be able to select one candidate (or perhaps more than one) from within a party (or sometimes from different parties)? The choice here will affect how candidates relate to voters and their party.

Rules within PR systems may allow the voter only to pick one party list, with no choice among the various candidates the party has nominated; this is called a "closed list," and is used, for example, in Israel and Portugal. The voter must take the slate as a whole, or leave it and vote for another party instead. In other systems, parties present lists, but voters may cast candidate votes, thereby having an intralist choice. Generically, these can be termed "preferential-list" systems (Shugart 2005a).

In the case of preferential-list systems, details vary in defining how the preference votes for specific candidates are counted in the process of determining winners. We explore these rules in greater detail in later chapters. One solution is the "open list," used for example in Brazil and Poland, where the identity of candidates elected from any party's list is determined entirely by the candidates' preference votes. Such rules may reduce the cohesion of the party, as candidates seek to court different groups of voters (Carey 2009). At the

[5] Readers wanting to explore these effects are referred to Carey (2009), Kam (2009), and Samuels and Shugart (2010).

same time it might broaden the appeal of the party by encouraging it to have candidates who can collect votes from these different groups.

Electoral systems may combine components in all sorts of ways. While this chapter is focused on "simple" systems that do not combine too many different components, it is worth noting that it is possible for a system to have all four of the basic components we just described all rolled into one. We can have a system in which some legislators are elected by FPTP yet with the overall nationwide seat allocation being proportional, thereby allowing voters to vote for both a candidate running in their local district and a party list. This complex combination typifies a type of electoral system called "mixed-member proportional," as used in Germany and, since 1996, New Zealand. These are complex systems, and are discussed in more detail in Chapter 3.

Components of Electoral Systems – Overview

Elections can apply to one position (e.g., president, governor), a few (local council) or several hundred (parliament in a large country). Voters may have to voice unqualified support for an entire party or for one or several candidates using a "categorical ballot." Alternatively, they may be able to give ranked preferences (first choice, second choice, etc.) over multiple candidates using an "ordinal ballot." There may be one or two "rounds" of voting,[6] by which voters are called back to the polls on a later date if the first round did not produce a decision, as defined by the rules.

We start by outlining the basic choices, and introducing some preliminary notation that will be needed for equations and formulas of later chapters. These basic components are: assembly size, districting arrangements, and allocation formula.

Assembly size. Some fundamental choices that pertain to elections are outside the electoral laws as such. Every democratic country needs electoral laws for the first or only chamber of its representative assembly. If a presidential regime is chosen, it also needs laws for presidential elections. If there is a two-chamber (bicameral) assembly and both chambers are elected, their electoral systems may be different in either small details or in more fundamental ways. To keep things tractable, this book focuses almost entirely on first (or sole) chambers. The book also focuses on *general elections*, rather than *primaries*, which are contests in which voters participate in the selection of candidates who will bear the party's label in the upcoming general election. Nonetheless, the basic choices of rules for general elections and first chambers also apply to primaries and second chambers.

6 In principle, there could be three (or more rounds), but in contemporary elections with public participation, we do not know of cases with more than two.

The first question is: How many seats should such a representative chamber have? Given that smaller assemblies offer fewer places over which parties can compete, the choice of *assembly size* (S) is critical. For this reason, assembly size is certainly part of the electoral system, yet before Lijphart (1994) analysis of the effect of electoral systems rarely included it as a factor. Even subsequent to Lijphart, few other than Taagepera (2007) and Li and Shugart (2016) have systematically recognized the importance of assembly size to electoral system effects.[7]

Even though the assembly size is a component of the electoral system, it is probably very rare for designers of electoral systems to consider assembly size directly when deciding what provisions to incorporate. It is generally a parameter that already has been set through separate, past decisions.

Large countries are almost bound to have more seats than smaller ones. Indeed, the *cube root law* for assembly size performs well as a statement of the relationship between population and assembly size, and has a quantitative logical model to back it up (Taagepera and Shugart 1989a):

$$S = P^{1/3}. \tag{2.1}$$

In words, the assembly size, S, equals the cube root of the country's population, P. In Figure 2.1, we see a data plot of the countries whose electoral systems will be the basis of the quantitative analysis in several later chapters of this book.

The solid diagonal line in Figure 2.1 represents Equation 2.1. At a glance, the best fit slope is steeper than the law predicts. However, focus on the cases above the horizontal line at S=50. The fit is better above that line. Below the line, all cases are English-speaking Caribbean island countries.[8] We do not have a theoretical reason for why these countries have systematically "undersized" assemblies, although this fact has been noted before by Lijphart (1990), who correctly noted that their small assemblies contribute to a tendency for very large seat bonuses for the largest party.

Figure 2.1 also differentiates presidential systems (triangles) from parliamentary systems (circles). There is some tendency for presidential systems also to have smaller assemblies than the cube root of their population.[9] It is unclear why.[10] When the Caribbean countries are omitted from a regression estimation (but presidential are included), we obtain an exponent of 0.349,

[7] For the given election, assembly size is fixed in advance, although under some electoral rules it can increase slightly depending on the outcome of the election.

[8] One such country, Jamaica, is above the line, with S>60 since the 1960s.

[9] For example, the United States (see Taylor, Shugart, Grofman, and Lijphart, 2014: 209–215). The undersized nature of the US House has drawn some scholarly notice, including Frederick's (2009) study of potential consequences of the populations of individual districts.

[10] To our knowledge, the question has not been addressed in any literature on institutional design.

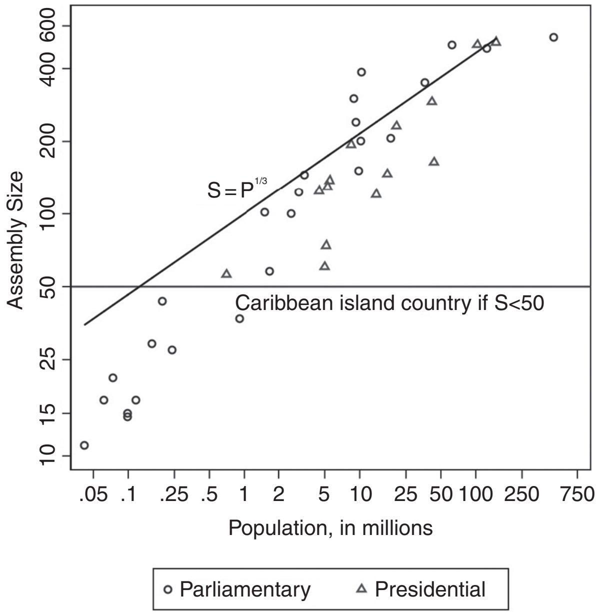

FIGURE 2.1 Population and assembly size

which is obviously a trivial difference from Equation 2.1.[11] Thus for the analysis of this book, we start off with a key parameter, assembly size, which can be thought of as, with some exceptions, exogenously set by the country's population size.

Districting arrangements. With assembly size thus essentially given, the next question is: Into how many electoral districts should the country be divided? Electoral districts mean the geographically bounded areas within which popular votes are converted into assembly seats. The number of seats allocated within a district is called *district magnitude* (M). It is arguably the single most important number for election outcomes. One can have numerous single-seat districts where M=1, or fewer multiseat districts where M>1. The limit is one nationwide district where district magnitude equals assembly size (M=S). All districts need not be of equal magnitude for the system to be simple; if there are overlapping districts, however, the system is complex. These matters are discussed further in Chapter 3.

[11] The regression has a statistically insignificant constant term of –0.189, and R^2=0.751. If the regression is run with presidential systems also dropped, the coefficient is 0.328! (And the constant becomes 0.02, also insignificant.) The insignificant constant is expected, because according to the logical-modeling techniques sketched in Chapter 1, we would be suspicious of a model with a significant constant. It would violate the obvious logical requirement that if there were only one person in a country, the assembly size would have to be S=1.

Many works speak of "single-member" and "multimember" districts. However, electoral rules allocate seats or memberships to candidates or parties – they do not allocate "members" as such. It is hence more logical to talk of single-seat and multiseat districts, and we shall do so throughout this book.

Allocation formula. A seat allocation formula defines the precise rules by which a party or candidate wins any one of a district's available seats, based on their votes. It is tied in with *ballot structure* and the number of *rounds* of voting, which we thus consider as subcomponents of the allocation formula.

The allocation formula stipulates how votes are to be converted into seats. At the one extreme, all seats in the district may be given to the party with the most votes (plurality rule). At the other extreme, one could use a PR formula in order to distribute the seats among multiple parties according to their votes. As we shall see, PR formula is itself a broad category including many specific variations that tend to produce more or less proportionality.

The voter may be given one or more votes. If only the first preferences are taken into account, the ballot is *categorical*. These are the simplest formulas, although more complex options exist. For instance, we already alluded to the possibility of *ordinal* (ranked) ballots. There are also electoral systems that allow voters to vote for two parties without ranking them – for instance a candidate of one party in their district and the list of another for seats allocated proportionally. We follow the suggestion of Gallagher and Mitchell (2005b, 2018) and call this a *dividual* ballot. As already noted, some allocation rules may require a second round – for instance, when no candidate obtains a majority of votes in the first round. These various features are discussed further in Chapter 3. For now, we keep the focus on simple systems.

What Makes an Electoral System Simple

Now we will begin to think about how the various components fit together, and what makes an electoral system simple or complex. Only categorical ballots and a single round of voting are simple, by our definition. The reason is that other ballot formats or multiple rounds, and the allocation rules that use them, may violate a basic criterion for simplicity of the translation of votes into seats: *the rank-size principle*. Under this principle, the relative sizes of parties in voting are reflected in the allocation of seats. That is, the first seat in a district always goes to the party with the most votes in the district. Then, if there are more seats ($M>1$), other parties that have sufficient votes may obtain seats, and they obtain them in order. That is, the next party to get a seat after the largest will be the one with the second most votes, and so on.

The rank-size principle may seem obvious. However, if an ordinal ballot is used, or there is a second round, it may be violated. For instance, under rules that use an ordinal ballot, the first seat may not go to the party that had the most first-preference votes. If there is a second round, the party that had the most votes in the first round may yet lose. There are other allocation formulas that likewise violate the rank-size principle. We discuss them in Chapter 3.

Assembly size, district magnitude, and *seat allocation formula* (plus the corresponding ballot structure) are the three indispensable features regarding which a choice cannot be avoided in designing a system to convert votes into seats. Taking the preceding factors together, a simple system is one in which:

(1) All seats are allocated in districts (as opposed to some being allocated in districts but others nationwide, for example); and
(2) The allocation rule respects the rank-size principle.[12]

Single-seat districts may look simple, but they still offer several choices for seat allocation – we will discuss some of them in Chapter 3. The simplest form of single-seat district is one in which the candidate with the most votes wins the seat – the plurality or FPTP system we mentioned earlier. This means a categorical ballot (vote for one) and a single round of voting, in which the candidate with the most votes wins the one seat. Even with these features, single-seat districts entail further complications arising from the need to delimit boundaries. We take these issues up briefly in the Chapter 3, along with other complicating factors.

In multiseat districts the options for how to elect representatives multiply; some of the options belong to the set of complex rules discussed in Chapter 3. Simple PR systems allocate seats among party lists. While it is possible to use plurality or another non-PR formula with party lists, such systems are not simple. We discuss them briefly here because they help focus our attention to the joint impact of our three essential components – assembly size (S), district magnitude (M), and allocation formula.

The plurality formula favors a large party for the very basic reason that it gives whatever number of seats are at stake in a district to the largest in that district. Usually when we think of a "plurality system" or FPTP, we mean that the country has S districts (one for every member of the assembly). The largest party in one district usually is not the largest in all districts, so two or more parties win seats in the assembly. Small parties, however, lose out unless they happen to be the largest in some districts. In Chapter 1, we gave the example from the Canadian election of 2008. In this election, under FPTP, four parties won seats in the assembly, but another party, the Green Party, won 6.8 percent of the votes nationwide, but not a plurality in any district, so no seats.

[12] This definition is only subtly different from that of Taagepera (2007: 19–20). The second criterion was not included in the previous definition, and Taagepera added closed lists and equal magnitude as criteria that he later relaxed.

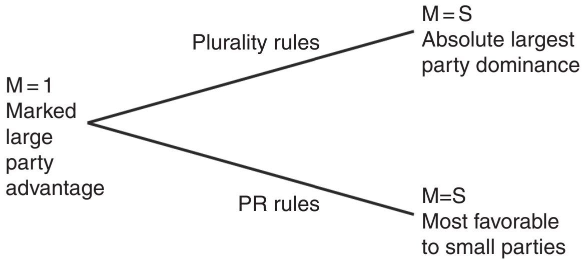

FIGURE 2.2 Contrasting effects of plurality and PR rules at the same district magnitude

Imagine instead that we had a single national district under plurality rule. Now we have $M=S$ so all the seats in the assembly go to the largest party: 100 percent of seats even if there was a close race among several parties to see who would be the largest. No country uses plurality in a nationwide district. Where the multiseat plurality formula is used, it is generally in low-magnitude districts. The main exception is the US Electoral College, where plurality is used statewide with some very high magnitudes (for instance, more than 50 in California).

When a PR formula is used, the effect of magnitude is reversed. The closest we can get to pure PR (for a given S) is when $M=S$. This means the entire country is one district. Two such examples are Israel and the Netherlands, both of which famously have many parties. With PR, a decreasing district magnitude increases the large-party advantage and hurts the small parties. The minimum district magnitude that is possible is one; and what happens when we use PR and $M=1$? Only the largest party in the district gets a seat. Thus single-seat plurality could just as easily be called single-seat PR! This is why we prefer to designate it as First-Past-The-Post (FPTP), a relatively neutral term between plurality and PR.

Figure 2.2 shows the overall picture for party lists in single- and multiseat districts. The contrast between plurality and PR allocation rules is extreme for a nationwide single district ($M=S$). Here plurality rule would assign all S seats in the assembly to the winning list, while PR rules would produce highly proportional outcomes. As the electorate is divided into increasingly smaller districts ($M<S$), the contrast between the outcomes of plurality and PR rules softens, until they yield the same outcome in the case of single-seat districts ($M=1$). We do not regard multiseat plurality as a "simple" system. Simple systems thus occur along the lower branch in Figure 2.1, ranging from FPTP to nationwide PR, with a wide range of districted PR systems in between.

In simple PR systems, those that use party lists, seats are allocated among the lists of parties competing in multiseat districts. While the focus of allocation across parties is on the votes earned by the lists, the ballot may allow voters to cast votes for individual candidates. These votes may affect which candidates win from the list – that is, they affect the intraparty dimension. For instance, instead of having only the choice of party lists (*closed lists*), voters may have the option to voice preferences for one or more candidates on a list (*open lists*), or they may be even required to vote for a specific candidate (as in the Polish case discussed in Chapter 1 and in Finland, discussed in Chapter 6). The so-called *panachage* (literally, cocktail) enables voters to mix candidates from across different lists – in effect, making up their own "free" list of preferred candidates.[13] Other allocation rules for multiseat districts that do not have party lists but have only candidates votes (whether categorical, ordinal, or dividual) are discussed in Chapter 3.

In sum, choosing an electoral system involves three inevitable choices (S, M, and allocation formula) and numerous optional ones. The ways to combine and mix them are infinite in principle and extremely numerous in practice. If many different options are mixed, the system may become quite complex. In the next section, we discuss in greater detail the basic types of simple systems, FPTP and PR.

FPTP AND PR AS "SIMPLE" SYSTEMS

In the preceding section we defined an electoral system as simple when (1) all S seats in the assembly are allocated solely within districts, and (2) when the formula follows the relative sizes of parties in votes as it allocates seats. The basic types of systems that meet the definition of "simple" are FPTP and any "districted PR" system, the systems on the lower branch of Figure 2.2. What these have in common is S seats allocated in districts, the number of which can range from one to S. If we have one district, we have nationwide proportional representation. If we have S districts, we have FPTP.

When the country is divided into single-seat districts, the only system that is truly simple is plurality, or FPTP, under which the winner in each district is the candidate with the most votes. Because any given party almost without exception nominates only one candidate per district when M=1, a vote for a candidate is also a vote for a party. Mechanically, it makes no difference whether we call this a system with only candidate votes, or a system of party-list votes in which each "list" contains only one name. The result is the same: the party with the most votes wins the seat, and its "first-ranked" candidate takes the seat. The ballot makes the choice categoric, and because there is only one round of voting, the winner is the one with the most votes, regardless of whether its total vote is more than half, and regardless of the margin over the runner-up.

[13] *Panachage* is clearly a complexity, but a few such systems (Luxembourg, Switzerland, and since 2005, Honduras) remain in our subsequent analysis of "simple" systems because their complexity primarily affects the intraparty dimension rather than the allocation of seats across lists.

The FPTP system is widely used in British-heritage countries. It is often associated with a two-party system, with a "bare-majority" (Lijphart 1984, 1999) governing party and a strong opposition. However, it would be hard to speak of a "law" because there are so many exceptions. Assembly size can make a difference. Large countries with large assemblies such as the United Kingdom (S around 650) and India (S around 540) can have more than two significant parties in the assembly.[14] If a country has significant regional variation, having many districts may help region-specific parties win. In small countries like Barbados (S=26) one party often has about 70 percent of the seats, leaving a weak opposition (Taagepera and Ensch 2006). Examples of FPTP systems are discussed in detail in Chapter 5.

Now, we turn our attention to multiseat districts, meaning those with magnitude greater than one (M>1). Under proportional representation, the goal is to make the seat shares of parties reasonably proportional to their vote shares. Simple PR means that each party presents a list of candidates, typically containing at least as many candidates as there are seats allocated in the district. Thus, under any party-list system "teams" of candidates run together, and seats are allocated first among competing lists, and only then to candidates (as we showed in the Polish example in Chapter 1). When, instead of party lists, the electoral system consists of voting for individual candidates in multiseat districts, the outcome, and the voting process itself become more complex. Thus we discuss candidate-based M>1 systems in Chapter 3, and focus on list systems here.

Under districted (simple) PR with M>1, the vote for candidate and for party is one act. If the vote is a categorical one for a single party list (closed list), the vote simply endorses the full slate of candidates, with the winners being determined based on the pre-election priority ranking on the party's list. When a PR system instead gives the voter a single categorical vote for a candidate (open list), it is similarly a vote for the party as well as for that candidate;[15] either way, parties receive seats on the basis of their total votes.[16] In this way, FPTP and list PR belong to the same family, and FPTP is just one endpoint of the continuum.

We discuss the distinctions among list systems in detail, using examples, in Chapter 6. The issue of intraparty seat allocation is important, e.g., for representation of women (Matland and Taylor 1997) and local interests

[14] In the case of India, the number of parties is especially large. However, many parties operate under the umbrella of two or three major alliances. We discuss this phenomenon in more detail in Chapters 5 and 15.

[15] Intermediary ways to allocate seats within the party (flexible lists) are also used; see Chapter 6.

[16] As mentioned earlier, some open-list systems allow the voter to vote for more than one candidate on a list (e.g., Peru). There are also the *panachage* (free-list) systems, in which voters may vote for candidates on several lists. What unites these systems under the rubric of "list" systems is that in all of them, seats are first allocated to lists, based on their collective vote totals, before turning to the intraparty allocation.

(Crisp et al. 2004, Shugart et al. 2005); we discuss a set of questions related to these intraparty considerations in Chapters 13 and 14. Our focus here will be on different formulas used to carry out PR among party lists (rather than allocation to candidates *within* lists); there are two broad subfamilies – *quota* formulas and divisor formulas.

Proportional Representation with Quotas

How should the seats be allocated among the parties?[17] Early advocates of PR came up with various quotas that still commonly bear their names: Hare (1859) and Droop (2012 [1869]). Hare's is perhaps the simplest of all, and hence is known by the name, *Simple Quota and Largest Remainders*, as well as *Hare-LR*. Under this quota, of the M available seats, a share 100%/M of the votes should entitle a party to a seat. This quota, also called exact or Hamilton quota, can be designated as q_0 for reasons of systematics that will become apparent later. Parties receive as many seats as they have full quotas of votes. These quotas are subtracted from the total vote shares, typically leaving vote remainders. Correspondingly, some seats remain unallocated; these are allocated to the parties with the largest remainders. Any party with a remainder of more than half-quota ($q_0/2$) is likely to receive such a remainder seat, but it depends on how the votes happen to be distributed among the other parties.

If we wish to have fewer remainder seats, we can reduce the quota. One might consider 100%/(M+1), designated here as q_1. Now it is possible to allocate more seats by full quotas – or even all of them. But one runs just a tiny risk of allocating more seats than the district has. Suppose M=4, so that q_1=20 percent. If parties' votes happen to be exactly 60.00, 20.00, and 20.00 percent, five seats would be allocated! To guard against this admittedly unlikely outcome, the *Hagenbach-Bischoff quota* adds 1 vote to the total votes, before calculating the quota, while *Droop quota* adds 1 vote to the quota itself (NOT 1 percent of all votes!), leading to q_{Droop}=100%/(M+1)+1 vote. This single vote makes overallocation of seats impossible. For practical purposes, the Droop and Hagenbach-Bischoff quotas are identical to q_1. If we are not concerned with overallocation, we can even use q_2=100%/(M+2) or q_3=100%/(M+3). Both have been referred to as *Imperiali quota* and were once used in Italy.

There also exist quotas larger than simple quota, e.g., q_{-1}=100%/(M−1), and so on, although they are not used in practice. Somewhat counterintuitively, *small quotas favor large parties*, while large quotas favor small parties. This is easiest to see when imagining extreme cases. Suppose again that M=4, and we decide to make use of q_{-3}=100%/(M−3) so that the quota is 100 percent. No one receives a quota seat. The four largest

[17] This subsection and the next draw liberally on Taagepera (2007).

 https://doi.org/10.1017/9781108261128.003

parties receive one remainder seat each, even when their vote shares are as unbalanced as 60, 30, 7, 2, and 1. Of course, q_{-3} is unrealistic, but unrealistic extreme cases offer a powerful conceptual tool in disciplines such as physics (Taagepera 2008). We will use extreme examples to illustrate general patterns of interest at many points in this book.

Subtraction is the basic philosophy for all PR formulas that are based on a quota and largest remainders: each time a seat is allocated to a party, a specified amount of votes is subtracted from its total votes. Alternatively, allocation can follow a *divisor* philosophy, which brings us to the second subfamily of PR formulas.

Proportional Representation with Divisors

Under a divisor formula, each time a party is allocated a seat, the rules state that its total votes be divided by a specified amount, before the allocation of the next seat is considered. The most widely used divisors are the *D'Hondt*[18] (Jefferson) divisors, 1, 2, 3, 4 … Table 2.1 shows how they work. Suppose M=6 seats are to be allocated, and the parties' percentages of the vote in the district are exactly 48, 25, 13, 9, 4, and 1. We divide all votes by the first divisor; because this is 1, it obviously does not alter the votes. Next, we allocate the first seat to the largest share, 48 percent, as indicated in Table 2.1. If this district had just one seat, we would be done! This reminds us of why FPTP is just PR (whether D'Hondt or another simple formula) in the smallest possible district. However, we have M=6 in our example. So now we need to divide the vote percentage of the party that won the first seat by 2. This reduces it to 24 percent for the next seat to be allocated. Now we compare this 24 percent to the other parties. The second largest party's 25 percent is larger, so it receives the second seat. Now its share is divided by 2. The next two seats go again to the largest party, with its share divided by 3 and then by 4.[19] The fifth and sixth seats go to the third and second largest parties, whose quotients (13 and 12.5, respectively) narrowly surpass the largest party's 12.

In this particular case, all major parties are overpaid: their seat shares exceed their vote shares. In general, however, D'Hondt divisors tend to favor the largest party. This advantage lessens as M increases. D'Hondt is one of the most widely used PR formulas. In addition to the Polish case discussed in Chapter 1, for example, Finland has used D'Hondt divisors for over 100 years (with mean M=13).

One can also use various other divisors. Faster increase in divisors reduces large party advantage. *Sainte-Laguë* (Webster) divisors (1, 3, 5, 7, …)

[18] Many works, including Taagepera and Shugart (1989) and Lijphart (1994), spell the name "d'Hondt"; however, the variant we are using here has been shown to be correct. See Gallagher and Mitchell (2005a: 632).

[19] A common mistake at this point is to divide 24 by 3, instead of dividing the original 48 by 3.

TABLE 2.1 *Allocation of six seats by D'Hondt divisors (1, 2, 3, …)*

Votes,%	48 (1st seat) 48/2=24 (3rd seat) 48/3=16 (4th seat) 48/4=12	25 (2nd seat) 25/2=12.5 (6th seat)	13 (5th seat) 13/2=7.5	9	4	1
Seats	3	2	1	0	0	0
Seats,%	50	33	20	0	0	0

Source: First shown in Taagepera (2007).

abolish this advantage, and the so-called Danish divisors (1, 4, 7, 10, …) actually favor the smaller parties.[20] To increase their seat shares, large parties might then split their lists strategically, but doing so might risk the splits becoming permanent.

In the other direction, the divisors could be increased more slowly. This would favor heavily the largest party. *Imperiali divisors* (1, 1.5, 2, 2.5, …) have been used. (Do not confuse them with the aforementioned Imperiali quotas!) The divisor series with the slowest increase would be 1, 1, 1, 1, … where the largest party wins all the seats. Thus, surprisingly, multiseat plurality rule surfaces as the extreme member of the divisor family of the PR formulas.

It is also possible to have divisors that tend to favor middle-sized parties. The *Modified Sainte-Laguë* divisors (1.4, 3, 5, 7, …) are used in some Scandinavian countries. Here the initial divisor 1.4 (instead of 1) makes it hard for small parties to receive their first seat. The quaintest divisors ever used might be the "modified D'Hondt" divisors used in Estonia: $1^{0.9}$, $2^{0.9}$, $3^{0.9}$, $4^{0.9}$, $5^{0.9}$, … They are equivalent to 1, 1.87, 2.69, 3.48, 4.26, …

Table 2.2 illustrates the effect of the various formulas from both quota and divisor families. It shows the allocations of seats when vote shares are again exactly 48, 25, 13, 9, 4 and 1 percent, as in Table 2.1, and the district has six seats.

The perfectly proportional seat share, shown at the top of Table 2.2, is fractional and can be only approximated. Visibly, allocations 3, 2, 1, and 3, 1, 1, 1, which occur in the center of the table, come closest to proportionality.[21] These two examples, D'Hondt and Sainte-Laguë, are the only divisor rules used fairly widely. Allocation formulas at the top of the table tend to overpay the largest party and are rarely used. Those at the bottom tend to overpay the small parties and are likewise rare.

[20] To see why this is so, consider an extreme case: let's increase the gap between divisors and use 1, 51, 101, 151 … then all *M* largest parties may win one seat each.

[21] Operational measures for deviation from proportionality will be presented in Chapter 4.

TABLE 2.2 *Allocation of seats in a six-seat district, by various quota and divisor formulas*

Votes,%	48	25	13	9	4	1
Perfectly proportional seat share	2.88	1.50	0.78	0.54	0.24	0.06
Steady divisors (1, 1, 1, 1, . . .) = plurality	6	0	0	0	0	0
Imperiali divisors (1, 1.5, 2, 2.5, . . .)	4	2	0	0	0	0
Modified D'Hondt (1, 1.87, 2.69, 3.48, . . .)	4	2	0	0	0	0
Imperiali quota q_3=100%/(M+3)=11.1%	4	2	1	[overallocation!]		
Imperiali quota q_2=100%/(M+2)=12.5%	3	2	1	0	0	0
D'Hondt divisors (1, 2, 3, 4, . . .)	3	2	1	0	0	0
Modified Sainte-Laguë div. (1.4, 3, 5, 7, . . .)	3	2	1	0	0	0
Droop/Hagenbach-B. quota q_1=14.3%	3	2	1	0	0	0
Sainte-Laguë divisors (1, 3, 5, 7, . . .)	3	1	1	1	0	0
Hare quota q_0=100%/M=16.7%	3	1	1	1	0	0
Danish divisors (1, 4, 7, 10, . . .)	3	1	1	1	0	0
Quota q_{-1}=100%/(M–1)=20%	3	1	1	1	0	0
Quota q_{-2}=100%/(M–2)=25%	2	1	1	1	1	0
Quota q_{-3}=100%/(M–3)=33.3%	2	1	1	1	1	0
Divisors 1, 51, 101, 151, . . .	1	1	1	1	1	1
Quota q_{-4}=100%/(M–4)=50%	1	1	1	1	1	1

Source: First shown in Taagepera (2007).

Since large quotas allocate all too many seats by largest remainders, while small quotas risk overallocation of quota seats, one may look for a *sufficient quota*. Start with Droop quota and reduce the quota gradually, until all seats are allocated by quota, with no need to consider the remainders. The result may surprise: *The remainderless quota is equivalent to D'Hondt divisors.*

So the D'Hondt formula occupies a central position on the landscape of PR formulas: it is at the crossroads of quota and divisor methods. This is how Thomas Jefferson actually came to define what later came to be called D'Hondt divisors in electoral studies (Colomer 2004a: 44). But Jefferson was concerned with seat allocation to the US states according to their populations. So were Alexander Hamilton, who first defined the simple quota, and Daniel Webster, who first defined the Sainte-Laguë divisors. All these approaches were reinvented in Europe when the need arose to allocate seats to parties according to their votes.

As district magnitude increases, the choice of the particular allocation formula makes less of a difference, because all formulas in the central range of Table 2.2 tend to produce seat allocations closer to perfect PR. In principle, these formulas can be applied to any district magnitude, up to and including nationwide allocation.

Going in the opposite direction, what happens if these formulas are applied to single-seat districts? All of them allocate the only seat at stake to the party with the most votes, and hence they become equivalent to FPTP. As seen in Figure 2.2, *FPTP is a limiting case of list PR*. Sharply distinguishing between FPTP and list PR in multiseat districts is artificial.

The official descriptions of seat allocation formulas in various countries often are confusing. For instance, the D'Hondt procedure can be speeded up by first allocating seats by full Hare, Droop, or Hagenbach-Bischoff quotas and then, instead of using largest remainders, switching to D'Hondt. The outcome is the same as D'Hondt which is, as we noted above, equivalent to a remainderless quota. Variants of this combination of some standard quota first, then D'Hondt on remainders, are found in Israel (Hazan et al. 2018), the Netherlands (Jacobs 2018), and Switzerland. Such procedures invite error in classification; these should not be classified according to their initial quota, but rather by their ultimate use of D'Hondt. It is thus valuable to understand the systematics of different allocation methods, as outlined in this chapter, in order to be able to put a given example into its broader family and context.

Conclusions

The central purpose of this book is to elucidate regularities in the impact of electoral systems on party systems. To that end, we started in this chapter with the simplest electoral systems, meaning all seats allocated within districts with a formula that allocates seats among lists of candidates according to the votes cast in the district. We noted that in single-seat (M=1) districts, a simple system means First-Past-the-Post (FPTP), also known as plurality. As district magnitude increases beyond one (M>1), a simple system implies the use of one of several proportional formulas; it is also possible to use plurality with M>1, but doing so takes us outside the family of simple rules.

Simple systems are important for several reasons. First of all, the very features that make them simple by our definition also make them simple for voters and party officials to understand. Simplicity is itself a virtue in electoral-system design. The more complications are added, the more difficult it may be for voters and other actors to learn how to make effective use of the system to obtain desired representation. Second, simple systems are the most straightforward for developing logical predictive models, a key aim of this book. In Chapter 3, we turn our attention to more complex rules, only some of which we are able to develop logical predictions about, in subsequent chapters.