

Notes

CHAPTER 2

1. It is notable that Arrow (1951/1963, 20–21) considered a game-theoretic approach to social choice. He did not pursue this, however, instead taking the social welfare function approach to its logical conclusion. The example Arrow gave was of noncooperative games, rather than the cooperative coalition-type games suggested earlier. “Thus, the model of rational choice as built up from pair-wise comparisons does not seem to suit well the case of rational behavior in the described game situation. It seems that the essential point is, and this being of general bearing, that, if conceptually we imagine a choice being made between two alternatives, we cannot exclude any probability distribution over those two choices as a possible alternative. The precise shape of a formulation of rationality which takes the last point into account or the consequences of such a reformulation on the theory of choice in general or the theory of social choice in particular cannot be foreseen; but it is at least a possibility, to which attention should be drawn, that the paradox to be discussed below might be resolved by such a broader concept of rationality.”

2. Rawls (1995/1996) accepts that the original position construction has no privileged status and is simply a device that has to be justified in actual discourse with real people. However, Rawls does not talk about the rules for such discourse.

3. “If we exclude the possibility of interpersonal comparisons of utility, then the only methods of passing from individual tastes to social preferences which will be satisfactory and which will be defined for a wide range of sets of individual orderings are either imposed or dictatorial.”

4. The Borda procedure has every voter rank order all the n alternatives, and then gives $n - 1$ points to each voter’s first choice, $n - 2$ to each voter’s second choice, etc. The scores for each alternative are then summed, and the alternative with the highest score is the winner.

5. Mathematically speaking, we do have a metric for the space of alternatives.

CHAPTER 3

1. For example, it would be possible to implement pure PR by distributing some seats by a lottery where each alternative’s chance of getting the seat is proportional to the difference between its vote share and the seat share it has received.

2. Subject to other institutional features. For example, Germany has a mixed-member system of this type. The results are approximately proportional, except that there is a 5 percent electoral threshold and that parties who win more district seats than their overall vote share would dictate are allowed to keep the “excess” seats.

3. With single transferable vote, a candidate requires a certain quota to be elected, typically the Droop quota = (number of voters / (number of seats + 1)) + 1. Once a candidate is elected, their excess votes are distributed to the candidates ranked second on the excess ballots. If there are still seats to be distributed, the candidate with the lowest vote is eliminated and their votes are redistributed to the candidates ranked second on the ballots. This procedure is repeated until all seats are filled.

4. If there are n candidates, the Borda procedure gives $n - 1$ points to each voter’s first choice, $n - 2$ to their second, etc. The scores are then totaled.

5. Strictly speaking, Rae (1967) uses three dimensions of classification: district magnitude, electoral formula, and whether the vote is categorical or ordinal (whether the voter gets a single vote or whether they get to rank-order the candidates). However, systems with ordinal vote typically use distinctive mechanisms, such as single transferable vote, so we can treat ordinal voting as a different kind of formula.

6. Strictly speaking, the Netherlands is divided into smaller electoral subunits. However, seats are allocated on a nationwide basis. See Gladdish (1991).

7. Strictly speaking, majority rule implies the winning candidate receives more than 50 percent of the vote against any other candidate. Plurality runoff violates this. It is possible that a candidate could beat either of his opponents in a head-to-head race but would lose in plurality runoff by being eliminated in the first round.

8. A variation of this is cumulative vote, which allows voters to award more than one vote to one candidate.

9. The debate between the two methods can be traced back to the dispute between Jefferson (who advocated a divisor method) and Hamilton (who advocated a quota method) over how to allocate House of Representative seats to the various states (see Balinski and Young 1982/2001; Nurmi 1999; Saari 2001).

10. Voters in the Netherlands have one vote, which they can cast for the head of a list (usually a party leader) or for any person on the list. Prior to 1998 the votes for the list head were distributed in list order, so list position overwhelmed personal vote, and very few candidates were elected on the basis of their personal following. Since 1998 personal votes can overturn list order, but only if a candidate receives votes equivalent to half a quota, a very high requirement.

11. I use the term *plurality* instead of *majoritarian* because there is no guarantee that a candidate elected from a district receives a majority of the votes in that district, nor is there any guarantee that a party that receives a majority of seats won a majority of the vote.

12. Party system fractionalization is defined as $\sum V_i^2$ or $\sum S_i^2$, where V_i is the vote share of party i and S_i is its seat share. The effective number of parties is defined as $1/\sum V_i^2$ or $1/\sum S_i^2$.

CHAPTER 4

1. To recap, anonymity means that the rule treats all voters equally—if we exchange the preferences of any pair of voters, it does not affect the outcome. Neutrality means all alternatives are treated equally—if one alternative gets a certain set of voters and wins, then if these voters all switch their support to another alternative, it must then win. Decisiveness simply means that the rule must produce a result, which may be a draw. Positive responsiveness means that if we have a draw, and one voter switches his support to alternative 1, then alternative 1 must win.

2. This amounts to saying that if someone prefers alternative *a* to *b*, then there is another alternative, very close (possibly infinitesimally close), that the person also prefers to *b*.

3. Strictly speaking, Schofield gives a characterization of the “heart” or locally uncovered set, which is a superset of the uncovered set.

4. Note that this definition of logrolling implies separable preferences over the issues in question. Logrolling involves people voting for measures they do not like. We could not talk about people liking or not liking specific measures unless their preferences are separable—that is, they like or dislike the measure regardless of how other measures are resolved.

5. That is to say, *some* representatives vote for things that they do not want in order to get things that they do, but there is no cycle or intransitivity because other representatives have a very particular kind of nonseparable preference that effectively cancels the cycle. Suppose that we change Representative C’s preferences. Representative C still wants his measure adopted, and prefers that measures A and B are both not adopted. However, he despises the situation where measure A is adopted and not B, to the extent of preferring that both measures be adopted to only one being adopted. (Perhaps C has a strong sense of equity, and thinks that adopting A but not B is unfair.) With this change, Representative C will no longer support a motion to go from funding measures A and B to just funding A. Therefore the cycle in figure 4.4 disappears, essentially because we have assumed an agent who despises the intermediate step in the cycle. This situation, incidentally, is morally troubling. Representative C does not get his measure, but he has to pay for A and B, precisely because A and B can exploit his sense of equity. In the previous version of the example, Representative C can defend himself against a coalition of A and B by trying to cut a deal with one of them; however, the new preferences we have given C make this line of defense impossible.

6. Riker (1982, 242–43) anticipates this objection, arguing that this randomness gives elected officials an even stronger incentive not to offend voters. However, this response is extremely unconvincing. Coleman and Ferejohn (1986) are surely right that if voters’ rejection is strictly random, officials can do nothing about it and will treat it as an act of God.

7. Przeworski (1999) does not rely on elections removing bad governments but rather argues that if elections were completely random, the exercise would still be worthwhile. Given that there is a probability for the incumbent to be removed

in the future, it is worthwhile for the opposition to continue to play the constitutional game, as opposed to taking up arms. It is also in the interest of the incumbent not to be too oppressive, as he may lose power in the future. Przeworski also argues that elections are useful as a measure of the relative strength (in the case of civil war) of the government and the opposition. I do not dispute Przeworski's claim that even minimalist democracy may have some value; the argument in this chapter is that a more expansive theory of democracy is viable.

8. Other justifications could be given for fairly distributing political resources, such as the symbolic value of fair representation. This argument is not pursued here, as it simply reinforces the case for majority rule as pure procedure.

9. Thus power indices, such as the Penrose, Shapley-Shubik, and Banzhaf indices, which are based upon the percentage of possible coalitions in which a player is crucial, are monotonic to the voting weight of the players. That is, as players get more voting weight, their power increases, all other things being equal. See Penrose (1946); Shapley and Shubik (1954); Banzhaf (1965).

10. Schwartz (1995) shows that a group of voters can become worse off by getting more representation. However, this argument relies on sincere voting. If their representatives vote strategically, they can at least reproduce the outcome they got when they were less represented.

11. Although majority rule is generally not transitive (it allows cycles), there is typically a great deal of transitivity within the web of preference relations. The covering relation can be restated as follows: a covers b implies that the social preference between a, b, and any other alternative is transitive. It is this transitivity information that allows us to reject certain alternatives as never being reasonable choices.

CHAPTER 5

1. This would rest on empirical evidence. As we will see, empirical evidence that supermajoritarian rules provide better rights protection does not exist.

2. See Madison's remarks to the Constitutional Convention on June 19, 1787 (Madison 1840/1966) and "Vices of the Political System of the United States" (Madison 1999).

3. On June 6 at the Federal Convention, Madison argued that extending the scope of the republic was the *only* way to protect minorities: "In a Republican Govt. the majority if united have always an opportunity. The only remedy is to enlarge the sphere, and thereby divide the community into so great a number of interests and parties, that in the 1st place a majority will not be likely at the same moment to have a common interest separate from that of the whole or of the minority; and in the 2nd place that in case they should have such an interest, they should not be apt to unite in the pursuit of it" (Madison 1840/1966, 77).

4. Madison most famously opposed equal representation of the states in the Senate. However, he also championed the supremacy of the national government over the states (including a national veto on all state legislation), supported a simple majority override of judicial review, and opposed the prohibition of legislators taking executive posts.

5. See also Federalist 22 (authored by Hamilton).

6. It is notable that Buchanan and Tullock (1962, 47) do not state how these rights come about but merely state that “it will be useful to ‘jump over’ the minimal collectivization of activity that is involved in the initial definition of human and property rights and the enforcement of sanctions against violations of these rights.”

7. Rae quotes Sen (1970a): “An economy can be optimal in this sense even when some people are rolling in luxury and others are near starvation as long as the starvers cannot be made better off without cutting into the pleasures of the rich. If preventing the burning of Rome would have made Emperor Nero feel worse off, then letting him burn Rome would have been Pareto-optimal. In short, a society or economy can be Pareto-optimal and still be perfectly disgusting.”

8. The term *core* is used differently here than in Laing and Slotznick (1987). We define the core as the set of points that cannot be defeated under the *q*-rule. This is equivalent to the set of points that cannot be overturned if that point is established as the status quo. Laing and Slotznick define the core as the set of points that are undefeated under the *q*-rule and dominate the current status quo. The definition of *core* used here is equivalent to Laing and Slotznick’s *heartland*.

9. Thanks to Donald Saari for pointing me to this source.

10. Rae (1975) gives the example of people building chimneys that pollute their neighbors in a jurisdiction that does not yet have effective regulation of pollution.

11. Rawls (1993/1996, 332–33) accepts that rights cannot be maximized and replaces “the most extensive possible scheme of equal basic liberties” in *A Theory of Justice* with “a fully adequate scheme.”

12. Sen’s illustration involves two people, one a prude, the other a libertine. The library has one copy of D. H. Lawrence’s *Lady Chatterly’s Lover*. The libertine wishes to read this, while the prude does not. If they are both free to decide whether to read the book, this will be the outcome. However, the libertine would get more pleasure from the prude reading it (and suffering embarrassment), while the prude would rather read it himself than see the libertine enjoy the book. Therefore both would prefer for the prude to read the book, even though the prude does not want to read the book and the libertine does.

13. The most commonly cited case of the U.S. Supreme Court extending basic rights is, of course, *Brown v. Board of Education* (1954), which outlawed segregation in public schools. This, however, did not protect a minority for a national legislative majority. Rather, it overturned the actions of state and local governments. Indeed, by 1954 there was already a clear national majority in favor of desegregation, and (ironically) the reason this could not be enforced by legislation was actually the supermajoritarian nature of the U.S. Congress, notably the filibuster in the Senate and the Southern domination of the committee system. Furthermore, the *Brown* decision only had practical impact because the most majoritarian branch of the U.S. government, the executive, chose to implement it by force.

CHAPTER 6

1. However, the deliberative polls organized by Fishkin do involve briefing by expert witnesses.

2. Essentially the argument is that people have a mutual interest in cooperation, but some forms of cooperation may suit some people better than others. When I talk, I try to convince you that society will coordinate on the outcome that suits me, and you will discount my communication somewhat. However, you will not discount it entirely, as it does provide some information about the outcome society is most likely to coordinate on.

3. Strictly speaking, Dryzek (1990) uses the term *discursive democracy* rather than *deliberative democracy*.

4. Johnson argues that communicative action involves persuasion. Persuasion, however, is necessarily perlocutionary and strategic, trying to change the opinion of another.

5. In *Political Liberalism* (1993/1996, 39–40) Rawls argues that political justice in a plural society requires an overlapping consensus on how reasonable claims can be reconciled. This consensus may originate historically from a compromise between different communities, but it eventually needs to become a consensus that can be justified on its own terms. See chapter 7 for a fuller account.

6. Knight and Johnson (1994) are a notable exception.

7. Barber uses the term *strong democracy* rather than *deliberative democracy*, but there is a strong emphasis both on collective will formation by discussion and on mass participation.

8. The folk theorem is that virtually any outcome that gives every player at least the minimum payoff they can guaranteed themselves can be an equilibrium in a repeated game with a low enough discount rate. It is so called because it was widely recognized before anyone published it.

9. The one partial exception of this was the Rabin government from 1993, which had Arab parties as part of the support coalition of a minority government.

10. Fearon (1998) actually prefers the term *discussion* to *deliberation*, as he feels that *deliberation* is too normatively loaded.

CHAPTER 7

1. Rae argues against Buchanan and Tullock on other grounds. See chapter 5.

2. See, in particular, *The Philosophical Discourse of Modernity* (1990b) for Habermas's reaction to the work of Derrida, Foucault, Bataille, and others.

3. The principles are: (1) everyone receives the greatest possible liberties compatible with the same liberties being universal; (2) all social goods are distributed by means of offices open to all, with inequality only being tolerated when it advantages the least advantaged.

4. See Harsanyi (1975), Hare (1973/1975), and Sen (1975) for critiques of the maximin principle as a solution concept. See Hart (1975) for a critique of the

elaborate “four-stage process” by which Rawls gets from principles of justice to their implementation.

5. Habermas does not, of course, argue that people engaged in communicative action lack their own action plans. However, they pursue them only on condition that their plans can be harmonized with others. Thus their action is oriented toward agreement first and goal attainment second (1984, 285–86).

6. Scanlon’s criterion is: “An act is wrong if its performance under the circumstances would be disallowed by any system of rules for the general regulation of behaviour which no one could reasonably reject as a basis for informed, unforced general agreement.”

CHAPTER 8

1. France’s 1958 coup d’état is ignored for this purpose.

2. Tsebelis (2002) uses two measures of stability. The first measure, based on Doering (1995), is the number of significant laws on working time and working conditions in the period 1981 through 1991. Apart from the concern about the generalizability of conclusions from one policy area, it is far from clear that the amount of actual change in “the law” is at all related to the number of laws passed. Tsebelis deals with this problem in part by correcting for Sweden (where every amendment is counted as a law) and by only considering “significant” laws to correct for legislative inflation. However, the judgment of which laws are significant are made by legal scholars, not by economists or political scientists. Indeed an economist might well suggest that labor policy is often impacted far more by social security rules and business regulation than by labor law per se. There is no attempt to validate the measure of legal change by correlating it with outcomes in the real economy such as changes in working hours, female participation, or part-time work. In addition to validating Doering’s measure, changes in these variables may actually be better measures of change in labor market conditions.

Tsebelis’s other measure of policy change is the Euclidean distance between the vectors of government spending in ten categories between consecutive years, based on work with Eric Chang (Tsebelis and Chang 2001). As a measure of policy change, this has considerable potential as it measures changes in the allocation of hard resources, as opposed to legislative gestures. However, Tsebelis uses the expenditure in each category as a percentage of total government spending. Thus, if a government doubles spending in every category, this counts as no change. Interestingly, Tsebelis does not use the distance measure to test the effect of the number of veto points, but only variables such as government alternation and the ideological cohesion of the governing coalition.