Electoral systems in the making


Daniel Bochsler, University of Copenhagen, Dept. of Political Science, and University of Zurich, Centre for Democracy Studies Aarau

Introduction

Most research on electoral systems – also in this handbook – deals with the effects of institutions. It expands on Duverger’s (1951) influential laws about how electoral systems alter party competition. This ‘Duvergerian agenda’ has been challenged by the question of where electoral systems originate. Grumm (1958) responded that the choice of institutions is not exogenous: political parties pick the electoral systems which fit them best. Grumm’s argument has highlighted that electoral systems are endogenous institutions with regards to the party system (Colomer, 2005; Rodden, 2007), and it serves as the starting point for a growing literature on the selection of electoral rules for national parties.

Altering Grumm’s view that political actors when designing electoral institutions only secure and optimise their representation in parliament (for a formal model: Benoit, 2004), the literature shows that lawmakers navigate between self-interest and multiple normative ideals. Even if the public debate about electoral reforms are dominated by the “political idealism and the search for theoretical perfection in electoral systems” (Carstairs, 1980: 4), the normative ideals connected to electoral systems lead to trade-offs against other principles. Eventually, the socio-economic context (Iversen & Soskice, 2006) or the political parties’ strategic calculations (Kreuzer, 2010a) will decide on which principles prevail. I summarise the normative debate by discussing three key questions, which constitution-makers need to address when drafting an electoral system (Benoit, 2004: 369-370).

1 I am grateful for invaluable comments and suggestions by Lucas Leemann, Marcus Kreuzer and the editors of this handbook.

2 Gallagher (2005: 568-571) offers a compact overview of the criteria of electoral systems in the literature. Carstairs relates these criteria to the historical discussion (1980: 2-3)
The first normative question regards the equal representation of citizens in parliament. The electoral system should treat all parties equally, and the party with most votes should win most seats (McGann, 2006: chapter 3). The parliament should also be a mirror of society, where all groups are ‘descriptively’ represented on an equal (proportional) basis (Pitkin, 1967: 60). Second, governments should have the power to govern, but also be accountable to citizens. Single-party (or small coalition) governments profit from a united parliamentary majority, are not stuck in coalition bargains, allow for a clear attribution of responsibility, and are responsive to the median voter (Powell, 2000: 5). Thirdly, voters can elect their representatives directly, or representation relies on political parties. Electoral systems vary in the degree to which they offer incentives for representatives to cultivate a personal vote (Carey & Shugart, 1995). A direct election not only allows voters to have a more direct link to parliament, and makes representatives accountable to voters, but also helps to promote the representation of local district interests.

While in many contexts the first two principles will require strong political parties, this stands against the principle of non-mediated, personal representation. And institutions which are also inclusive for political minorities (equal representation), will not be best in terms of governability (Powell, 2000: 5-17). In the 1990s, mixed electoral systems were briefly hailed in the literature as a perfect electoral system – presumably performing well on all three ideals – there is no convergence of views among scholars or political agents about the ‘best’ institutional design (Farrell, 2001: 181-183). Rather, the importance of the goals differs between different contexts: In big, homogeneous nation-states, the principle of governability tends to prevail. Small states and culturally diverse countries instead place an emphasis on inclusive representation.

This chapter discusses how the historical and the political context shapes the principles that are emphasised, and accordingly the actors’ choices of electoral systems. The discussion is limited to national parliaments, and to key features of electoral laws, with regards to their effects on representation and type of mandates (universal suffrage, formula, magnitude, personal votes). For
a discussion of electoral system effects, we refer to Taagepera and Shugart (1989), Farrell (2001) or the IDEA Handbook (Reynolds, Reilly, & Ellis, 2005).

The next section offers a brief introduction to the origins of electoral systems in the early industrialised world. This is followed by a section on each of the main theoretical perspectives on electoral system choice, structure and agency. Section four sheds light on the choice of electoral systems in the process of political transition and consolidation of democracy, and section five discusses the personalisation of electoral systems.

**The origins of electoral systems in the early industrialised world**

While the choice of an electoral formula is highly driven by normative ideals and party strategies, constitution-makers select a formula from those familiar to them, and by considering the challenges to representation that a country faces in a given period. Thus, the choice is path-dependent on a century-old history of electoral laws (Bol, Pilet, & Riera, 2015; Blais, Dobrzynska, & Indridason, 2005). This section briefly outlines their evolution, with a focus on the early industrialised world.

Following Colomer’s (2004a) comprehensive wrap-up of the history of electoral laws, the world of electoral systems was divided between the plurality vote and its cousins in the Anglo-Saxon sphere, and the majority vote and proportional representation on the European continent and its colonies. This division has its roots in medieval decision-making, both in church and in politics. Unanimity rules were once the norm, but became impracticable as the decision-making bodies became more diverse, and politically split. Due to political conflicts in the Holy Roman Empire, as of the 12th century, the election body could no longer agree on appointing successors, resulting in the nomination of kings and anti-kings (Colomer, 2004a: 17). The alternatives to the unanimity rule which were developed included elections by lot, qualified majorities (usually two thirds, so that in order to change the decision, ‘the losers would have to persuade a majority of the winner’s original supporters to change their mind’ (Colomer, 2004a: 17), and – since late
medieval and early modern times – majoritarian decision-making. An English statute issued before the mid-15th century introduced the rule that the candidate with the “greatest number” of supporters should be elected (Colomer, 2004a: 28). Absolute majority rule was used for parliamentary elections in France and six further European countries (Colomer, 2004a). Plurality or majority rule was the practice widely used for the election of parliaments, often mixing single-seat and multi-seat districts.

Electoral systems followed step with the increased political importance of representative institutions, and the (slow) extension of suffrage. In parallel to the extension of suffrage, the electoral laws needed to establish conditions allowing voters to express their vote freely, in particular free from economic pressure by their employer or vote buying, and in a transparent and fair process (Rokkan, 1970: 153-154). This included election day procedures and measures against clientelism and election fraud. For instance, the establishment of vote secrecy, which prevents potential vote buyers from controlling how the votes they bought were cast (Kreuzer, 1996; Lehoucq & Molina, 2002; Mares, 2015).

With suffrage extension, elections became more competitive. This, in combination with plurality or majority rule in multi-seat districts lead to unpredictable and disproportional election results, and produced the demand for more proportional electoral systems. Parties representing the upper and middle classes were afraid that newly enfranchised voters would vote for the Socialists, meaning that they might lose elections (see subsequent sections).

One institutional strategy to encourage a more inclusive representation – primarily chosen in the English-speaking world – was splitting the districts into single-seat districts. This also allowed the old parties to strategically manipulate district boundaries in order to remain represented (Carstairs, 1980; Colomer, 2007; Ahmed, 2013: 65-78,96-97; Alesina & Glaeser, 2004: 111). In addition, Great Britain put the limited vote into practise in a few districts in 1867: in multi-seat districts, voters have a restricted number of votes, fewer than the number of seats available. This allows minorities to win representation. The system was introduced by economic elites in cities,
in order to guarantee their own representation in the course of the mobilisation of the working classes (Mitchell, 2005: 158-159; Bogdanor, 1981: 100-102). Going even further down the path towards proportional representation, in 1856 Denmark introduced the Single-Transferable Vote (STV) for a special council for the relations between Denmark and Schleswig-Holstein in 1856, although combined with a very restrictive suffrage. The STV does not rely on the logic of party list elections, instead voters rank candidates. Still, it pursues to the proportional representation of political minorities (Elklit, 2002: 30). Proportional representation (PR) became the most popular option for electoral reform in the 20th century.

**Why different countries need different institutions (structural theories)**

The move to proportional representation, or related systems, rendered representation more inclusive. The principle of political inclusion and broad representation became more salient in some countries than in others. This section addresses structural factors, which address these differences, and highlight aspects of political economy and social diversity, which motivate actors to pass electoral reforms.

The industrial revolution, and the enfranchisement of new voters, which led in some cases to the electorate growing tenfold in very few years, altered the structure of the party systems. Previously, social and economic communities were territorially segmented, and local notables were elected to parliament in order to represent the interests of their local community, primarily those of the local economic elites (Caramani, 2004: 151-153) With the industrial revolution and the enfranchisement of the working classes, the social conflict between labour and capital became dominant (Caramani, 2004: 212-213). This class cleavage – a ‘functional cleavage’ – no longer cut across territory, but lead to political competition within territorial units. This changed patterns of representation fundamentally. Territorial politics, with proto parties that were often based on clientelistic links, and on the representation of local interests through local notables, were
replaced by mass parties organised along social cleavages with a stable organisation and national outreach (Rokkan, 1970: 226; Caramani, 2004; Cusack, Iversen, & Soskice, 2010).

Rokkan’s theory is still prominently represented in the literature: in the run-up to and at the end of WWI, suffrage was extended to all male citizens, and in some countries also to women, and new social classes were politically mobilised. Proportional representation allowed the accommodation of new groups of voters, while ensuring that old parties would not lose their representation. The new electoral formula was attractive to the right: in the context of a growing left, it eliminated the need for electoral coordination among the right. In some countries, both the Socialists and the old established parties approved proportional representation and lower thresholds to win representation, as the Socialists wanted equal access to representation (Rokkan, 1970: 157) (see also the section on agency). In other contexts, the established parties were afraid of the revolutionary threat from the streets, and saw PR as a concession to the left (Alesina & Glaeser, 2004).

*Political Economy*

Going beyond Rokkan’s the focus on suffrage extensions, political economists investigate the introduction of PR in an economic context, with particular focus on labour relations3 (Cusack, Iversen, & Soskice, 2007). With majoritarian institutions, the economic right initially remained in power, and could control political decision-making. This was particularly rewarding in countries where economic interests were weakly organised (Cusack et al., 2007). In countries with highly specialised industries, where workers invested in industry-specific, specialised skills, the degree of economic organisation and coordination of the labour market was higher (Cusack et al., 2007; Kreuzer, 2010a: 370). According to some accounts, this also lead to the desire of political stability and autonomy from the state, which is best achieved through proportional representation and parliamentary systems (Rogowski, 1987: 212). Although, the historical evidence for this

---

3 The political economy literature on the choice of electoral systems has grown as a response to an earlier finding that proportional representation leads to more social spending (Rodden, 2007).
explanation relies on a limited set of countries in Western Europe, and is controversial (Kreuzer, 2010a).4

Katzenstein’s corporatist model (1985: 34-35,136-156), however, emphasises that the encompassing “historical compromise between business and labour” occurred during the 1930s and 1940s, after the small European states had introduced proportional representation. Thus, inclusive governments and proportional representation “facilitated the corporatist compromise of the 1930s” (Katzenstein, 1985: 156).

Cultural diversity and conflict

The literature on corporatism addresses in particular the Scandinavian democracies and a few Western European democracies. The latter – Belgium, the Netherlands, and Switzerland, are not only small, but also used to be divided by deep cultural cleavages along confessional, economic, and linguistic lines. They developed highly inclusive models of government. Proportional representation and corporatism were only two pillars; apart from this their governments included all major parties representing the main social pillars or segments in society. Decision-making was based on broad compromises, instead of majority decisions, often complemented with super-majority decision rules or veto rights for minorities, and federalism or political autonomy for the social segments. Arend Lijphart (1969, 1977) branded this inclusive type of political systems “consociational democracies”, and posited them as an alternative to the majoritarian ‘Westminster’ model in the Anglo-American part of the ‘free’ world. And he argued that consociationalism is the only system to produce stable democracy in divided societies. Re-labelled as ‘power-sharing’, this institutional package became the winning formula for conflict-solution in divided societies (McGarry & O’Leary, 2006: 48-54).

In divided societies, proportional representation of identity-based groups can be achieved in different ways. Lijphart (1986) defined proportional representation no longer as the electoral

4 See the following exchange between Cusack et al. (2010: 401) and Kreuzer (2010b). Further, Leemann and Mares (2014) have use detailed census data at the district level to operationalise the skill variable. At this level, they do not find an association, according to Cussack et al’s argument, and the vote to adopt PR.
formula which allocates seats in accordance to the vote share of parties, but as any electoral system which produces a parliament where cultural communities are represented (roughly) along their share of the population. Power-sharing arrangements today include electoral systems as diverse as the Single-Transferable Vote (STV) in Northern Ireland or the Single Non-Transferable Vote (SNTV) in Afghanistan. “Nongeographic” electoral districts, defined along ethnic lines, prevailed in Cyprus (Lijphart, 1986) and Montenegro (Bochsler, 2010). Lebanon in the 1960s defined an ethnic key, according to which lists of candidates needed to be defined in heterogeneous districts (Lijphart, 1986: 118-119). In Kosovo, proportional representation is complemented by strong affirmative action quotas for minorities (see Bird, 2014 for a comprehensive overview).

Deviating from the Lijphartian logic, several divided societies in the Australian sphere of influence have introduced majoritarian rules with ranked ballots (on these, and further examples see also the section on democratisation below) (Reilly, 2006: 101). They are advised by the school of centripetalists, who warn that proportional representation leads to a segmentation of political representation along ethnic lines, whereas majoritarian rules and/or preferential ballots encourage vote-pooling across ethnic-sectarian boundaries (Reilly, 2006; Horowitz, 2003; but see Fraenkel & Grofman, 2006a; Horowitz, 2006; Fraenkel & Grofman, 2006b).

Agency

While structural theories explain why the principle of inclusive representation is more important in some countries than in others, structural factors alone are insufficient to explain electoral reforms. Kreuzer’s (2010a) careful analysis of historical evidence for electoral reforms shows that cases, where structural factors alone have motivated both Socialists and right-wing parties to jointly support the introduction of PR, are very rare (e.g. Norway or Finland), and they were not economically motivated.

Electoral reforms, per se, are puzzling. Parliaments in almost all countries can not only propose electoral reforms, but also veto them. We would expect that governing parties either
have a preference to keep the rules which led to their election stable (Benoit, 2004), or to narrow representation, and produce clearer majorities, in line with the governability principle. However, in the empirical world, reforms occur which mostly strengthen the principle of inclusion, rather than governability. This section asks why governing parties, as legislators, enact reforms which (at first sight) shift power to the opposition and to smaller parties.

Reforms often occur in contexts when long-time losers occasionally become the majority, when radicals threaten democracy, or in periods of major changes to party systems.

*The parachute effect of proportional representation*

In a first set of cases of electoral reform, incumbents alter the rules to avert expected losses in the future.

“Since electoral laws are determined by policymakers, we should expect that the ruling political parties, anticipating the (varying) effects of different electoral regimes, choose the regime that maximises their chances of staying in power” (Boix, 1999: 611).

In anticipation of major electoral losses, incumbents will move from a more restrictive electoral systems (majority or plurality vote, related to a small party system) to more permissive ones, such as proportional representation. These allow for the representation of smaller parties according to their vote shares. This allows the incumbent to reduce the expected losses, as another party will gain in votes. In the literature, this explanation is usually related to the political consequences of the massive extension of suffrage around the turn of the 20th century. The inclusion of new voters also altered the preferences of enfranchised citizens (Boix, 1999: 611). Along with the suffrage extension in the UK in the 19th century, moves towards the representation of minorities were motivated by the fear of the Conservatives of losing their dominance in urban places:
“Giving the vote to ‘large masses in our large towns’, would result in ‘the great bulk of the property and intelligence’ being ‘found of necessity in the minority and not in the majority’” (Lord Cairns, cited in Bogdanor, 1981: 102).

In the context of the growing mobilisation of the working class, the ‘old parties’ of the centre-right introduced proportional representation (instead of plurality or the majority vote) under two conditions: first, they evaluated the mobilisation potential of the Socialists. The move only occurred if the Socialist party was strong, and perceived as a threat to the incumbents. Second, the move depended on the capacity of the incumbent parties to coordinate in elections. Facing a Socialist threat, the voters of the centre-right could try to coordinate strategically, and vote for the party which is most likely to win against the Socialist enemy. A problem of coordination emerged in cases where the right-wing bloc was split evenly between similarly strong parties (Boix, 1999: 112). Or, in the view of Blais et al., it was introduced in countries using the majority vote, where elections in multi-party systems lead to quite complex strategic coordination problems (2005: 184-185). Although, as Blais et al. (2005) object, most European countries which switched to PR previously used the majority vote with a runoff round and not the (single-round) plurality vote. In two-round electoral systems with a runoff, coordination between the main right-wing parties is easier.

**Critiques and amendments: districts, history, actors**

Recently, the literature has refined the “left threat” thesis, bringing the district perspective in, and looking more closely at the historical process and the actors involved in the introduction of PR.

Theoretically, the mechanics of electoral systems and party strategies play at the level of districts. This change of the unit of analysis allows scholars to investigate and control for possible confounders which are related to geography. The political cleavages separating the right-wing parties in the 19th and early 20th century were related to conflicts between the centre and the periphery, rural and urban areas, or to conflicts about the role of the church in society (Rokkan,
Both the centre-periphery and the rural-urban conflicts (and to a lesser extent the state-church conflict) are territorial, i.e. the parties have their constituencies in different parts of the country’s territory (Caramani, 2004). In the presence of such territorial conflicts, the electoral district design affects who wins and loses in the vote-seat translation. Typically, peripheral regions are overrepresented (Calvo, 2009: 272-281). Also, parties with a regional stronghold will select majoritarian electoral systems or those with small districts, while those with dispersed voters might instead opt for PR with large districts (Brady, 1992; Calvo, 2009). Accordingly, some of the old parties’ deputies, who could rely on strongholds in their districts, wanted to conserve the majoritarian systems (e.g. Ahmed, 2013: 172-173; Kreuzer, 2010a).

The strategic district component also leads to different views within parties on electoral reforms. Incumbents will seek re-election in very different electoral contexts: some are running in districts where their seats are safe, whereas others face fierce competition from new entrants in their districts. Hence, incumbents are cross-pressured from two (or three) perspectives, the electoral vulnerability of their party (i.e. in the future could the party gain or lose from the reform), possibly the position of their state party branch, and their own electoral vulnerability in their district (Leemann & Mares, 2014).

Historically, the move to PR came in several cases long before the Socialists constituted an electoral threat (Ahmed, 2013: 11). In many of the cases analysed, the move to PR was not a preventive action on the part of right-wing parties afraid of the newly mobilised voters, but instead resulted from Socialist support for the reforms, motivated by their under-representation due to uneven district sizes (malapportionment) (Blais et al., 2005). Furthermore, thinking in terms of election strategies, the reaction of the ‘old parties’ to the Socialists’ entry might also be related to the positioning of the parties in the space, i.e. the degree of hostility between the right-

---

5 Examples in the literature include the US, and how individual MPs voted on the Reapportionment Act 1929 (Bowler & Donovan, 2008: 98-100), the introduction of proportional representation in Germany (Leemann & Mares, 2014) and in Switzerland (Lutz, 2004: 286), vote secrecy in Germany (Mares, 2015). In Japan, incumbent deputies of the governing party blocked a reform which would have introduced single-seat districts (McElwain, 2008).

6 Examples include the move to PR in Denmark (in the shape of a MMP system) in 1915, after the Social Democrats and the Social liberals gained a majority (Elklit, 2002: 35-37).
wing parties, and the radicalism of the Socialists (Ahmed, 2013). Ahmed’s (2013) analysis differentiates the agency perspective and the Socialist threat argument. Thereafter, the right-wing parties differentiated their strategy between radical and moderate labour movements. They were primarily afraid of radical, revolutionary Socialists. After the introduction of PR in Italy, the left did not need to accommodate the median voter, and kept it both radical, but also weak (Ahmed, 2013: 78-81). This strategy also kept the left-wing parties weak. A counter-example with a moderate labour movement was Australia. Hence, “it could be treated as just another entrant into the party system, rather than as a socialist threat” (Ahmed, 2013: 72). The Liberals saw Labour as an equal suitable coalition partner as the Conservatives, and rather than opting for PR, they formed a joint government with Labour in 1903.

A last critique argues that much of the innovation in electoral systems is not based solely on domestic actors, but is driven by learning processes across borders (Blais et al., 2005; Bol et al., 2015).

**Narrowing representation**

While the first explanation – protection against future losses – leads to more inclusive electoral systems, a second explanation relates to a different principle of representation, governability, and leads to more restrictive rules. Governability is enhanced by the presence of clear, single-party majorities, which leads to clear responsibility and accountability. Electoral reforms occur when large political parties aim to limit the representation of political minorities. Highly fragmented party systems can render the formation of majority coalitions difficult, leading to unstable cabinets. Thus, in the presence, or after the historical experience of a proliferation of small parties, lawmakers can try to engineer an electoral system with high thresholds of representation (Kreuzer, 2004). Restrictive electoral systems can also be a means against the representation of parties that are judged as unacceptable for government inclusion: extremist parties or anti-system parties.
Either way, the majority in parliament can pick electoral rules which help to fulfil this goal. In the history of electoral system reforms, major moves from proportional to more majoritarian rules remain rather rare. Exceptions are France and Italy which are notorious for their frequent, and short-lived electoral reforms (Elgie, 2005; Renwick, 2010: 111-128,169-178). Some of these reforms were motivated by a partisan engineering of the parliamentary composition of parliaments.

In 1953 Italy introduced a law which granted a substantial seat bonus to the majority coalition. This was not only supposed hit the Communists, Monarchists, and Fascists (who were actually represented no real threat, as they were not able to gain a majority, or form a coalition), but was intended to allow the moderate coalition around the Christian Democrats to govern with a clear majority. The design of the reform was considered controversial even within the coalition of moderates. The smaller parties understood that strong majoritarian elements would eventually allow the Christian Democrats to govern with a single-party majority, and thus insisted on a more limited majority bonus. In the end the coalition did not gain enough votes in order to win the majority bonus, possibly because of the accusation of manipulation of electoral law (Renwick, 2010: 113-117).

French lawmakers were more successful with two reforms of 1951 and 1986. They sought to engineer a parliament that minimised the impact of radical parties. In 1951, a coalition of parties implemented (short-lived) mixed rules in order to contain the Communists and former General de Gaulle’s Rassemblement du Peuple Français (RPF), a political outsider party, which aimed to change the constitutional order. The rules were tailored to the regional distribution of the votes by the outsiders. The seats in the Paris metropolitan area, where the Communists were strong, were elected by PR, whereas a conditional block vote formula with multi-seat districts was introduced in all other regions. This gave the moderate coalition parties an important advantage (Renwick, 2010: 95-97). In 1986, after several rounds of reforms, and a one-election-intermezzo by proportional representation the French parliament agreed to return to runoff majority rule, in
light of the rising popularity of the radical right Front National (FN). The runoff system allowed
the old parties of the left and right to form a ‘cordon sanitaire’: in districts, where the FN is
sufficiently strong to enter with its candidate in the second round, the ‘old parties’ supported
each other’s candidates. With very few exceptions, this has prevented the FN from winning any
seats in parliament (Elgie, 2005: 120).

Other electoral systems also provide effective tools to establish an electoral bulwark against
rising polarising parties. Some electoral rules allow voters to express their dislike for a political
party or candidate – for instance systems with negative votes, or systems with ranked ballots,
where candidates gain a score corresponding to their relative rank (Borda Count). However, their
application is rare. In Nauru, where the Borda Count has been introduced for national
parliamentary elections in 1971, the legislators did not express any intention of excluding
extremists from representation (Reilly, 2002b: 363-366).

In countries with proportional representation, parties introduce high thresholds to limit the
access of small parties. Numerous democracies in Central and Eastern Europe have introduced
legal thresholds under PR in order to counter the proliferation of small political parties and the
entry of new competitors, mainly in the second or third multi-party elections (Moraski &
Loewenberg, 1999: 159-161; Shvetsova, 2003: 202), but the highest thresholds have been applied
in Turkey and in Greece. Other means of partisan engineering include, among others,
redistricting in electoral systems with small electoral districts, or changes to compensation
formulas in mixed electoral systems. In combination, in 2014, these two reforms allowed the
Hungarian government to win a two-third majority with a vote share below 45 percent (Csató,
2015).

**Systemic failure**

*Systemic failure and legitimacy - a phantom?*
The third explanation of agency-induced reform refers to crises of legitimacy of the representative system, which offer the opportunity for reforms. The process of electoral reform in New Zealand, moving from the plurality vote to a MMP system, is a case of this. It replaced a system built on single-party governments with one with broad inclusion and multi-party governments.

Plurality rule was widely supported in New Zealand, as it is related to a high governability and representation of the median voter. Two parallel developments brought it into disrepute. On the one hand, the Labour government broke its electoral promise on economic politics (Vowles, 2008: 177). This undermined the public belief that the plurality vote was conducive to high policy responsiveness to median voters. On the other hand, in two subsequent elections, 1978 and 1981, the strongest party in terms of votes (Labour) only came second, whereas the National party won an absolute majority of seats with a slightly lower vote share. The legitimacy of the elected parliament and of the government was at stake. However, as Shugart (2001, 2008) objects, legitimacy deficits – here a partisan bias and reversed majorities – are not sufficient to produce an electoral reform. Reforms only happens if one of the actors, or a coalition thereof, who are put at disadvantage by the current system, or hope to improve their situation in the future, control the decision-making institutions. Once Labour came to power in New Zealand in 1984, it initiated a process to review the electoral system. This process took on its own momentum, and in the end, citizens decided to change to MMP. At this point only the supporters of the National party (profiting from a partisan bias in the plurality vote) were opposed to the change (Vowles, 2008).7

So far, studies have failed to show how the loss-of-legitimacy argument can be generalised beyond specific country settings. There are other country-based studies that also find that losses in public support for democratic institutions were followed by institutional reform (Reed & Thies, 2001: 156). However, it seems not to be applicable in cross-national accounts. Both Norris (2011) and Bedock (2016) fail to find a correlation between the public evaluation of democracy,

---

7 However, with the parliamentary majority being the key actor to initiate a reform process, the referendum was the product of rather peculiar circumstances, and of Labour as the main loser of the previous system coming to power.
or even between drop in the legitimacy accorded to a democracy, and reforms. Renwick and Pilet (2016: 70-73) find that dissatisfaction with democracy leads to a move towards more personalised electoral systems (see below). The comparative analysis of electoral systems in Latin America shows that reforms mainly occur in periods of high electoral volatility, i.e. in periods of high electoral uncertainty, but also when established political parties become de-stabilised (Remmer, 2008).

*Are parties able to predict their electoral fortunes?*

A highly rational view sees political parties as agents who tailor institutions with the goal of increasing their own political power. However, political parties can not always rely on information to foresee their electoral stakes in the future (Renwick, 2010: 56-57), let alone to predict how this might impact their economic interests (Kreuzer 2010:17). While parties might sometimes rely on simulations of the short-time electoral impact of a decision (Leemann & Mares, 2014: 468), in political transition periods, the political insecurities are so large that accurate forecasts are close to impossible (Shvetsova, 2003; Andrews & Jackman, 2005).

In the process of post-communist transitions throughout the region, parties miscalculated their electoral prospects, and were surprised that the effects of the rules they supported often did not play in their favour (Moraski & Loewenberg, 1999: 161, 168-169). The Communist party, expected to perform well, but these hopes rarely materialised. PR, originally advocated by their adversaries, helped the Communists to remain represented (Moraski & Loewenberg, 1999), - and is the safest choice in the absence of reliable information about further electoral performance (Andrews & Jackman, 2005). In the Polish case, some of the parties that supported the introduction of a high electoral threshold under PR, in order to keep new parties out of competition, became the victim of their own rules.

However, political parties can also fail to realise that they need to reform a system, in order to remain represented in parliament. Non-reforms can end tragically. The Liberals of the United Kingdom have missed the last opportunity to change to a more proportional system. At the end
of WWI their vote in parliament was pivotal to deciding on a proposal that would have introduced PR. However, the Liberals, who were internally divided, the proposal was rejected (Bogdanor, 1981: 128-134). Soon after, the electoral gains of the Labour party inverted the picture, and under the plurality vote, the Liberals lost almost all their seats in parliament.

*When parties undermine their own legislation*

While in transition contexts, the outcomes come as surprises, elsewhere the surprises are a consequence of loopholes for strategic behaviour, which the legislators – willingly or unaware – built into the electoral systems. Essentially, *parties as legislators* of electoral reforms have not always considered their own future behaviour as *competitors* under the new electoral rules. The electoral consequences reforms depend heavily on the behaviour of political parties in electoral competition (Bochsler & Bernauer, 2014).

The Limited Vote (LV) and the Single-Non-Transferable Vote (SNTV) prevent large parties from winning all seats in elections, leaving some of the seats for the minority. While the SNTV should in theory work as a quasi-proportional system, the LV allows large minorities to gain some seats. However, this only occurs in reality if parties manage to engage in a complex coordination of candidates. They need to anticipate their electoral strength, nominate the right number of candidates, and make sure they all achieve roughly equal vote shares. Nominating too many candidates might result in none of them obtaining sufficient votes to be elected.\(^8\) The need of excessive ex-ante information and strategic coordination was also an obstacle when the Limited Vote was put in practice. When the strategy is played imperfectly, the result is a dramatic distortion of the representation principle, with minorities winning large majorities of seats (Bogdanor, 1981: 103-104; Mitchell, 2005: 158-159).

Some electoral rules also create loopholes, allowing parties in electoral competition to play them against the intention they had as parties as legislator. The most prominent loophole is provided by the Borda Count, where parties running with multiple (although chanceless)

---

\(^8\) For the Japanese case, see Cox (1996).
candidates can increase the electoral fortune of their main candidate. A reason why Jean-Charles de Borda, the father of the system, called it as “intended only for honest men” (Black, 1958: 182), and possibly one of the reasons, why it was seldom applied. Exceptions are Nauru, where an amended Borda Count was introduced in 1971 for national elections (Reilly, 2002b: 363-366), and Slovenia where it is used for the election of the two parliamentary seats reserved for the ethnic Italian and the ethnic Hungarian minorities.

Not unlike the Borda Count, mixed-member proportional systems are vulnerable to being derailed by the nomination strategy of political parties. Under usual circumstances, their election formula leads to a (roughly) proportional allocation of seats. However, if parties want, they can outsmart the electoral mechanism, and produce a distorted (disproportional) seat allocation in their own favour. In several elections in Italy, Lesotho, Venezuela, and Albania, the largest parties have presented their candidates under two different party labels. One label has been reserved for the candidates in the single-seat districts, the other for the proportional (party) list.9 The effect of this strategy is that the compensation mechanism – ensuring proportionality under usually circumstances – no longer works. Strategically acting parties win a disproportional large number of mandates both in the district tier and with their party vote, leading to massive overrepresentation (Katz, 2006: 296; Bochsler, 2012; Elklit, 2008). Italy, Albania and Venezuela abandoned their MMP systems such incidences.

Other actors: Referendums and courts

The emphasis on parties and parliaments is clear, given that they can initiate and veto institutional reforms, and given the theoretical puzzle behind why incumbents change electoral systems. However, on some occasions, the courts or the citizens – through direct legislation – have a say.

Despite the role of courts in initialising electoral reforms, there is surprisingly little comparative literature across countries. In the United States, the Supreme Court played a crucial

---

9 Under some legislations, this only works if the party presents two lists in the PR tier.
role in the implementation of the Voting Rights Act, aimed at abolishing discriminatory practices regarding the access of racial minorities to the polls, and it initiated a series of redistricting reforms for the Congress elections (Cox & Katz, 2002: 12-22,66-86; Bowler, Donovan, & Brockington, 2003: 15-18). In Europe, courts have intervened in Germany and Italy to demand for electoral reforms (Nuñez & Jacobs, forthcoming). Although there may be no case of courts having successfully brought about full systemic change (Renwick, 2010: 14), their rulings can initiate major constitutional earthquakes. In Bosnia and Herzegovina, twice in one decade courts demanded substantial alterations of the consociational order – based on the inclusion of three constitutionally recognised ethnic groups. In 2000, the Constitutional Court forced the subnational entities to adopt more inclusive institutions, and lead to the introduction of ethnic quotas (Merdzanovic, 2015: 288-289). A few years later, the European Court of Human Rights ruled that quotas for the election of the three-head-presidency violate the passive voting rights of citizens who do not belong to one of the three constitutional ethnic groups. The sentence might not only alter or unravel the entire political system of the country, but may also limit the capacity to establish quota-based consociational agreements elsewhere (McCrudden & O'Leary, 2013). However, at the time of writing, the implementation of the rule is still pending. In Switzerland, the Federal Court required that the last of 26 Swiss cantons introduce women’s suffrage in 1991, after men had repeatedly rejected it at the polls. The same court later pushed several cantons which practiced PR with tiny districts to move towards a lower natural threshold in PR elections. Following the court rules, today some of them apply probably the most proportional PR formulas worldwide.10

In several countries, the electoral law can be subject to direct legislation, even when this circumvents or blocks decisions made by the parliament. Again, Switzerland stands out. Both at the subnational level and at the federal level, Swiss men have repeatedly rejected the introduction of female suffrage, and only granted this right in 1971 in a mandatory referendum for the

---

10 They use a two-tiered compensatory system, with the Sainte-Laguë formula (Pukelsheim, 2009), and no legal threshold. Schwyz elects a parliament of 100 seats, Schaffhausen of 60 seats.
necessary amendment of the constitution. However, when the Swiss men decided their own
voting rights, they were more progressive. Already in 1918, they supported a proposal launched
by the political left for the introduction of PR (after three earlier attempts had failed) (Lutz, 2004:
286-287). An other example of the introduction of a proportional electoral system by referendum
is New Zealand. The decision to switch from FPTP was taken in several stages in a referendum
in 1992-3. Voters had the choice between four alternative electoral system, and opted for to a
mixed-member proportional system.

More numerous are reforms which failed in referendums. Voters in Ireland voted twice (in
1959 and 1968) against replacing the Single-Transferable Vote with a plurality vote system, and
voters in United Kingdom rejected a proposal to switch to the Alternative Vote in 2011. Italian
voters (in 1999 and 2000) and Romanians (in 2007) were asked to decide on electoral reforms,
although they did not turn out in sufficient numbers. Both countries use participation quora, i.e.
referendum results are only valid if at least half the registered voters participate (Renwick, 2010:
122).

**Beyond the industrialised world**

Many of the most prominent contributions in the field – even those making general claims –
limit themselves to the institutional choices of a few industrialised democracies in Western
Europe and North America, and possibly Australia. Even Central Europe, where several in the
period crucial for the choice of PR prior to WWII several countries were democratic, and thus
suitable cases for comparisons, is often omitted. This also affects the results of studies. Kreuzer’s
(2010a) replication study shows that the results of some prominent comparative studies would be
different had they used a more inclusive selection of cases. Beyond this, Central and Eastern
Europe and Latin America are both interesting laboratories for the study of the emerging of new
democratic regimes since the 1970s, and the process of institution building.

*The choice of rules for founding elections*
In founding elections, in the process of political transition, the political actors can often design their rules (almost) from a tabula rasa. The diversity of electoral systems employed worldwide has grown substantially since the end of WWII, and thus nowadays there are many more solutions legislators can chose from. When available, they sometimes reinstate the institutions from an earlier democratic order (Elster, Offe, & Preuss, 1998), but mostly little remains from earlier practice, and a democratic model is selected from those which are known to the legislators. Legislators can take the form of representative institutions which have survived from the previous regime, but often no such institutions exist or they have ceased to function, so that the new institutional design is selected by agreement between the parties at a roundtable, or simply decreed, as in Russia in 1993 by president Boris Yeltsin (Remington & Smith, 1996).

**International actors**

National actors are not the only ones to play a role in electoral engineering in the course of the political transition. Often, former colonial powers pass their own institutional traditions on to their former colonies (Blais & Massicotte, 1997; Lundell, 2010: 59-67). After the Cold War, the United Nations with its machinery to administrate elections, and its presence at the negotiation table with domestic constitution-makers became an important actor in the choice of electoral systems in post-conflict societies. Proportional representation was part of the standard prescription of UN-negotiated peace agreements (Reilly, 2002a: 129-130; Reilly, 2006: 78-79).

**Asia**

International impacts – in two ways – explain parts of the electoral systems map in Asia. Former British colonies (e.g. India, Malaysia, Singapore) inherited the plurality vote – albeit adjusted to the structure of their societies, for instance with mechanisms of ethnic engineering and opposition representation in Singapore. The system evolved into the Alternative Vote in the Australian influence sphere (Australia, Fiji for a short duration with Horowitz’ interference, 11 Colonial legacies also play a role in Europe: the representative body of Iceland, in 1844, was elected along the lines of the Danish electoral system (Hardarson, 2002: 112) and the single-transferable vote of Ireland and Malta is a legacy of British colonial rule.
Papua New Guinea, and variations of the system on small Pacific Islands). Indonesia inherited proportional representation from the Dutch.

In Cambodia, the UN peace mission UNTAC not only organised the first elections, but also drafted the electoral law, introducing proportional representation (Doyle & Suntharalingam, 1994). Other electoral systems which were introduced during international interventions, are either PR (Iraq), or have strong PR components: SNTV in Afghanistan, and a mixed electoral system with a very strong PR component in East Timor. Since the 1990s, mixed-member electoral systems have been popular in Asia, Latin America, and Central and Eastern Europe.

**Latin America**

Until the early 20th century, the development of democracies in Latin America broadly followed the Western European pattern: in the 19th century, the democratisation of elections proceeded in rather slow and hesitant steps. Several countries used electoral colleges or indirect elections in order to keep representation under the control of the elites (Colomer, 2004b: 84). The move from plurality vote, limited ballot, or cumulative ballot (both with quasi-proportional effects) to PR came after WWI. As of the third wave of democratisation (1974-), all Latin American democracies were electing their parliaments by PR (Colomer, 2004b: 86) or by mixed systems (see next section), although these were often overshadowed by presidential elections by a two-round majority. Some countries have altered their rules more than a dozen times in the 40 years since WWII (Remmer, 2008: 7). These changes occurred in periods of electoral instability. However, electoral reforms fail to lead to a consolidation of the party system. They do not decrease the number of parties (neither do more permissive reforms increase it), and instead, they are a catalyst for electoral volatility in subsequent elections (Remmer, 2008).

**Africa**

Comparative studies on electoral system choice on the African continent highlight three main aspects. First, bans on ethnic, regional or religious parties, or regulations requiring political parties or candidates to pool votes across districts are widespread. There is no other continent where
they play an equally important role (Bogaards, 2003; Bogaards, Basedau, & Hartmann, 2010; Moroff, 2010). Second, divided societies, particularly those with previous conflicts, often opt for list-PR. South Africa is the paradigmatic case, but it also has also spurred institutional choices in Burundi (complemented by quota) (Vandeginste, 2009), Angola, Mozambique, Sierra Leone, and Namibia (Mozaffar, 1998: 89-91). And third, the choice of electoral rules is often a legacy of the colonial history. British colonies mainly selected plurality rule; French colonies opted either for the majoritarian vote or for multi-seat districts (Mozaffar, 1998).

**Central and Eastern Europe**

To a considerable extent, the choice of the electoral system in the course of the transition to democracy in Central and Eastern Europe was the result of the form of transition – pact or coup – and who participated in the negotiations (Benoit & Schiemann, 2001). The communist parties were discredited, and thus ran the risk of significant losses if they were to run under a party label. However, with strongly institutionalised parties, they also held important posts throughout the territory that functioned as clientelist networks. The logical choice was to move to majoritarian or plurality vote, which should allow the communist nomenclatura to win seats in their districts. The reformers often preferred PR, rarely for idealistic reasons (Elster et al., 1998: 115-116), but because without proper territorial organisation, and with a split between different parties, the reformers expected to do better (Geddes, 1996). In Estonia, where constitution-makers were well informed even about rare types of electoral systems, the logical compromise was the (short-lived) Single-Transferable Vote, a PR-system operating with candidates, not parties (Taagepera, 1997). All around the region, mixed electoral systems resulted primarily in countries where the political transition was agreed between the regime and reformers at roundtables (also addressed as ‘pacted transitions’) (e.g. Hungary and Bulgaria) (Elster et al., 1998: 112-114; Shvetsova, 2003). While proportional representation or mixed systems prevailed, and still prevail today, several heterogeneous countries (Macedonia, Bulgaria) adopted PR only later on, and others impose high thresholds or bans to exclude or curtail parties of minorities (Bochsler, 2010).
Squaring the circle

In the 1990s and early 2000s, the search for the ‘best’ electoral system met with new hope. The common trend of electoral reforms in Asia, Latin America, Africa and in new democracies in Central and Eastern Europe pointed towards mixed electoral systems. Mixed electoral systems combine two types of mandates, one part (usually at least half the parliament) is elected by plurality or majority rule in single-seat districts, the other part by proportional rule from regional or national lists. In electoral system research, this shifted the emphasis from simple electoral systems (PR or plurality majority vote) to more sophisticated rules, and it brought the personalisation dimension onto the electoral reform agenda. Mixed electoral systems were considered by practitioners and academics alike to resolve the initially emphasised trade-off between the principles of representation, governability, and personalisation, combining positive properties on all three dimensions. Their proportionality allows for broad representation, while the majoritarian districts prevent a too excessive fractionalisation of the party system and/or encourage parties to align in pre-electoral coalitions. Additionally, they combine a direct, localised, and personalised election in districts with national list mandates (Shugart, 2001). Such systems were originally used in Denmark, and in post-war Germany.

Personalisation of elections

In the post-cold war period, political parties and legislatures elected by list-PR experienced a crisis of legitimacy, which put the direct election of individual MPs on the agenda (Renwick & Pilet, 2016: 62; D’Alimonte, 2005: 255). Proportional representation can in many ways be complemented by personal votes (Carey & Shugart, 1995). Historically, in Western Europe the degree of party-centrism versus personalisation of the electoral system, including the move to PR (Manow & Schröder, 2014), but also how important individual candidates are in PR systems as opposed to political parties was less a matter of the legitimacy, but rather of the strength of parties.

12 A high degree of personalisation of the electoral system also increases the salience of personal candidate attributes in the election (Shugart, Valdini, & Suominen, 2005).
Strong parties – especially countries with powerful Social Democrats – opted for more power to the parties, and closed lists. In countries with weak parties and with low degrees of party polarisation, PR with open- or flexible lists (e.g. Italy, Finland) or the STV, the proportional system that also operates without parties, were adopted (Renwick & Pilet, 2016: 86-87). In open-list proportional representation (e.g. Finland, Switzerland), voters cast their votes for candidates who are affiliated with a party list. Each party list wins a number of seats proportional to its votes, although the candidates winning most votes are elected. In flexible-list PR systems, voters can cast preferential votes for candidates, but the candidate order, which is pre-established by the party, also has some weight.

Since the late 1980s, countries electing their parliaments by list-PR experienced a push towards a higher personalisation of their elections (Renwick & Pilet, 2016: 44-49), mainly towards flexible-list PR.

When personal votes do not lead to personal representation

Once flexible list PR is introduced, it is subject to repeated adjustments of its minor technical elements: reforms of flexible-list PR systems are more frequent than other reforms (Renwick & Pilet, 2016: 72). In many ways flexible lists are always subject to the tension between parties and voters, who both want to gain control over the candidate elections (Renwick & Pilet, 2016: 26-27). In Iceland, while voters have the right to express their candidate preferences, the parliament found it “undemocratic that a small minority of voters could change who was elected when the vast majority had left the order of names unchanged (Renwick & Pilet, 2016: 116).

There are numerous ways to curtail the impact of the personal votes. In Iceland, since 1959, parties retain the right to co-influence who gets elected (Renwick & Pilet, 2016: 91-92, 115-116). In some countries, the votes of voters who did not make the effort to alter the candidate order on the list, are counted in support of the pre-established order (Belgium, Austria until 1970, Iceland). Other countries only move candidates up the list if they reach a threshold of votes (e.g. Austria since 1970, Czech Republic, Netherlands, Sweden since 1997, Slovakia). If voters can
only cast a single preferential vote, and if districts are large or candidates are allowed to run in several districts, then preferences are often concentrated on very few prominent candidates.

Mixed-member and multi-tier systems suffer from a different problem: they combine single-seat districts (mixed systems) or an open- or flexible-list part (multi-tier systems) with a closed national list (e.g. Estonia, Poland 1991-7, Slovenia 1992-2000; Italy 1993-2005). If candidates are allowed in both tiers, then party elites can obtain a safe passage into parliament on the closed (national) list, even if they run simultaneously in their district, and fail to secure their seat (Mikkel & Pettai, 2004: 339-340). This makes the candidate vote appear toothless. When double candidacy is not allowed, party elites might run on safe spots on the closed national list, so that the most prominent personnel is not directly accountable to the voters (Nikolenyi, 2004: 1046).

In contrast to the European trend, in the 1990s some Asian countries decided to shift power back from voters to parties. Japan, Thailand and the Philippines suffered from an over-personalisation of elections and clientelism under the Single-Non-Transferable-Vote or the multi-seat plurality vote (Reed & Thies, 2001; Reilly, 2006).

“In the Philippines, the combination of candidate-centred elections with oligarchic family dynasties, enduring patron-client links, and weak political parties was widely seen as having subverted party system development and undermined coherent public policy”, paving the transition to the authoritarian Marco regime in 1972 (Reilly, 2006: 100).

An agenda for future research

Research on the origins of electoral systems gained scientists’ attention partly due to concerns that Duverger’s laws might work upside-down: it is not institutions that shape party systems, but parties that tailor the institutions to their needs (Grumm, 1958). Indeed, as this chapter has discussed, the saliency of the different ideals underlying the choice of electoral systems depends heavily on two aspects: structural factors (such as cleavages, economic conditions), which are also related to the format of the party systems, and parties themselves as actors. The principle of inclusion matters most in small states, fragile states (i.e. post-conflict states or ethnically divided
societies), and in periods of political uncertainty. Governability is emphasised in countries with a British heritage, and when governing parties want to reduce the threat by (radical) outsiders. The principle of personalisation has driven electoral reforms since the 1990s, but has also played an important role when countries with weak party systems have selected proportional electoral systems.

More research is needed to investigate electoral reforms and party system changes jointly (e.g. Colomer, 2005), i.e. to study whether electoral reforms actually produce the party systems they intend to. So far, from this perspective the evidence does not speak for the skills of political parties as lawmakers. Most often, the purpose of their reforms seems not to be achieved (Shvetsova, 2003; Remmer, 2008).

While studies on the effect of electoral systems increasingly move towards comprehensive cross-regional samples, the literature on the origins of electoral rules with few exceptions – notably edited volumes (e.g. Shugart & Wattenberg, 2001; Colomer, 2004b) – remains largely segmented regionally, and in terms of historical periods.

References


