

Essays on Power, Freedom, and Success
Concepts, Measurement, and Applications



Essays on Power, Freedom, and Success

Concepts, Measurement, and Applications

Matthew Braham, M.Sc (University of Oxford, 1996)

Lindenstrasse 25, D-21521 Aumühle

Dissertation Thesis

Accepted by the Faculty of Economics in partial fulfilment
of the requirement for the degree of

Doctor Rerum Politicarum

at the

UNIVERSITY OF HAMBURG

Hamburg • 15 December 2004

Thesis Committee

Chairman: Prof. Dr. Lothar Streitferdt, Faculty of Economics and Business Administration, University of Hamburg, Germany

1st Examiner: Prof. Dr. Manfred J. Holler, Faculty of Economics and Business Administration, University of Hamburg, Germany

2nd Examiner: Prof. Dr. Martin van Hees, Faculty of Philosophy, University of Groningen, The Netherlands

Date of Defence (*Disputation*): 07 September 2004

Contents

Preface	vii
Acknowledgements	ix
Chapter 1: Introduction: Power, Freedom, and Success	1
Chapter 2: Preferences and the Measurement of Power	9
Chapter 3: The Measurement of Freedom	27
Chapter 4: The Success of a Chairman	48
Chapter 5: Voting Rules in Insolvency Law	77
References	100

Preface

This dissertation comprises a selection of essays about power, freedom, and success that have been written during the period April 2000 and May 2004 while a research fellow at the Institute of SocioEconomics, Faculty of Economics, University of Hamburg.

Each of the essays is a self-contained article and have, over the years, been presented in some form or another at international conferences and workshops:

Chapter 2 (published in the *Journal of Theoretical Politics*) was presented at the Second Joint Hamburg–Siena Workshop in Applied Game Theory, 6–9 June 2002, Hamburg; ‘New Approaches to Rights, Freedoms and Power’, European Consortium of Political Research 31st Joint Sessions of Workshops, 28 March – 2 April 2003, Edinburgh; and the 2003 European Public Choice Society Conference, 26–28 April 2003, Aarhus, Denmark.

Chapter 3 was presented at the Workshop on Freedom and Power, University of Bayreuth, 12 February 2004; University of Turku Political Science Workshop held in Tartu, Estonia, 15–17 August 2003; the Third Joint Hamburg–Siena Workshop in Applied Game Theory, Siena, 11–13 April 2003; the European Consortium of Political Research 31st Joint Sessions of Workshops (‘New Approaches to Rights, Freedoms and Power’), Edinburgh, 28 March – 2 April 2003; the 2002 Annual Meeting of the European Public Choice Society (EPCS2002), Belgirate 4–7 April 2002; and the conference ‘Analysis and Measurement of Freedom: Theoretical, Empirical, and Institutional Perspectives’, Palermo, 27–29 September 2001.

Chapter 4 (forthcoming in *Social Choice and Welfare*) was presented at the University of Turku Political Science Workshop held in Tartu, Estonia, 15–17 August 2003.

Chapter 5 (published in the *International Review of Law and Economics*) was presented at 6th Spring Meeting of the Young Economists, 30 March – 1 April 2001, Copenhagen; the 18 Annual Conference of the European Association of Law and Economics, 13–15 September 2001, Vienna; and the 2002 Annual Meeting of the European Public Choice Society, 4–7 April 2002, Belgirate, Italy.

Acknowledgements

I must express my gratitude to the people who have helped me en route. In particular, I am deeply indebted to my supervisor Manfred J. Holler who provided the intellectual environment and encouragement – nay, freedom, as he posed no obstacle and actively helped remove any that came my way – in which I could develop my interest in the subject matter of this dissertation, as well as to Frank Steffen, my co-author and colleague at the Institute of SocioEconomics, with whom I have written a number of other papers on power that I have not included here and who undertook the arduous task of double-checking the results of Chapter 4. Special appreciation also goes out to Sebastiano Bavetta, René van den Brink, Ian Carter, Keith Dowding, Martin van Hees, Ines Lindner, Moshé Machover, Peter Morriss, Stefan Napel, Hannu Nurmi, and Georg von Wangenheim who have all intensively commented on various versions of the essays that comprise this dissertation. I must also thank Ad van Deemen for suggesting the idea of Chapter 4 while I was struggling with a different approach and Steven J. Brams who confirmed a central result in that essay. And of course I must thank Santi Chakravarty, Rudy Fara, Daniel Friedrich, Christian List, Peter Weikard, Mika Widgrén, Stefano Vannucci for discussions and comments during my research. Finally I would like to thank the editors and referees of the *Journal of Theoretical Politics*, *International Review of Law and Economics*, and *Social Choice and Welfare* for also helping to improve three of the chapters of this thesis.

I should also thank the Deutsche Forschungsgemeinschaft (DFG) for financial support while first as a full-time member (October 1998 – December 2000) and then as an associate member (January 2001 – June 2002) of the Graduate Programme in Law and Economics, University of Hamburg. Hans-Bernd Schäfer provided the access to this support.

Last, but not least, Dietlind, Uriel, and Raphael need to be mentioned for their great patience and endurance over the past five years.

M.B.

Aumühle bei Hamburg, 15 December 2004

Introduction: Power, Freedom, and Success

Beginning as it did in the writings of philosophers, teleologists, pamphleteers, special pleaders, and reformers, economics has always been concerned with problems of public policy and welfare.

— P.A. SAMUELSON, *Foundations of Economic Analysis* (1947: 203)

1. *Fundamental Concepts*

Ask any one to provide a description of a ‘state of society’ with the intention to define the best one and it goes with saying that the description will not only include reference to income, wealth, and happiness, but will also include in some way a mention of ‘powers’ and ‘freedoms’ as an indication of the possibilities and constraints on individual action and wellbeing. The collection of essays that comprise this dissertation are about various aspects of the content, measurement, and application of these two fundamental concepts plus a third related – but lesser discussed – concept of ‘success’.¹ Given that these concepts are central for normative assessment and ranking of social states, each essay can be seen as an exercise in that broad field that we call *welfare economics*.

The concepts of power and freedom and success require brief explication in order to understand how the four essays in this dissertation, each of which deals with specialist topic, fit together. Although the concepts are often mixed up and used interchangeably in ordinary language as well as in some of the formal social choice and related political science literature (Dowding 2004), they are in fact analytically distinct.

The basic distinctions which will, in particular, be clarified in chapters 2 and 3, are, broadly, as follows:

Power The *ability* to force an outcome.

Freedom A state of being *unprevented* to do or have something.

Success Getting what one *wants*.

¹ In some of the voting power literature, ‘success’ is also called *satisfaction*. See, for instance, Brams and Lake (1978) and Straffin *et al.* (1982).

With regard to the first two concepts, the relation is obvious: if I have the ability to force some outcome in which I perform a specific action or have a particular item (*power*), then I am unprevented from doing so (*freedom*); but I may be unprevented from undertaking an action (*freedom*) or possessing an item but do not have the ability to force that outcome (*power*) in which I am performing that action or possessing that item. Reformulated: the power to do x implies the freedom to do x , but not vice versa.

The concept of *success*, which more or less entered our conceptual vocabulary with Barry's (1980a, 1980b) important essay 'Is it Better to be Powerful or Lucky?' (although it has its formal precursor in Penrose (Penrose 1946) and Rae (Rae 1969)) is, however, a shade different than either power or freedom, although a particular version of it is equivalent to freedom. Success is about getting what one wants. Obviously, if you want to do or have x , and you can do or get it, then you are free to have or do it because nothing is preventing you having it or doing it. Hence, if I am successful it implies I am free; and for the same reason that I can be free without being powerful, I can be successful without being powerful. I can get what I want merely as a matter of 'luck' or at the behest of another. Success and freedom differ, however, in that I can be free (unprevented) to do or have x even if I do not want or desire it. Where the two concepts converge is in the counterfactual dimension of freedom: if I *were* to attempt to do or obtain x and *would* be unprevented from doing or having it, then I am free and thereby can be said to be successful in my attempt. This is a major issue in Chapter 3 and there is no need to discuss it in detail here. The success–power relation is also clear: if I have the power to do or have x , then I will be successful if I were to want to have or do x ; and obviously if I am successful does not imply that I am powerful.

Thus, power, taken as the ability to force outcomes – the ability to see to it that I perform or obtain x – is the more generic concept of the three, although not necessarily the most important from a welfaristic standpoint, i.e. individual well-being determined merely by utility levels (Sen 1979). The slave whose preferences perfectly track those of his master will always be successful, although he may be entirely powerless. Yet, that power, and even freedom, may not be important from a welfaristic standpoint does not detract from their central importance for welfare economics and social ethics more generally. The ranking of social states merely by personal utilities is one, but not the only, way of making social welfare judgements, as Amartya Sen (1970a) has been arguing for ever since the publication of his *Collective Choice and Social Welfare* and which led to the birth of a now mammoth literature in economics, political science, and philosophy.

Freedom in particular can be said to have a *non-specific value*, meaning that 'The love of liberty can be something more than just the love of being free to do certain things' (Carter 1999: 32), and that this is as important for evaluating states of affairs as utility levels. As Isaiah Berlin (1969: xliii) eloquently penned it in a footnote in his *Four Essays on Liberty*: 'A man struggling against his chains or a

people against enslavement need not consciously aim at any further social state. A man need not know how he will use his freedom; he just wants to remove the yoke. So do classes and nations.'

Power may also have such non-specific value, although to the best of my knowledge this has not been analysed properly and is something that has to be undertaken (but I will not do this here or in any of the other chapters). That there are social contexts in which power can easily be said to have a non-specific value is easy to see. The allocation of voting power in committees is such an instance. When designing voting rules, we are generally concerned about how voting power (the potential to change a winning situation into a losing one or vice versa) is allocated without being concerned about the issues that will be voted upon or whether the members of the committee even want this power. On the other hand, the non-specific value of power presents its problems for making social judgements. As Barry (1980a: 184) has perceptively and colourfully pointed out, 'a committee made up entirely of people who had no interest in pursuing some particular outcome but were fascinated by the process as such would be as frustrating as a brothel all of whose customers were voyeurs'.

It is indispensable to underscore that I make no pretence as to the universal acceptability of the definitions of power, freedom, and success that I use. I neither claim, nor assume, that these are the fixed usages of the terms. Many writers will undoubtedly find the definitions somewhat anaemic.

The concept of *power* as merely the 'ability to force outcomes' is not what many take power to be. It would not be far of the mark to say that power is more widely conceived as the potential 'to affect others' and not the mere 'ability to effect outcomes'. To be powerful is to have 'power over' some person or group of persons. The locution of power is usually 'A has power over B' (Benn 1967, Dahl 1957, Harsanyi 1962b, 1962a, March 1955, Oppenheim 1960, Simon 1953) and not 'A has power to do x'. Why then, do I concentrate on the anaemic 'power to' when the 'power over' is richer? Consider, for instance, Stanley Benn's (Benn 1967: 424) famous definition (which, except for the phrasing, is the same as Dahl's):

A, by his power over *B*, successfully achieved an intended result *r*; he did so by making *B* do *b*, which *B* would not have done but for *A*'s wishing him to do so; moreover, although *B* was reluctant, *A* had a way of overcoming this.

The answer to this question is in the citation itself. *A* has power over *B* – can get *B* to do something he would not otherwise do – because he has a means to do so. *A* can force the outcome in which *B* performs *b*. Reformulated: 'power over implies 'power to', although the converse does not hold. I may have the naked ability to throw a rock of a given dimension and weight under specified environmental conditions 20 meters, but this does not imply that I have power over anybody. But if I do not have the ability to prevent *B* from not doing *b*, it could not be said that I have power over *B*. I have something to say on this issue in Chapter 2.

Similarly, many writers on *freedom* will object to the traditional liberal conception of ‘pure negative liberty’ that I make use of. ‘Freedom’, many will argue, is more than just being ‘unprevented to do or have *x*’. Republican writers such as Philip Pettit (1997) will consider this too weak; for Pettit the unpreventedness must be ‘robust’. Others, writing within a ‘moral’ view, such as David Miller (1983) or Kristján Kristjánsson (1996) require that a responsibility condition is used to filter out the different types of prevention; not all states of prevention can be considered as states of unfreedom. Then there are those who will hold, either alone or in conjunction with the negative view, a positive conception of freedom; freedom is not just about absence of constraints, it is about doing things in certain ways, or achieving some state of being. In his essay, ‘Rights and Capabilities’, Amartya Sen (1984: 318) writes that ‘First, freedom is concerned with what one *can* do, and not just with what one does do. Second, freedom is concerned with what one can *do*, and not just with utility that doing leads to’ (emphasis in the original). What Sen is saying is that when we make freedom ascriptions we must also look at the *functioning* of a person. A person’s positive freedom, given in terms of ‘capabilities to function’, is as important – or if not more so – than their negative freedoms.

Again, this is not the place to defend the conception of freedom that I use. I do not deny the relevance and importance of the different conceptions, but I choose a specific conception because of a particular problem at hand. This is fully elaborated in Chapter 3. If there is any basic defence, it is that freedom as mere unpreventedness is analytically the simplest and most basic definition that we can work with.

Although *success* – ‘getting what one wants’ – is analytically less ambiguous than power or freedom, it is not without its problems. A central issue is that a success ascription must in some way account for what we mean by ‘wants’, i.e. it has to take into account the source of the preference. When *A* points a loaded gun at *B* and demands ‘your money or your life’ to which *B* responds by handing over his wallet and is not shot, one reading of success is that *B* is *successful* because by handing over his wallet he can continue to live, which he desires (otherwise he would not hand over his wallet); that is, he can be satisfied with the outcome, given his strategy choice (handing over the wallet). The problem here is how to evaluate the success of *B* not handing over his wallet and is shot. A standard revealed preference approach would say that he was successful because otherwise he would not have chosen the action. This is not so odd. Maybe *B* no longer had a desire to live because he had a terminal illness and this was a better way of dying because medical euthanasia was not a feasible option; or maybe he had embezzled his company’s pension funds and the detectives were hot his trail so that death was better than a public scandal. On the other hand, he may not be successful because he thought he could call *A*’s bluff and failed (he wanted to live and have his

money). Success ascriptions require, therefore, consideration of preferences and circumstance.

I confess that the nature of success ascriptions is an important issue and one that requires careful attention. However, I need not concern myself with it here as the precise meaning of ‘getting what you want’ is further clarified in the appropriate essays (chapters 3 and 4).

2. Preview

With the forgoing introductory sketch of our three fundamental concepts at hand, a brief tour of the contents of the four self-contained essays will indicate the scope of this dissertation, as well as how each of the essays are conceptually and methodologically linked. As a initial pointer, it should be said that the essays have a common core of ideas: n -person game theory and coalitions.

The first essay (Chapter 2), ‘Preferences and the Measurement of Power’,² concerns a recent debate in the literature on voting power indices in which classical measures such as the Banzhaf (1965), Shapley-Shubik (Shapley and Shubik 1954), and Public Good indices (Holler 1982) have been criticized on the grounds that they do not take into account preferences. The argument is that because these indices are based on simple games they are blind to preferences and therefore miss a vital component of power, namely strategic interaction. This has led Steunenberg *et al.* (1999) and Napel and Widgrén (2004) to the develop so-called strategic power indices on the basis of non-cooperative game theory. The essay argues that the criticism is unfounded and that attempts to develop preference-based power indices are doomed to conceptual failure because it will clash with the elementary notion of power itself, which, after Morriss (1987/2002), we call a generic ability: ‘the ability to effect outcomes’ (see section 1, above).

An important step in our analysis is that we show that any attempt to ditch the notion of power as a ‘generic ability’ has the unfortunate result that it ditches the concept of power as a ‘potential’. This is very problematic, primarily because ‘power’ loses its meaning; if power is not a potential, then what is it? We like to think of this as a kind of conceptual ‘impossibility result’ that is germane to the theory of power generally. Power, we claim, resides in, and only in, a game form and not in a game.

Apart from the substantive conclusion that power is preference-free, a further contribution of this essay is that it brings conceptual and semantic analysis to bear on a literature that is largely dominated by game theoretical formalities. It is part of a small but growing literature that attempts to bridge the gap between concep-

² Written in collaboration with Manfred J. Holler and forthcoming in *Journal of Theoretical Politics* 17: 137–158 (2005) under the title ‘The Impossibility of a Preference-based Power Index’.

tual argument and formal technique (van Hees and Wissenburg 1999, Dowding 2004). That is, the essay demands that a power index must be consistent with a ‘grammar of power’. Another important aspect of this essay is that it brings the notion of power conceptually into line with that of freedom, which is also considered to be preference-free (Carter 1992, van Hees 2000).

Chapter 3, ‘The Measurement of Freedom’,³ continues and develops the methodological approach of combining conceptual analysis and formal modelling that is represented in the previous chapter. This time, I examine how we can measure specific freedom, which is the freedom of an agent to undertake some particular action. The essay builds on a recent paper by Dowding and van Hees (2003), in which they discuss the need for, and general form of, a ‘freedom function’ that assigns a value between 0 and 1 to a right or freedom and that describes the expectation that a person may have about being in a position to exercise (‘being free to perform’) that freedom or legal right. The usefulness of such a function is that in principle it could be used to define threshold values for indicating whether or not a person has a particular freedom or legal right and therefore for making non-welfaristic judgements about social states or to design the assignment of rights related to government policy, public regulation, or legal rules. An examination of the literature, however, shows that such a measure has never been properly defined. Based on the framework of a game form, I develop a very simple and natural measure of specific freedom that fulfils Dowding and van Hees’ properties and that turns out to be formally equivalent to the conditional variant of ‘success’, a measure that we know from Penrose (1946), Rae (1969), and Barry (1980a, 1980b) in the voting power literature.

The new ground that I break in this essay is not merely the construction of a freedom function by bringing conceptual analysis and game theory together, but that in the process of constructing such a function I can show (i) that there is a conceptual and formal link between the concepts of freedom and success (alluded to above), (ii) from a purely logical standpoint, individual freedom is a collective property, and (iii) power is a more basic concept than freedom. This essay, therefore, starts to lay the foundations for a conceptual and formal synthesis of our notions of power and freedom.

In Chapter 4, ‘The Success of a Chairman’,⁴ I re-examine the so-called ‘chairman’s paradox’ that was first noticed by Farquharson in his path breaking tract on sophisticated voting, *Theory of Voting* (1969). The essay has a close conceptual link to Chapter 3, on two accounts. First, it is about *success*. Second it has a link to freedom via a new paradox that is uncovered in the analysis that has bearing upon

³ Under review with *Economics and Philosophy* under the title ‘Freedom, Power, and Success: A Game Theoretic Perspective’.

⁴ A revised version of this paper, written together with Frank Steffen, is forthcoming in *Social Choice and Welfare* under the title ‘The Chairman’s Paradox Revisited’.

the freedom of choice literature. However, it differs on a central aspect both from Chapter 3 and the others essays: the analysis is based on games and not merely on game forms, which means that success is ascribed according to a solution concept, which is not the case in Chapter 2.

To the essay itself. The Chairman's paradox is concerned with the case of a three member committee in which a particular player who has a regular and a tie-breaking vote – the 'chairman' – not only will do worse in specific instances under the plurality procedure for three alternatives than if he did not have such a vote, but will also do worse overall. That is, the chairman's *a priori* probability of success ('getting what one wants') for all possible games with *linear* (strict) preference orders is lower than that of the two regular members. I demonstrate that this result, which comes about if voters act strategically rather than sincerely, is not as robust as it has been thought to be. By merely replacing the standard assumption of *linear* preference orders that do not allow for players to be indifferent with *weak* preference orders, which allow for indifference, we can escape from the paradox for the canonical case of three players and three alternatives. With weak preference orders, the *a priori* success of the chairman is now greater than that of the other two players.

The contribution that this essay makes goes beyond merely demonstrating an escape route for a well-known paradox. I also unearth a previously unrecognized paradox of sophisticated voting that for all intents and purposes is even more bizarre than the canonical chairman's paradox itself. This new paradox says that, *ceteris paribus*, if the chairman raises an alternative in his preference order, then it can be the case that the set of equilibrium outcomes contain neither his most preferred alternative or that which he has raised in his profile, whereas previously it contained his most preferred alternative. That is, an expansion of your choice set can be extremely harmful in a strategic setting. This is the link to the literature on freedom of choice. Increasing the number of options you value does not imply an increase in your freedom if freedom means being unprevented from obtaining these options.

A further contribution of this paper is that I also demonstrates that weak orders introduces a whole new dimension of complexity in plurality voting games and calls for the introduction of an alternative solution concept that I call *ultimately admissible coalition-proofness* which is a refinement of iterated weak dominance using Bernheim *et al.*'s (1987) concept of coalition-proofness.

The final essay of Chapter 5, 'Voting Rules in Insolvency Law',⁵ is an applied exercise that examines the power relations of a collectivity, and, as I will be briefly indicate below, the freedom and success relations of individuals as determined by legal rules. The essay concerns our understanding of different voting rules in

⁵ Written in collaboration with Frank Steffen and published in the *International Review of Law and Economics* 22 (2003): 421–442.

insolvency law in Canada, Germany, the UK, and the US, a chief characteristic of which is the provision for ‘workouts’ or ‘schemes of arrangement’ by which insolvent companies can attempt to rehabilitate the business. If such a choice is made, the debtor has to devise a plan of action which will be voted upon by claimants. The voting rules, however, differ in each jurisdiction to a greater or lesser extent and as yet have not been analysed in any rigorous manner. The new contribution of this essay is that it provides an approach based upon the theory of simple games to analyse the rules in terms of the ease which each of these regimes can pass (or hinder) plans and how these rules distribute value among claimants. That is, the analysis make systematic use of Coleman’s concept of the ‘power of a collectivity to act’. In particular the role of classification of creditors into separate classes and voting bodies and how this effects the probability of reorganization comes under the microscope. A main result is that we show that classification, which is driven by an equity criterion, decreases the odds of a reorganization plan being approved. In a link to freedom and success, this can be interpreted as saying that equity considerations reduce the probability that the individual who has proposed a reorganization will be unprevented (free) to implement this plan; or simply the probability of success will fall.

3. *Notation*

As a final word in this introductory chapter, the reader should be aware that because of the self-contained nature of the essays there is an unavoidable amount of repetition of concepts and formal apparatus (i.e. a game form, a simple game, power indices). The reader can, however, be confident about notation, which I have taken care to keep consistent. Except where otherwise stated, N is always a set of players, S is always a subset of N (coalition), W is the set of winning coalitions, A_i is the is the set of feasible actions or strategies, g is a game form, etc.

Preferences and the Measurement of Power

1. *Introduction*

In recent years it has become customary to say that *classical power indices*, such as the Shapley-Shubik and the Banzhaf indices suffer a major drawback in that they do not take into account player preferences. Particularly sharp criticism has come from Tsebelis and Garrett (1996), Garrett and Tsebelis (1996, 1999a, 1999b, 2001) and Steunenberg *et al.* (1999), while a more conciliatory critique has been penned by Napel and Widgrén (2001, 2002, 2004).¹ The corollary of this ‘lack of preferences’ criticism is that classical power indices – because they are based on *simple games* which are rooted in cooperative game theory and not in non-cooperative game theory – are insensitive to the strategic aspects of power and therefore are inappropriate for a positive analysis of the distribution of power in institutional structures.

The upshot of the criticism has been the development of so-called ‘strategic’ power indices based on non-cooperative games as a way to fill this apparent lacuna (Steunenberg *et al.* 1999) as well as renewed attempts to introduce preferences or strategic considerations into the classical indices (Napel and Widgrén 2001, 2002, Hosli 1997, 2002) and even attempts to find a unified framework that brings together cooperative and non-cooperative approaches (Napel and Widgrén 2002). Another strand of the debate initiated by Rusinowska and de Swart (2002, 2003) has been to examine and develop a little known index known as the Hoede-Bakker index (Hoede and Bakker 1982) that is similar to the classical indices but purportedly takes into account player inclinations.

The aim of this essay is to explore this ‘absence of preference’ criticism and examine the following question: is it conceptually meaningful for *any* measure of power – not just voting power – to include the preferences of the player whose power is being measured? The question is fundamental and any serious conceptual or applied analysis of power must explicitly or implicitly deal with the role of preferences in power relations.

¹ The basic criticism is very widespread. See, for instance, Lane and Maeland (1995), Hosli (1997), Bilal *et al.* (2001), Colomer and Hosli (2002).

Although there have been a number of studies that touch upon the role of preferences within the context of voting power indices,² a perusal of the literature suggests that there is no equivalent study that singularly focuses on the problem and pushes the analysis as far as we attempt to do so here. Thus while make liberal use of the philosophical semantic analysis of power conducted in particular by Goldman (1972, 1974) and Morriss (1987/2002), we will do more in this essay than just restate their respective positions. We will actually sharpen some of their original insights and express them more forcefully within the context of study of power indices. In particular, we express our main result that the basic concept of power as a potential or capacity cannot accommodate the preferences of the player's whose power we are measuring in what we, with an intended abuse of terminology, loosely christen as the 'core theorem of the measurement of power'.³ As the title of our essay suggests, we like to think of our core theorem – which is not a theorem in the formal sense of the term – as a kind of conceptual impossibility result that is germane to the theory of power generally.

The corollary of our theorem says that a player's power resides in, and only in, the strategies available to her given by the *game form* and not in the way that she plays the *game*. This implies that power is a value-independent concept. The upshot is that the core theorem renders unintelligible any attempt to formulate a measure of power in terms of the equilibrium of a non-cooperative game – the very idea of strategic power indices. Put bluntly, assessing how a player may play a game does not help us answer such questions as 'is Smith more powerful than Jones?' or 'what is the extent of Smith's power?' because power concerns what player may be *able to do*, not the actions they *may* or *do* take. It must not, therefore, be thought that we are rehashing old philosophical debates. Rather we are bringing the semantics of power into the centre of the debate about how to measure power. The fact that there seems to be a quite a widespread belief about the need to develop preference-based measures of power indicates that there is still a general confusion regarding the nature of a power ascription.⁴ This calls for an awareness of a philosophical analysis and not simply more formal modelling.

The remainder of this article is in six sections. In the next section we briefly recap the definition of a power index which forms the source of the controversy that we examine. In the third section we set out the argument in favour of devel-

² See, for example, Barry (1980a, 1980b), Miller (1982), Dowding (1991), and Morriss (1987/2002).

³ We do not concern ourselves in this essay with equally interesting question of the role of player's *j*'s preferences in a measure of a player *i*'s power. This would require a far more comprehensive essay than can be undertaken here.

⁴ The belief that power – not just voting power – ought to be analysed in a preference-based framework is extremely prevalent. See Nagel (1975) for an authoritative analysis of power as causation of outcomes by preferences.

oping a preference-based power index by recourse to two examples. In the fourth section we lay out the meaning of a generic power ascription. The fifth section is the heart of the essay. Here we state and defend our core theorem. In section six we discuss the ‘power to’–‘power over’ distinction. Section seven concludes.

2. Power Indices

A *power index* assigns to each player of an n -person *simple game* – a game in which each coalition that might form is either all powerful (*winning*) or completely ineffectual (*losing*) – a non-negative real number which purportedly indicates a player’s ability to determine the outcome of the game. This ability is a player’s power in a game given the rules of the game.

Let $N = \{1, 2, \dots, n\}$ be the set of players. The power set $\wp(N)$ is the set of logically possible coalitions. The simple game ν is characterized by the set $W(\nu) \subseteq \wp(N)$ of *winning coalitions*. $W(\nu)$ satisfies $\emptyset \notin W(\nu)$; $N \in W(\nu)$; and if $S \in W(\nu)$ and $S \subseteq T$ then $T \in W(\nu)$. In other words, ν can be represented as a pair (N, W) . It should be noted that ν can also be described by a characteristic function, $\nu: \wp(N) \rightarrow \{0, 1\}$ with $\nu(S) = 1$ iff $S \in W$ and 0 otherwise.

Weighted voting games are a special sub-class of simple games characterized by a non-negative real vector (w_1, w_2, \dots, w_n) where w_i represents player i ’s voting weight and a quota of votes necessary to establish a winning coalition. We call this quota a decision rule, d , such that $0 < d \leq \sum_{i \in N} w_i$. A weighted voting game is represented by $[d; w_1, w_2, \dots, w_n]$.

A power ascription in a simple game is given whenever a player i has the ability to change the outcome of a play of the game. A player i who by leaving a winning coalition $S \in W(\nu)$ turns it into a losing coalition $S \setminus \{i\} \notin W(\nu)$ has a *swing* in S and is called a *decisive member* of S . Coalitions where i has a swing are called *critical coalitions with respect to i* . A concise description of ν can be given by a set $M(\nu)$, which is the set of all $S \in W(\nu)$ but no subset of S is in $W(\nu)$, i.e. all members of S are critical. We call such a coalition a *minimal winning coalition* (MWC). Further, we denote by η_i the number of swings of player i in a game ν . A player i for which $\eta_i(\nu) = 0$ is called a *dummy* (or *null player*) in ν , i.e. it is never the case that i can turn a winning coalition into a losing coalition (it is easy to see that i is a dummy iff it is never a member of an MWC; and i is a *dictator* if $\{i\}$ is the sole MWC).

Numerous power indices based upon the framework of a simple game have been proposed down the years, notably by Penrose (1946), Shapley and Shubik (1954), Banzhaf (1965), Coleman (1971), Deegan and Packel (1978), Johnston (1978), and Holler (1982). For illustrative purposes, the *Shapley-Shubik index*, which is a special case of the Shapley value for cooperative games (Shapley 1953), measures power as the relative share of pivotal (‘swing’) positions of a player i in a simple game ν . It is assumed that all orderings of players are equally probable. The

idea (or ‘story’) is that the players line up to vote ‘yes’ and the player that turns a losing coalition into a winning coalition is the pivot (‘swing’). It is given by:

$$\phi_i(v) =_{def} \sum_{\substack{S \in W \\ i \in S \\ S \setminus \{i\} \notin W}} \frac{(|S|-1)!(n-|S|)!}{n!} \quad (2.1)$$

In contrast, the *absolute Banzhaf index* for a player i in a game v measures the frequency in which player i is a decisive member of a coalition, i.e. the ratio of the number of swings to the number of coalitions in which i is a member:

$$\beta'_i(v) =_{def} \frac{\eta_i(v)}{2^{n-1}} \quad (2.2)$$

Holler’s (1982) Public Good Index (PGI) measures the share of swings in MWCs. The motivation for this index is that if the outcome of the vote is a public good; then this fact together with rationality of the players means that only MWCs should be taken into account (oversized coalitions include free riders and will only form by chance). The PGI is given by:

$$h_i(v) =_{def} \frac{|M_i(v)|}{\sum_{j=1}^n |M_j(v)|} \quad (2.3)$$

3. Feasible Coalitions and Credible Swings

The apparent shortcoming of a classical power index is that because the underlying framework of a simple game only classifies the subsets of players (coalitions) into ‘winning’ and ‘losing’, such an index is insensitive to the strategic aspects of power relations. This can be captured by two elementary examples that have been discussed in the recent literature.

Example 3.1 Consider the three player simple game with winning coalitions $\{a,b,c\}$, $\{a,b\}$, and $\{a,c\}$. Assume, as Napel and Widgrén (2001, 2002) do, that a is in a position to make an ultimatum offer to either player b or c : accept almost no share of the spoils or be prevented from taking part in a winning coalition. Player a could be the federal government that requires the approval from one of two provincial governments to pass laws; or a could be a major shareholder that requires the support of a minority shareholder in order to determine corporate policy. If the players are rational and have utility functions that are monotonic in the spoils, and that there is no way to credibly enforce a blocking coalition $\{b,c\}$

which could extract concessions from a , then the non-cooperative game theoretic equilibrium will be that whichever of players b or c that a approaches first will accept a 's pittance of an offer. Something, however small, is, after all, better than nothing for *homo oeconomicus*. Drawing on cooperative game theory, Napel and Widgrén further point out that the core of this game is $\{(1,0,0)\}$. The conclusion that Napel and Widgrén come to is that given the pittance or nothing at all that b or c will receive under these two solution concepts, it is only reasonable to deduce that they must be more or less powerless because both of these players are robbed of the power commonly associated with their swing. In contrast, the absolute Banzhaf index, the Shapley-Shubik index, and the PGI, for instance, yield power vectors of $(\frac{3}{4}, \frac{1}{4}, \frac{1}{4})$, $(\frac{2}{3}, \frac{1}{6}, \frac{1}{6})$, and $(\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$ respectively, values that are radically at odds with a competitive analysis.

Example 3.2 Consider committee of seven players, $N = \{a, b, c, d, e, f, g\}$ in which each member has one vote and a $5/7$ majority rule. Assume a preference configuration $abcdefg$ which ranks the players in a uni-dimensional policy space according to their ideal points. Suppose there is a proposal χ located between e and f but which is closer to e than f , and suppose further that the status quo q is located to the left of a (see Figure 3.1). Now the question is, given the spatial configuration, what are the possible outcomes of this voting game? Take the spatial MWCs $S_1 = \{a, b, c, d, e\}$, $S_2 = \{b, c, d, e, f\}$, and $S_3 = \{c, d, e, f, g\}$. Inspection of S_1 indicates that it cannot be a MWC in a spatial sense because given the locational assumptions, if the players in S_1 accept the proposal, then so too will players f and g : if a finds χ acceptable, then any player more in the vicinity of χ than a will do so as well. Now consider S_2 . By the same argument, S_2 is also not a spatial MWC because b accepting χ implies that g will accept it. The third case, S_3 , is a spatial MWC. What this reasoning implies for the measurement of power is twofold. Firstly, it says that certain coalitions will not form, viz. S_1 and S_2 , and consequently should be ignored in calculating the power of a player. Secondly, not every swing in a spatial MWC should be taken into account in a (descriptive) measure of power. In coalition S_3 only c 's swing should count because c is the only player that can apparently make a credible threat to actually exercise the choice of leaving the coalition (causing it to become losing). The argument is that because c 's position is equidistant between q and χ , c is indifferent as to whether q or χ prevails. Note, that if q would happen to be a little more to the right, then even the credibility of c 's swing could be doubted as it prefers χ to q .

Both examples deal with one and the same fundamental issue: not all coalitions are rationally feasible (although they are logically possible), and not every swing is 'credible', i.e. *will* be exercised by a rational agent. Why, the argument runs, consider states of the world which will not occur when analysing power relations and measuring a player's power? Surely, a descriptive measure of power must filter the set of logically possible coalitions and swings for their feasibility

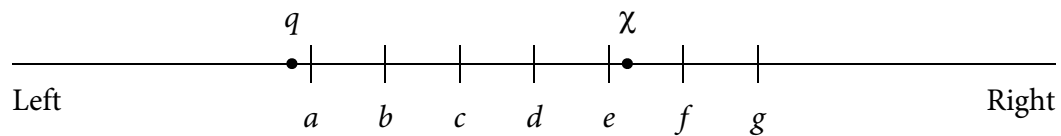


Figure 3.1

and credibility. That is, because real players are generally faced by choices and seek to maximize their utility, coalitions – to borrow Garrett and Tsebelis (1996: 278) similitude – ‘do not form like a motion of gas molecules in a container’. If player b and c in Example 3.1 both stand to lose by rejecting a ’s overture, why should we account for the power denoted by their swing in either $\{a,b\}$ or $\{a,c\}$? By the same token, in Example 3.2 we should not only ignore the swings of d , e , f , and g in S_3 , we should even ignore the swings of the players in S_1 and S_2 because these coalitions, while *logically* possible are not, what Rescher (1975: 146) would call, ‘*genuine* or *real* possibilities’. A valid descriptive measure of power must, therefore, make reference to the preferences of the player whose power we wish to measure.

Various proposals for a strategic or non-cooperative power index have been put forward, as noted above, by Tsebelis and Garrett (1996), Garrett and Tsebelis (1999a, 1999b, 2001), Steunenber *et al.* (1999), and Napel and Widgrén (2001, 2002, 2004). For our purposes, however, we need not concern ourselves with a presentation and discussion of the technical details of these indices because we actually want to take issue not with *how* one should incorporate preferences of player i in a measure of i ’s power, but *if* we should do so at all. It is sufficient to say that the basic intuition of a strategic power index is that the power index value assigned to a player should be related to how that player values the outcomes. In the case of Steunenber *et al.* (1999) this is measured in terms of the proximity of the equilibrium outcome to a player’s ideal point; while in Napel and Widgrén (2002, 2004) strategic power is taken, loosely speaking, as the expected contribution of a player to the equilibrium outcome.

4. *Generic Ability: The Fixed Core of Meaning*

Despite the intuitive appeal of the criticism encapsulated in the Examples 3.1 and 3.2 above, it is fundamentally mistaken. The reason hinges on a conceptual issue: what we mean by a power ascription.

Ordinarily speaking, a ‘power’ ascription refers to a person’s ability: what a person is *able* to do.⁵ In the game theoretic context that we are discussing, the abil-

⁵ See in particular Goldman (1970, 1972); Harré (1970), Morriss (1987/2002).

ity in question is to *effect outcomes* (i.e. ‘force’ or ‘determine’ outcomes) of the game. That is, a player has a strategy which if chosen will make a decisive difference to the outcome. This basic definition is the same for a power index based upon a simple game and one that is ostensibly based upon a non-cooperative game. The difference lies in the specification of the ability. In a simple game the ability is turning a winning coalition into a losing coalition or vice versa and thereby being decisive for the acceptance or rejection of a bill; while in a non-cooperative game, the ability is specified in terms of shifting the equilibrium in ones own favour.

Here lies the heart of the problem. The intuition behind applying non-cooperative games to the analysis of power is that if a player i is able to determine the outcome of a game but only by playing a dominated strategy, i.e. playing a strategy that is not a best reply to any other, it makes no sense to ascribe ‘power’ to this player for this ability. In other words, if a player can change the outcome only by doing something that he or she would never rationally choose to do, it is equivalent to saying that the player cannot determine the outcome with that strategy. Ergo, the player is powerless for this scenario. This is exactly what is hinted at in the Examples 3.1 and 3.2 above. Our claim, which we will now elaborate, is that this last conclusion is false, given what we customarily mean by ‘ability’. And if that is the case, then preference-based power of player i is an unintelligible concept.

To explain the problem, we can, without loss of generality, reduce our analysis in the first instance to the one-person case. Consider a player i who has a set of actions or strategies $A_i = \{a_1, a_2\}$ which is mapped onto a set of outcomes $X = \{x_1, x_2\}$ such that if i chooses a_1 , x_1 is the outcome; and if i chooses a_2 , x_2 is the outcome. In keeping with the standard assumptions of game theory, i is free to choose any element of A_i , i.e. i is not unfree to choose either a_1 or a_2 . This structure is what is meant by a *game form* (which is a game in which the utility functions (preferences) of the players remain unassigned).⁶

Now, in this game form, we can observe two ‘abilities’. The first is that i is able to choose an element of A_i . This is simply the trivial fact that the elements of the action set are feasible. The second, and more relevant, ‘ability’ is that by choosing an element of A_i , i is able to determine or force the outcome. Thus we can say that it is within i ’s power to see to it whether x_1 or x_2 occurs; or i possess power with respect to x_1 and x_2 . That is, by having available a strategy which can effect an outcome, we should ascribe power to i . Conversely, if a player does not possess a strategy that effects an outcome, then that player has no power (is powerless).

⁶ A game form g consists of four main features: a set of players N , a set of actions or strategies A_i for each player $i \in N$, a set X of feasible outcomes, and an outcome function π that yields some single outcome x for any given n -tuple $[a_i]$ of strategies, one strategy $a_i \in A_i$ for each player i . That is, $g = (N, \{A_i\}_{i \in N}, \pi)$. See Gibbard (1973).

This account of power in terms of a one-person game form could be slated on the grounds that as a social concept power clearly has more to it than an actor just being able to do what he chooses to do (*i* seeing to it whether x_1 or x_2 occurs). In a strategic context, the result of adopting a particular strategy may also depend on what the others can do. What Robinson could do on his proverbial desert island may have depended upon Friday's choices.

It would be mistaken, however, to believe that strategic interaction qualitatively affects the meaning of power as 'an ability to effect outcomes'. Strategic interaction only means that the sets of outcomes that a player can effect may not be singletons (i.e. one member subsets of the outcome set X). All that matters is that the set of outcomes that a player's strategies can effect different subsets of X . There is nothing in the meaning of 'ability' or power that says that a player must be able to realize specific elements of X to have power. Only a dictator can guarantee this; and one can have power without being 'all powerful'. For a player to be ascribed power it is sufficient that the state of the world would be different in absence of that player's intervention.⁷

The spirit of a power ascription is, then, as follows: *If* player *i* wanted a particular outcome or set of outcomes, and that *i* has an action (or sequence of actions) such that the performance of these actions under stated or implied conditions will result in that outcome or set of outcomes and would not result if *i* would not perform this action (or sequence of actions), *then* player *i* would perform this action (or sequence of actions) and the specified outcome or set of outcomes would obtain. That is, *i* is essential or non-redundant for an outcome or set of outcomes.⁸ When we ascribe power to a player we are, therefore, (1) making a claim about what a player is *able to do* under specified conditions irrespective of the occurrence of these conditions and thus, (2) describing a *capacity* or *potential* of a player, i.e. what a player *could do if* the specified conditions were manifest.

Power in this 'general sense' of the term is, to deploy Morriss' (1987/2002) terminology, a *generic ability* because it involves a *capacity* to do things that have an effect. This notion of a generic ability is what we take to be the natural 'fixed core of meaning' of power.

There are three important properties of power as a generic ability that we need to be cognisant of and which essentially rule out the inclusion of *i*'s preferences in a measure of *i*'s power. Firstly, a power ascription is indelibly *categorical*: it is 'like a promissory note, we need only believe that it is not [logically] impossible for it to be cashed' (Harré 1970: 91). Secondly, the *subjunctive nature* of a power ascription

⁷ The basic formal tools for analyzing the power of players in game forms are known as *effectivity functions*. See Peleg (1984) and Moulin (1983). Vannucci is an accessible formal review (2002).

⁸ Note that this corresponds to the definition of a swing in a simple game. See section 2, above. This framework is also used by Goldman (1974) in his analysis of the Shapley-Shubik index.

leaves the matter of what i wants undefined. And thirdly, a power ascription does not say *how much* power i has, only that there exist circumstances in which i is non-redundant for the outcome; a measure of power – a power index – aggregates these ascriptions of non-redundancy in some way.⁹

5. *The Core Theorem*

Now, the central claim of this essay, which we have outlined in the introduction, is as follows:

Core Theorem of the Measurement of Power If power is the ability of i to affect an outcome, then a measure of i 's power must exclude any reference to i 's preference (behavioural content) with respect to affecting that outcome.

There are three basic reasons for excluding i 's preferences in a measure of i 's power, when taken as a generic ability. These are: (1) being disinclined to do something does not imply the inability to do it; (2) psychological states such as desires and wants are not normally applied to the concept of ability; (3) the exercise of an ability is not to be conflated with its possession.

5.1 *Disinclination and Inability*

The definitional framework of a game form may appear somewhat trivial, but its implications for the notion of preference-based power are not. Assume that i has a preference relation on X , i.e. i might prefer one outcome to the other or be indifferent between them, i.e. we have a game. How will this shape i 's ability to effect an outcome?

For the sake of simplicity and, as we have indicated above, without loss of generality, let us return to our one-player case. Let us say that our player i is a lonesome Robinson on his proverbial desert island. He is fortunate enough to have available two possibilities to entertain himself to while away the hours between fishing and collecting coconuts: he can either read the e-book version of *Treasure Island* (outcome x_1) or watch the Paricival DVD (outcome x_2) on his solar-power notebook that by some good fortune he has with him and which is still fully functional. Robinson definitely prefers the peace of reading Stevenson under the shady palms than suffering the musical torments of Wagner – it wasn't him who actually packed the DVD in the notebook in the first place. The question to be answered is, in what way does Robinson's preference for reading Stevenson (x_1)

⁹ It might be objected at this point that we have ignored the 'power to'–'power over' distinction in our analysis. For the sake of continuity of argument, we deal with this in section 6.

make him able to perform those actions that result in him reading Stevenson (a_1) but unable to perform those actions (a_2) that result in him watching Wagner?

The question may appear to be a little odd, but it is one of the central conceptual issues that a preference-based power index must confront, but also one that has up to now been ignored. Recall Example 3.1. In their discussion of this case, Napel and Widgrén (2001, 2002) claim that players b and c are ‘robbed of their swing’.¹⁰ That is, b and c are effectively unable to choose an element of their strategy set, in this case leaving a coalition with a and forming a blocking coalition $\{b, c\}$ because they prefer a pittance of payoff to none at all. It is this fact that effectively makes these players dummies.¹¹ The same problem holds for all bar c in the MWC $\{c, d, e, f, g\}$ in Example 3.2: because all are better off under the proposal χ than under the status quo q , they are all effectively unable to change the outcome and therefore have no power in this scenario.

In order to find an answer to our question, let us take a step back and ask ourselves how can the non-performance of an action be explained. As far as we can determine, and here we follow the philosophical literature on possibility and counterfactuals, the non-performance of an act may have resulted from two quite distinct factors: either (1) our *inability* to do so, or (2) our *disinclination* to do so.¹² While for (1) we can say that there are prior conditions that necessitated that we are *unable* to perform an action – it was made impossible – and therefore cannot be an element of A_i (because A_i is the set of possible actions), this is not so for (2). Even though there may exist prior conditions that necessitate our being *disinclined* to perform an action – it is too painful – and therefore ‘necessitate’ that we do not undertake the action, it does not follow that we are *unable* to do so. To put it bluntly, if it is not impossible for i to perform, say a_2 , and i either would never conceivably perform a_2 because i prefers a_1 or is observed not to have performed a_2 , it is not because i is *unable* to perform a_2 , but because i does or did not want to.

Let us return to Robinson on his lonesome desert island. If Robinson does not watch Wagner because it is too gloomy and prefers to read Stevenson instead, surely it is absurd to conclude that that he is *unable* to watch Wagner. Given that Robinson’s notebook computer and the requisite programs are in working order, that the DVD disc is present, that Robinson knows how to operate the computer and its programs, etc., it follows that Robinson is able to watch Wagner by performing the requisite actions. That fact that he *will* not, or ultimately *does* not perform these actions, in no way vitiates his ability to do so. Moreover, even if

¹⁰ In Napel and Widgrén’s (2004: 23) this is rephrased as a player ‘having a swing that matters to the outcome’.

¹¹ Actually Napel and Widgrén classify them as *inferior players*, which are players who have swing but are as effective as a dummy.

¹² The literature on possibility and counterfactuals is vast. Here we follow both Goldman (1970) and Rescher (1975).

Robinson's not watching Wagner is necessitated by prior events, for instance an aversion to Wagner because it reminds him of his father and Robinson suffers an Oedipal complex, there is still no reason for saying that he is unable to watch Wagner.

In other words, and following Goldman (1970: 198–199), we take the necessary and sufficient condition for ascribing an ability to perform an action (choose a strategy in game theoretic terms) to be that the action is possible (not impossible). Taking into account i 's preference in a measure of i 's power violates this condition for ascribing power by conflating disinclination with inability.

5.2 Phobias and Strategies

Although we believe that the fact that a preference-based power ascription conflates a disinclination with inability is sufficient to rule out a preference-based power index, we have to deal with a subtle argument that could be thought of as a way around this problem. In his analysis of the term 'ability', Goldman points out that some may consider the 'possibility criterion' as too weak for a reasonable power ascription. Taking our Robinson Crusoe example as a starting point, some theorists might say that the fact that a_2 is possible (not impossible) does not *really* entail that i is *able* to perform it. Instead, it might be contended – and this seems to be what a preference-based power index is getting at – i must also *want* or have an *inclination* to choose a_2 , however remote that want or inclination may be.¹³ In Example 3.1 players b and c are 'robbed of their swing' precisely because as members of the species *homo oeconomicus* they will *never* have the inclination to reject a 's offer. For Robinson to *really* be able to see to it that he watches Wagner – perform those acts which result in him sitting in front of his notebook computer and the DVD spinning away – he must either have some desire to watch Wagner or at least be *able to want* to watch Wagner. The idea is that if, as *homo oeconomicus*, b and c are *unable to want* naught instead of an epsilon of payoff, why attribute them the power to reject a 's offer? Similarly, if due to his Oedipal complex Robinson is *unable to want* to watch Wagner because doing so will generate memories of his father and any memory of his father will so psychologically incapacitate him to the degree that he cannot even collect coconuts and fresh water so he will die of starvation thirst, and like any living being Robinson is genetically programmed to try and survive, then why attribute him the power to see to it that he watches Wagner instead of reading Stevenson?

The straightforward answer to these questions is simply that power ought to be attributed because the concept of ability is not ordinarily applied to psychological and behavioural states such as wants, desires, or preferences.

¹³ Napel and Widgrén (2002: 336), for instance, write 'for a player to be truly powerful, his preferences should matter in terms of outcome, i.e. a small change in preferences should lead to a small change in outcome'.

The obvious counterargument from adherents to preference-based power would be to say that just because the concept of ability is not ordinarily applied to behavioural states does not imply that it cannot be done. One could, of course, try to introduce a notion of ability that encompasses wants in order to conceptually shore up a preference-based power index. However, as we will easily demonstrate, the most that such a notion can do is to play a role in defining the game form, and not in analysing the game form or game itself and therefore it can only affect a power ascription in an indirect way.

What, then, does it mean to say that a person is 'unable to want to do something'? As far as we can make out it refers to the extreme case of a person having a phobia; that is, having an abnormal or morbid fear or aversion to some action or experience such that the performance of the action or undergoing the experience is in a very significant and psychological sense, impossible. That is, the phobia means that the person can under no circumstance voluntarily choose the action or experience. Or, put another way, the person's constitution predisposes them to be unfree to select a particular action from their act repertoire or choose a particular experience even though that act or experience is for all intents and purposes perfectly feasible.

It should be clear that a 'phobiafied' strategy brings us face-to-face with a basic inconsistency in a game theoretic analysis of power and one that has not yet been acknowledged by the preference-based power theorists. Such a strategy cannot be considered as a strategy at all. Recall that we said that the standard assumption of a game form is that i is free (not unfree) to choose any element of the strategy set, A_i . If it really is that case that i is free to choose any element of A_i , it would seem correct to say that it is possible (not impossible) for i to choose any element of A_i . If it were not the case, we could not say that i is free to choose any element of A_i . In which case, if we want to maintain the assumption of a game form that i is free to choose any element of A_i , we would actually have to eliminate the impossible strategy from A_i with the upshot that we have actually redefined the game form. By consequence, if a i is unable to want to do or experience something, then we are forced to remove the relevant strategy from A_i . At most it appears that the notion of ability that encompasses wants can be used to define a game form and therefore only indirectly affect a player's power. This would say that a Robinson with an Oedipal complex does not have the strategy in his strategy set that results in him watching Wagner, but a Robinson who merely find Wagner boring does. Hence the latter Robinson can be said to be more powerful than the former because he can do something that his alter ego cannot; but then again, the two Robinsons are in reality playing different games because they in fact have different strategy sets.

The point is plain. If we define power in a fundamental sense of 'a generic ability' (capacity or potential), which is what most, if not all, definitions of power (in a social context) do in some form or another, then the direct (assigning utility

functions) or indirect (eliminating strategies from the strategy set) inclusion of i 's preferences or behavioural states in a measure of i 's power has some peculiar, if not absurd, conceptual consequences. We are reminded of a joke of Lewis Carroll's in his *Sylvie and Bruno Concluded* which is worth quoting in full:

'Well, how much have you learned, then?'

'I've learned a little tiny bit,' said Bruno, modestly, being evidently afraid of overstating his achievement. 'Can't learn no more!'

'Oh Bruno! You know you *can* if you like.'

'Course I can, if I *like*,' the pale student replied; 'but I can't if I don't like!'

What we wish to say with this quote is that to believe in preference-based power is to be like Bruno who doesn't like learning no more; because he doesn't like learning no more he claims he can't. To wit, Bruno has 'lost' his ability to learn because of his preference. In case it is thought that this conclusion is inapplicable to the more strategic n -person scenario, one should think again. We need only consider the reply of all the players bar c in the spatial MWC S_3 in Example 3.2 when asked if they have the ability to make their coalition losing (i.e. change the social outcome)? 'No, sorry,' they say, aping Bruno emphatically, 'we can't because we don't like' (= 'have been robbed of our swing'), despite the fact that should any one of them actually leave the coalition would no longer be winning. This interpretive absurdity holds true for players b or c in Example 3.1 or in fact *any* index that is based on the notion of a spatial swing or pivot such. Napel and Widgrén's (2004) recent contribution is an example.

5.3 Dispositions and the Exercise Fallacy

The third, and possibly the most fundamental, reason for eschewing reference to i 's preferences in a measure of i 's power is related to the class of concepts that power belongs. Conceived of as a 'potential', 'capacity', or 'ability' makes power a *dispositional* concept akin to terms such as 'soluble', 'brittle', 'flammable', etc.¹⁴ One of the basic characteristics of such a dispositional ascription is that it is categorical due to the fact dispositions are independent of their manifestation or exercise. This immediately rules out the idea of spatial swings and pivots and even the idea of a preference or choice 'tremble' in the sense of Selten's (1975) perfectness concept that Napel and Widgrén (2004) use in order to give members of an MWC who do not have a spatial pivot some power.

The point is not difficult to make. Simply put, a particular grain of salt remains soluble even if it never happens that it is immersed in a sufficient amount of water

¹⁴ The dispositional nature of a power ascription is discussed in detail in Morriss (1987/2002) as well as in Dowding (1991, 1996). See also Oppenheim (1961), Goldman (1972), and Wrong (1979).

and for sufficient period for it to dissolve. Likewise, the United States Congress has the power (ability) to pass bills vetoed by the President (by passing them with two-thirds majority), even if the President never vetoes a bill. Thus when we say that a player has the capacity to effect an outcome (or do certain things) we are stating that *if* given conditions obtain, *then* that player can effect the outcome (or do those things). In the case of voting games, *if* (1) a given coalition S obtains in a voting game v , and (2) $S \in W(v)$ but $S \setminus \{i\} \notin W(v)$, *then* (3) i has power in S because i has the capacity to see to it that S is winning or losing. Whether i will exercise this swing is dependant upon other factors (such as having a reason to do so, i.e. i 's preferences). To use an almost Marxian turn of phrase, power in the game theoretic setting (as against 'natural' powers to do things like my picking up a suitcase, throwing a rock etc.) is a 'structural capacity' and does not therefore wax and wane with its exercise. The inclusion of i 's preferences or behaviour in a measure of i 's power via spatial swings or pivots as in Example 3.2 or Selten-like trembles not only does not wash with the categorical nature of a dispositional ascription, but it also conflates an ascription of the *possession* of a disposition (*having* power) with its *exercise*. This happens to be an instance of what is called the *exercise fallacy* (Morriss 1987/2002: 15–18).

Committing this fallacy is not to be taken lightly (even though it is prevalent in political science and some illustrious philosophers committed it).¹⁵ The problem is twofold: (1) it leads to false statements about who has power and how much, and (2) like the effect of the notion of ability applied to wants it renders it impossible to make independent positive or normative judgements about how institutional arrangements (rules of the game) and resources determine the distribution of power. Epistemologically speaking, a power index that commits the exercise fallacy is an unreasonable index.

6. Digression: 'Power to' and 'Power Over'

Before tying up this essay, we have to make a brief digression to do away with a possible suspicion that the notion of power as a generic ability is not congruent with the notion of power used in Examples 3.1 and 3.2 that underpins the intuition of strategic power. In current terminology, power as a generic ability is what is called 'power to' or 'outcome power' while the examples we discuss are ostensibly based on the idea of 'power over' or 'social power'.¹⁶ The difference is essen-

¹⁵ In his *A Treatise of Human Nature*, Hume (1739 [1951]) #1144] #1144] famously reduced power to its exercise. Some equally eminent philosophers sought to nail the fallacy. See, for instance, Ryle's discussion in *The Concept of Mind* (1949: 124).

¹⁶ The distinction, which was originally made by Oppenheim (1961, 1981), is discussed in Morriss (1987/2002: 32–35) and Dowding (1991: 48–51), from where the terms 'outcome power' and 'social power' come from. The distinction can also be found in Brams and Affuso (1976). A

tially that ‘power to’ concerns an actor’s ‘ability to bring about or help to bring about outcomes’ (generic ability), while ‘power over’ concerns ‘the ability of an actor to change the incentive structure of another actor or actors to bring about, or help bring about outcomes’ (Dowding 1991: 48) (i.e. ‘power over’ is an asymmetric relation between two or more actors). The case being argued for in these examples is that ‘swings’ are ‘power to’ and that this form of power does not entail ‘power over’, taken as the ‘ability to extract concessions’.¹⁷

In the light of this distinction one might be tempted to conclude that the core theorem is valid for a ‘power to’ ascription but not for a ‘power over’ one (it is preference-based). We believe that this rather neat and simple conclusion is false on two accounts.

First, it is not true that there is a disjuncture between the ‘power to’ of the players and their the ability to extract concessions in these examples. Consider once more Example 3.1, which is a game of imperfect information. Players *b* and *c* have a potential to extract concessions from *a* (have ‘power over’) on the grounds that if *a* makes an offer to *b*, not knowing the full history of the moves *b* might conclude that *c* has rejected a similar entreaty from *a*. Player *a* would then reason that if he were to approach *c*, *c* would reason likewise; hence *a* would accept a counter offer (make a concession) from *b* (or *c*). By this reasoning, any counter offer from *b* (or *c*) can be an equilibrium; so that the potential represented by *b*’s and *c*’s swings (‘power to’) implies an ability to extract concessions (‘power over’) because *a* would never make a concession to a dummy player. A similar argument can be applied to Example 3.2.

Second, even if the game of Example 3.1 was one of perfect information which would allow us to validly conclude that *b* and *c* cannot extract any concessions on the grounds of sub-game perfection (as in the usual ultimatum bargaining game), this does *not* imply that *b* and *c* are so because of their preferences. Rather, they are rendered powerless in this sense because of the information set: the ‘power to’–‘power over’ distinction relates not to the game but to the *game form*. My ability or inability to extract concessions from another does not hinge upon what I or they like; it hinges upon what I know. If the information set changes, so too does my ‘power over’. Thus, even if we find that ‘power to’ does not necessarily translate into ‘power over’ in a strategic setting, it does not follow to say that ‘power over’ is a preference-based concept and that the core theorem is inapplicable.

more recent and similar distinction in the voting power literature is that of Felsenthal and Machover’s (1998) ‘I-power’ (power as influence) and ‘P-power’ (power as prize).

¹⁷ It should be fairly obvious, however, that ‘power over’ implies ‘power to’: *A* cannot extract any concessions from *B* (has ‘power over’) unless *A* has the ‘power to’ do something to *B* that *B* cannot do to *A* (such as killing or wounding *B*). Or in the context of voting games, a dummies (null players) cannot extract concessions.

The distinction, then, between the two concepts of power may boil down to be related to nothing more than two components of the game form: (1) the size of the player set, ‘power over’ (social power) ascriptions necessarily involve at least two actors, while ‘power to’ (outcome power) ascriptions do not (Dowding 1991: 50–51); and (2) the informational structure.¹⁸ Neither of these elements can be said to affect the validity of the core theorem.¹⁹

7. Concluding Remarks

We wish to tie up this essay by enunciating the corollary of the core theorem, the usefulness of which is that it acts as an aid for making meaningful power ascriptions.

Corollary (Game Forms) The power of player i is given by the game form and not in the way i plays the game.

The corollary follows from the core theorem because by ignoring i ’s preferences we are, from the perspective of i , restricted to the game form. In other words, and to again borrow a little Marxist terminology, a valid power index is independent of the ‘use value’ of a player’s ability. Ergo, when making a power ascription we must separate off discussions about *the ability of an individual to do something* – shape the state of affairs (which includes extracting a concessions) – from discussions about *the value to an individual of this ability*. We also take it that power is independent of a player’s ‘exchange value’, as Hobbes already told us in Chapter X of *Leviathan*:

The *Value*, or Worth of a man, is as of all other things, his Price; that is to say, so much as would be given for the use of his Power: and therefore is not absolute; but a thing dependant on the need and judgement of another. An able conductor of Souldiers, is of great Price in time of War present, or imminent; but in Peace not so.

We do not dispute the obviously uncontroversial claim that that whole point of wanting or valuing power is to bring about outcomes that we like in much the same way as the whole point of wanting money is to buy things that make us happy. What we do dispute is that the ‘use value’ and ‘exchange value’ of my

¹⁸ We also suspect that set inclusion with respect to effecting outcomes is also a necessary demarcation criterion, but we cannot pursue this here.

¹⁹ It is worth noting that there is actually a gap in the literature regarding a detailed and comprehensive analysis of the various ways in which ‘power over’ is used. For instance, it is possible to construe Dowding’s use of ‘power over’ (social power) as a form of ‘power to by means of acting through others’. We suspect that a comprehensive ‘grammar of power’ could be adequately analysed in terms of game forms.

power is necessarily a proxy for my power. The fact that Hobbes' 'able conductor of Souldiers' commands a very low price in times of peace does not mean that he has lost his power to wage war.²⁰ Evidently, preferences are relevant for determining the value of power, for predicting what an agent may do with her power, or for relating it to concepts such as well-being; but from this it does not follow that preferences are necessary for the analysis of power *per se*.

As a coda to this essay we would like to mention two further points. First, if our core theorem is sound, then it transpires that we can head off any recrudescence of the usual criticisms trotted out against the framework of classical power indices that they ignore the preferences of the player whose power we wish to measure. Classical power indices satisfy the core theorem and its corollary because a simple game is actually a game form. In contrast, the notion of strategic power violates the core theorem and hence it cannot be taken to be saying anything intelligible about power in the fundamental sense of a 'generic ability'.²¹

Second, there is a very significant methodological by-product and contentious issue of our analysis that needs explicating. By demonstrating that power of a player is best analysed in a non-preference-based framework, we challenge the default presupposition of economists that ignoring preferences is methodologically unsatisfactory. Like its sister concept of freedom, power does not sit comfortably with preferences. A notable parallel with the freedom literature is also the relatively recent switch from the preference-based social choice theoretic framework to the preference-free framework of game forms.²² When studying power, we must always bear in mind that power is about events or potential outcomes themselves, not the utility attached to these outcomes.

We also hope that our analysis indicates that progress in the modelling and measurement of power requires an understanding of the conceptual issues involved. The present standoff between classical power indices and strategic power indices exists because a basic component of the debate has been missing.

²⁰ The 'power-value' distinction is another issue in the conceptual analysis of power that is still open. In a personal communication, Peter Morriss noted that no one has yet established the conditions under which there is a connection between power and value and what the nature of that connection is.

²¹ Our claim that classical power indices are in accordance with the core theorem does not mean we are suggesting that the power index sense of 'power' is the *only* correct use of the term.

²² See, for example, Sugden (1985), Gaertner *et al.* (1992), Dowding and van Hees (2003).

The Measurement of Freedom

1. Introduction

This paper is about the measurement of specific freedom – the freedom of an agent to undertake some particular action. In this regard, its general subject matter is not new. In a recent paper, Dowding and van Hees (2003) discuss, for example, the need for, and general form of, a ‘freedom function’ that assigns a value between 0 and 1 to a right or freedom and that describes the expectation that a person may have about being in a position to exercise (‘being free to perform’) that right or freedom. The usefulness of such a function is that in principle it could be used to define threshold values for indicating whether or not a person has a particular freedom or legal right and therefore for making non-welfaristic judgements about social states or to design the assignment of rights related to government policy, public regulation, or legal rules.

Much light, however, still needs to be shed on the actual nature of such a function. In their contribution, Dowding and van Hees leave the matter more or less open, claiming only that extent to which a person is free to perform a particular type of action or right depends only on the probabilities with which each of the relevant instances of the action or right will not be prevented. A straightforward example is that of determining our ‘freedom of expression’. According to Dowding and van Hees, this is given by the probability that shouting ‘Down with the Government’ at Whitehall at a given time and date and doing the same thing at Piccadilly Circus, etc. will go unprevented.

Dowding and van Hees refer to the recent and burgeoning literature on measuring freedom for a hint as to how such a function could variously be defined (Arrow 1995, Carter 1999, Dowding 1992, Pattanaik and Xu 1990, Pattanaik and Xu 1998, Sugden 1998, Rosenbaum 2000). A perusal of this literature indicates, however, that as yet there is no agreed upon framework for defining this function as the ‘probability of being unprevented’. The papers by Arrow, Dowding, Pattanaik and Xu, Rosenbaum and Sugden are all concerned with ‘freedom of

choice' rather than with the 'freedom to do x ' *per se* (freedom *simpliciter*).¹ Even Carter's (1999) extensive analysis of measuring overall freedom as an aggregation of the probability of being unprevented to do x , y , and z does not suggest an explicit model for determining the 'input probabilities' into a freedom function. Instead, they enter into his measure as an exogenous variable.² It is, therefore, still an open question about (i) the source of the input probabilities and by implication (ii) how to aggregate these probabilities into a value as suggested by Dowding and van Hees. This paper provides a tentative answer to both issues.

In this paper it will be argued that the value describing i 's freedom to perform an action can be identified with the 'conditional probability of success'. This model makes an agent's freedom a function of the propensities of other agents to choose a strategy that does not oppose the agent performing an action *and* the 'decision rule', which is a function that maps strategy choices into a unique outcome. The basic idea is that an agent is free (is unprevented) to perform a specific action if she belongs to a subset of agents (a coalition) that can guarantee the performance of the action. In a slogan, 'freedom is membership of powerful coalitions',³ and a measure of specific freedom is the probability of being a member of a such a coalition. This clearly gives a twist to the meaning of 'success' which was independently introduced by Penrose (1946), Rae (1969), and Barry (1980a, 1980b) in the voting power literature.

In the process of constructing a freedom function I make four other contributions of general theoretical importance. First, I unearth an unrecognized link between the concepts of freedom and success.

¹ For a discussion of the importance of maintaining the distinction between freedom *simpliciter* and freedom of choice, see Carter (2004) and Kramer (2003a). However, one can – as van Hees (1998) has done – interpret the concept of an *opportunity set*, which underpins the freedom of choice literature, as expressing the extent of an person's specific freedom.

² In his review of Carter's measure of overall freedom, van Hees (2000) does not tackle this issue either. To the best of my knowledge, the only two papers that come anywhere close to hinting at a reasonable model for a freedom function are Sugden (1978) and Bavetta (1999). Both implicitly assume a game form. I will not discuss these contributions here because they are in fact only very suggestive; neither actually defines a freedom function in a precise way.

³ This gives additional substance to Steiner's (1994: 39) slogan that 'Freedom is the possession of things'. That is, membership of powerful coalitions is the condition for 'possession of that action's physical components'. I am grateful to Ian Carter for pointing out this extension of Steiner's claim. This 'coalitional' understanding of freedom can also be seen to give another perspective on the 'capability' or 'material wherewithal' view of freedom associated with the work of Amartya Sen and Philippe van Parijs. Explicitly figuring in this line of thinking is clearly an important task, but one that cannot be pursued here. Readers familiar with this literature should not take my lack of attention to it as a sign that I have put it exclusively into the 'exercise' and 'positive' category of freedom. Clearly, the coalitional understanding of freedom can be used to model what is meant that an individual is 'empowered' with capabilities or material resources. See footnote 20.

Second, I provide an answer to the age-old question of the relationship between power and freedom. Starting from a basic opportunity concept of freedom I am able to show that a specific freedom derives from a power structure and therefore power is the more basic of the two concepts. This conclusion itself hinges on a demonstration that an individual-agent based definition of negative freedom appears to be logically untenable. Generically speaking, individual (negative) freedom is a collective property, although under special cases it can be given an ‘individualistic’ expression.

Third, I address the issue of how to measure freedom in a strategic rather than the parametric setting of social choice theory that developed since Sen’s (1970b) seminal contribution. Although a number of writers have, for some time, considered this to be a necessary step (Nozick 1974, Gärdenfors 1981, Sugden 1985, Gaertner *et al.* 1992, Pattanaik and Suzumura 1996, van Hees 2000), it is still a largely underdeveloped area (Deb 2004).

Fourth, I add to the nascent literature that seeks to develop formal models of freedom on an explicit philosophical framework (Steiner 1983, Carter 1999, Dowding and van Hees 2003, Bavetta 2004). In other words, I do not take for granted a particular notion of freedom, but rather base my measure on a philosophically grounded generic concept and syntax of freedom.

The remainder of this paper is organized as follows. In the next section I set out in more detail the type of freedom concept that I will work with. In the third section I discuss in some detail a formal definitional framework of specific freedom. In section 4 I present the game theoretic measure of specific freedom. Section 5 is a discussion of a number of conceptual issues relating to the measure that I have constructed. Section 6 concludes.

2. *The Concept of Specific Freedom*

2.1 *Opportunity and Exercise Concepts*

When constructing a measure of freedom, it is essential to be clear from the outset about the type of freedom that we want to deal with. Primarily, this means distinguishing between an ‘opportunity’ and an ‘exercise’ concept of freedom.⁴ As Carter (2004), who employs this distinction in his dissection of Pattanaik and Xu’s (1990) axioms of freedom of choice, puts it, ‘where freedom is treated as an opportunity concept, it means the *possibility* for an agent of performing some action or

⁴ The opportunity–exercise distinction originates with a classic essay by Taylor (1979) in which he studies Berlin’s (1969) distinction between *negative* and *positive* liberty. Taylor argues that the gamut of views of negative liberty fall into either the opportunity or exercise concepts but positive views are only ever exercise concepts. That is, opportunity and exercise concepts are generic categories into which any concept of freedom can be classified.

actions' (Carter 2004: 64), where 'possibility' is understood as meaning a lack of constraints of various kinds. Taken in this sense, freedom is concerned with actions that *might* be performed, given the absence of constraints, at some moment subsequent (or identical) to that at which the agent possesses the freedom in question. On this view, freedom is a matter 'of what we can do, of what it is open to us to do, whether or not we do anything to exercise these options' (Taylor 1979: 177).

In contrast, freedom as an 'exercise concept' concerns the performance by an agent of some action or actions; it is 'to *do* certain things or to achieve certain outcomes in a certain way' (Carter 2004: 64). On this view, freedom usually involves exercising control over one's life, so that one is free to the 'extent that one has effectively determined oneself and the shape of one's life' (Taylor 1979: 177). Clearly the Hobbes–Bentham notion of negative freedom as simply the 'absence of external physical or legal obstacles' (Taylor 1979: 176) is an opportunity concept, while the Rousseau–Marx notions of positive freedom as 'self realization' or 'collective self-government' is an exercise concept.

Given that the aim of this paper is to define a freedom function that is applicable, although not solely restricted to, the language of legal rights, the type of freedom that we are interested in here is that of an opportunity concept. I make no apology for this restriction because the language of rights generally concerns the 'opportunity' to do things (voting, protesting, reading) and not 'exercising'. If I have a right to read a certain book, then I have that right whether or not I ever read it; or whether or not I read it as a Marxian 'species being'.

2.2 MacCallum's Syntax

Following MacCallum's (1967) now canonical analysis, we can define the opportunity conception of freedom as a triadic relation between agents, constraints (preventing conditions), and possible actions: '... freedom is thus always *of* something (an agent or agents), *from* something, *to* do, not do, become, or not become something ...' (p. 314). MacCallum summarizes this relation in the format of: ' x is (is not) free from y to do (not do, become, not become) z ,' where x ranges over agents, y ranges over 'preventing conditions' such as constraints, restrictions, interferences, and barriers, and z ranges over actions ('doings') or conditions of character (Marxian 'self-fulfillment' or 'realization of one's true nature') or circumstance ('becoming angry'). As discussed by MacCallum, disagreements about different conceptions of freedom boil down to different views about the content of the range variables, x , y , and z (e.g. whether the agent (x) is to be conceived as an individual or collectivity; whether the obstacles (y) are only external to the agent; and whether *any* action or condition are to be counted).⁵

⁵ For an application of MacCallum's triadic syntax to different measures of 'freedom of

For the purposes of defining a freedom function, it is clear that MacCallum's rather opened-ended syntax needs honing down. As Bavetta (2004: 34) has cogently noted, a conception of freedom requires a syntax *and* a set of arguments to fill in the content of the range variables in some specific way; 'it cannot coincide with the syntax itself.' Hence, a freedom function that can be used to analyse the extent of what Dowding and van Hees' (2003) call my *material* as against my *formal* rights or freedoms concerns not just *any* opportunity, but those that are *social*.

Obviously, socially determined opportunities restrict, without much ado, the domain of x to natural or juridical persons (or groups) and that of the preventing conditions (y) to those that are inflicted by the actions of other such agents or groups of agents.⁶ External conditions of natural origin in the 'wildest sense' are to be weeded out as are 'internal' psychological or neurobiological states of mind. This means – uncontroversially, I believe – that the form of the freedom function need not be applicable to determining the freedom of the mountaineer who has become physically stuck in a crevasse or the person who is hindered from performing an action because of a morbid fear or phobia, depression, or lack of awareness, etc. no matter how figuratively correct it may be to speak of their conditions in terms of freedom or lack of it. The language of freedom and rights, generally construed as a social relation, makes no sense in these circumstances.⁷

While restricting the domain of x and y is a relatively straightforward affair, the z variable requires a little more philosophical consideration. The issue is whether or not we should, like Carter (1999: 16–17), narrow it down to only possible actions or 'doings' or should we, like Kramer (2003b: 156–169), be more expansive and include 'beings' and 'becomings'? The answer, it turns out, cuts both ways. However, because the main contribution of this paper hinges on the y variable, I will ignore the issue and side with Carter on the grounds that by restricting ourselves to actions we obtain a concept of specific freedom 'on which all liberals, in a broad sense, can agree', by which he means, 'at least libertarians like Friedrich von Hayek and Robert Nozick and liberal egalitarians like John Rawls and Ronald Dworkin'. Kramer's position, while unquestionably valid for a fully fledged analysis of freedom, can be safely ignored in this context because it takes us into the Byzantine intricacies of the philosophy of action without adding anything to my own contribution.

choice', see Bavetta (2004).

⁶ In order to keep the analysis tractable, I pass over the valid philosophical problems of determining the *range* of afflictions or misfortunes imposed on x that qualify as *actions* (y) of other agents. Essentially this amounts to providing justification for the elements of the 'strategy sets' that are alluded to in section 3 and discussed more fully in section 4.

⁷ In a slightly weaker formulation, it could be said that I am addressing freedom in a 'political sense' as against freedom in a 'physical sense' or 'psychological sense'.

3. Formal Definitions

3.1 Basic Framework

Having elaborated the concept of specific freedom and flected around the extensions to be assigned to the variables in MacCallum's syntax (and more or less settled on a broadly liberal concept of negative freedom), we can now state a pair of definitions for making ascriptions of specific freedom and specific unfreedom respectively, i.e. the freedom for an agent x to perform an action z . We will use these definitions in constructing a freedom function. (To avoid notational confusion, I will denote agents by lowercase Latin letters and the z variable in the triadic syntax by φ).

Definition 3.1 (*Specific freedom*) ' i is free to φ ' if no non-empty set of agents prevents i from φ -ing.

Definition 3.2 (*Specific unfreedom*) ' i is unfree to φ ' if some non-empty set of agents prevents i from φ -ing.

Before proceeding further, we need to train some careful scrutiny on a basic aspect of these pair of definitions, in absence of which misunderstandings can easily arise. Although the definitions 3.1 and 3.2 appear nearby indistinguishable from Carter's (1999: 27) Kramer's (2003b: 3), and in part Steiner's (1994: 8) definitions – 3.2 in particular – they do in fact differ in a small but highly significant way. And it is this difference upon which the main contribution of this paper pivots.

The usual method, which Carter, Kramer, and Steiner employ, is to define prevention in terms of the action(s) of natural or juridical *individual* agents. Indeed, a review of the literature on specific freedom indicates a preoccupation with dyadic relations: j 's preventing or not preventing i from φ -ing. In the framework presented here, however, prevention arises not from the action of an individual agent, but from the combined actions of a non-empty *set* of individual agents, i.e. from individual agents who 'belong together' by dint of a common characteristic, that of the choice, coordinated or otherwise, of a specified action (or omission) that opposes ('is against'), but not necessarily one that can alone 'prevent', some other agent performing φ . In n -person game theory, these sets are referred to as *coalitions*.⁸ That is, we can regard a coalition (a group) as an index for a certain collection of actions by individuals.

⁸ Note that the literature that makes use of n -person game theory implicitly accounts for coalitions in our understanding of freedom, although this is not explicitly marked by the authors as being the *generic* formulation of a freedom ascription. See, among others, Gärdenfors (1981), Deb (1994), Peleg (1998), van Hees (1995, 2000). In a different context, Pettit (1996, 1997: 52) allows for 'collective agents' such as coalitions in his definition of freedom, but he neither discusses

It may be asked why we need to leap from agents to coalitions. The answer is that it is a solution to a logical conundrum that afflicts an individual agent-based definition. In order to convince the reader that we have the correct concept, let us consider the following example. Suppose we have a society made up of four members, denoted by the set $N = \{a, b, c, d\}$. Suppose further that there is some action φ that if a were to desire to perform it she would require the consent of at least two others. Now, if we apply an individual agent-based definition of specific freedom and unfreedom we will find that there is a configuration of agents who together prevent a from performing φ (by not giving their consent) but none of these agents can be said to be doing the preventing as such. This is the case when all other agents, i.e. the subset $\{b, c, d\}$, are against a doing φ .

To establish this, let us, without any loss of generality, take Carter's (1999: 27) definitions as our point of departure. (I do not take Steiner's because he only formally defines unfreedom and I do not take Kramer's because he includes a condition for an agent's personal ability to do something which in this context only complicates the issue without adding anything.) Carter says that an agent is free to φ 'if every other agent refrains from preventing her doing it' and she is unfree to φ 'if some other agent prevents her doing it'. Consider the configuration $\{b, c, d\}$, the members of which have chosen, either jointly or severally, not to consent to a performing φ . Obviously a cannot be free because, assuming that by 'refrains from preventing' means the same as 'not preventing', it is not the case that 'every other agent does not prevent a from φ -ing'.⁹ But by Carter's account, a cannot be said to be unfree either because this would require that 'some other agent' (at least one) is preventing a from φ -ing, which is not the case. This can be established as follows. If some other agent is preventing a , it means, *ceteris paribus*, that there is at least one agent who, if she were to decide otherwise, would see to it that a was free to φ . If we hold the decisions of c and d constant, and ask whether a would be free to φ if b were to consent, the answer is 'no' (because at least two agents must do so). So it cannot be said that b is doing the preventing. A similar question can be asked of c and d and in both cases the answer is also 'no'. Thus, while the non-fulfilment of

the necessity for doing so, the relationship between individual agents and the 'collectivity', nor the implications that follow.

⁹ The assumption that 'refrains from preventing' means the same as 'not preventing' is crucial, otherwise the fulfilment of Carter's condition for specific freedom is not straightforward. In a personal communication, Carter indicated that 'not preventing' is what he had in mind because he was not assuming anything about the opportunities or potential to prevent, i.e. it is not to be thought that this definition is referring to some action or strategy called 'refraining from prevention'. Note, therefore, that in the event of b and c consenting – i.e. the coalition $\{b, c\}$ is in favour – to a performing φ , this fulfils Carter's condition for the ascription of a specific freedom, even though not all other agents are consenting: a is free to φ because by consenting b and c are 'not preventing' and nor is d , who, despite being in opposition to a performing φ , cannot prevent it.

Carter's conditions for a specific unfreedom (it is *not* the case that some agent j, k, \dots, n is preventing i) logically entails the fulfilment of his conditions for a specific freedom (all other agents j, k, \dots, n are not preventing i) this entailment does not necessarily imply that an agent will possess a specific freedom, as our example unambiguously demonstrates.¹⁰

Hence, while there can be no dispute that the *logical* relationship between Carter's two definitions is correct (his use of the universal quantifier for freedom and existential quantifier for unfreedom assures this) what is disputed is the *acceptability* of *both* his definitions because they fail to pick up cases where there is absence of 'agental prevention'. By defining prevention in terms of coalitions instead of individual agents we rid ourselves, in a very natural and simple way, of the pathology of logically ascribing a the freedom to φ when she is in fact unfree to do so: a 'preventing coalition' can *always* be identified.

In rounding up this section, a general point seems in order. The source of the difficulty with the individual agent-based definition appears to be its very *strong* assumption about the nature of *power* relations: that they are individualistic. That is, every social state can be forced by *some* (at least one) individual agent (either i can see to it that she performs φ or there is some j, k, \dots, n who can see to it that she does not perform φ). As we have seen, this is neither logically nor empirically true. To belabour the point, if 'agental prevention' exists this is simply the special case of the singleton set. In the case of $\{b, c, d\}$ preventing a from φ -ing there is no such 'agental prevention', because none of $\{b\}, \{c\}, \{d\}$ can see to it that a is free to perform φ ; but there is if $\{c, d\}$ is the preventing coalition, because $\{c\}$ and $\{d\}$ can see to it that she is free to do so. In a social context, then, power is a property to be ascribed to coalitions and not to individuals – tempting as it may be, i is not to be confused with $\{i\}$ (Holler and Widgrén 1999).¹¹ To confuse the two is to commit what is best called the 'individualistic fallacy'.

3.2 A Modified Framework

Although definitions 3.1 and 3.2 can, and will, be used in constructing a freedom function, some may be quick to point to a weakness which appears to be very counterintuitive. I want to discuss this in this section and show how the definitions can be tightened up with a natural modification, albeit one that is not without its own logical peculiarity. The issue is important because the modification leads to a different form for the freedom function.

¹⁰ Obviously if by 'agent' Carter – or for that matter, any other theorist working within a similar framework – would include 'collective agents' such as coalitions within the meaning of 'agent', then this criticism would not hold. But a close reading of Carter (and the work of others) suggests that by 'agent' he (and others) means an individual 'person'.

¹¹ This is not the place to elaborate, but it should be noted that this result poses serious problems for the 'responsibility view' of freedom (Miller 1983, Kristjánsson 1996).

The problem to be addressed is that because the definitional framework is purely behavioural, in the sense that it concerns only the presence or absence of prevention and not the hypothetical choices of the agent whose freedom we are interested in, definition 3.1 will ascribe i the freedom to φ even if she were compelled by the *force* (not coercion or threats) of another to do so (which she could not resist). In the now proverbial case of a 's freedom to φ , this would be if $\{b,c,d\}$ or any of its two player subsets could not only see to it that a performs φ if a were to attempt to do so, but also see to it that she do so even if she were not to make such an attempt. That is, definition 3.1 ascribes a the freedom to φ even if $\{a\}$ cannot prevent the outcome in which a performs φ . This is obviously the favourite theme laboured by G. A. Cohen (1979: 9): 'that one is free to do what one is forced to do' by dint of the fact that one cannot, even by the force of another, do that which one is not free (unprevented) to do.

If we believe, and I am inclined to, that it is unnatural to ascribe as an instance of my 'freedom of expression' the case in which I am dragged by others to the gates of Downing Street and made to hold up a placard, then the natural solution is to impose a subjunctive restriction in definitions 3.1 and 3.2 that represents a minimal element of *personal agency* or *doing* (Cohen 1988: 245).¹² That is, freedom ascriptions require that we *do* something; when I am dragged by others to the gates of Downing street I am *not* doing anything. Hence, we arrive at:

Definition 3.1* (*Specific freedom*) ' i is free to φ ': if i were to attempt to φ , then no non-empty set of agents prevents i from φ -ing.

Definition 3.2* (*Specific unfreedom*) ' i is unfree to φ ': if i were to attempt to φ , then some non-empty set of agents prevents i from φ -ing.

To observe this restriction at work, consider once more the case of a 's freedom to φ . The restriction tells us to ascribe a the freedom to φ only in those instances that *if* she were to attempt it she would be unprevented from doing so, despite it being true that she can be forced to φ ; and likewise the unfreedom to φ in those instances in which she is prevented *if* she were to attempt it. Thus, even if it were true that $\{b,c,d\}$ can force a to φ , the subjunctive restriction says we *ignore* this case as a component of assessing a 's freedom to φ – but this is not to deny the truth of her being free to φ in this instance.¹³

¹² Although not necessarily in the sense of 'acting freely' (Dworkin 1970), which is an exercise and not an opportunity concept of freedom.

¹³ For those familiar with the literature on voting power, this heuristic step is akin to the one taken by Holler (1982) in the definition of his power index, the Public Good Index. Holler *ignores* the 'oversized' coalitions – he does not deny they may form – on the grounds that they are formed by 'luck' (in the sense of Barry (1980a, 1980b)) and therefore not relevant for making power ascriptions because they do not contribute to what a coalition can achieve.

Technically speaking, the subjunctive restriction says that the coalitions that we examine to determine a 's freedom *must* include a , i.e. $\{a,b,c,d\}, \{a,b,c\}, \{a,b,d\}, \{a,c,d\}, \{a,b\}, \{a,c\}, \{a,d\}, \{a\}$. This is the basis of the slogan introduced at the beginning: 'freedom is membership of powerful coalitions' – because in the technical jargon we will use later, the coalitions which are sufficient for a to φ are denoted as 'powerful'.

The logical peculiarity of this step that has been alluded to is that while the subjunctive restriction is quite natural and undemanding, definitions 3.1* and 3.2* can imply a state of logical limbo for i 's freedom because the antecedents may not be true (if i were *not* to attempt to φ) although the implications are. This is the state of affairs in which a does not attempt to φ but $\{b,c,d\}$ or any of its two member subsets nevertheless would not prevent her if she were to attempt it and cannot force her to do it either. Hence, i is logically neither free nor unfree to φ because the conditions in definitions 3.1* or 3.2* are not satisfied.¹⁴ As I wish to shy away from discussing the 'bivalent' (one is either free or unfree) and 'trivalent' (one can be free, unfree, or neither free nor unfree) views of freedom,¹⁵ I will simply side with the more intuitive bivalent position by *assuming* that an agent always attempts to φ , which means that the antecedent is always fulfilled.¹⁶

4. A Game Theoretic Measure

4.1 Types and Tokens

Having established (i) what we mean in a very weak sense by specific freedom (unfreedom) and (ii) the conditions for making a such an ascription, we can now turn our attention to the principle problem of constructing a freedom function that describes i 's expectation that she is free (unfree) to φ . Following Dowding and van Hees (2003), we want this expectation to reflect the different instantiations r_1, \dots, r_n , called *act-tokens*, of performing a particular type of action R , called an *act-type*, given by:

$$\Gamma_i(R) = \Gamma(p(r_1), \dots, p(r_n)) \quad (4.1)$$

¹⁴ In a sense this is a specious problem (but I mention it in order to ward off anticipated criticism) because the subjunctive captures the counterfactuality of the case when i does not attempt to φ , by saying what happens *if* i were to. So the state of affairs *when* i does not is irrelevant to the freedom ascription. I am grateful to Keith Dowding for pointing this out to me.

¹⁵ On this see Steiner (2001) and Kramer (2003b: 41–60).

¹⁶ I am aware that assuming that the antecedent is fulfilled (i always attempts) does not preclude the possibility of force being applied to i . My attempting to leave a room does not preclude that you will, and can, drag me out of the room at the same time. I ignore this possibility because it does not alter the fact that my attempt goes unprevented.

where $p(r_i)$ is the probability that an act-token r_i will not be prevented (the agent is free to perform an instance of the act-type or right, R).¹⁷ The basic idea, is that while the formal existence of a class of acts (an *act-type*), R , is given by the possibility that at least one of its instantiations, the *tokens* r_1, \dots, r_n , is *possible*, we want to determine how *probable* each of these tokens or instantiations are and from this derive a probabilistic judgement about the extent to which R can be said to *materially* and not just *formally* exist.

In the language that I have been using, an action φ can be taken as either an act-type, R , or an act-token, r_i , because a ‘specific freedom’ can be more or less ‘specific’ (Carter 1999, Steiner 1994, van Hees 2000). To use the example of ‘freedom of expression’ again, this is an act-type R that can be instantiated in the different ways we have said: r_1 is shouting ‘Down with the government’ at Whitehall at a particular time and date, r_2 is doing so at Piccadilly Circus, and so on. Each of these tokens can be specified further as act-types themselves: shouting ‘Down with the government’ at Whitehall alone or doing so with others, etc. An action to which there is unique corresponding event is an act-token; it is an action in which all spatiotemporal and physical components are specified. Thus in the example of a performing φ , each of the coalitions $\{a, b, c, d\}$, $\{a, b, c\}$, $\{a, b, d\}$, $\{a, c, d\}$ are the instantiations (tokens) r_1, \dots, r_n of a performing the act-type φ . Hence, given our definitional framework, we arrive at the central contribution of this paper: the natural way to define $\Gamma(\cdot)$ is on the domain of possible coalitions. From this we will demonstrate that specific freedom can be related with the notions of success and power.

4.2 Game Forms

To define $\Gamma(\cdot)$ on the domain of coalitions in a systematic manner, we have to skip through some game theoretic preliminaries. The basic concept that we need is that of a *game form* (all of which has been implicit in our example of a ’s freedom to φ). A game form is a specification of a finite set of outcomes X , a finite set of individuals (or players) $N = \{1, \dots, n\}$, a finite set of feasible actions or strategies A_i for each $i \in N$,¹⁸ and an outcome function π (or decision rule) that yields some single outcome x for any given n -tuple $[a_i]$ of strategies, one strategy $a_i \in A_i$ for each i , i.e. $g = (N, \{A_i\}_{i \in N}, \pi)$. A game form can be said, therefore, to specify the ‘rules of the game’.

¹⁷ Note that most of the recent philosophical literature on the measurement of freedom discusses types and tokens in some detail. See Steiner (1994), Carter (1999:), van Hees (2000), and Kramer (2003b).

¹⁸ If we would be interested in freedom under legal rules, then the set of feasible strategies should be restricted to those which are admissible. See Fleurbaey and Gaertner (1996) and Fleurbaey and van Hees (2000).

For our purposes, we are interested in a particular game form in which the outcome set, X , has two elements, either i can perform φ (φ_i) or cannot perform φ ($\neg\varphi_i$), i.e. $X = \{\varphi_i, \neg\varphi_i\}$ and in which each player (including i) has two possible strategies: to either agree that i should be free to φ or not, which we designate as $A_i = \{yes, no\}$. To be clear, by ‘strategy’ is not necessarily meant a particular action as such, but rather a ‘bundle of actions’; they should be seen as courses of actions. Depending upon the context of the specific freedom, the act of agreeing to or hindering a ’s freedom to φ may involve different things. It could be as minor as a nod or a wink or providing a signature; or it could involve moving a heavy object; or it could even be an ‘omission’ in the sense of not doing something that is required, either consciously or unconsciously. In any case, what is involved are many actions (to provide a signature I must pick up a pen, put the pen to paper, hand over the signed form, etc), each of which I must be free to perform.¹⁹ Note, then, that under this construction a specific freedom or unfreedom presupposes other prior specific freedoms and that the specific freedom or unfreedom in question is the outcome of a combination of such bundles of actions as determined by a ‘decision rule’, π .

Now, according to our definitional framework, π defines the subsets of agents, $S \subseteq N$, called coalitions, that can force an outcome in X . That is, we are looking at a game form with a very sharp distribution of power: a coalition S , which is a collection of members of N who have made the same strategy choice, has either *full power* (is ‘winning’) or *zero power* (is ‘losing’). Thus, in our example, a has the support of b and c in $\{a, b, c\}$ and this coalition has the power to see to it that a can φ , while its complement, $\{d\}$, is powerless (cannot prevent a from φ -ing); while a only has the support of b in $\{a, b\}$, which because it is not enough is therefore powerless to see to it that a can φ , while its complement $\{c, d\}$ is powerful (can prevent a from φ -ing).

Such a game form is also called a *simple game* and can be represented by a non-empty set $W \subseteq 2^N$ consisting of the winning coalitions. We assume, as is usual, that W satisfies three basic conditions: (i) $\emptyset \notin W$, otherwise all coalitions would be winning and no player could prevent anything; (ii) $N \in W$, i.e. the grand coalition is powerful; and (iii) if $S \in W$ and $S \subseteq T$, then $T \in W$, i.e. if a coalition is winning then additional support will not alter the outcome. Note that the first condition, $\emptyset \notin W$, guarantees the freedom game to be non-trivial because if the empty set is winning then every set would be one because every subset includes the empty set, which would imply that both φ_i and $\neg\varphi_i$ and would be the outcome. Note also that the non-emptiness of W implies that the specific freedom formally exists, i.e. it is possible.

¹⁹ This point is discussed in detail in Chapter 2: an element of A_i by definition presupposes that a player is free to perform that strategy; otherwise it would not be in A_i and not part of the game form.

To complete the preliminaries, there are four further definitions that we will need. The first concerns what we mean by saying that a player has *power* in a game form: we take it to simply be the ability of a player to bring about an outcome by changing a losing coalition to a winning coalition and vice versa. Formally, we say that i exerts power in S if $i \in S \in W$ but $S \setminus \{i\} \notin W$ (or power outside S if, and only if, $i \notin S \notin W$ but $S \cup \{i\} \in W$). Such instances are known as ‘swings’ or a player’s ‘decisiveness’. (Aggregating the swings gives us what are known as power indices – a frequently used one is the absolute Banzhaf score.)

The next two definitions concern types of players: a *dummy* or *powerless* player is one that can never effect an outcome (never has a swing) – for all $S \subseteq N$, $S \setminus \{i\} \notin W$ and $S \cup \{i\} \notin W$; a *dictator* is one in which $\{i\} \in W$ and all other players are powerless; and finally, a *veto* player (blocker) is a player that is a necessary member of $S \in W$, i.e. for all $S \subseteq N$, if $i \notin S$, then $S \notin W$ (or simply, $\{i\}$ is a blocking coalition).

4.3 The Conditional Probability of Success

Having identified the set of winning coalitions for a given specific freedom φ , it can be said that we have a freedom game form, $W(\varphi)$. We need make no other assumption as regards the decision rule, π ; we take it to be ‘natural’ in the sense that no social law or convention need be contained in it.²⁰

The essence of a ‘freedom game’ is that i ’s attempt to φ is, figuratively speaking, made in the form of a ‘proposal’ to the members of N who either accept or reject it by – again, figuratively speaking – choosing ‘yes’ or ‘no’ from their respective strategy sets; i ’s attempt to φ is registered by i choosing ‘yes’ from her strategy set.²¹ That is, as a heuristic device we assume that S and $N \setminus S$ form. With this apparatus at hand, Definitions 3.1* and 3.2* reincarnate as:

Definition 4.1* (*Specific freedom*) ‘ i is free to φ ’ if, for some S , $i \in S \in W(\varphi)$ ($N \setminus S$ cannot prevent i from performing φ , i.e. it is not a blocking coalition).

Definition 4.2* (*Specific unfreedom*) ‘ i is unfree to φ ’ if, for some S , $i \in S \notin W(\varphi)$ ($N \setminus S$ can prevent i from performing φ , i.e. it is a blocking coalition).

²⁰ Note: $W(\varphi)$ can be a weighted game, i.e. a game in which there are non nonnegative weights (w_1, \dots, w_n) attached to the players and a *quota* $0 < q \leq \sum_{i \in N} w_i$ such that $S \in W$ iff $\sum_{i \in N} w_i > q$. The weights can be taken to represent resources such as money, social status, or authority. They can also be seen as a way of operationalizing a ‘capability’ and ‘material wherewithal’ views of freedom (see footnote 3). Obviously this is a subject of future research.

²¹ This may not be so ‘figurative’ for a broad class of situations. My attempt to go from A to B may require that I ask my wife for her car keys; even the act of buying a bus ticket is a ‘proposal’.

Note: (i) In accord with definitions 3.1* and 3.2* $i \in S \in W(\varphi)$ implies that no set of agent prevents i from φ -ing because $S \setminus \{i\}$ is not preventing and neither is $N \setminus S$ because it is powerless; and if $i \in S \notin W(\varphi)$ then at least one set of agents is preventing because $N \setminus S$ has the power to do so. (ii) The subjunctive restriction in 3.1* and 3.2* is captured by the conditions $i \in S \in W(\varphi)$ and $i \in S \notin W(\varphi)$; without these restrictions we would have the much weaker $S \in W(\varphi)$ and $S \notin W(\varphi)$ respectively. (iii) Logically speaking, the conditions are sufficient but not necessary for freedom *per se*, because, as we discussed above, it may be true that i is free if $i \notin S \in W(\varphi)$ (this is discussed again below); if, on conceptual grounds we rule out ‘being free while being forced’, then the conditions can be taken to be necessary and sufficient. (iv) To avoid indeterminacy of the freedom ascription it is assumed that if $S \in W(\varphi)$ then i is free to φ , i.e. for each $S \in W(\varphi)$ we exclude the possibility that the decisive members of S (those with a swing) will not in fact permit i to φ (exercise their swing)²²; if this would be the case then these members by definition belong to $N \setminus S$ (it is also assumed that those members of N not in S are in $N \setminus S$).

Thus to speak of i ’s freedom to φ in a freedom game form $W(\varphi)$ is to speak of membership of a powerful coalition. Following the idea that $\Gamma(\cdot)$ is to be an aggregation of the probabilities of act-tokens r_i of an act-type R as in (4.1), then $\Gamma(\cdot)$ is precisely the probability of such an $i \in S \in W(\varphi)$.²³ To define $\Gamma(\cdot)$ in a more precise fashion we need some additional structure and notation because calculating the probability of an $i \in S \in W(\varphi)$ requires a probability model for S . This means incorporating a minimal, but necessary, amount of behavioural information. That is, for any coalition S that may arise we may either know, are able to estimate, or make a reasonable *a priori* judgement as to the probability $p(S)$ that the players in N will choose an element of their strategy set such that S occurs.²⁴ In other words, $\Gamma(\cdot)$ is made up of two components, the 2^N elementary events denoted by each $S \subseteq N$ and a probability distribution $p: 2^N \rightarrow \mathbb{R}$ that associates each S with its probability of occurrence $p(S)$. That is, $p(S)$ gives the probability that players in S consent to i performing φ (by choosing ‘yes’ from their strategy set A_i) and those in $N \setminus S$ will not (by choosing ‘no’ from their strategy set A_i). (As is usual, $0 \leq p(S) \leq 1$ for any $S \subseteq N$, and $\sum_{S \subseteq N} p(S) = 1$). Our freedom function $\Gamma(R)$ is, then, specified by the pair $(W(\varphi), p)$.

²² Note that i ’s freedom to φ does not imply the absence of the power of others to prevent her performing φ . This issue is discussed in more detail in section 5.

²³ Note that it may be more reasonable to restrict S to the set of *minimal winning coalitions* (MWC), W^m , where $S \in W^m$ if $S \in W$, but $T \subseteq S \notin W$. It is questionable if excess sized coalitions add to the freedom of i to perform φ . In our example, if $\{a, b, c\}$ is sufficient for a to φ , in what way does $\{a, b, c, d\}$ contribute to a ’s freedom? This is a question that can not be answered here.

²⁴ This does not include, however, information about intentions. The precise meaning of $p(S)$ is an open question. On the one hand it can be taken to reflect personal preferences; on the other it can be taken to reflect social structure and conventions. For a detailed discussion, see Braham and Steffen (2002).

With this basic set-up, the natural form for $\Gamma(\cdot)$ is given by a conditional probability. Note that we have said that the existence of an act-type R is given by $W(\varphi)$, so that for a given $W(\varphi)$ and p :

$$\Gamma_i(W(\varphi), p) =_{def} \text{Prob}\{\text{outcome is } \varphi_i \mid i \text{ chooses } \varphi_i\} = \frac{\sum_{S: i \in S \in W(\varphi)} p(S)}{\sum_{S: i \in S} p(S)} \quad (4.2)$$

To put it in another way, a player's specific freedom is simply a conditional variant of the notion of 'success' independently introduced by Penrose (1946) and Rae (1969) and more fully discussed by Barry (1980a, 1980b). It means, in a loose sense, getting the outcome you want, in this case φ -ing (recall the heuristic assumption accompanying 3.1* and 3.2*), without necessarily being able to force it (being powerful) – cashing out 'want' in the sense of choice and not in the sense of what you 'truly want'.

Before moving on to discuss some interpretive issues, I would like to briefly point out the significant effect that the subjunctive restriction in definitions 3.1* and 3.2* has. We have already seen that in its absence we would use the weaker conditions $S \in W(\varphi)$ and $S \notin W(\varphi)$. We would therefore write definitions 3.1 and 3.2 as 'free if $S \in W(\varphi)$ ' and 'unfree if $S \notin W(\varphi)$ '. This yields a different measure altogether: the probability of a winning coalition. For a given $W(\varphi)$ and p :

$$\Gamma_i(W(\varphi), p) =_{def} \text{Prob}\{\text{outcome is } \varphi_i\} = \sum_{S: S \in W(\varphi)} p(S) \quad (4.3)$$

Those familiar with the voting power literature will recognise (4.3) to be none other than Coleman's 'power of a collectivity to act' (Coleman 1971).

It should be obvious, however, that if we model the freedom game slightly differently and say that i is an 'outsider' and her attempt to φ is registered by her 'proposing' the action to the remaining $N \setminus \{i\}$ players in a game $W_{-i}(\varphi)$ (where the $-i$ in subscripts represents the reduced player set) and not by her strategy choice in that 'voting game', then (4.2) and (4.3) are formally equivalent. They are also conceptually equivalent because it is now as if we have two games $W_{\{i\}}(\varphi)$, with a player set $\{i\}$, and $W_{-i}(\varphi)$ forming a bicameral system $W_{\{i\}}(\varphi) \wedge W_{-i}(\varphi)$ and because $W_{-i}(\varphi)$ only has one player, the bicameral meet is equivalent to adding a single new veto player to $W_{-i}(\varphi)$. This has the result of saying that a winning coalition must contain i . Such 'veto power' was not assumed in (4.4) but it is implicit in the idea that we ignore all $i \notin S \in W(\varphi)$. Thus, under this alternative model we can still say that 'freedom is membership of powerful coalitions'.

4.4 Three Interpretations

Up to now I have not properly addressed the issue of what $\Gamma_i(W(\varphi), p)$ can really be said to accomplish.²⁵ Saying that it is a representation of $\Gamma(\cdot)$ which itself is a ‘measure of specific freedom’ does not say all that much because it is likely that any formal quantification of a concept such as freedom will not be without different shades of meaning. Those familiar with power indices will know this all too well. Any of the traditional power indices can be seen as having two primitive interpretations: as a direct quantification of an ‘ability’ called ‘voting power’ or as a ‘reasonable expectation’ of possessing this ‘ability’ (Holler 1998). It is not really any different here.

There are three immediate interpretations of $\Gamma_i(W(\varphi), p)$ that come to mind. The first is that it can be said to measure ‘the degree to which a freedom-type exists’ and is brought forward by the expression ‘measurement of specific freedoms’ with which this paper begins. This interpretation is problematic because as according to Steiner (1983), Carter (1999: 233ff), and Kramer (2003b: 169ff) it makes no sense to speak of degrees of existence; existence is binary, taking values from the set $\{0,1\}$, and *not* scalar, taking the values from the interval $[0,1]$. The existence of an act type is but the mere possibility of one of its tokens. If we submit to this the notion of existence then Steiner *et al.* are correct; and it finds its parallel in the non-emptiness of $W(\varphi)$. So for the moment, at least, we can eschew this interpretation.

What Steiner *et al.* would say is that as a probabilistic judgement $\Gamma_i(W(\varphi), p)$ captures ‘being probably free’ to perform a specific action, and not the extent of this freedom. And this is what is more properly meant by the locution ‘material existence’ that Dowding and van Hees use. This, however, does not solve the interpretive quibble because it yields two further options. The locution ‘being probably free’ can be cashed out as ‘the probability that *i* possesses an act-type freedom’ or ‘the overall degree of freedom that the act-tokens represent’, which is van Hees and Dowding’s position. The first of these is a grander claim than the second, but in absence of a more extensive probability model I am not sure that this can be defended. To do so, it would have to be shown that $\Gamma_i(W(\varphi), p)$ gives a value for all possible probability distributions over this act-type. Hence we are left at this point with saying that $\Gamma_i(W(\varphi), p)$ is closer to the more modest Dowding and van Hees interpretation.

5. Discussion

No great claims are being made for the freedom measure, $\Gamma_i(W(\varphi), p)$. From a

²⁵ I am indebted to Martin van Hees for drawing my attention to this issue.

theoretical point of view it is not an overly surprising result, even if it has not been obvious up to now; and from a practical point of view it is clearly of limited use – except in very specific institutional contexts where the game form is clearly defined or easily determined, it is unrealistic to believe that we can actually calculate an agent's specific freedom.²⁶ However, the measure does have a fair amount of conceptual and programmatic value.

First, the fine-grained process of constructing the $\Gamma_i(W(\varphi), p)$ has churned up a significant conceptual finding: an individualistic definition of specific freedom taken in a broadly negative sense appears to be logically unsustainable. Individual freedom is generically a collective property. Once this is accounted for we are able to bring the notions of freedom, success, and power into a single conceptual and formal framework. We have found that an agent's specific freedom can be identified with her expected success of performing an action; and this is distinct from their social power, although dependant upon the structure and distribution of such power. By way of construction of the measure we have obtained an answer to the age old and controversial problem of how power and freedom relate to each other in a conceptual hierarchy. Power is the more basic concept. Even the freedoms presupposed in the strategy sets A_i in a game form g presuppose a power structure: each element of the strategy set is itself an outcome of a prior (maybe trivial one-player) game form in which i is a 'dictator'. With respect to each element of A_i , a player can unequivocally see to it that she performs this strategy or not.

It would be mistaken to believe that this conceptual result is 'planted' by the given formalism. It is not: the game theoretic apparatus has been chosen as a way of solving a logical problem inherent in the Carter-type individualistic framework: that it can result in an ascription of freedom where there is in fact unfreedom. How general the result is, that is, whether it extends to other notions of freedom (such as Sen's capability view), is something for future investigation, although I suspect that the generality and flexibility of my game theoretic model would allow for this. This can be seen in the brief discussion of the republican conception of freedom *à la* Pettit (1997) below.²⁷

²⁶ Bureaucracies are a clear case of such an area of application. Here we generally find clear permission structures and decision rules. $\Gamma_i(W(\varphi), p)$ could also be used to give conceptual and empirical content to the management science literature on empowerment, which is often taken as meaning the 'freedom to do something' in an organisation. See, for example, Conger (1988), Gal-Or and Raphael (1998), Spreitzer (1995, 1996), Pfeffer (1992).

²⁷ An equally interesting line of investigation would be to consider how the taxonomy of freedom and power found canonical texts such as Hohfeld's *Fundamental Legal Conceptions as Applied in Legal Reasoning* (1923/1978) fits within the game theoretic framework presented here. See Dowding(2004) for an analysis that puts Hohfeld's typology into the context of the formal analysis of power and freedom.

Second, $\Gamma_i(W(\varphi), p)$ provides a very natural and simple answer to an outstanding and significant question in the literature on the measurement of freedom and rights; and it is an answer that sits comfortably with the increasingly accepted game form approach to freedom and rights. It is also an answer that satisfies a number of intuitively appealing properties. This is not the place to go into the matter in detail, but it is not difficult to see that the measure satisfies the following: (i) a *dictator* property (if i is a dictator with respect to φ , then i has maximal freedom to φ , i.e. $\Gamma_i = 1$); (ii) a *powerless player (dummy)* property (the addition of powerless players to N (or N_{-i}) does affect i 's freedom to φ); and (iii) a *resource monotonicity* property (if in a game form g , i has more of the requisite resources needed for φ -ing than she has in a game form g' , then i is at least as free to φ in g as in g'). The proof of (i) and (ii) are more or less trivial, following from the definitions of a 'dictator' and a 'powerless (dummy) player' respectively. Because a dictator is a member of every winning coalition the probability that i is free to φ is 1. In contrast, a dummy never effects an outcome so adding a dummy to the set of players merely doubles the number of winning coalitions: for each old winning coalition S you now have two, S and $S \cup \{\text{new dummy}\}$, so the probability of obtaining a winning coalition is unchanged. The proof of (iii) is a little more involved. For my purposes here, we need only say that it is a form of 'global monotonicity' (Levínský and Silársky 2001), and can be derived from the proof of Proposition 3 in Laruelle and Valenciano (2004a).

Third, inspection of $\Gamma_i(W(\varphi), p)$ indicates an interesting relationship between the distribution of power among the members of N_{-i} and i 's freedom to φ . Roughly speaking, as we move from a 'democratic' game in which each of the N_{-i} players has an equal chance to determine the outcome (whether or not i performs φ) to games in which power is increasingly concentrated in the hands of fewer and fewer players, i 's freedom will tend to decline.²⁸ While this is certainly unsurprising, what may be surprising is the fact that from a purely *a priori* perspective (i.e. when we make judgements behind a quasi-Rawlsian 'veil of ignorance' in which we apply the principle of insufficient reason and assume each player will say 'yes' or 'no' with equal probability), i 's freedom to φ reaches its nadir not when all power is concentrated in the hands of a single player – a dictator – but when it is concentrated in the hands of a few – an oligarchy.²⁹ Formally, an oligarchic game is one in which there is a set of veto players that is also a

²⁸ A systematic examination of this relationship would plot $\Gamma_i(W(\varphi), p)$ on the vertical axis and an index of inequality of power (such as those axiomatized by Laruelle and Valenciano's (2004b) on the horizontal axis. The idea would be for a fixed N and quota of weights q (see footnote 20) one would find different power distributions yielding 0.1 increments in the inequality index.

²⁹ The restrictiveness of an oligarchy was recognized by Oppenheim (1961: 198): 'The degree to which X is unfree to do x is a function of the number of actors Y who limit his freedom'; but he did not comment on the fact that, *ceteris paribus*, a dictator provides more freedom.

winning coalition; while a dictatorial game is one in which there is a single individual that is also a winning coalition. This is not a mystery because an oligarchic game is in fact a unanimity game – a game in which the only winning coalition is N – on a reduced player set: the set of veto players minus all other players (who are dummies).³⁰ Needless to say, gaining the consent of one person (a dictator) is easier, *ceteris paribus*, than having to gain the consent from two or more players all of whom *must* consent (an oligarchy).³¹ The basic point is that probability of a specific freedom does not *necessarily* decline monotonically with increasing inequality of power in N_{-i} . In fact, from a strictly *a priori* perspective, if i is interested in her pure negative freedom she would be indifferent between a ‘democratic’ and a ‘dictatorial’ game on N_{-i} (or a game $W(\varphi)$ in which there is one other veto player): in both cases the odds are even that i will be free to φ , $\Gamma_i(W(\varphi), p) = 0.5$.

To many, the parity between a democratic and dictatorial game form might be considered an unacceptable pathology of $\Gamma_i(W(\varphi), p)$ as an index of specific freedom. But this would be mistaken for it is merely indicative of their insensitivity to procedural aspects of a game form: that the democratic one is an ‘anonymous’ decision rule (what matters is *how many* are in favour of i performing φ , not *who* is in favour), while a dictatorial one is not. Clearly, one could argue that a reasonable measure of freedom ought to account for such qualitative differences on the grounds that we may intuitively feel that ‘procedures matter’ in much the same way as we may feel that ‘choices which matter’ contribute more to our freedom than ‘choices which do not matter’, or ‘matter less’. Or looking at it the other way, we may feel that the degree to which our freedom is curtailed depends somehow on the procedure by which it happens.³²

To argue this way is, however, to posit a conception of freedom different to that which I have used here. Hence, it is not the measure that is procedure insensitive but the notion of negative freedom itself. As already alluded to in Section 2, from the standpoint of freedom as mere unpreventedness, it is entirely irrelevant whether the preventing conditions are inoperative because of a particular agent or because of a group of some agents. Notice that this is not something that main-

³⁰ For a review of the basic properties of oligarchic games, see van Deemen (1997: 126–129).

³¹ If the N_{-i} players act independently, and each have the same probability, p , to consent to i performing φ , but not necessarily 0.5, then this ‘oligarchy result’ holds for any $p < 1$ (assuming of course that in the dictator game the dictator has the same probability to consent as the members of the oligarchy). The robustness of this result under an asymmetric constellation of player propensities is a matter for future investigation.

³² Obviously procedural aspects can be of relevance to how people *evaluate* and act in social circumstances. It is well known in experimental bargaining, for example, that the outcome of ultimatum games is sensitive to the presence of face-to-face communication (Roth 1995). This does not change the fact that there is no necessary connection between i ’s freedom to φ and how that freedom has been obtained. See also Gaertner and Xu (2004).

stream liberals like Isaiah Berlin would object to. ‘Liberty’, remarked Berlin in his *Four Essays on Liberty* (1969), ‘... is principally concerned with the area of control, not with its source’ (p. 129), and, he continues, ‘The answer to the question “Who governs me?” is logically distinct from the question “How far does government interfere with me?”’ (p. 130). He goes on to tell us that it is conceptually possible for there to be more freedom in a dictatorship than in a democracy.³³ The question ‘By whom am I ruled?’ (p. 130), Berlin says, belongs to the domain of the ‘positive’ conception of freedom. The point is that the desire to be governed in a particular way (by myself in particular) may be ‘as deep a wish as that of a free area of action ... But it is not a desire for the same thing’ (p. 130). To take issue, then, with this property of the measure is tantamount to a rejection of the negative conception of freedom, not of the measure itself.

To complete our discussion, it would seem necessary to make a brief excursus on the matter of a republican conception of freedom *à la* Pettit (1997). Republicans of Pettit’s mould would probably have difficulty endorsing $\Gamma_i(W(\varphi), p)$ on the grounds that its definitional framework is too weak in the sense that it only requires that potential constraints are inoperative for making a freedom ascription; not that they do not exist. For republicans, what definitions 3.1, 3.1*, and 4.1* lack is a reference to *i*’s *power* to φ (i.e. the ability of *i* to φ despite potential opposition by others) because they ascribe *i* the freedom to φ even if constraints *could* have been operative had *i* chosen to φ , but *would not have been*, because the set of agents that *could* have made the constraints operative would not have done so (because the members of such a set had no common desire to do so). The unacceptable upshot is that these definitions allow us to say *i* is free to φ even if it is at the grace and favour of some set of agents who could, at will and with impunity, make *i* unfree to φ (this is what Pettit means by ‘domination’ or ‘power’). For Pettit, such a situation is hardly deserving of the badge of freedom.

In Pettit’s (1996, 1997) terminology, what the definitional framework is lacking is a criterion of ‘non-domination’ or ‘antipower’. For *i* to be free to φ in a non-domination sense is to say that *i* cannot be prevented from φ -ing by *any* set of agents, viz. is immune from any interference. Only then, Pettit says, can we say that *i* is free to φ because there is no need for ‘luck, cunning, or fawning’ for unpreventedness. Freedom in this view is to enjoy ‘noninterference resiliently’ (Pettit 1996: 589); and it is this form of noninterference (in our terminology, ‘unpreventedness’) that Pettit calls ‘republican freedom’.

The extent to which Pettit’s conception of freedom is sustainable is something that we can probe here. On the one hand it seems to be plausible because we may find it counterintuitive to say that the slave is free to φ if it is only possible at the behest of his master; on the other, it is equally counter-intuitive to say that he is unfree to φ given that his master has granted him permission. However, what is

³³ See, in particular, footnote 3 on p. 129 of *Four Essays on Liberty*.

important is that even if it is a conceptually sustainable notion, it has no bearing upon the general form that we have given $\Gamma(\cdot)$. To see this, we need only note that Pettit's notion of non-domination ascribes i the freedom to φ *only* in those instances in which the set of preventing coalitions is empty. To be free in this sense is not only being free in the *actual* world where there is no effective resistance ($\{b,c,d\}$ do not oppose a from φ -ing), but also to be free even if there were resistance ($\{b,c,d\}$ or any of its subsets do oppose her doing it). Freedom as non-domination means freedom in *all nearby possible* worlds (List 2004), which in accord with our definitions happens to be the set of all possible coalitions on the set of agents. In Dowding's (1991: 48) terminology, this is having full 'outcome power' (in the technical jargon we have used, it is being a *dictator* with respect to φ); and this is completely covered by definitions 3.1, 3.1*, and 4.1*. It is trivially true that if $\{a\}$ has unqualified power – immune from any interference or encroachment – with respect to a φ -ing, then the conditions contained in these definitions are always satisfied because whatever the configuration of other agents, no coalition ever prevents a from performing φ . The point to emphasize, then, is that a person can have a specific freedom in an unpreventedness sense without having that freedom in a republican sense; but if a person has a freedom in republican sense then she also has it in an unpreventedness sense. We do not need, therefore, an alternative framework to account for a republican conception of freedom; we only need to place a restriction on the set of winning coalitions, $W(\varphi)$.³⁴

6. Conclusion

In this paper I have outlined and defended an interpretation of specific freedom as 'membership of powerful coalitions' (or 'freedom as conditional success') and from this constructed an elementary and natural function for its measurement. In wrapping up I would like to briefly point to two lines of further investigation that have been opened up by this idea.

The first is to examine the possibility of using $\Gamma_i(W(\varphi), p)$ to construct a measure of *overall freedom*. This involves two tasks. The straightforward task is to determine the 'overall freedom of i ' as the expectation that i is unprevented to perform *all* the specific freedoms that i can conceivably have; while the less straightforward one is to aggregate the overall freedom for each i into an measure for all members of N , i.e. a measure that would allow us to answer the question 'how free in an overall sense is society A compared to society B ?' (Carter 1999: 28–29). While intuitively it would seem that the first task can be accomplished by sim-

³⁴ In the same way as we have suggested in footnote 23 that maybe we should only take into account MWCs. Note also that different conceptions of negative liberty can be contrasted on the domain of coalitions.

ply summing over the value for each specific freedom and then dividing by the number of such freedoms in the spirit of Steiner and Carter, it is unclear how to proceed with the second task. For it is not obvious – to me at least – that one can add the probability that i will be unprevented to φ to the probability that j, k, \dots, n will be unprevented to φ in a conceptually meaningful way.

The second line of investigation concerns developing a game theoretic method for assessing the robustness an agent's specific freedom or the *sensitivity* of the freedom to φ to potential changes in the behaviour of others. This line of thought, would take as its starting point a particular winning coalition and ask how sensitive this coalition is to a marginal behavioural change by its members (Napel and Widgrén 2004). Such a measure would allow an agent to form a reasonable expectation of being left unhindered to φ once that freedom has been granted. It would seem natural to say that, *ceteris paribus*, that I have more freedom to φ in a game g than in a game g' if my freedom in g is less sensitive to changes in the behaviour of others than in g' . This would be the basis of developing a freedom function that can be said to determine the probability that an agent possesses an act-type freedom. From this we would then be able to speak in more precise terms about the degree to which a freedom is respected and can be said to materially exist in a *robust* sense.

The Success of a Chairman

1. *Introduction*

The purpose of this paper is to re-examine a voting paradox that was first discussed in a brief way by Farquharson (1969) in his pioneering analysis of sophisticated voting. The paradox, which Farquharson christened ‘The Paradox of the Chairman’s Vote’ (p. 51), is concerned with the case of a committee in which a player who has a regular and a tie-breaking vote – the ‘chairman’ – can do worse under the plurality procedure than if he had only a regular vote.

Despite the fact that this seemingly ‘bizarre’ (Brams and Zagare 1977: 260) result cuts to the bone of our intuitions about two fundamental concepts we find in the vocabulary of political scientists and social choice theorists, that of *power*, and *success*, it has received surprisingly scant attention.¹ Only Brams and his colleagues (Brams *et al.* 1986, 1988, Brams and Zagare 1981, Brams 1976), and Niemi *et al.* (1983) have studied the paradox in detail. There have been two main branches to this research. One branch, pursued by Brams *et al.* (1986, 1988) has been to examine the existence of preference profiles that generate instances of the paradox for different rules (chairman only has a tie-break vote) and procedures (approval voting). The other branch concerns ‘escape routes’ from the paradox. Here there have been two approaches. One approach taken by Brams and Zagare (1977, 1981) is to investigate what they call ‘counter strategies’ of the chairman for three players and three alternatives by looking for conditions under which the chairman can induce his most preferred alternative in those cases where sophisticated voting leads to his least preferred alternative. To do this Brams and Zagare relax the basic assumptions of complete information, assuming instead that the chairman, by virtue of his unique position, can obtain information about the preferences of the other two players but not vice versa. This creates a class of what they call ‘deception vulnerable’ voting games and this can be to the advantage of the chairman. In contrast, the approach taken by Niemi *et al.* (1983) is to examine the exis-

¹ Although it has found its way into elementary textbooks on modern political theory, e.g. Ordeshook (1992) and Taylor (1995). Moulin’s (1983) more advanced text also briefly discusses the paradox in the context of dominance solvability.

tence and probability of the paradox under McKelvey and Niemi's (1978) binary procedures model of sophisticated voting. They find that under this construction the paradox disappears for the case of three players and three alternatives.

The objective of this paper is to pursue a more parsimonious escape route from the paradox. In place of assuming that a player's preferences are *linear* (or strict) orders, i.e. rankings which do not allow for indifference, which the chairman's paradox is based upon, we will assume them to allow for indifference, i.e. preferences are *weak* (or *pre-*) orders. It turns out that this small alteration to the setup is sufficient to rid us of the paradox for the canonical case of three players and three alternatives. It is worth remarking here that this modification of the canonical setup is in line with recent analysis of the classical 'paradox of voting' (Condorcet paradox) using weak orders (van Deemen 1999, Lepelley and Martin 2001, Tsetlin *et al.* 2003).

There are two additional contributions of this paper that need to be mentioned. First, we have had to introduce an alternative – albeit not entirely original – solution concept for chairman games, which we call *ultimately admissible coalition-proof Nash equilibria*. This is necessary because the standard method of *iterated elimination of weakly dominated strategies* (iterated weak dominance) is a too blunt an instrument given that a large number of the games are not *dominance solvable* (i.e. have multiple equilibrium outcomes after iterated weak dominance). Ultimately admissible coalition-proofness is essentially a refinement of iterated weak dominance by applying Bernheim *et al.*'s (1987) concept of *coalition-proofness* to shrink the set of equilibria.

Second, we bring to light a previously unseen paradox of sophisticated voting that afflicts chairman games that is seemingly more perverse than the classical chairman's paradox itself. This new paradox concerns the case in which, *ceteris paribus*, if a player (not necessarily the chairman) raises an alternative in his preference order, then the set of equilibrium outcomes may not contain either his most preferred alternative or that which he has raised in his preference order, whereas previously it contained his most preferred alternative. This paradox is not dependant upon the coalition-proof refinement of iterated dominance but occurs for a particular case of a dominance solvable game.

The remainder of this paper is organized as follows. In section 2, we briefly recap the canonical chairman's paradox; in section 3, we introduce the basics of weak preference orders and discuss its strategic implications; in section 4, we discuss the solution concept; in section 5, we calculate the probability of success for the case of three players and three alternatives; section 6 discusses the main result of this paper and points to the new paradox of sophisticated voting. To keep the exposition tractable, methodological and computational details are relegated to three appendices.

2. The canonical paradox

To illustrate the chairman's paradox more precisely, consider a committee consisting of three players $N = \{1, 2, 3\}$, a set of three outcomes $X = \{x, y, z\}$, and for each $i \in N$ a set of feasible strategies (or actions) of voting for non-empty singleton subsets of X , $A_i = \{\{x\}, \{y\}, \{z\}\}$. Adding the decision rule, or outcome function π that yields a single outcome for any given n -tuple of strategies, one strategy for each player, $\pi: \prod_{i=1}^n A_i \rightarrow X$, we have a *game form*, $g = (N, \{A_i\}_{i \in N}, \pi)$. A *chairman game form*, is a game form in which the decision rule is a *plurality procedure* (the alternative with the most votes wins and voting is simultaneous) with ties being broken by the vote of a specified or 'privileged' player (the *chairman*).

Next, for every fixed preference profile $\mathbf{u} = (u_1, u_2, u_3)$, with u_i a binary strict preference order on X , we have a normal form chairman *game* $g(\mathbf{u})$, taking as payoff functions utility functions that represent the preferences of the players.

Next, we say that for a player i strategy a_1 *dominates* strategy a_2 , or $\{a_1\} D_i \{a_2\}$, if, for any contingency, a_1 is at least as good as a_2 , and in at least one instance better.² A strategy that, for any contingency, is at least as good as *any* other strategy and in at least one instance better is said to be a *weakly dominant* strategy. A strategy that is *undominated* at \mathbf{u} is also called as *primarily admissible strategy* (a terminology that will be used throughout the analysis). Note that in Farquharson's terminology, a *unique* primarily admissible strategy of voting for a most preferred alternative is called a *straightforward strategy*.

Finally, we say that a player behaves *sincerely* if he does not try to anticipate how the others may vote and always votes for his most preferred alternative, and behaves *sophisticatedly* (strategically) if he tries to anticipate the behaviour of the others and based on this anticipation will iteratively eliminate strategies that could lead to the adoption of his least preferred outcome. The sophisticated strategies that survive this process of iterated elimination of weakly dominated strategies are known as *ultimately admissible*. To complete the preliminaries, it is assumed that for the class of games that we will examine, information is complete such that each player is aware of the whole profile \mathbf{u} as well as the decision rule.³

Now, suppose we have the canonical paradox of voting (Condorcet paradox) profile. We will designate player 1 as the chairman (which will always be the case unless otherwise stated):

² As a matter of clarification, D_i will also be used to denote domination at each iteration, i.e. reduction of the normal form game. There is no need to introduce new notation for this process because we are not concerned here with the sequence of iterations.

³ This is a standard assumption in the literature on sophisticated voting. It is usually defended on the grounds that it is at least as good, or better, than assuming incomplete information because the latter requires ad hoc assumptions about the nature and quantity of incomplete information (Felsenthal *et al.* (1988)).

Example 2.1 Let $g(\mathbf{u})$ be a chairman game, where:

$$\mathbf{u} = (x y z, y z x, z x y)$$

In this representation the triples are in order of the players 1, 2, 3 such that player 1 prefers x the most, y the next most, and z the least, and so forth. If each player behaves *sincerely* the outcome of $g(\mathbf{u})$ will be a tie among the three alternatives, so that the final outcome will be x , the chairman's most preferred alternative. If, however, voting is *sophisticated* (strategic) the outcome will be z – the chairman's least preferred alternative.

The sophisticated result comes about as follows. First, and in line with Farquharson's reasoning, we ask if a player has a primarily admissible strategy. In this example, the chairman (player 1) has such a strategy. To see this, suppose that players 2 and 3 both make the same strategy choice, they both choose $\{y\}$ or both choose $\{z\}$. In such an instance, the chairman can do nothing to improve his lot by choosing anything other than $\{x\}$. But suppose now that players 2 and 3 vote for different alternatives. In which case, if player 1 chooses $\{x\}$, then x will always win. Hence, the chairman can never do better, and may do no worse, than choosing $\{x\}$. Voting for $\{x\}$ is the chairman's unique primarily admissible strategy. If such a strategy exists and the player is rational, it is assumed that he will adopt it. (Note that a straightforward strategy is sincere, although the converse is not necessarily true.)

Second, having found a straightforward strategy for the chairman we can press on in a similar fashion for the other two players. It is easy to confirm that neither of them have such a strategy. That is, neither player 2 or 3 have a unique primarily admissible strategy, only the strategy of voting for their least preferred alternatives, $\{x\}$ and $\{y\}$, being respectively dominated. (It is a basic result of sophisticated voting that the strategy of voting for one's worst alternative – as well as abstaining – is always weakly dominated by voting for one's most preferred alternative.⁴) The situation therefore reduces to the following matrix given by the primarily admissible strategies of each player (subscripts refer to the players):

	$\{x\}_1$	
	$\{y\}_2$	$\{z\}_2$
$\{x\}_3$	x	x
$\{z\}_3$	x	z

⁴ Because (i) there will be at least one situation for which voting for one's first preference will result in that alternative but voting for one's worst alternative will result in another alternative; and (ii) one can never obtain a better result by voting for one's worst alternative than by voting for one's most preferred alternative. The formal proof is quite trivial and can be found in Brams (1994).

The third step now requires the complete information of players 2 and 3. Given player 1's choice, inspection of the above matrix shows that player 3 has a weakly dominant (the 'sincere' strategy), $\{z\}$. If player 2 knows the complete profile \mathbf{u} and assumes that player 3 will not play a dominated strategy in this reduced form game in which the chairman plays $\{x\}$, then y can no longer be elected so for him, $\{z\}D_2\{y\}$ and $\{z\}D_3\{x\}$. Hence the unique pure strategy Nash equilibrium in *ultimately* admissible strategies at \mathbf{u} will be $(\{x\}_1, \{z\}_2, \{z\}_3)$, giving outcome z . Notice that this result is purely non-cooperative in that there is no explicit cooperation between players 2 and 3 because the behaviour of each player is at once predictable.

The essence of the chairman's paradox is that if we assume that players have linear preference orders (as Farquharson and others do) then solving the $6^3 = 216$ possible games given by the set of all possible linear preference profiles in this way will mean that the *a priori* probability that the chairman gets his most preferred alternative is smaller than that of the other two players. This comes about because of the 48 possible first choice ties (there are 6 such ties each corresponding to 8 preference profiles), 12 are cyclical profiles as in Example 2.1 in which the chairman always gets his last choice. Of the remaining 36 profiles, each player gets his first choice on twelve occasions.⁵ The success for the 168 non-tied profiles is divided equally among the three players. Thus under the most widely used probability model (the IC assumption) in which each profile \mathbf{u} is assumed to be equally likely, the probability that the chairman gets his first choice is 0.61 while for players 2 and 3 it is 0.64.⁶ Farquharson summed up this paradox by saying: 'Therefore to get your way, seek extra power in a sincere committee, but shun extra power in a sophisticated committee' (p. 63). By 'power' he meant the number of 'votes' and *not* the more appropriate notion of 'decisiveness' that underpins indices of voting power.⁷ We will now show that this conclusion no longer holds if we allow for indifference.

⁵ One qualification must be made here. Of these 36 profiles, 12 lead to indeterminate or 'battle of the sexes' type games between the two non-chairman. Here we need to assume that each equilibrium is chosen with equi-probability. This is discussed in more detail in section 4.

⁶ There are two different probability models for this calculation, the Impartial Culture assumption (IC) and Impartial Anonymous Culture (IAC) assumption. See section 5 and Appendix 2.

⁷ As is well known from indices of voting power (e.g. Banzhaf, Shapley-Shubik, and Public Good indices), the number of votes (voting weight) is a poor proxy for the ability of a voter to alter outcomes in a committee. In this sense the chairman's paradox is more properly one of *success* – getting your preferred outcome – and not of power.

3. The effect of indifference

3.1 Basics

Our basic point of departure is the aforementioned binary *preference* relation, which we will denote as \succsim_i ('at least as good as'), that describes the opinion of each $i \in N$ over the set of alternatives X . *Strict preference*, \succ_i ('is better than'), is defined as $x \succ_i y \leftrightarrow [x \succsim_i y \ \& \ \neg(y \succsim_i x)]$; and *indifference*, \sim_i ('as good as'), as $x \sim_i y \leftrightarrow [x \succsim_i y \ \& \ y \succsim_i x]$. Note that up to now we did not make use of \succ_i , writing instead xyz . Where appropriate, we will use this latter notation, where the indifference relation $x \sim_i y$ will be written as (xy) .

A preference relation \succsim_i is said to be:

- (i) *reflexive* if $\forall x \in X : x \succsim_i x$;
- (ii) *complete* if $\forall x, y \in X : x \succsim_i y$ or $y \succsim_i x$;
- (iii) *transitive* if $\forall x, y, z \in X : [x \succsim_i y \ \& \ y \succsim_i z] \rightarrow x \succsim_i z$;
- (iv) *anti-symmetric* if $\forall x, y \in X : [x \succsim_i y \ \& \ y \succsim_i x] \rightarrow x = y$.

A preference relation is a *weak preference order* if it satisfies (i)–(iii); and it is a *linear preference order* if in addition it satisfies (iv). This condition says that if two alternatives are considered as good as each other then they are the same alternative. Note that a linear order is also a weak order. By $\mathcal{W}(X)$ we denote the set of weak orders over X . For the case of three alternatives that we will examine in this paper, the members of $\mathcal{W}(X)$ are:

1. xyz 2. xzy 3. yxz 4. yzx 5. zxy 6. zyx 7. $(xy)z$ 8. $(xz)y$
9. $(yz)x$ 10. $x(yz)$ 11. $y(xz)$ 12. $z(xy)$ 13. (xyz)

Next, for each $i \in N$ define a set $\mathcal{C}_i(X, \succsim_i)$, called the *choice set* of i , the elements of which are those elements of X that are not dominated by others with respect to \succsim_i , i.e. $x \in \mathcal{C}_i(X, \succsim_i)$ iff $\neg[\exists y : (y \in X \ \& \ y \succ_i x)]$.⁸ For clarification, note: (i) If the preference relation \succsim_i is a weak order, then each element of $\mathcal{C}_i(X, \succsim_i)$ is to be *weakly preferred* to all elements of X not in $\mathcal{C}_i(X, \succsim_i)$ and *indifferent* to all elements of $\mathcal{C}_i(X, \succsim_i)$. (ii) If \succsim_i is a linear order, then $\mathcal{C}_i(X, \succsim_i)$ is a singleton; if it is a weak order then it may have more than one element.

⁸ $\mathcal{C}_i(X, \succsim_i)$ is traditionally used to define the concept of choice because 'it describes a functional relationship in that it assigns a choice to each possible environment' (Arrow 1963: 15). In this regard, Arrow calls $\mathcal{C}_i(X, \succsim_i)$ a choice function, considering it to be a straightforward generalization of the demand function as it appears in consumer choice theory and under perfect competition, with the sets X there being budget planes. However, as we will see, we should not be misled by this analogy in a strategic context.

Although the assumption of linear orders is common in the social choice and game theoretic literature on voting procedures,⁹ it is actually hard to defend. First, the condition of anti-symmetry, which rules out indifference, is itself fairly strong: we may rank two candidates in an election as being as good as each other but that does not make them the same person.

Second, and more importantly, while linear orders may be thought of as a way of reducing complexity, it is erroneous to believe that it is a ‘harmless simplifying device’ (Niemi and Frank 1982: 152); it is simply untrue to say that ‘Little would be gained if one assumed that players may be indifferent between any two or all three alternatives’ (Felsenthal *et al.* 1988: 145). The claim that if $x, y \in \mathcal{C}_i(X, \succsim_i)$, then i would choose $\{x\}$ and $\{y\}$ with equal probability (Niemi and Frank 1982, Felsenthal *et al.* 1988) may be true in a parametric setting but – as we will see below – it does not necessarily hold in a strategic one.¹⁰ Failure to see this is probably the reason for dismissing the relevance of indifference. The next section will correct this misperception.

3.2 Strategic implications

To demonstrate the gamut of strategic effects of indifference, let us consider three cases.

Example 3.1 Let $g(\mathbf{u})$ be a chairman game where:

$$\mathbf{u} = ((x\ y)z, (x\ z)y, z(x\ y))$$

In this game, player 3 has a straightforward strategy, $\{z\}$; while the sets of primarily admissible strategies of players 1 and 2 are $\{\{x\}, \{y\}\}$ and $\{\{x\}, \{z\}\}$ respectively. The reduced game can be represented by the following matrix:

		$\{z\}_3$	
		$\{x\}_2$	$\{z\}_2$
$\{x\}_1$	x	z	
$\{y\}_1$	y	z	

⁹ For classic examples, see McKelvey and Niemi (1978) and Moulin (1983). For a very recent one, see Dhillon and Lockwood (2004).

¹⁰ One cannot even defend the use of linear orders for the ‘special case’ of three alternatives on the grounds that it is reasonable to expect that we can strictly rank so few alternatives. While it is certainly true that a strict ranking becomes increasingly difficult in a behavioural sense as the number of alternatives increase, this does not imply that I cannot be indifferent when faced with only three alternatives. For, I may only care about avoiding what I consider to be a worst outcome and am therefore indifferent between the other two alternatives; or I may know what I most prefer, and do not care what I get if I do not get my most preferred alternative.

Note that $\{z\}D_2\{x\}$ even though $x, z \in \mathcal{E}_2(X, \succ_2)$, leaving us with two strategy profiles which are also Nash equilibria: $(\{x\}_1, \{z\}_2, \{z\}_3)$, $(\{y\}_1, \{z\}_2, \{z\}_3)$. That is, while it is true that in $g(\mathbf{u})$ the chairman has no reason to prefer $\{x\}$ or $\{y\}$ because regardless of his choice the equilibrium outcome is the same, and may well choose one or the other with equal probability as Niemi *et al.* (1982) and Felsenthal *et al.* (1988) suggest, this is not true for player 2.

The next example concerns a game in which players 1 and 2 will randomise:

Example 3.2 Let $g(\mathbf{u})$ be a chairman game, where:

$$\mathbf{u} = ((x\ y)z, (x\ z)y, y(x\ z))$$

As in Example 3.1, player 3 has a straightforward strategy, $\{y\}$, while the set of primarily admissible strategies for players 1 and 2 are $\{\{x\}, \{y\}\}$ and $\{\{x\}, \{z\}\}$ respectively. The game therefore reduces to the following matrix:

		$\{y\}_3$	
		$\{x\}_2$	$\{z\}_2$
$\{x\}_1$		x	x
$\{y\}_1$		y	y

All four strategy profiles are Nash equilibria so that in contrast to Example 3.1 both players 1 and 2 are indifferent as to their strategy choices. There does not seem any intuitive way to single out one or the other equilibria as more ‘meaningful’ or ‘reasonable’ because the chairman always gets what he wants and the other two players have conflicting interests, cancelling out any possible incentive for the players to coordinate.

In the next example we examine a game with multiple equilibria that leaves room for equilibrium refinement and therefore the need for an even richer solution concept than iterated weak dominance.

Example 3.3 Let $g(\mathbf{u})$ be a chairman game, where:

$$\mathbf{u} = ((x\ y)z, x\ z\ y, z\ y\ x)$$

As in the two earlier examples, the chairman does not have a straightforward strategy. But unlike these two cases, this does not imply that he will randomise between $\{x\}$ and $\{y\}$. We contend that he will choose $\{x\}$ with certainty.

To prove this proposition, we start by representing the game in its reduced form (after deleting inadmissible strategies) given by the following matrix:

		$\{x\}_1$		$\{y\}_1$	
		$\{x\}_2$	$\{z\}_2$	$\{x\}_2$	$\{z\}_2$
$\{y\}_3$		x	x	y	y
$\{z\}_3$		x	z	y	z

Now we have $\{z\}D_3\{y\}$, leaving us with two Nash equilibria: $(\{x\}_1, \{x\}_2, \{z\}_3)$ and $(\{y\}_1, \{z\}_2, \{z\}_3)$. The first of these is an equilibrium in sincere strategies, the second is not; in that strategy profile player 2 plays a sophisticated strategy, $\{z\}$.

Should we take both equilibria into account when determining the success of the players? That is, is it legitimate to claim that the outcome is indeterminate and take it that each equilibrium is equiprobable? We believe not. For us $(\{x\}_1, \{x\}_2, \{z\}_3)$ is the more ‘meaningful’ outcome. If we make the usual assumption that an outcome of a game will be an equilibrium, then the chairman can guarantee x by choosing $\{x\}$ within the set of equilibria. Given that this choice is fully coincident with the interests of player 2, it would seem unreasonable for him to consider $\{y\}$ and chance it to end up with his worst alternative, z , which would occur because $(\{y\}_1, \{x\}_2, \{z\}_3)$ which has an outcome y is not a Nash equilibrium.

To support this argument we must allow for unlimited (but non-binding) pre-play communication.¹¹ If player 1 and player 2 were to meet and try and hammer out a coordinated strategy, they would agree on $(\{x\}_1, \{x\}_2, \{z\}_3)$ because their interests are perfectly coincident and for both of them it payoff dominates $(\{y\}_1, \{z\}_2, \{z\}_3)$. Note that because player 3 has an unique ultimately admissible strategy his decision is a fully decentralized one, ruling out any coordination problem with player 2. This implies that if player 2 were to seek to coordinate his actions with another, he would only do so with the chairman, who has no incentive to deviate from an agreement to play $\{x\}$ because he is indifferent between x and y , and as just mentioned $(\{y\}_1, \{x\}_2, \{z\}_3)$ is not a Nash equilibrium.

This reasoning allows us to actually say something stronger about $(\{x\}_1, \{x\}_2, \{z\}_3)$. Because no subset of players (a coalition) can make (i) a mutually beneficial *and* (ii) self-enforcing deviation from $(\{x\}_1, \{x\}_2, \{z\}_3)$, but can do so from $(\{y\}_1, \{z\}_2, \{z\}_3)$, the former is a *coalition-proof Nash equilibrium* (Bernheim *et al.* 1987), while the latter is not. Hence, because $(\{x\}_1, \{x\}_2, \{z\}_3)$ is Pareto efficient within the class of self-enforcing agreements the chairman will choose $\{x\}$ with certainty despite (i) $x, y \in \mathcal{E}_1(X, \succsim_1)$ and (ii) it is *not* the case that $\{x\}D_1\{y\}$.

¹¹ Actually it is questionable whether communication *per se* is necessary because all the players need to do is base their strategy choices on subjunctive reasoning of the form ‘if it were the case that ..., then ...’.

We leave to Appendix 1 a demonstration that the strategic effect of weak orders is not peculiar to chairman games but afflicts plurality voting games more generally.

Two preliminary results can be stated at this point. The first is that tempting as it may be, for plurality voting games it is imperative that we avoid conflating $\mathcal{C}_i(X, \succsim_i)$ with the set of *ultimately admissible strategies*. The proposition

$$\forall x, y \in \mathcal{C}_i(X, \succsim_i): x \sim_i y \rightarrow \neg\{x\}D_i\{y\} \ \& \ \neg\{y\}D_i\{x\}$$

is *false*. It is this error which mistakenly purports that indifference is irrelevant for sophisticated voting in plurality voting games.

The second result is that solving chairman games with weak preference orders requires a richer solution concept than that which is required for the canonical setup with linear preference orders. For the latter we need only ask whether the players are doing their best while acting *alone*. With weak preference orders this changes. By Example 3.3 we see that we now have to ask if players are doing their best given the possibilities to form a self-enforcing *coalition* with another. The introduction of coalition-proofness as a refinement device when iterated weak dominance fails to pin down a unique equilibrium outcome appears to be the natural way to do this.

4. Solution concept

Although the elementary components of the solution concept that we will use to solve the set of chairman games have been introduced – iterated weak dominance and coalition-proofness – it is necessary to explain why it is reasonable in the light of possible alternatives.

As a starting point, it is vital to note that there are in fact few alternatives at hand. Despite it being well-known that plurality voting games are not in general dominance solvable,¹² there have been surprisingly few refinement efforts. De Sinopoli (2000) is a recent attempt, although his results are not overly helpful because (i) he only relates classical refinements such as perfection, properness, and Mertens stability to iterated dominance; and (ii) of these three refinements he only characterizes perfection, which he shows to be wanting: at most it can exclude the Condorcet loser (i.e. the alternative that would be defeated by all other candidates in pairwise comparisons). To the best of our knowledge, the only other attempts at

¹² In fact Dhillon and Lockwood's (2004) characterization shows that the class of dominance solvable plurality voting games is quite small.

refinement are two unpublished papers by Feddersen (1993) and Messner and Polborn (2002). Both papers essentially introduce the same refinement that we do.

Let us now define our solution concept:

Definition 4.1 *Ultimately admissible coalition-proof Nash equilibrium* A strategy profile $A^* = (a_1^*, a_2^*, \dots, a_n^*)$ is an ultimately admissible coalition proof Nash equilibrium if the following two conditions hold:

- (1) the strategy profile survives the iterated elimination of weakly dominated strategies, and
- (2) the strategy profile is coalition-proof,

where coalition-proofness is defined recursively as (Bernheim *et al.* 1987):

- (2a) for any game $g(\mathbf{u})$ with player set N , $a^* \in A^*$ is *self-enforcing* if, for all proper subsets $S \subset N$, a_s^* is a coalition-proof Nash equilibrium in the game $g(\mathbf{u})_{a_{-s}^0}$, where $g(\mathbf{u})_{a_{-s}^0}$ is the game induced on the subgroup S by the actions of the complement of S , given by $a_{-s}^0 \in A^{-s}$ (in which $-S$ denotes the complement of S);
- (2b) for any game $g(\mathbf{u})$ with player set N , $a^* \in A^*$ is a *coalition-proof* Nash equilibrium if it is self-enforcing and there does not exist another self-enforcing strategy vector $a \in A$ such that $u_i(a) > u_i(a^*)$ ($\forall i \in N$).

The logic of (2) is clear and it has already been stated in the discussion of Example 3.3: an agreement between the members of S is *coalition proof* 'if it is efficient within the class of self-enforcing agreements because no coalition can benefit by deviating in a self-enforcing way' (Bernheim *et al.* 1987: 6).

To defend the solution concept, recall Example 3.3. This scenario amply highlights the important role that the failure of players to coordinate has for the multiplicity of equilibria. If we assume that a feasible decision of a committee is, and only is, a Nash equilibrium, then if players 1 and 2 fail to align their strategy choices, a lesser preferred alternative for both of them could win. Given that in small committees such coordination failure among players with coincident interests seems unrealistic, it is quite plausible to treat the multiplicity of Nash equilibrium outcomes in part as a coordination problem which is solved by coalitional arrangements. It is essential to emphasize that while the chairman game is one of a simultaneous moves, this should not be thought of as ruling out pre-play communication in which free discussion of strategies is permitted but in which binding commitment is not possible. Although pre-play communication is not part of the canonical setup there is no good reason to rule it out.

The question, then, is only how to incorporate collective non-binding agreements that result from pre-play communication into the solution concept. Here there appears only two bona fide possibilities. One is Aumann's (1959) *strong*

Nash equilibrium which requires that a strategy profile is immune to deviation by every conceivable coalition. The problem with this concept is, as Bernheim *et al.* (1987: 6) point out, that it is ‘too strong’ because it permits coalitions complete freedom so that the deviations may not be self-enforcing and therefore credible because they may lack the Nash best-response property. (Note that as a result of its conditions strong Nash equilibria almost never exist.) This then leaves us with Bernheim *et al.*’s coalition-proofness which corrects for this problem by stipulating that deviations are internally stable such that a deviation from a strategy profile would not trigger further deviations by sub-coalitions by the deviating coalition.

Having pinned down an appropriate coalitional based refinement of Nash equilibria, we are faced however, with a new problem. Characterizing committee decisions as a coordination problem that is in part solved by coalition-proofness no longer necessarily implies that we must accept condition (1). First, it can be argued that iterated weak dominance is now too strong; all that is needed is *primary* admissibility and not *ultimate* admissibility. Consider, for instance:

Example 4.1 Let $g(\mathbf{u})$ be a chairman game, where:

$$\mathbf{u} = ((x y)z, (x z)y, y z x)$$

Without going through the motions, there is a single coalition proof Nash equilibrium that survives iterated weak dominance: $(\{y\}_1, \{z\}_2, \{y\}_3)$. The problem, here, is that the process of elimination deletes a seemingly reasonable coalition proof strategy profile, $(\{x\}_1, \{x\}_2, \{y\}_3)$. That is, player 2 paradoxically rules out this possibility because after the first round of eliminations, $\{z\} D_2 \{x\}$. This is certainly hard to bring into line with the idea that committee decisions are coordination games. There appears no strategic reason for player 2 to rule out an agreement with the chairman that they jointly choose $\{x\}$ in the face of the fact that the only other coalition proof outcome is y . Player 2 has nothing to lose in this procedure from playing a weakly dominated strategy and therefore ultimate admissibility would seem not to consistently blend well with coalition-proofness.

Second, given the well-known fact that the order of elimination of weakly dominated strategies generally matters – although for the set of games we examine here it does not – replacing condition (1) with the weaker condition of primary admissibility would avoid this problem for a more general setting.

While clearly the second point has a good ring to it, the first does not. If we refer back to Example 3.1, which is a dominance solvable game we see that $\{z\} D_2 \{x\}$ so that the coalition-proof Nash equilibrium strategy profile $(\{x\}_1, \{x\}_2, \{z\}_3)$ is eliminated. But in contrast to Example 4.1, there is no advantage for player 2 in playing his weakly dominated strategy and in fact it is risky to do so if the chairman were to make a mistake at the ballot. That is, despite there being

no incentive for the chairman to deviate from $(\{x\}_1, \{x\}_2, \{z\}_3)$, it is strategically unstable.

We have visibly come face-to-face with a basic conceptual clash: weakening the solution concept allows too much; strengthening it excludes too much. Our hunch is that the restriction to ultimately admissible strategies is preferable on account that assuming players will *only* not play an inadmissible strategy at u is methodologically arbitrary. If we start with the reasonable assumption that no player chooses an inadmissible strategy at u , then it is equally logical to consider what might happen with the reduced set of strategies. In other words, there is nothing to stop players recalculating which strategies are weakly dominated for them given that others will also not play an inadmissible strategy. Hence, it appears to be a very weak stipulation that a solution concept for plurality voting games should be based on iterative weak dominance. We can square this with the exclusion of $(\{x\}_1, \{x\}_2, \{y\}_3)$ in Example 4.1 on the grounds that given $\{x\}$ is weakly dominated for player 2, the chairman, maintaining consistency of belief that no one will play a weakly dominated strategy will simply proceed to coordinate his strategy choice with player 3. While this leaves us with two coalition proof Nash equilibria $(\{y\}_1, \{z\}_2, \{y\}_3)$ and $(\{y\}_1, \{x\}_2, \{y\}_3)$ if player 2 happens to play his dominated strategy, it makes no difference to the outcome.¹³ In a sense we can ignore $(\{y\}_1, \{x\}_2, \{y\}_3)$ on the grounds that it occurs by ‘luck’ and not as a result of systematic strategic considerations by *all* the players.¹⁴

5. Paradox lost

As already indicated, the essence of the chairman’s paradox is not simply the *existence* of sophisticated outcomes in which the chairman ends up with his worst alternative because he possesses a tie-break vote, but his *overall success* given by the probability that he gets his most preferred outcome.¹⁵ That is, whether or not a

¹³ Note that ultimately admissible coalition proofness does not completely rid us of indeterminacy in plurality voting games; it merely reduces it. A fully fledged solution concept is beyond the scope of this paper.

¹⁴ This heuristic step is akin to the one taken by Holler (1982) in the definition of his power index, the Public Good Index. Holler *ignores* the ‘oversized’ coalitions – he does not deny they may form – on the grounds that they are formed by ‘luck’ (in the sense of Barry (1980a, 1980b)) and therefore not relevant for making power ascriptions because they do not contribute to what a coalition can achieve. Here we say that equilibria in weakly dominated strategies may form, but because they would do so only by ‘luck’, they are not relevant for ascriptions of success.

¹⁵ Note that we are explicitly using a very narrow definition of *success*. It could be argued that that it is also a success to avoid least preferred outcomes. This, however, creates conceptual difficulties which we cannot discuss here. For instance, suppose that i obtains her first choice while j his second. Does this make j as successful as i ?

player should seek both a regular and tie-breaking vote depends on the net advantage, or *overall success*, that this brings for *all* possible preference profiles; in other words, considering all possible ‘chairman games’ for a given set of players. It is composed of his *specific success* $\sigma_i(\mathfrak{s}, \mathbf{u})$ which is the probability that he will obtain his most preferred alternative in a *specific game* $g(\mathbf{u})$. A measure of *overall success*, $\Phi_i(\mathfrak{s})$, is, therefore, the weighted sum of *specific success*, given by:

$$\Phi_i(\mathfrak{s}) = \sum_{\mathbf{u} \in \mathcal{W}(X)^N} p(\mathbf{u}) \cdot \sigma_i(\mathfrak{s}, \mathbf{u}) \quad \forall i \in N \quad (5.1)$$

where \mathfrak{s} is a solution concept for normal form games, $\mathcal{W}(X)^N$ is the set of all possible weak order preference profiles \mathbf{u} on the set of voters N and set of alternatives X , and $p(\mathbf{u})$ is the probability of each $\mathbf{u} \in \mathcal{W}(X)^N$, i.e. each $g(\mathbf{u})$. A formal derivation is given in Appendix 2.

Note that before we can calculate $\Phi_i(\mathfrak{s})$ we need to define a probability model for $p(\mathbf{u})$. There are two standard models to be found in the social choice literature: the so-called *Impartial Culture* (IC) assumption and the *Impartial Anonymous Culture* (IAC) assumption (Gehrlein and Fishburn 1976, Fishburn and Gehrlein 1980, Lepelley and Martin 2001).

The IC model assumes that each possible profile \mathbf{u} is equally likely, i.e. $1/|\mathcal{W}(X)^N|$. The IAC model is more complicated. It assumes that each *anonymous preference profile*, or *a-profile*, is equally likely. The difference between an *a-profile* and a profile \mathbf{u} is that in the latter, the identities of the players are attached to each preference order u , while in the former this is not the case; in an *a-profile* we only know how *many* – not *who* – have a particular order u .¹⁶ Consequently, some *a-profiles* correspond to very few \mathbf{u} -profiles and therefore each $\mathbf{u} \in \mathcal{W}(X)^N$ may be weighted differently. It should be noted that whereas the IC model assumes that voters act in a completely independent way, the IAC model implies some degree of dependence between the voters. Additional explanations and computational details are given in Appendix 2.¹⁷

Finally, for the purpose of clarification, the determination of the equilibria that underpin our analysis was undertaken without the assistance of a computer algorithm. The rather daunting prospect of analysing $13^3 = 2197$ ($= |\mathcal{W}(X)^N|$)

¹⁶ That is, the profiles $\mathbf{u} = (x y z, x y z, y z x)$, $\mathbf{u} = (x y z, y z x, x y z)$, $\mathbf{u} = (y z x, x y z, x y z)$ are all instances of the same *a-profile*, while the profile $\mathbf{u} = (x y z, x y z, x y z)$ is an *a-profile* (or rather, the *a-profile* in which all players have the same preference order u has only one instance).

¹⁷ Because it is purely a matter of personal taste or analytical convenience whether one works with the IC or IAC models we compare the results for both models. Traditionally speaking, however, social choice theory has mostly made use of IC. The relative merits of both models are briefly discussed in Gehrlein and Fishburn (1976). Note, however, the IC assumption maximizes the probability of majority cycles (Tsetlin *et al.* 2003).

Table 5.1: Results

	(1)		(2)		(3)	
	Linear preference orders		Weak preference orders		Weak concerned preference orders	
	Players		Players		Players	
	1	2, 3	1	2, 3	1	2, 3
	$\Phi_i(\mathfrak{s})$	$\Phi_i(\mathfrak{s})$	$\Phi_i(\mathfrak{s})$	$\Phi_i(\mathfrak{s})$	$\Phi_i(\mathfrak{s})$	$\Phi_i(\mathfrak{s})$
IC	0.6111	0.6389	0.7298	0.7106	0.7080	0.7030
IAC	0.6786	0.6964	0.7238	0.7018	0.7138	0.7005

games is deceptive, because it is a fairly straightforward statistical exercise to cut it back to a manageable 273 games. Appendix 3 describes the method of reduction.

Table 5.1 summarises our results for three players and three alternatives. As it can be seen, equipping players with weak preference orders is sufficient to escape from the chairman's paradox for both the IC and the IAC model. Of note is that the chairman's overall success compared to the other players is relatively higher under the IAC model. The reason for this derives from the fact that under this model, a particular profile \mathbf{u} in which each player has a different preference order u is less likely than one in which two players have the same profile u . These are precisely the \mathbf{u} -profiles in which the chairman will suffer at the hands of sophisticated voting.

To round up. Although we have, strictly speaking, found a natural escape route from the chairman's paradox, it still remains in a weak sense because despite what our intuition may lead us to expect, there is no *substantial* advantage in terms of overall success in being the chairman under weak preference orders (column 2). The difference in overall success between the chairman and the other players is measly: 1.92 percentage points under the IC model and 2.20 percentage points under the IAC model. For all intents and purposes there is little practical difference in terms of the success of committee members between having a chairman as a tie-breaker and flipping a coin.¹⁸ On the other hand, the difference under linear preference orders (column 1) is not that great either: here the chairman was comparatively less successful than the two regular players by 2.78 and 1.78 percentage points for the IC and IAC models respectively. Any real edge, at least in the three players and three alternatives case, that a chairman may actually be observed to have will come from factors that lie outside the pure

¹⁸ From a normative standpoint this result can be considered as quite attractive, as a designated tie-breaker is not put at a disadvantage and receives a small fillip for bearing the burden of any additional responsibilities associated with this position.

structure of the plurality voting game – such as having privileged information (Brams and Zagare 1977, 1981) or being an agenda setter.

Having technically rid ourselves of the Chairman's paradox in its overall form (that is, for all possible games with weak preference orders), it still remains to be explained why allowing for indifference creates a qualitatively different result. The most prominent effect, although not the only one, is the advantage that a casting vote has in games with a player who is completely indifferent among all alternatives (called an 'unconcerned player'). In the class of such games in which the chairman is a concerned player, having the casting vote *always* gives him a greater specific success than not having such a vote. In some cases this is marginal, as in the game on $\mathbf{u} = ((x\ y)z, (x\ yz), (x\ yz))$. The specific success vector for this game is $\sigma(\mathfrak{s}, \mathbf{u}) = (0.889, 1.0, 1.0)$ if player 1 is the chairman, while if one of the others is the chairman, it is $\sigma(\mathfrak{s}, \mathbf{u}) = (0.875, 1.0, 1.0)$. In other cases, the effect is more substantial, as in the game on $\mathbf{u} = (x\ yz, (x\ yz), (x\ yz))$. The specific success vectors for the two scenarios are $\sigma(\mathfrak{s}, \mathbf{u}) = (0.778, 1.0, 1.0)$ and $\sigma(\mathfrak{s}, \mathbf{u}) = (0.5, 1.0, 1.0)$ respectively. This means that for the class of games in which the chairman is unconcerned, the concerned player(s) can never do better than the chairman were we to interchange their preferences. The extent of this effect can be gauged by looking at the overall success vector for the set of concerned profiles, i.e. the case of $12^3 = 1728$ preference profiles. The results of this exercise are given in column 3 of Table 5.1. Here we see that if unconcerned profiles are excluded (although there is no valid reason to do so) the paradox also disappears as the chairman has whisker of an advantage under both probability models (0.50 and 1.34 percentage points respectively).

6. Discussion

We have no illusion that the main achievement of this paper is far from glorious: three players and three alternatives can hardly be considered 'general'. At this stage it is difficult to speculate what will happen for more than three players even keeping the number of alternatives constant. In particular, there will probably be 'order-of-deletion effects' (i.e. the order in which dominated strategies are deleted) which will make solving the games more complicated. Another unknown effect is that of Condorcet cycles, which we know from the examples we have discussed is sufficient but not necessary to produce an instance of the chairman's paradox. With three alternatives and weak preference orders the probability of a Condorcet cycle increases as the number of players increase (Lepelley and Martin 2001). Whether this will be balanced out by other strategic effects of indifference we cannot say. Nor can it be conjectured at this point what will happen under alternative procedures such as approval voting. These are clearly topics for future investigation.

This said, the lack of generality should not detract from the significance of our main result. First, it must be kept in mind that the canonical setup with its assumption of linear preference orders is even narrower than the one used here. This has provided a very parsimonious escape route from an established paradox.

Second, weak preference orders for the ‘special case’ of three players and three alternatives requires a more sophisticated solution than for linear orders.

Third, one of the contributions of this paper pertains to the *significance* of indifference in individual preference relations in sophisticated voting. In a sense, the fact that we have been able to escape the canonical paradox does not atrophy its importance. Rather, its disappearance under weak preference orders brings the canonical paradox into stark relief: a paradox generated by strict preference orders can be thought to be qualitatively ‘stronger’ than if generated by weak preference orders (particularly with only three options).¹⁹ This said, however, it should not be thought that life in an environment of weak preference orders is uncomplicated. It is not. Using weak orders has meant that we have been able to adduce a new ‘choice set paradox’. Consider, for instance:

Example 6.1 Let $g(\mathbf{u})$ and $g(\mathbf{u}')$ be two chairman games, where:

$$\mathbf{u} = (x y z, (x z) y, z y x) \quad \text{and} \quad \mathbf{u}' = ((x y) z, (x z) y, z y x)$$

$g(\mathbf{u})$ has two equilibria after applying our solution concept (although coalition proofness is in fact redundant in this case): $(\{x\}_1, \{x\}_2, \{z\}_3)$ and $(\{x\}_1, \{z\}_2, \{z\}_3)$. That is, the chairman has even odds of obtaining his most preferred outcome in *this* game. However, in $g(\mathbf{u}')$ this opportunity disappears. This game is dominance solvable, with z being the unique equilibrium outcome given by $(\{x\}_1, \{z\}_2, \{z\}_3)$ and $(\{x\}_1, \{z\}_2, \{z\}_3)$ (again the coalition-proof refinement is redundant).

What we observe here is that despite the chairman’s unchanged support for x , raising of a losing alternative, y , in his preference order, *ceteris paribus*, results in neither x nor y . That is, his success completely collapses in the move from $g(\mathbf{u})$ and $g(\mathbf{u}')$. Expansion of the individual choice set $\mathcal{E}_i(X, \succsim_i)$ can be downright harmful in a strategic environment. This is not something that is restricted to a chairman. Consider:

Example 6.2 Let $g(\mathbf{u})$ and $g(\mathbf{u}')$ be two chairman games, where:

$$\mathbf{u} = ((x y) z, x z y, y z x) \quad \text{and} \quad \mathbf{u}' = ((x y) z, (x z) y, y z x)$$

¹⁹ I am grateful to an unnamed reviewer pointed out this interpretation.

$g(\mathbf{u})$ has four ultimately admissible coalition proof Nash equilibria, viz., $(\{x\}_1, \{x\}_2, \{y\}_3)$, $(\{x\}_1, \{x\}_2, \{z\}_3)$, $(\{y\}_1, \{x\}_2, \{y\}_3)$, $(\{y\}_1, \{z\}_2, \{y\}_3)$; but the move to $g(\mathbf{u}')$ (which is the same as that in Example 4.1) results in a single such equilibrium: $(\{y\}_1, \{z\}_2, \{y\}_3)$.

It is constructive to be aware that Example 6.2 is a cause for additional consternation because it presents us with a doubly counterintuitive effect. Here we find that not only does player 2's success completely collapse by raising, *ceteris paribus*, up a loser, z , in his preference order, but the loss is a result of his *own* rational calculus. The outcome x becomes infeasible by our solution concept because by iterated weak dominance $\{z\}D_2\{x\}$. In contrast, in Example 6.1 the collapse of the chairman's success is due to the rational calculus of *another* player, viz. in $g(\mathbf{u}')$ x becomes infeasible because by iterated weak dominance $\{z\}D_2\{x\}$. Phrased in another way, in Example 6.2 we have a situation in which the expansion of a player's choice set results in a rational calculus in which voting for a losing alternative weakly dominates voting for an equally preferred alternative that can win in $g(\mathbf{u})$.

What mileage there is in this paradox we cannot say because generalising on it here would lead us astray from the focus of this paper. It is, however, worth noting that: (1) The counterintuitive effect of moving from $g(\mathbf{u})$ to $g(\mathbf{u}')$ in both examples depends on moving an alternative *up* in a player's preference order, not *down*. That is, in Example 6.1 for instance, if y is not in the set of equilibrium outcomes and I reduce my estimation of x to that of y , it is not particularly surprising that I get nothing. However, it is surprising if I now raise my estimation of y to that of x .²⁰ (2) A cursory check shows that with three players and three alternatives the paradox is not restricted to plurality rule games as it holds up under approval voting even without a tie-breaking chairman (while it does not appear to afflict plurality rule with a random tie-breaker).²¹ (3) It appears to be a 'sophisticated' cousin of the monotonicity paradoxes that can be found in the social choice literature (Fishburn 1982, Nurmi 1999), and one that has gone unnoticed in the literature on sophisticated voting because of the preoccupation with linear preference orders. And finally, (4) the paradox appears to be particularly relevant to the freedom of choice literature: in a strategic setting, 'more can be downright harmful'.

There are two final points on which we would like to close this paper. The first concerns the relevance of the chairman's paradox. Although it could be argued

²⁰ A unnamed reviewer pointed out that the paradoxical effect of moving from $g(\mathbf{u})$ to $g(\mathbf{u}')$ may be dependant on the form of g because for this referee the collapse of success observed in Examples 6.1 and 6.2 does not appear counterintuitive under positional voting rules.

²¹ I am grateful to Steven Brams for pointing this out. Brams also noted that the fact that the paradox holds up under approval voting even without a chairman as tie-breaker makes it a 'purer' form of paradox.

that it is largely of academic interest because the *explicit* use of a chairman game form is rare in real-life,²² one could argue that it is more common than we believe. Zagare (1979), for instance, argues that even in plurality rule games without an explicit chairman, if the rule is such that in the event of a tie the status quo prevails, then the player who favours the status quo is effectively a ‘chairman’. On this interpretation, Zagare claims that the 1954 Geneva conference on Indochina exhibited an instance of the canonical chairman’s paradox (see also Brams (1990: 252–258)).

Second, we must underscore that our main result in no way depends upon a particular conception of *power* or *voting power* – as ‘control over resources’ (in this case voting rights or weights), as ‘control over actors’, or ‘control over events or outcomes’ (Brams *et al.* 1986, Hart 1976). The important concept throughout has been *success*. Whatever view of power one holds, the focus has been whether or not a player obtains his most preferred outcome, and not whether he has been instrumental in obtaining it.²³ The question of whether the chairman has more voting power in the sense of being more decisive for outcomes requires that we have an appropriate extension of power indices for plurality voting games. This is a task that still has to be accomplished.²⁴

Appendix 1

This appendix demonstrates that the strategic effects of indifference described in section 3.2 apply to plurality voting games in general. Let $g(\mathbf{u})$ be a plurality voting game with a random tie-breaker where:

$$\mathbf{u} = ((x\ y)z, (x\ z)y, y(x\ z))$$

The following matrix represents the reduced form game, where slashes indicate a tie:

²² As Niemi *et al.* (1983) note, the chairman game form is used in Swedish parliament (*Riksdag*).

²³ That is, you can get what you want without being powerful at all: you are simply ‘lucky’ (Barry 1980a, 1980b). And as Barry points out, ‘With enough luck you can get everything you want without any power’, where power is ‘opportunity to change outcomes from what they would otherwise have been’. Note that defining power in terms of an equilibrium is a very problematic issue, see Chapter 2.

²⁴ One approach has been suggested by Haunsperger and Melville (1996), but this is not convincing because it suffers from the conceptual problems discussed in Chapter 2. Under their approach different preference orders may result in different amounts of power. This would seem to be a measure of success, not power. Similar problems arise with the approach suggested in Brams and Fishburn (1983: 73–92).

$$\begin{array}{c} \{y\}_3 \\ \hline \{x\}_2 \quad \{z\}_2 \\ \hline \begin{array}{l} \{x\}_1 \\ \{y\}_1 \end{array} \left| \begin{array}{cc} x & x/y/z \\ y & y \end{array} \right. \end{array}$$

Now, if we assume that:

$\forall x, y \in X : x \succ_i y \rightarrow [\{x\} \succ_i \{x, y\} \& \{x, y\} \succ_i \{y\}]$ (where the second binary relation is over subsets of X and not its elements).

then player 1 has a weakly dominant strategy after the first reduction, viz. $\{y\}$. This assumption is eminently reasonable because all it asserts is that if a player prefers an alternative x to a alternative y , then he prefers the subset $\{x\}$ to the subset of tied outcomes $\{x, y\}$ which in turn is preferred to $\{y\}$ regardless of how the tie is broken (e.g. a flip of a coin), if the player believes that x and y each have a positive probability of being elected. This leaves us, then, with two pure strategy equilibria: $(\{y\}_1, \{x\}_2, \{y\}_3)$ and $(\{y\}_1, \{z\}_2, \{y\}_3)$ (both of which are also coalition-proof). Consequently, while player 1 will *not* randomise in $g(\mathbf{u})$, player 2 may well do so, because regardless of his choice it has no effect on the outcome, y .

Appendix 2

This appendix derives simple, exact expressions for a measure of *specific success*, $\sigma_i : g(\mathbf{u}) \rightarrow \mathbb{R}$; and *overall success*, $\Phi_i = f(\sigma_i(g(\mathbf{u}_1)), \sigma_i(g(\mathbf{u}_2)), \dots, \sigma_i(g(\mathbf{u}_n)))$, where $\mathbf{u}_1, \dots, \mathbf{u}_n$ are the elements of $\mathcal{W}(X)^N$, as the aggregation of specific success as discussed in section 5.

A2.1 Measures

To construct such measures, let \mathfrak{s} be a solution for normal form games. Then $\mathfrak{s}(g(\mathbf{u}))$ is the set of strategy profiles assigned by this solution in a chairman game $g(\mathbf{u})$, as defined in section 4. We discussed the solutions and combination of solutions to be used for such games with weak orders in section 4. Now, for each $i \in N$ define $\mathcal{E}_i(\mathfrak{s}, \mathbf{u})$ as the set of equilibria in $\mathfrak{s}(g(\mathbf{u}))$ such that the outcomes belong to the choice set, $\mathcal{E}_i(X, \succ_i)$. Then,

$$\sigma_i(\mathfrak{s}, \mathbf{u}) = \frac{|\mathcal{E}_i(\mathfrak{s}, \mathbf{u})|}{|\mathfrak{s}(g(\mathbf{u}))|} \quad \forall i \in N \quad (\text{A2.1})$$

is a natural measure of the *specific success* of player i in $g(\mathbf{u})$. Equation (A2.1) expresses the fact that for every $g(\mathbf{u})$ all equilibria are assumed to occur with equal probability. By

$$\sigma(\mathfrak{s}, \mathbf{u}) = (\sigma_1(\mathfrak{s}, \mathbf{u}), \sigma_2(\mathfrak{s}, \mathbf{u}), \dots, \sigma_n(\mathfrak{s}, \mathbf{u})) \quad (\text{A2.2})$$

we denote the specific success vector of a game $g(\mathbf{u})$, where subscripts denote the players.

Next, to obtain a general form of a measure of *overall success* of a player i given the game form g , we aggregate his specific success for all possible games, $g(\mathbf{u})$. The set of such games is merely the set of all possible preference profiles \mathbf{u} on the set of voters N and set of alternatives X with respect to the preference relation \succsim_i . In our context, it is $\mathscr{W}(X)^N$. But now we require a probability model because for any $g(\mathbf{u})$ that may arise we may either know, are able to estimate, or make a reasonable *a priori* judgement as to the probability $p(\mathbf{u})$ of each $\mathbf{u} \in \mathscr{W}(X)^N$ that generates $g(\mathbf{u})$. Thus $p(\mathbf{u})$ is given by a probability distribution, which is represented by a map that associates each profile \mathbf{u} with its probability of occurrence p :

$$p: \mathscr{W}(X)^N \rightarrow \mathbb{R} \quad (\text{A2.3})$$

As usual, $0 \leq p(\mathbf{u}) \leq 1$ for any $\mathbf{u} \in \mathscr{W}(X)^N$ and $\sum_{\mathbf{u} \in \mathscr{W}(X)^N} p(\mathbf{u}) = 1$. It is natural, therefore, to define $\Phi_i(\mathfrak{s}, \mathscr{W}(X)^N)$ as the weighted sum of specific success of a player i :

$$\Phi_i(\mathfrak{s}) = \sum_{\mathbf{u} \in \mathscr{W}(X)^N} p(\mathbf{u}) \cdot \sigma_i(\mathfrak{s}, \mathbf{u}) \quad \forall i \in N \quad (\text{A2.4})$$

This gives us the *overall success* vector of a game form g :

$$\Phi(\mathfrak{s}) = (\Phi_1(\mathfrak{s}), \Phi_2(\mathfrak{s}), \dots, \Phi_n(\mathfrak{s})) \quad (\text{A2.5})$$

where subscripts denote the players.

In absence of any other relevant information, the social choice literature (Gehrlein and Fishburn 1976, Fishburn and Gehrlein 1980, Lepelley and Martin 2001) has developed two standard *a priori* assumptions that can be made w.r.t. (A2.3): the so-called *Impartial Culture* (IC) and the *Impartial Anonymous Culture* (IAC) models. We will deal with each in turn.

A2.2 Probability models

A2.2.1 Impartial culture (IC)

Definition A2.1 (*Impartial Culture*) A randomly selected player $i \in N$ is equally likely to have each of the possible preference orders u from the set of preference orders $\mathcal{W}(X)$, i.e. $p_i(u) = 1/|\mathcal{W}(X)|$, where $p_i(u)$ is the probability of player i having order u satisfying the usual conditions of $0 \leq p(u) \leq 1$ for any $u \in \mathcal{W}(X)$ and $\sum_{u \in \mathcal{W}(X)} p(u) = 1$ (hence ‘impartial culture’). (In the three players and three alternatives case, $|\mathcal{W}(X)| = 13$.)

In this model,

$$p(\mathbf{u}_j) = p(\mathbf{u}_k) = \dots = p(\mathbf{u}_n) > 0 \quad (\text{A2.6})$$

implying that:

$$p(\mathbf{u}) = \frac{1}{|\mathcal{W}(X)^N|} \quad \forall \mathbf{u} \in \mathcal{W}(X)^N \quad (\text{A2.7})$$

which means that (A2.4) can be expressed as:

$$\Phi_i(\mathbf{s}) = \frac{1}{|\mathcal{W}(X)^N|} \sum_{\mathbf{u} \in \mathcal{W}(X)^N} \sigma_i(\mathbf{s}, \mathbf{u}) \quad \forall i \in N \quad (\text{A2.8})$$

For three players and three alternatives, $p(\mathbf{u}) = 1/2197 = 0.0004551$.

A1.2.2 Impartial anonymous culture (IAC)

Definition A2.2 (*Impartial Anonymous Culture*) Each possible *anonymous preference profile*, called an \mathbf{a} -profile, is equally likely.

An \mathbf{a} -profile, which has also been called a *return profile*, and *pattern* (Gehrlein and Fishburn 1976), is a function that assigns a non-negative number of players to each potential preference order $u \in \mathcal{W}(X)$ such that the sum of the assigned integers equals n , the number of players in N . Thus, the only data that is used in an \mathbf{a} -profile is *how many* players have a particular preference order u , not *who* (hence, ‘anonymous’).

Denote the members of $\mathcal{W}(X)$ as given in section 3.1 by k_1, k_2, \dots, k_m , then the set of \mathbf{a} -profiles, \mathcal{A} , is:

$$\mathcal{A} = \{(k_1, k_2, \dots, k_m) : k_i \in \{0, 1, \dots, n\}\} \quad \forall i \in N \quad \& \quad \sum k_i = n \quad (\text{A2.9})$$

In general, many different profiles $\mathbf{u} \in \mathcal{W}(X)^N$ – which retain voter identities – map onto the same \mathbf{a} -profile, and any two profiles \mathbf{u}_j and \mathbf{u}_k that map into the same \mathbf{a} -profile bear the same simple majority relation on the set of alternatives (although not the same equilibrium in a normal form game). Consequently, some \mathbf{a} -profiles correspond to very few \mathbf{u} -profiles and therefore each $\mathbf{u} \in \mathcal{W}(X)^N$ may be weighted differently. An \mathbf{a} -profile (k_1, k_2, \dots, k_m) therefore corresponds to

$$\frac{n!}{k_1!k_2!\dots k_m!} \quad (\text{A2.10})$$

distinct \mathbf{u} -profiles. For three players and three alternatives, each \mathbf{a} -profile corresponds to either 1, 3, or 6 \mathbf{u} -profiles. That is, all three players have the same preference order (1 permutation), two players have the same preference order (3 permutations), all three have different orders (6 permutations).

The cardinality of \mathcal{A} is given by (Lepelley and Martin 2001):

$$|\mathcal{A}| = \frac{(n+1)(n+2)\dots(n+(|\mathcal{W}(X)|-1))}{(|\mathcal{W}(X)|-1)!} \quad (\text{A2.11})$$

Taking $p(\mathbf{a})$ as the probability of an \mathbf{a} -profile defined by a distribution $p: \mathcal{A} \rightarrow \mathbb{R}$ satisfying the usual conditions of $0 \leq p(\mathbf{a}) \leq 1$ and $\sum_{\mathbf{a} \in \mathcal{A}} p(\mathbf{a}) = 1$ and by definition of IAC,

$$p(\mathbf{a}_1) = p(\mathbf{a}_2) = \dots = p(\mathbf{a}_n) > 0 \quad (\text{A2.12})$$

then,

$$p(\mathbf{a}) = \frac{1}{|\mathcal{A}|} \quad (\text{A2.13})$$

Now, because unlike with the classical paradox of voting (Condorcet paradox), the set \mathcal{A} cannot be partitioned into exclusive subsets in which chairman's paradox occurs and the other in which it does not, application of the IAC model means that we have to weight each game $g(\mathbf{u})$ by the probability of an \mathbf{a} -profile given by (A2.13) and the probability of one of its instantiations, \mathbf{u} . The total number of \mathbf{u} -profiles that instantiate a particular \mathbf{a} -profile is given by (A2.10). Let \mathbf{u}_a be a \mathbf{u} -profile that corresponds to a particular \mathbf{a} -profile, then, $p(\mathbf{u}) = p(\mathbf{a}) \cdot p(\mathbf{u}_a)$. Assuming that it is equally likely that each player has a preference order k_i that constitutes that \mathbf{a} -profile ('impartiality') and denoting the set of \mathbf{u}_a -profiles by \mathcal{U}_a , we have:

$$p(\mathbf{u}) = \frac{1}{|\mathcal{A}|} \cdot \frac{1}{|\mathcal{U}_a|} \quad (\text{A2.14})$$

Plugging (A2.14) into (A2.4), we get:

$$\Phi_i(\mathfrak{s}) = \left(\frac{1}{|\mathcal{A}|} \cdot \frac{1}{|\mathcal{U}_a|} \right) \sum_{\mathbf{u} \in \mathcal{U}(X)^N} \sigma_i(\mathfrak{s}, \mathbf{u}) \quad \forall i \in N \quad (\text{A2.15})$$

as the measure of player i 's overall success under the IAC model.

Note that in contrast to the IC model the IAC model, despite its 'anonymity', implies a degree of dependence between the players. It is easy to see from (A2.15) that the probability that players will share the same preference order u is higher than if they have different orders.

For three players, three alternatives:

	linear orders $ \mathcal{A} = 56$	weak orders $ \mathcal{A} = 455$
3 players have the same preference order u :	$p(\mathbf{u}) = 0.01785$	$p(\mathbf{u}) = 0.00219$
2 players have the same preference order u :	$p(\mathbf{u}) = 0.00595$	$p(\mathbf{u}) = 0.00073$
all players have different preference orders u :	$p(\mathbf{u}) = 0.00298$	$p(\mathbf{u}) = 0.00037$

Appendix 3

This appendix describes the method of reducing the 2197 chairman games for three-players and three alternatives with weak preference orders.

First, we reduce the set of games to the logically generic orders that are repeated systematically throughout the entire set of profiles. To illustrate, consider the two sets of cyclical preferences orders that generate sophisticated outcomes in which the chairman ends up with his least preferred alternative, i.e. $\{x y z, y z x, z x y\}, \{x z y, y x z, z y x\}$. In each set there are six permutations of the rankings among three players giving us a total of twelve preference profiles that generate specific instances of the chairman paradox. For these 12 profiles, there are 4 profiles in which each of the triple x, y, z will be in the choice set for each player, $\mathcal{C}_i(X, \succ_i)$. Therefore, to simplify the task at hand we need pick only, say, those profiles in which $x \in \mathcal{C}_i(X, \succ_i)$ for player 1, i.e.

- (i) $x y z$; (ii) $x z y$; (iii) $x(y z)$; (iv) $(x y)z$ or $(x z)y$; (v) $(x y z)$.

We can further reduce the set by restricting our attention to only one of (i) or (ii) or (iii), plus (iv) and (v) because we know that the chairman always has a straightforward strategy (see section 2), and therefore what holds for games in which the chairman has a profile $x y z$ holds, ceteris paribus, for a game in which he has $x z y$ or $x(y z)$, i.e. the games on $\mathbf{u} = (x y z, y z x, z x y)$, $\mathbf{u}' = (x z y, y z x, z x y)$, and $\mathbf{u}'' = (x(y z), y z x, z x y)$ are strategically equivalent. For each of these generic preference orders there are $13^2 = 169$ preference profiles (13 preference orders for each of the other two players), which means that we have reduced the number of games to $3 \cdot 169 = 507$.

Second, by dint of (i) the symmetric voting rights of the two non-chairman and (ii) the *symmetry invariance* property of a Nash equilibrium (Harsanyi and Selten 1988) we can reduce the remaining 507 games by just under a half. To do this, note that the symmetry invariance property says that a Nash equilibrium is independent of strategically irrelevant features such as names and numbers used to distinguish players, agents, and choices. Hence, once we have determined the equilibrium for the profile \mathbf{u} in Example 2.1 we also have determined the equilibrium for \mathbf{u}' in which we have interchanged the names of players 2 and 3. Thus, for each generic order we need only determine the equilibria for *distinct* profiles. Applying (A2.10) gives us 91 such profiles for each of the generic orders, bringing us to $3 \cdot 91 = 273$ chairman games. These profiles can easily be picked from the space of \mathbf{u} profiles by constructing a table such as Table A3.1.

Finally, we obtain the precise result by multiplying up as follows: we take the result for one of the generic rankings (i)–(iii) above and multiply this by 9. We then multiply the result for generic ranking (iv) by 3. Finally we add these values to the results of generic ranking (v), i.e.

Table A3.1

1. $x y z$ $x y z$ $x y z$	1. $x(y z)$ $x y z$ $x y z$	1. $(x y z)$
2. $x y z$ $x y z$ $x z y$	2. $x(y z)$ $x y z$ $x z y$. .
3. $x y z$ $x y z$ $y z x$	3. $x(y z)$ $x y z$ $y z x$. .
. $x y z$ $x y z$.	. $x(y z)$ $x y z$.	. .
. $x y z$ $x y z$.	. $x(y z)$ $x y z$.	. .
. $x y z$ $x y z$.	. $x(y z)$ $x y z$.	. .
13. $x y z$ $x y z$ $(x y z)$	13. $x(y z)$ $x y z$ $(x y z)$	13. .
1. $x y z$ $x z y$ $x y z$	1. $x(y z)$ $x z y$ $x y z$	1. $(x y z)$
2. $x y z$ $x z y$ $x z y$	2. $x(y z)$ $x z y$ $x z y$	2. .
3. $x y z$ $x z y$ $y z x$	3. $x(y z)$ $x z y$ $y z x$	3. .
. $x y z$ $x z y$.	. $x(y z)$ $x z y$.	. .
. $x y z$ $x z y$.	. $x(y z)$ $x z y$.	. .
. $x y z$ $x z y$.	. $x(y z)$ $x z y$.	. .
13. $x y z$ $x z y$ $(x y z)$	13. $x(y z)$ $x z y$ $(x y z)$	13. . etc...
etc...		

9 · (i) [or (ii) or (iii)] + 3 · (iv) + (v) .

A list of the ultimately coalition proof Nash equilibria for the 273 games is given in Table A3.2.

Table A3.2: Generic Games

	Preference Profile			Equilibria		Preference Profile			Equilibria
	1	2	3			1	2	3	
1.	xyz	xyz	xyz	(x,x,x)	47.	xyz	zxy	zxy	(x,z,z)
2.	xyz	xyz	xzy	(x,x,x), (x,x,z), (x,y,x), (x,y,z)	48.	xyz	zxy	zyx	(x,z,z)
3.	xyz	xyz	yzx	(x,x,y), (x,x,z)	49.	xyz	zxy	x(yz)	(x,x,x), (x,z,x)
4.	xyz	xyz	yxz	(x,x,x), (x,x,y)	50.	xyz	zxy	y(zx)	(x,x,y), (x,z,y)
5.	xyz	xyz	zxy	(x,x,x), (x,x,z), (x,y,x), (x,y,z)	51.	xyz	zxy	z(xy)	(x,z,z)
6.	xyz	xyz	zyx	(x,x,y), (x,x,z)	52.	xyz	zxy	(xy)z	(x,x,x), (x,x,y), (x,z,x), (x,z,y)
7.	xyz	xyz	x(yz)	(x,x,x), (x,y,x)	53.	xyz	zxy	(xz)y	(x,z,x), (x,z,z)
8.	xyz	xyz	y(zx)	(x,x,y)	54.	xyz	zxy	(yz)x	(x,z,z)
9.	xyz	xyz	z(xy)	(x,x,z)	55.	xyz	zxy	(xyz)	(x,z,x), (x,z,y), (x,z,z)
10.	xyz	xyz	(xy)z	(x,x,x), (x,x,y)	56.	xyz	zyx	zyx	(x,z,z)
11.	xyz	xyz	(xz)y	(x,x,x), (x,x,z), (x,y,x), (x,y,z)	57.	xyz	zyx	x(yz)	(x,x,y), (x,x,z)
12.	xyz	xyz	(yz)x	(x,x,y), (x,x,z)	58.	xyz	zyx	y(zx)	(x,y,y)
13.	xyz	xyz	(xyz)	(x,x,x), (x,x,y), (x,x,z)	59.	xyz	zyx	z(xy)	(x,z,z)
14.	xyz	xzy	xzy	(x,x,x)	60.	xyz	zyx	(xy)z	(x,y,x), (x,y,y)
15.	xyz	xzy	yzx	(x,x,y), (x,x,z)	61.	xyz	zyx	(xz)y	(x,z,x), (x,z,z)
16.	xyz	xzy	yxz	(x,x,x), (x,x,y), (x,z,x), (x,z,y)	62.	xyz	zyx	(yz)x	(x,y,y), (x,z,z)
17.	xyz	xzy	zxy	(x,x,z)	63.	xyz	zyx	(xyz)	(x,y,x), (x,y,y), (x,z,y), (x,z,z)
18.	xyz	xzy	zyx	(x,x,z)	64.	xyz	x(yz)	x(yz)	(x,x,x)
19.	xyz	xzy	x(yz)	(x,x,x), (x,z,x)	65.	xyz	x(yz)	y(zx)	(x,x,y)
20.	xyz	xzy	y(zx)	(x,x,y), (x,z,y)	66.	xyz	x(yz)	z(xy)	(x,x,z)
21.	xyz	xzy	z(xy)	(x,x,z)	67.	xyz	x(yz)	(xy)z	(x,x,x), (x,x,y)
22.	xyz	xzy	(xy)z	(x,x,x), (x,x,y), (x,z,x), (x,z,y)	68.	xyz	x(yz)	(xz)y	(x,x,x), (x,x,z)
23.	xyz	xzy	(xz)y	(x,x,x), (x,x,z)	69.	xyz	x(yz)	(yz)x	(x,x,y), (x,x,z)
24.	xyz	xzy	(yz)x	(x,x,y), (x,x,z)	70.	xyz	x(yz)	(xyz)	(x,x,x), (x,x,y), (x,x,z)
25.	xyz	xzy	(xyz)	(x,x,x), (x,x,y), (x,x,z)	71.	xyz	y(zx)	y(zx)	(x,y,y)
26.	xyz	yzx	yzx	(x,y,y)	72.	xyz	y(zx)	z(xy)	(x,y,z)
27.	xyz	yzx	yxz	(x,y,y)	73.	xyz	y(zx)	(xy)z	(x,y,x), (x,y,y)
28.	xyz	yzx	zxy	(x,z,z)	74.	xyz	y(zx)	(xz)y	(x,y,x), (x,y,z)
29.	xyz	yzx	zyx	(x,y,y), (x,z,z)	75.	xyz	y(zx)	(yz)x	(x,y,y)
30.	xyz	yzx	x(yz)	(x,y,x), (x,z,x)	76.	xyz	y(zx)	(xyz)	(x,y,x), (x,y,y), (x,y,z)
31.	xyz	yzx	y(zx)	(x,y,y)	77.	xyz	z(xy)	z(xy)	(x,z,z)
32.	xyz	yzx	z(xy)	(x,z,z)	78.	xyz	z(xy)	(xy)z	(x,z,x), (x,z,y)
33.	xyz	yzx	(xy)z	(x,y,x), (x,y,y)	79.	xyz	z(xy)	(xz)y	(x,z,x), (x,z,z)
34.	xyz	yzx	(xz)y	(x,z,x), (x,z,z)	80.	xyz	z(xy)	(yz)x	(x,z,z)
35.	xyz	yzx	(yz)x	(x,y,y), (x,z,z)	81.	xyz	z(xy)	(xyz)	(x,z,x), (x,z,y), (x,z,z)
36.	xyz	yzx	(xyz)	(x,y,x), (x,y,y), (x,z,x), (x,z,z)	82.	xyz	(xy)z	(xy)z	(x,x,x), (x,x,y), (x,y,x), (x,y,y)
37.	xyz	yxz	yxz	(x,y,y)	83.	xyz	(xy)z	(xz)y	(x,x,x), (x,y,x), (x,x,z), (x,y,z)
38.	xyz	yxz	zxy	(x,x,x), (x,x,z), (x,z,x), (x,z,z)	84.	xyz	(xy)z	(yz)x	(x,x,y), (x,y,y)
39.	xyz	yxz	zyx	(x,y,y)	85.	xyz	(xy)z	(xyz)	(x,x,x), (x,x,y), (x,x,z), (x,y,x), (x,y,y), (x,y,z)
40.	xyz	yxz	x(yz)	(x,x,x), (x,y,x)	86.	xyz	(xz)y	(xz)y	(x,x,x), (x,x,z), (x,z,x), (x,z,z)
41.	xyz	yxz	y(zx)	(x,y,y)	87.	xyz	(xz)y	(yz)x	(x,x,z), (x,z,z)
42.	xyz	yxz	z(xy)	(x,x,z), (x,y,z)	88.	xyz	(xz)y	(xyz)	(x,x,x), (x,x,y), (x,x,z), (x,z,x), (x,z,y), (x,z,z)
43.	xyz	yxz	(xy)z	(x,y,x), (x,y,y)	89.	xyz	(yz)x	(yz)x	(x,y,y), (x,z,z)
44.	xyz	yxz	(xz)y	(x,x,x), (x,x,z), (x,y,x), (x,y,z), (x,y,z)	90.	xyz	(yz)x	(xyz)	(x,y,x), (x,y,y), (x,z,x), (x,z,z)
45.	xyz	yxz	(yz)x	(x,y,y)	91.	xyz	(xyz)	(xyz)	(x,x,x), (x,x,y), (x,x,z), (x,y,x), (x,y,y)
46.	xyz	yxz	(xyz)	(x,y,x), (x,y,y), (x,y,z)					

Preference Profile			Equilibria	Preference Profile			Equilibria		
1	2	3		1	2	3			
			$(x,y,z), (x,z,x), (x,z,y), (x,z,z)$	143.	$(xy)z$	zxy	$(xy)z$	$(x,x,x), (x,x,y), (y,x,x), (y,x,y)$	
92.	$(xy)z$	xyz	xyz	$(x,x,x), (y,x,x)$	144.	$(xy)z$	zxy	$(xz)y$	$(x,x,x), (x,z,z), (y,x,x), (y,z,z)$
93.	$(xy)z$	xyz	xzy	$(x,x,x), (y,x,x)$	145.	$(xy)z$	zxy	$(yz)x$	$(x,z,z), (y,z,z)$
94.	$(xy)z$	xyz	yzx	$(x,x,y), (y,x,y)$	146.	$(xy)z$	zxy	(xyz)	$(x,x,x), (x,x,y), (x,z,x), (x,z,y), (x,z,z), (y,x,x), (y,x,y), (y,z,y), (y,z,z)$
95.	$(xy)z$	xyz	yxz	$(x,x,y), (y,x,y)$	147.	$(xy)z$	zyx	zyx	$(x,z,z), (y,z,z)$
96.	$(xy)z$	xyz	zxy	$(x,x,x), (y,x,x)$	148.	$(xy)z$	zyx	$x(yz)$	$(x,y,x), (y,y,x), (x,z,x), (y,z,x)$
97.	$(xy)z$	xyz	zyx	$(x,x,z), (y,x,z)$	149.	$(xy)z$	zyx	$y(zx)$	$(x,y,y), (y,y,y)$
98.	$(xy)z$	xyz	$x(yz)$	$(x,x,x), (y,x,x)$	150.	$(xy)z$	zyx	$z(xy)$	$(x,z,z), (y,z,z)$
99.	$(xy)z$	xyz	$y(zx)$	$(x,x,y), (y,x,y)$	151.	$(xy)z$	zyx	$(xy)z$	$(x,y,x), (x,y,y), (y,y,x), (y,y,y)$
100.	$(xy)z$	xyz	$z(xy)$	$(x,y,z), (y,y,z), (x,x,z), (y,x,z)$	152.	$(xy)z$	zyx	$(xz)y$	$(x,z,z), (y,z,z)$
101.	$(xy)z$	xyz	$(xy)z$	$(x,x,x), (x,x,y), (y,x,x), (y,x,y)$	153.	$(xy)z$	zyx	$(yz)x$	$(x,y,y), (x,z,z), (y,y,y), (y,z,z)$
102.	$(xy)z$	xyz	$(xz)y$	$(x,x,x), (y,x,x)$	154.	$(xy)z$	zyx	(xyz)	$(x,y,x), (x,y,y), (x,z,x), (y,z,z), (y,y,x), (y,y,y), (y,z,x), (y,z,y), (y,z,z)$
103.	$(xy)z$	xyz	$(yz)x$	$(x,x,y), (x,x,z), (y,x,y), (y,x,z)$	155.	$(xy)z$	$x(yz)$	$x(yz)$	$(x,x,x), (y,x,x)$
104.	$(xy)z$	xyz	(xyz)	$(x,x,x), (x,x,y), (x,x,z), (y,x,x), (y,x,y), (y,x,z)$	156.	$(xy)z$	$x(yz)$	$y(zx)$	$(x,x,y), (y,x,y)$
105.	$(xy)z$	xzy	xzy	$(x,x,x), (x,x,z), (x,z,x), (y,x,x)$	157.	$(xy)z$	$x(yz)$	$z(xy)$	$(x,x,z), (y,x,z)$
106.	$(xy)z$	xzy	yzx	$(x,x,y), (y,x,y), (y,z,y)$	158.	$(xy)z$	$x(yz)$	$(xy)z$	$(x,x,x), (x,x,y), (y,x,x), (y,x,y)$
107.	$(xy)z$	xzy	yxz	$(x,x,y), (y,x,y)$	159.	$(xy)z$	$x(yz)$	$(xz)y$	$(x,x,x), (y,x,x)$
108.	$(xy)z$	xzy	zxy	$(x,x,x), (y,x,x)$	160.	$(xy)z$	$x(yz)$	$(yz)x$	$(x,x,y), (x,x,z), (y,x,y), (y,x,z)$
109.	$(xy)z$	xzy	zyx	(x,x,z)	161.	$(xy)z$	$x(yz)$	(xyz)	$(x,x,x), (x,x,y), (x,x,z), (y,x,x), (y,x,y), (y,x,z)$
110.	$(xy)z$	xzy	$x(yz)$	$(x,x,x), (y,x,x)$	162.	$(xy)z$	$y(zx)$	$y(zx)$	$(x,y,y), (y,y,y)$
111.	$(xy)z$	xzy	$y(zx)$	$(x,z,y), (y,z,y), (x,x,y), (y,x,y)$	163.	$(xy)z$	$y(zx)$	$z(xy)$	$(x,y,z), (y,y,z)$
112.	$(xy)z$	xzy	$z(xy)$	(x,x,z)	164.	$(xy)z$	$y(zx)$	$(xy)z$	$(x,y,x), (x,y,y), (y,y,x), (y,y,y)$
113.	$(xy)z$	xzy	$(xy)z$	$(x,x,x), (x,x,y), (y,x,x), (y,x,y)$	165.	$(xy)z$	$y(zx)$	$(xz)y$	$(x,x,x), (x,x,z), (y,y,x), (y,y,z)$
114.	$(xy)z$	xzy	$(xz)y$	$(x,x,z), (x,z,x), (x,x,x), (y,x,x)$	166.	$(xy)z$	$y(zx)$	$(yz)x$	$(x,y,y), (y,y,y)$
115.	$(xy)z$	xzy	$(yz)x$	$(x,x,y), (x,x,z), (y,x,y)$	167.	$(xy)z$	$y(zx)$	(xyz)	$(x,y,x), (x,y,y), (x,y,z), (y,y,x), (y,y,y), (y,y,z)$
116.	$(xy)z$	xzy	(xyz)	$(x,z,x), (x,z,y), (x,x,x), (x,x,y), (x,x,z), (y,x,x), (y,x,y)$	168.	$(xy)z$	$z(xy)$	$z(xy)$	$(x,z,z), (y,z,z)$
117.	$(xy)z$	yzx	yzx	$(y,y,z), (y,z,y), (x,y,y), (y,y,y)$	169.	$(xy)z$	$z(xy)$	$(xy)z$	$(x,z,x), (x,z,y), (y,z,x), (y,z,y)$
118.	$(xy)z$	yzx	yxz	$(x,y,y), (y,y,y)$	170.	$(xy)z$	$z(xy)$	$(xz)y$	$(x,z,z), (y,z,z)$
119.	$(xy)z$	yzx	zxy	(y,y,z)	171.	$(xy)z$	$z(xy)$	$(yz)x$	$(x,z,z), (y,z,z)$
120.	$(xy)z$	yzx	zyx	$(x,y,y), (y,y,y), (y,y,z)$	172.	$(xy)z$	$z(xy)$	(xyz)	$(x,z,x), (x,z,y), (x,z,z), (y,z,x), (y,z,y), (y,z,z)$
121.	$(xy)z$	yzx	$x(yz)$	$(x,z,x), (y,z,x), (x,y,x), (y,y,x)$	173.	$(xy)z$	$(xy)z$	$(xy)z$	$(x,x,x), (x,x,y), (x,y,x), (x,y,y), (y,x,x), (y,x,y), (y,y,x), (y,y,y)$
122.	$(xy)z$	yzx	$y(zx)$	$(x,y,y), (y,y,y)$	174.	$(xy)z$	$(xy)z$	$(xz)y$	$(x,x,x), (x,y,x), (y,x,x), (y,y,x)$
123.	$(xy)z$	yzx	$z(xy)$	(y,y,z)	175.	$(xy)z$	$(xy)z$	$(yz)x$	$(x,x,y), (x,y,y), (y,x,y), (y,y,y)$
124.	$(xy)z$	yzx	$(xy)z$	$(x,y,x), (x,y,y), (y,y,x), (y,y,y)$	176.	$(xy)z$	$(xy)z$	(xyz)	$(x,x,x), (x,x,y), (x,x,z), (x,y,x), (x,y,y), (x,y,z), (y,x,x), (y,x,y), (y,x,z), (y,y,x), (y,y,y), (y,y,z)$
125.	$(xy)z$	yzx	$(xz)y$	(y,y,z)	177.	$(xy)z$	$(xz)y$	$(xz)y$	$(x,x,x), (x,x,z), (x,z,x), (x,z,z), (y,x,x), (y,z,z)$
126.	$(xy)z$	yzx	$(yz)x$	$(y,y,z), (y,z,y), (x,y,y), (y,y,y)$	178.	$(xy)z$	$(xz)y$	$(yz)x$	$(x,z,z), (y,z,z)$
127.	$(xy)z$	yzx	(xyz)	$(x,y,x), (x,y,y), (x,z,x), (y,y,x), (y,y,y), (y,y,z), (y,z,x), (y,z,y)$	179.	$(xy)z$	$(xz)y$	(xyz)	$(x,x,x), (x,x,y), (x,x,z), (x,z,x), (x,z,y), (x,z,z), (y,x,x), (y,x,y), (y,z,y), (y,z,z)$
128.	$(xy)z$	yxz	yxz	$(x,y,y), (y,y,y)$	180.	$(xy)z$	$(yz)x$	$(yz)x$	$(x,y,y), (x,z,z), (y,y,y), (y,y,z), (y,z,y), (y,z,z)$
129.	$(xy)z$	yxz	zxy	$(x,y,x), (y,y,x)$	181.	$(xy)z$	$(yz)x$	(xyz)	$(x,y,x), (x,y,y), (x,z,x), (x,z,z), (y,y,x), (y,y,y), (y,y,z), (y,z,x), (y,z,y), (y,z,z)$
130.	$(xy)z$	yxz	zyx	$(x,y,y), (y,y,y)$	182.	$(xy)z$	(xyz)	(xyz)	$(x,x,x), (x,x,y), (x,x,z), (x,y,x), (x,y,y), (x,y,z), (x,z,x), (x,z,y), (x,z,z), (y,x,x), (y,x,y), (y,x,z), (y,y,x), (y,y,y), (y,y,z), (y,z,x), (y,z,y), (y,z,z)$
131.	$(xy)z$	yxz	$x(yz)$	$(x,y,x), (y,y,x)$	183.	(xyz)	xyz	xyz	$(x,x,x), (y,x,x), (z,x,x)$
132.	$(xy)z$	yxz	$y(zx)$	$(x,y,y), (y,y,y)$	184.	(xyz)	xyz	xzy	$(x,x,x), (y,x,x), (z,x,x)$
133.	$(xy)z$	yxz	$z(xy)$	$(x,x,z), (y,x,z), (x,y,z), (y,y,z)$	185.	(xyz)	xyz	yzx	$(x,x,y), (y,x,y), (y,y,y), (z,y,y)$
134.	$(xy)z$	yxz	$(xy)z$	$(x,y,x), (x,y,y), (y,y,x), (y,y,y)$					
135.	$(xy)z$	yxz	$(xz)y$	$(x,y,x), (x,y,z), (y,y,x), (y,y,z)$					
136.	$(xy)z$	yxz	$(yz)x$	$(x,y,y), (y,y,y)$					
137.	$(xy)z$	yxz	(xyz)	$(x,y,x), (x,y,y), (x,y,z), (y,y,x), (y,y,y), (y,y,z)$					
138.	$(xy)z$	zxy	zyx	$(x,z,z), (y,z,z)$					
139.	$(xy)z$	zxy	$x(yz)$	$(x,x,x), (y,x,x)$					
140.	$(xy)z$	zxy	zxy	$(x,z,z), (y,z,z)$					
141.	$(xy)z$	zxy	$y(zx)$	$(x,x,y), (y,x,y), (x,z,y), (y,z,y)$					
142.	$(xy)z$	zxy	$z(xy)$	$(x,z,z), (y,z,z)$					

Preference Profile			Equilibria	Preference Profile			Equilibria		
1	2	3		1	2	3			
186.	(xyz)	xyz	yxz	(x,x,x), (x,x,z), (y,x,y), (z,x,x)	227.	(xyz)	yxz	(yz)x	(x,y,y), (y,x,y), (y,y,y), (z,y,y)
187.	(xyz)	xyz	zxy	(x,x,x), (x,x,z), (y,x,x), (z,x,z)	228.	(xyz)	yxz	(xyz)	(x,x,x), (x,x,z), (x,y,x), (x,y,y), (x,y,z), (y,x,y), (y,x,z), (y,y,x), (y,y,y), (y,y,z), (z,x,x), (z,x,z), (z,y,y), (z,y,z)
188.	(xyz)	xyz	zyx	(x,x,y), (x,x,z), (y,x,y), (y,x,z), (y,y,y), (y,y,z), (z,x,z), (z,y,z)	229.	(xyz)	zxy	zxy	(x,z,z), (y,z,z), (z,z,z)
189.	(xyz)	xyz	x(yz)	(x,x,x), (y,x,x), (z,x,x)	230.	(xyz)	zxy	zyx	(x,z,z), (y,z,z), (z,z,z)
190.	(xyz)	xyz	y(zx)	(x,x,y), (y,x,y), (y,y,y), (z,y,y)	231.	(xyz)	zxy	x(yz)	(x,x,x), (x,z,x), (y,x,x), (z,z,x)
191.	(xyz)	xyz	z(xy)	(x,y,z), (y,y,z), (z,y,z), (x,x,z), (y,x,z), (z,x,z)	232.	(xyz)	zxy	y(zx)	(x,x,y), (y,x,y), (z,x,y), (x,z,y), (y,z,y), (z,z,y)
192.	(xyz)	xyz	(xy)z	(x,x,x), (x,x,y), (x,y,x), (y,x,x), (y,x,y), (y,y,y), (z,x,x), (z,y,y)	233.	(xyz)	zxy	z(xy)	(x,z,z), (y,z,z), (z,z,z)
193.	(xyz)	xyz	(xz)y	(x,x,x), (y,x,x), (z,x,x)	234.	(xyz)	zxy	(xy)z	(x,x,x), (x,z,x), (y,x,x), (z,z,x)
194.	(xyz)	xyz	(yz)x	(x,x,y), (y,x,y), (y,y,y), (z,y,y)	235.	(xyz)	zxy	(xz)y	(x,x,x), (x,z,x), (x,z,z), (y,x,x), (y,z,z), (z,x,z), (z,z,x), (z,z,z)
195.	(xyz)	xyz	(xyz)	(x,x,x), (x,x,y), (x,x,z), (x,y,x), (x,y,z), (y,x,x), (y,x,y), (y,x,z), (y,y,y), (y,y,z), (z,x,x), (z,x,z), (z,y,y), (z,y,z)	236.	(xyz)	zxy	(yz)x	(x,z,z), (y,z,z), (z,z,z)
196.	(xyz)	xzy	xzy	(x,x,x), (y,x,x), (z,x,x)	237.	(xyz)	zxy	(xyz)	(x,x,x), (x,x,y), (x,z,x), (x,z,y), (x,z,z), (y,x,x), (y,x,y), (y,z,y), (y,z,z), (z,x,y), (z,x,z), (z,z,x), (z,z,y), (z,z,z)
197.	(xyz)	xzy	yzx	(x,x,y), (x,x,z), (y,x,y), (z,x,y), (z,x,z), (z,z,y), (z,z,z)	238.	(xyz)	zyx	zyx	(x,z,z), (y,z,z), (z,z,z)
198.	(xyz)	xzy	yxz	(x,x,x), (x,x,y), (y,x,y), (z,x,x)	239.	(xyz)	zyx	x(yz)	(x,y,x), (y,y,x), (z,y,x), (x,z,x), (y,z,x), (z,z,x)
199.	(xyz)	xzy	zxy	(x,x,x), (y,x,x), (y,z,z), (z,z,z)	240.	(xyz)	zyx	y(zx)	(x,y,y), (y,y,y), (y,z,y), (z,z,y)
200.	(xyz)	xzy	zyx	(x,x,z), (y,z,z), (z,x,z), (z,z,z)	241.	(xyz)	zyx	z(xy)	(x,z,z), (y,z,z), (z,z,z)
201.	(xyz)	xzy	x(yz)	(x,x,x), (y,x,x), (z,x,x)	242.	(xyz)	zyx	(xy)z	(x,y,y), (y,y,y), (y,z,y), (z,z,y)
202.	(xyz)	xzy	y(zx)	x,z,y), (y,z,y), (z,z,y), (x,x,y), (y,x,y), (z,x,y)	243.	(xyz)	zyx	(xz)y	(x,z,z), (y,z,z), (z,z,z)
203.	(xyz)	xzy	z(xy)	(x,x,z), (y,z,z), (z,x,z), (z,z,z)	244.	(xyz)	zyx	(yz)x	(x,y,y), (x,z,z), (y,y,y), (y,z,y), (y,z,z), (z,y,z), (z,y,z), (z,z,y), (z,z,z)
204.	(xyz)	xzy	(xy)z	(x,x,x), (y,x,x), (z,x,x)	245.	(xyz)	zyx	(xyz)	(x,y,x), (x,y,y), (x,z,x), (x,z,z), (y,y,x), (y,y,y), (y,z,x), (y,z,y), (y,z,z), (z,y,x), (z,y,z), (z,z,x), (z,z,y), (z,z,z)
205.	(xyz)	xzy	(xz)y	(y,x,x), (z,x,z), (z,z,z)	246.	(xyz)	x(yz)	x(yz)	(x,x,x), (y,x,x), (z,x,x)
206.	(xyz)	xzy	(yz)x	(x,x,z), (y,z,z), (z,x,z), (z,z,z)	247.	(xyz)	x(yz)	y(zx)	(x,x,y), (y,x,y), (z,x,y)
207.	(xyz)	xzy	(xyz)	(x,x,x), (x,x,y), (x,x,z), (x,z,x), (x,z,y), (y,x,x), (y,x,y), (y,z,y), (y,z,z), (z,x,x), (z,x,y), (z,x,z), (z,z,y), (z,z,z)	248.	(xyz)	x(yz)	z(xy)	(x,x,z), (y,x,z), (z,x,z)
208.	(xyz)	yzx	yzx	(x,y,y), (y,y,y), (z,y,y)	249.	(xyz)	x(yz)	(xy)z	(x,x,x), (y,x,x), (z,x,x)
209.	(xyz)	yzx	yxz	(x,y,y), (y,y,y), (z,y,y)	250.	(xyz)	x(yz)	(xz)y	(x,x,x), (y,x,x), (z,x,x)
210.	(xyz)	yzx	zxy	(x,z,z), (y,y,z), (z,y,z), (z,z,z)	251.	(xyz)	x(yz)	(yz)x	(x,x,y), (x,x,z), (y,x,y), (y,x,z), (z,x,y), (z,x,z)
211.	(xyz)	yzx	zyx	(x,y,y), (x,z,z), (y,y,y), (z,z,z)	252.	(xyz)	x(yz)	(xyz)	(x,x,x), (x,x,y), (x,x,z), (y,x,x), (y,x,y), (y,x,z), (z,x,x), (z,x,y), (z,x,z)
212.	(xyz)	yzx	x(yz)	(x,z,x), (y,z,x), (z,z,x), (x,y,x), (y,y,x), (z,y,x)	253.	(xyz)	y(zx)	y(zx)	(x,y,y), (y,y,y), (z,y,y)
213.	(xyz)	yzx	y(zx)	(x,y,y), (y,y,y), (z,y,y)	254.	(xyz)	y(zx)	z(xy)	(x,y,z), (y,y,z), (z,y,z)
214.	(xyz)	yzx	z(xy)	(x,z,z), (y,y,z), (z,y,z), (z,z,z)	255.	(xyz)	y(zx)	(xy)z	(x,y,y), (y,y,y), (z,y,y)
215.	(xyz)	yzx	(xy)z	(x,y,y), (y,y,y), (z,y,y)	256.	(xyz)	y(zx)	(xz)y	(x,y,x), (x,y,z), (y,y,x), (y,y,z), (z,y,x), (z,y,z)
216.	(xyz)	yzx	(xz)y	(x,z,z), (y,y,z), (z,y,z), (z,z,z)	257.	(xyz)	y(zx)	(yz)x	(x,y,y), (y,y,y), (z,y,y)
217.	(xyz)	yzx	(yz)x	(x,y,y), (x,z,z), (y,y,y), (y,y,y), (y,z,z), (z,y,y), (z,y,z), (z,z,z)	258.	(xyz)	y(zx)	(xyz)	(x,y,x), (x,y,y), (x,y,z), (y,y,x), (y,y,y), (y,y,z), (z,y,x), (z,y,y), (z,y,z), (z,y,y), (z,y,z), (z,z,x), (z,z,z)
218.	(xyz)	yzx	(xyz)	(x,y,x), (x,y,y), (x,z,x), (x,z,z), (y,y,x), (y,y,y), (y,z,x), (y,z,y), (y,z,z), (z,y,x), (z,y,y), (z,y,z), (z,z,x), (z,z,z)	259.	(xyz)	z(xy)	z(xy)	(x,z,z), (y,z,z), (z,z,z)
219.	(xyz)	yxz	yxz	(x,y,y), (y,y,y), (z,y,y)	260.	(xyz)	z(xy)	(xy)z	(x,z,x), (x,z,y), (y,z,x), (y,z,y), (z,z,x), (z,z,y)
220.	(xyz)	yxz	zxy	(x,x,x), (x,x,z), (x,y,x), (x,y,z), (y,y,x), (y,y,z), (z,x,z), (z,y,z)	261.	(xyz)	z(xy)	(xz)y	(x,z,z), (y,z,z), (z,z,z)
221.	(xyz)	yxz	zyx	(x,y,y), (y,y,y), (y,y,z), (z,y,z)	262.	(xyz)	z(xy)	(yz)x	(x,z,z), (y,z,z), (z,z,z)
222.	(xyz)	yxz	x(yz)	(x,x,x), (x,y,x), (y,y,x), (z,x,x)	263.	(xyz)	z(xy)	(xyz)	(x,z,x), (x,z,y), (x,z,z), (y,z,x), (y,z,y), (y,z,z), (z,z,x), (z,z,y), (z,z,z)
223.	(xyz)	yxz	y(zx)	(x,y,y), (y,y,y), (z,y,y)	264.	(xyz)	(xy)z	(xy)z	(x,x,x), (x,x,y), (x,y,x), (x,y,y), (y,x,x), (y,x,y), (y,y,x), (y,y,y), (z,x,x), (z,x,y)
224.	(xyz)	yxz	z(xy)	(x,x,z), (y,x,z), (z,x,z), (x,y,z), (y,y,z), (z,y,z)	265.	(xyz)	(xy)z	(xz)y	(x,x,x), (y,x,x), (z,x,x)
225.	(xyz)	yxz	(xy)z	(x,x,x), (x,y,x), (x,y,y), (y,x,y), (y,y,x), (y,y,y), (z,x,x), (z,y,y)	266.	(xyz)	(xy)z	(yz)x	(x,y,y), (y,y,y), (z,y,y)
226.	(xyz)	yxz	(xz)y	(x,x,x), (x,y,x), (y,y,x), (z,x,x)	267.	(xyz)	(xy)z	(xyz)	(x,x,x), (x,x,y), (x,x,z), (x,y,x), (x,y,y), (x,y,z), (y,x,x), (y,x,y), (y,x,z), (y,y,x)

Preference Profile			Equilibria	Preference Profile			Equilibria		
1	2	3		1	2	3			
			(y,y,y), (y,y,z), (z,x,x), (z,x,z), (z,y,y), (z,y,z)				(y,y,y), (y,y,z), (y,z,x), (y,z,y), (y,z,z), (z,y,x), (z,y,y), (z,y,z), (z,z,x), (z,z,y), (z,z,z)		
268.	(xyz)	(xz)y	(xz)y	(x,x,x), (x,x,z), (x,z,x), (x,z,z), (y,x,x), (y,z,z), (z,x,x), (z,x,z), (z,z,x), (z,z,z)	273.	(xyz)	(xyz)	(xyz)	(x,x,x), (x,x,y), (x,x,z), (x,y,x), (x,y,y), (x,y,z), (x,z,x), (x,z,y), (x,z,z), (y,x,x), (y,x,y), (y,x,z), (y,y,x), (y,y,y), (y,y,z), (y,z,x), (y,z,y), (y,z,z), (z,x,x), (z,x,y), (z,x,z), (z,y,x), (z,y,y), (z,y,z), (z,z,x), (z,z,y), (z,z,z)
269.	(xyz)	(xz)y	(yz)x	(x,z,z), (y,z,z), (z,z,z)					
270.	(xyz)	(xz)y	(xyz)	(x,x,x), (x,x,y), (x,x,z), (x,z,x), (x,z,y), (x,z,z), (y,x,x), (y,x,y), (y,z,x), (y,z,z), (z,x,x), (z,x,z), (z,x,z), (z,z,x), (z,z,y), (z,z,z)					
271.	(xyz)	(yz)x	(yz)x	(x,y,y), (x,z,z), (y,y,y), (y,y,z), (y,z,y), (y,z,z), (z,y,y), (z,y,z), (z,z,y), (z,z,z)					
272.	(xyz)	(yz)x	(xyz)	(x,y,x), (x,y,y), (x,z,x), (x,z,z), (y,y,x),					

Voting Rules in Insolvency Law

1. Introduction

A chief characteristic of modern insolvency law in Canada, Germany, the UK, and the US is the provision for ‘workouts’ or ‘schemes of arrangement’ by which insolvent companies (or individuals) can attempt to rehabilitate the business (or their personal affairs) instead of going into liquidation. In the case of liquidation a court usually appoints a trustee to sell the firm’s assets, either piecemeal or as a going concern, to outside buyers. The proceeds from this sale are divided among those who have rights against the corporation, with the division made according to the legal priority of these rights.

In reorganization, however, there is no sale to third parties. Rather, there is a ‘hypothetical sale’ of the firm to existing ‘participants’. That is, those who hold claims and rights against the bankrupt entity exchange these for claims and rights against the ‘new’ corporation. For example, a bankrupt company may emerge from reorganization with all its debt cancelled and with the former debtholders holding some or all of the equity of the reorganized company.

If the court supervising a reorganization could observe the value of the reorganized firm, it would allocate the value among the participants according to the legal priority of their claims. But because a court cannot determine accurately an objective figure for this value, the law leaves the division of the reorganized company’s value to a process of bargaining among participants who are generally divided into classes of equal types and priority of claims.

Under the bankruptcy codes in the countries that we have mentioned, each class of equity holders and debtholders whose interests are impaired must vote to approve a reorganization plan, which would include a division of value of the ‘new’ or reorganized firm among claimants. What is interesting, however, is that although the voting rules for approving a reorganization plan are essentially concerned with the same substantive issue in each of the countries that we have mentioned, the rules differ in detail to a greater or lesser extent (see Table 1.1).¹ For

¹ One could argue that while the voting rules may be dealing with essentially the same

Table 1.1: Voting Rules

<i>Country</i>	<i>No. of classes</i>	<i>Classes</i>	<i>Requisite majority</i>
UK ^a	2	Unsecured creditors Shareholders (or owners)	> 75% of face-value of claim of those voting > 50% in value of those voting
Germany ^b	Min. 2	Secured creditors Unsecured creditors	> 50% of face-value of claim of those voting and > 50% in number of those voting As for secured creditors
Canada/US ^c	Min. 3	Secured creditors Unsecured creditors Shareholders (or owners)	≥ 66.66% of face-value of claim of those voting and > 50% in number of those voting As for secured creditors ≥ 66.66% of value of those voting

Source ^aHalsbury's Statutory Instruments (1991, pp. 257-8); ^bBalz and Landfermann (1999); ^c(Henry 1999, pp. 133-4).

example, in the UK there are only two classes of claimants: unsecured creditors and shareholders (owners or 'members'). The plan is approved if, of those voting, more than 75 percent in value of unsecured claims vote in favour and more than 50 percent in value of shareholders vote in favour.

At the other end of the spectrum, we have the Canadian and US rules which are legally and strategically the most complex of all four codes. Here we see a minimum of three classes – each defined by their legal entitlement to the firm's assets – shareholders (owners) and two classes of creditors (secured and unsecured) with ample room for the formation of sub-classes within the creditor class based upon economic interests (e.g. all secured creditors who have a claim on real

substantive issue it is mistaken to isolate them from other formal legal procedures of bankruptcy which may provide the reasons for the differences. The counter argument to this position is that by isolating the rules we can create a benchmark for comparing the formal properties of the rules which would then allow us to more easily determine the influence of other legal requirements.

estate, employees, etc.). The rules also introduce a 'double' requirement for the creditor classes. What this means is that a plan is approved if, of those voting, two-thirds in value of each creditor class (qualified majority) and a simple majority in number vote in favour (the latter condition is known in American legal parlance as a 'numerosity requirement'); and two thirds (qualified majority) in value of shareholders vote in favour (i.e. no numerosity requirement).

The German rules stand mid-way between the UK and Canadian/US rules. Here we also see the inclusion of a numerosity requirement as well as room for a more complex class structure than in the UK although less so than in the US. The plan is approved only by the creditors (shareholders have no voting rights), who are broken up into a minimum two classes, based on type of claim. For a plan to be approved, each class must achieve a simple majority in value of claims and persons voting. The formation of a more differentiated class structure that includes other interest groups such as employees is not excluded.

The question is, then, what are the principle structural differences in these rules? Although discussion of voting as a legal procedure in corporate reorganizations is a common in the literature, none of it really pays any attention to the rules *per se*, other than for being merely descriptive.² And for the small number of papers that contain some substantial analysis, the rules are either embedded in bargaining models (Baird and Picker 1991, Bebchuk and Chang 1992) or considered in a much wider context of how they work together with debtors exclusivity period, distribution (or liquidation) floors, and priority rules (Kordana and Posner 1999). To our knowledge there is no paper that studies the principal differences of the character of, for example, the German and US rules. That is, the literature is highly specific in that it concerns mainly the concrete reality of the rules and in particular those of Chapter 11 of the US Bankruptcy Code. It provides no general criteria for comparing the properties of one type of majority requirement or another, or the effect of the different number of classes.

We fill this lacunae by comparing the 'resistance' of the rules (the *a priori* probability of a plan being rejected) and how the rules distribute the value of the reorganized firm.

This essay differs from the prior literature in two ways:

First, we introduce the formal apparatus of the theory of simple voting games (SVG) that goes back to the monumental work of von Neumann and Morgenstern (1944), and hitherto has not found its way into the literature on voting rules in insolvency law. This approach strips the analysis down to the absolute bare essentials: the rules themselves; we abstract away from both preferences and the legal details such as priority rules, distribution floors, equal treatment, cram down,

² Most papers usually describe the voting rules. For a small sample of the literature see, for example, Aghion, Hart and Moore (1992), Franks, Nyborg and Torous (1996), LoPucki and Triantis (1994), Norberg (1998), and White (1989).

automatic stay, etc., as well as from the strategic aspects such as who proposes a reorganization plan or gerrymandering.

In technical terms, our analysis is based upon what Gibbard (1973) has defined as a *game form*. And for generality we ignore the matters such as whether or not reorganization is frequently used or whether, for example, German firms have few or many unsecured creditors. An SVG concerns the *a priori* structure of the rules which does not include any of this information. While this abstraction can be seen as a point of weakness, it has the advantage that it allows us to clarify the structure of the rules, in absence of which ‘richer’ analyses are prone to error. Put another way, we describe and characterise the structural environment into which agents enter and which then influences the outcomes by creating constraints on choices because the SVG places a restriction on which collection of the bankrupt firm’s ‘participants’ win the game.

Second, we concentrate exclusively on two aspects of the rules: the majority quota and the classification of claimants. By way of example and formal proof we demonstrate the effects of the different quotas and degree of classification on (a) the probability that a reorganization plan will be rejected, and (b) the distribution of value.³

Finally a word of warning. In contrast to much of the law and economics literature on insolvency law we do not focus on the normative aspects of either the voting rules or our results: which rule minimizes the cost of credit, a criterion that is taken in much of the literature to be the desideratum of bankruptcy law. While this, too, can be seen as a shortcoming, it must be made clear that it is beyond the intention and scope of this present essay, which concerns only on a minimalist characterization of the voting rules and the methodological implications that flow from it.

The essay is organized as follows. In section 2 we introduce the formal structure of SVGs, providing the basic definitions that we will make use of in our analysis. Given that the theory behind SVGs is more or less foreign to the law and economics literature we provide additional remarks and methodological comments on the formal apparatus in order to ease the comprehension of our analysis and technique. In section 3 we compare and contrast the different voting rules in terms of their ‘resistance’ or the *a priori* probability that a reorganization plan will be rejected. In section 4 we consider the *a priori* distributional aspects of the different rules under a very mild theory of coalition formation. Section 5 contains some concluding remarks.

³ After the completion of an earlier version of this paper, we came across the paper by Kordana and Posner (1999) which coincidentally contains the question about the effect of classification. Kordana and Posner do not, however, provide any formal structure or proof of their claims (which are similar to ours). We thank Lucian Bebchuk for directing our attention to Kordana and Posner’s paper.

2. Simple Voting Games and the Structure of Voting Rules

In this section we state the basic definitions from the theory of *simple voting games* that we will require for our analysis. We refer the reader to Shapley (1962), Felsenthal and Machover (1998), and Taylor and Zwicker (1999) for additional background and results.

The most important definition that we require is that of a *decision rule* which we will first formulate informally as follows. Let a n -member decision-making body be denoted by a set $N = \{1, \dots, n\}$. A decision rule specifies which subsets of N can ensure the acceptance of a proposal. Formally:

Definition 2.1 (i) A *simple voting game* (SVG) is a pair (N, W) where W is a collection of subsets of a finite set N , satisfying the following three conditions: $\emptyset \notin W$; $N \in W$; and (monotonicity) if $S \in W$ and $S \subseteq T$ then $T \in W$. (ii) By N is meant an *assembly* (or voting body) and is the largest set in W , its members are *voters*, and its subsets are *coalitions*. A coalition S is said to be *winning* or *losing* according to whether $S \in W$ or $S \notin W$. (iii) A coalition S is called a *minimal winning coalition* (MWC) if $S \in W$, but no subset of S is in W . The set of MWCs is denoted by W^m .

Remark 2.2 (i) An SVG can be represented by W because N is uniquely determined by W (its largest member). We will therefore follow this notation throughout this essay. (ii) Monotonicity of W and the finiteness of N imply that W^m completely determines W .

For certain purposes – which we will come to in sections 3 and 4 but which will be made evident in Rem. 2.4 – it is convenient to represent an SVG using a characteristic function:

Definition 2.3 Let W be an SVG on an assembly N . The characteristic function of W is a mapping $v: 2^N \rightarrow \{0, 1\}$ such that for any coalition S , $v(S) = 1$ if $S \in W$ and 0 otherwise.

Remark 2.4 (i) Using the characteristic function an SVG W can be represented as a pair (N, v) that satisfies the conditions of Def. 2.1. (ii) For the unfamiliar reader, the characteristic function v defines the value or worth of each coalition S , i.e. it can be taken as the total amount of transferable utility that the members of S could earn without any help from the voters outside of S . In this sense when we define an SVG by v we are not merely distinguishing winning from losing coalitions by attaching to them their respective values of 1 and 0 as arbitrary labels; rather $v(S)$ represents the total payoff that the members of S earn when $S \in W$. That is, by winning a vote S captures the spoils or prize of victory (e.g. a lump sum of money to be won), which it then divides among its members. The formation of the winning coalition and the distribution of the spoils is not simply

a random process, but rather consequent upon the process of bargaining. (iii) The characteristic function form represents the voting rules as an n -person cooperative game for which all the basic assumptions are met in the case reorganization plans: voters can communicate before each play of the game, side payments are permissible and feasible, and utility is transferable.

Comment 2.5 It is easy to see that defining an SVG in terms of v captures the character of those type of voting games which are clearly about the division of a fixed purse. It represents the conflict of interests between winning and losing coalitions which is characterized by the zero-sum (or more generally, constant sum) nature of the outcome: the winner takes all. The voting by claimants on a reorganization plan is a particular case of such a game as it is all about the division of the hypothetical value of the reorganized firm over and above its liquidation value. In other words, the representation of an SVG by its characteristic function describes an SVG as an allocation mechanism because under a certain set of additional assumptions (see section 3.2), the party (debtor or claimant) that proposes the plan can simultaneously put together a winning coalition (each member being given a portion of the value above the liquidation floor).

Comment 2.6 A representation of an SVG by its characteristic function can also be considered as an instance of a *game form*. This is a term introduced by Gibbard (1973) to describe games in which individual utilities are not yet attached to possible outcomes (in effect a game without payoff functions). That is, a game form is a system which allows each individual his choice among a set of *strategies*, and makes an *outcome* depend, in a determinate way, on the strategy each individual chooses. Obviously, for an SVG the strategies are $\langle \text{yes, no} \rangle$ and the outcome is an element of $\{0,1\}$.⁴

We apply this distinction between a game form and a game in section 3 where we show that results for game forms do not necessarily hold for games. This is a methodological point of fair importance because as we will show, if we assume coalition formation – as for example Kordana and Posner (1999) do in their analysis of Chapter 11 voting rules – it is erroneous to draw positive and normative conclusions based upon the game form.

We will also make use of the following definitions:

Definition 2.7 In an SVG W , let S be a coalition and i a player. We say that: (i) i is *critical* in S if $S \in W$ but $S \setminus \{i\} \notin W$; (ii) i is a *dummy* if i is never critical; (iii) i is a *dictator* if $\{i\} \in W$ and all other voters are dummies; (iv) i is a *vetoer* if $\{i\} \in W$ and $N \setminus \{i\} \notin W$.

⁴ See Miller (1982) for a brief introduction to game forms.

Remark 2.8 It is easy to see that (i) a coalition is a MWC iff each of its members is critical; (ii) a player is a dummy iff it is never a member of a MWC; and (iii) i is a dictator if $\{i\}$ is the sole MWC.

We can define two types of SVGs, a *simple majority* SVG and a *unanimity* SVG.

Definition 2.9 (Felsenthal and Machover 1998) Let $I_n =_{def} \{1, \dots, n\}$ be a ‘canonical assembly’. Then for any positive integer k such that $k \leq n$, define $M_{n,k}$ as the SVG whose winning coalitions are just those subsets of I_n that have at least k members. As a matter of shorthand, we denote (i) the simple majority SVG as $M_{n,(n/2)+1}$, and (ii) the unanimity SVG as $M_{n,n}$.

As described in the introduction, a reorganization plan is approved if the claims of those who voted in favour of the plan aggregate to a certain proportion of the total claims. This means that we need the definition of that important subclass of SVGs known as *weighted voting games* (WVG):

Definition 2.10 (i) An SVG W is a WVG if there are nonnegative weights (w_1, \dots, w_n) allocated to the voters and a *quota* $0 < q \leq \sum_{i \in N} w_i$ such that $S \in W$ iff $\sum_{i \in N} w_i > q$. (ii) A WVG can be represented by:

$$[q; w_1, \dots, w_n] \tag{2.1}$$

Comment 2.11 In insolvency law, the weights w_i are the certified claims of each claimant, i.e. one euro one vote (similar to shareholders voting rights of one share one vote).

A further definition we need in order to analyse the voting rules in insolvency law is that of a *compound* or *composite* SVG. In all the countries that we have mentioned in the introduction (Canada, Germany, US, and UK), the creditors are divided into classes and each class votes upon the reorganization plan and in some of these classes there are two decision rules: a weighted rule (Def. 2.10) and a head count rule (in US parlance, the ‘numerosity’ requirement) (Def. 2.9). Thus the *game* is composed from each of the separate SVGs both within and between the classes. In order to model these rules we need a specific type of composite known as a *meet*.

Definition 2.12 (Felsenthal and Machover 1998) (i) Let m be a positive integer and let W^* be an SVG with assembly $I_m = \{1, \dots, m\}$. For each $i \in I_m$, let W_i be an arbitrary SVG. We now define an SVG

$$W^*[W_1, \dots, W_m] \tag{2.2}$$

called the *composite* of W_1, \dots, W_m under W^* . Next, define the assembly N of $W^*[W_1, \dots, W_m]$ as the union of the assemblies N_i of the W_i : $W_i: N =_{def} \bigcup_{i=1}^m N_i$. We refer to W^* as the ‘top’ and W_i as the i -th ‘component’ of the composite SVG $W^*[W_1, \dots, W_m]$.

(ii) A *meet* of the W_i is denoted by

$$W_1 \wedge W_2 \wedge \dots \wedge W_m \quad (2.3)$$

which is equal to a unanimity SVG $M_{n,n}$ on $[W_1, \dots, W_m]$, i.e. each component game must be won. If the assemblies N_i are pairwise disjoint then the meet is called a *product* of W_i and is denoted by,

$$W_1 \times W_2 \times \dots \times W_m \quad (2.4)$$

Remark 2.13 (i) The ‘top’ is the game that is made up of all the component games: the decisions of each W_i are fed to W^* which is a rule for collating the m lower level decisions into a final decision, i.e. W^* is the collection of subsets that assures the acceptance of a proposal.

(ii) A meet (2.3) can be used to model a multi-cameral system, that is when a motion has to be approved by different ‘chambers’, ‘committees’ or ‘houses’. If the voters are taken to be individual persons, and no person can be a member of more than one chamber, then the meet becomes a product (2.4).

(iii) For the case where all N_i coincide (the voters in each W_i are the same), then (2.3) can be given by:

$$W_1 \cap W_2 \cap \dots \cap W_m \quad (2.5)$$

Expression (2.5) can be used to represent SVGs with a weighted rule and a numerosity rule. SVGs of this kind are also known as ‘weighted games with consensus’ (Carreras and Freixas 2001) or ‘voting by count and account’ (Peleg 1992), which is what we find in the Canadian, German, and US insolvency rules, and which can be represented as

$$[q_1; w_1, \dots, w_n] \cap [q_2; \underbrace{1, \dots, 1}_n] \quad (2.6)$$

Note that the basic properties of (2.6) are that in such a game there is never (a) a dictator even if there exists a $w_i \geq q$, and (b) a dummy because a simple majority SVG guarantees that each voter i is in at least one MWC. In other words, (2.6)

provides a much more equal distribution of voting or bargaining power than (2.1) alone.⁵

Comment 2.14 (i) Def. 2.12 is extremely important for the analysis of the voting rules in insolvency law. Failure to comprehend it can result in inappropriate analysis. Kordana and Posner (1999), for example, call (2.6) a ‘bicameral system’. This is not entirely accurate as it is only the ‘bicameralism’ within each component game W_i in (2.2). Bicameralism – or better, the multi-cameralism – of the rules is actually represented by (2.4): claimants are divided into separate classes, with each creditor belonging to only one class. Put another way, splitting an assembly into different chambers is different than having the same set of voters vote twice.

The implication of this distinction is not without consequence because it is not clear that for WVGs an analysis based on the assumption of only one class, i.e. $W^*[W_1]$, is generalizable to the case of multiple classes, $W^*[W_1, \dots, W_m]$ – an assumption that appears to be made in the small amount of literature that deals with voting rules in insolvency law. This is especially the case when investigating voting or bargaining power: Kordana and Posner, for example, attempt to examine the effect of the different components of (2.6) for stylised examples with only one class of voters. While we will not engage in this question in this essay, it is suffice to say that Kordana and Posner completely miss the fact that a voter i can exert power if, and only if, he is critical (see Def. 2.7) in a coalition that wins at the ‘top’. This could dramatically alter their results.

(ii) Expression (2.4) indicates that at the ‘top’ all classes or interests will be represented in the reorganization plan. That is, the combination of classification and a product game guarantees additional ‘consensus’ in the outcome, reducing exploitation of particular classes of claimants, although this will lead to an increase in bargaining costs as each class now obtains veto power. As we will prove below (Lemmas 4.2 and 4.4), (2.4) can also imply that as the number of components in (2.2) increases, more claimants can receive some value above the distribution or liquidation floor, i.e. the amount that claimants would receive if the insolvent company was liquidated instead of reorganized. This insight has also been made by Kordana and Posner (1999, pp. 211–212) but without formal proof. They also claim – again without formal proof – that as the number of components of (2.2) increases it becomes more difficult to approve a reorganization plan. We will show that this is generally true only under the assumption that all winning coalitions are equally probable (i.e. when we consider the game form). This conflicts with their implicit assumption made throughout their analysis that only minimal winning coalitions will form – which in fact turns out to be a very reasonable assumption (section 3.2). In the latter case the effect of partitioning claimants into different

⁵ See, for example, Peleg (1992).

classes is not clear cut: it does not necessarily make it harder to pass a plan, although it will generally still increase the number of creditors obtaining value in excess of the liquidation floor.

3. Resistance

3.1 A Priori Resistance

In this section we introduce a very simple, but useful, metric for comparing decision rules: the ‘power of a collectivity to act’ (Coleman 1971) or its linear transformation, a *coefficient of resistance* (Felsenthal and Machover 1998, p. 62). One can think of these two measures in the following sense. In an assembly of n members, there are 2^n distinct partitions into the set of ‘yes’ and ‘no’ voters, then the ‘power of a collectivity to act’ is defined as the *a priori* probability that a proposal that is put before the assembly will be approved.

Definition 3.1 (Coleman 1971) We put

$$\mathcal{A}(W) =_{\text{def}} \frac{|W|}{2^n} \quad (3.1)$$

Remark 3.2 \mathcal{A} measures the *a priori* probability that a proposal will be adopted by the decision-making body rather than rejected by it; i.e. it is a measure of success or ‘complaisance’ of W . Under the unanimity rule, the power of the collectivity is at a minimum because only the grand coalition can approve a proposal, i.e. $1/2^n$. If the assembly has an odd number of members and the decision rule is simple majority, then exactly half the coalitions will be equal to, or greater than, the number required for approval, so $\mathcal{A} = 0.5$; for large n and an even number of members, \mathcal{A} is slightly less than 0.5.

Felsenthal and Machover (1998, p. 62) define a linear transformation of \mathcal{A} which they call a *coefficient of resistance*.

Definition 3.3 (Felsenthal and Machover 1998) We put

$$\mathcal{R}(W) =_{\text{def}} \frac{2^{n-1} - |W|}{2^{n-1} - 1} \quad (3.2)$$

Remark 3.4 (i) It is easy to see that $\mathcal{R} = 0$ for a simple majority SVG and obtains its highest value $\mathcal{R} = 1$ for a unanimity SVG. For large values of n , \mathcal{R} approximates to Coleman’s \mathcal{A} as $\mathcal{A} \approx 1 - \mathcal{R}/2$.

(ii) \mathcal{R} measures the dual of success, namely the propensity of an SVG to favour a negative outcome.

(iii) \mathcal{A} and \mathcal{R} can be regarded as measuring the quality of decision-making in terms of the trade-off between errors of not approving proposals that should have been accepted and approving bad proposals that should have been rejected. Using an analogy from the classical theory of statistical inference, this corresponds to the trade-off between Type-I and Type-II errors respectively. For example, as \mathcal{R} rises (\mathcal{A} falls) one decreases Type-II errors but only at the expense of increasing Type-I errors.

Comment 3.5 Applying Coleman's \mathcal{A} and Felsenthal and Machover's \mathcal{R} for comparing the different voting rules would clearly be very informative. The problem, however, is that in contrast to the voting rules in a political body, say the European Commission,⁶ we know neither N nor w_i nor the components of $W^*[W_1, \dots, W_m]$. Thus it would seem that \mathcal{A} and \mathcal{R} are of little use unless we were to construct stylised examples, which while informative may contain little or no generality. There is, however, a property of \mathcal{A} and \mathcal{R} that is of particular interest and which allows us to make a qualified judgement.

Recall that the voting rules for reorganization plans are weighted rules, i.e. WVGs. A very important fact about WVGs that is apparently not widely realized is that if q is pegged at a constant percentage of the sum of weights, and if the percentage is greater than 50 percent, then as the number of voters increases \mathcal{R} tends to its maximal value (\mathcal{A} approaches its minimal value) very quickly.⁷

We can therefore make the following claim:

Claim 3.6 *If we would have a set of voters that are classified into secured and unsecured creditors and any further partition thereof, then (a) the probability of a reorganization plan being approved under the German rules is never less than under the Canadian or US rules, and (b) the resistance of the Canadian and US rules will tend to be higher than the German rules as N increases.*

Proof Follows directly from (a) the fact that for the German rules, each component game is $[50; w_1, \dots, w_n] \cap M_{n, (n/2)+1}$ while for the Canadian and US rules each component game is $[66; w_1, \dots, w_n] \cap M_{n, (n/2)+1}$ and (b) Com. 3.5. ■

Remark 3.7 (i) Note the claim is (a) weak, i.e. $|W^*|_{\text{Germany}} \geq |W^*|_{\text{US/Canada}}$ because if there is a voter $i \in N$ with $w_i = 67$, then both rules produce the same set of win-

⁶ See Felsenthal and Machover (2001) for an example of the use of \mathcal{R} to compare different voting rules before and after the Treaty of Nice.

⁷ This proposition is given in Felsenthal (2001: 456) but without formal proof. For a proof and further investigation of this property of WVGs, see Lindner (2003). As a graphic illustration, for the simple case of one-man-one-vote with $q \geq 67$, \mathcal{A} falls below 0.05 by the time we have 28 players.

ning coalitions.

(ii) Also note that our claim is limited to the case where claimants can be classified in the same way. The German rules give no voting rights to equity, so we have omitted this class.

(iii) Note further that we have not made a comparison with UK rules because these rules do not have a numerosity requirement and the majority quota's for both classes are different: greater than 75 percent for unsecured creditors and greater than 50 percent for equity. While a comparison with the US rules would be possible if we assume away secured creditors, we cannot make any general comparison about the cardinality of the set of winning coalitions. The UK rules may produce more, the same, or less winning coalitions than the US rules and therefore the resistance of the UK rules may be lower, the same, or higher than the US rules.

Example 3.8 Consider the games,

$$[75; 75, \underbrace{1, 1, \dots, 1}_{25 \text{ times}}] \times [50; 66, \underbrace{1, 1, \dots, 1}_{34 \text{ times}}]$$

and,

$$([66; 75, \underbrace{1, 1, \dots, 1}_{25 \text{ times}}] \cap [14; \underbrace{1, 1, \dots, 1}_{26 \text{ times}}]) \times [66; 66, \underbrace{1, 1, \dots, 1}_{34 \text{ times}}]$$

The first game represents the UK rules, the second the US rules and where the first component is the class of unsecured creditors and the second the class of equity. Note that for the UK rules *any* coalition containing the first voter and at least one other voter is a winning coalition. For equity it is the same. Under the US rules, however, a winning coalition for the unsecured creditors must contain at least 14 voters. This is the source of the resistance that is not present in the UK rules, which will produce many more winning coalitions and therefore be less resistant than the US rules *for this game*. It is not difficult to construct an example in which the converse is true.

Comment 3.9 In Claim 3.6 we compared the resistance of the rules in terms of the majority quotas. There is, however, another source of resistance that we have already indicated in Com. 2.14: the effect of the number of components in the composite game (i.e. the effect of classification). The basic intuition, and as it turns out one that is straightforward to prove, is that as the number of components of a product game increases, there is a fall in number of winning coalitions at the top and therefore \mathcal{R} will increase (\mathcal{A} decreases). As a matter of illustration, consider the SVG $M_{n, (n/2)+1}$. By Rem. 3.4, $\mathcal{R} = 0$. Now partition the SVG into n components so that we have n committees each with a single player. If we apply a product game

at the top, then we have an equivalent to a unanimity SVG and by Rem. 3.4, $\mathcal{R} = 1$. Thm 3.10 is a general proof.⁸

Theorem 3.10 *Let W be a $M_{n,k}$ with $n > k \geq (n/2) + 1$ and $q =_{\text{def}} k/n$. If N is partitioned into m disjoint non-empty chambers playing a product game W^* with the same q in each chamber and $q < 1 - (1/n_h)$ in at least one chamber h , then $|W^*| < |W|$.*

Proof Partition N into two disjoint non-empty subsets N_1 and N_2 playing the product game $W^* = W_1 \times W_2$. Let each game apply the same quota q . The idea is to show that there exists at least one coalition $S \in N$ that is winning in W , but losing in W^* if W_1 and W_2 have the same quota as W . To achieve this take a MWC in each of the chambers and subtract a player from a one of the MWCs so that it is now a losing coalition in this chamber and then add a player to a the MWC in the second chamber so that by monotonicity of W (Def. 2.1) it is still winning. Doing this we obtain a coalition winning in W , but losing in W^* .

Formally, we start by demonstrating weak inclusion, $W^* \subseteq W$. Let $S_h \subseteq N_h$, then $S_1 \in W_1$ and $S_2 \in W_2$ implies $S_1 \cup S_2 \in W^*$. As $|S_1| > q \cdot n_1$ and $|S_2| > q \cdot n_2$, $|S_1| + |S_2| > q(n_1 + n_2) = q \cdot n$ and thus each $S_1 \cup S_2 \in N$ winning in W^* is also winning in W from which the weak inclusion follows.

We now demonstrate strict inclusion, $W^* \subset W$. Let $\lfloor q \cdot n_h \rfloor$ be the largest integer such that $q \cdot n_h \geq \lfloor q \cdot n_h \rfloor$. If $n_1 > \lfloor q \cdot n_1 \rfloor + 1$ then there exists an S_1 with $|S_1| = \lfloor q \cdot n_1 \rfloor + 1$ and an $N_1 \setminus S_1 \neq \emptyset$. Let S be given by $S = S_1 \cup S_2$ with $|S_2| = \lfloor q \cdot n_2 \rfloor + 1$. Then take an $i \in N_1 \setminus S_1$ and a $j \in S_2$. Now $(S_1 \cup \{i\}) \cup (S_2 \setminus \{j\}) \in W$, but $(S_2 \setminus \{j\}) \notin W_2$. Finally, note that $(1 - q)n_1 > 1$ implies $n_1 > \lfloor q \cdot n_1 \rfloor + 1$. This proves $W^* \subset W$ and $|W^*| < |W|$ as claimed for $m \geq 2$ because we can apply the same reasoning for each partition of chamber h . ■

Corollary 3.11 *For $M_{n,(n/2)+1}$ Thm 3.10 holds if $n > 2$.*

Note that the voting rules on reorganization plans for Canada, Germany, and the US contain an intersection between a WVG and $M_{n,(n/2)+1}$. While the $M_{n,(n/2)+1}$ will reduce the number of winning coalitions, we must check that this is not off-set by the WVG. As it turns out, the partition effect also holds for this class of games, albeit under a very small restriction.

Theorem 3.12 *Let W be a WVG $[q; w_1, \dots, w_n]$ with $\sum_{i \in N} w_i / 2 < q < \sum_{i \in N} w_i$. If N is partitioned into m disjoint non-empty chambers playing a product game W^* with the same q in each chamber and if for at least one chamber h an $S_h \in W_h$ with*

⁸ Note that we are ignoring the legal restrictions on the ability to classify; this does not form part of the *a prioristic* and structural analysis of this paper. This has no effect on our results in any way because they concern the parameters in which the law operates.

$S_h \subset S \in W$ fulfils $\sum_{j \in N_h \setminus S_h} w_j \geq w_i'$ where w_i' denotes the smallest critical player $i \in S \setminus S_h$, then $|W^*| < |W|$.

Proof The proof is similar to Thm. 3.10. Partition N into two disjoint non-empty subsets N_1 and N_2 playing the product game $W^* = W_1 \times W_2$. Let W_1 be $[q; w_1, \dots, w_l]$ and W_2 be $[q; w_{l+1}, \dots, w_m]$. We need to show that a losing coalition in W^* is winning in W .

Again, we first demonstrate weak inclusion, $W^* \subseteq W$. Let $S_h \subseteq N_h$, then $S_1 \in W_1$ and $S_2 \in W_2$ implies $S_1 \cup S_2 \in W^*$. Then $\sum_{i \in S_1} w_i > q \sum_{i \in N_1} w_i$ and $\sum_{i \in S_2} w_i > q \sum_{i \in N_2} w_i$ and therefore $\sum_{i \in S_1 \cup S_2} w_i > q \sum_{i \in N_1 \cup N_2} w_i$ from which weak inclusion follows.

We now demonstrate strict inclusion, $W^* \subset W$. Without loss of generality, let w_i' denote the weight of the smallest critical player $i \in N_1$. In order to find this player determine in N_1 every coalition with at least one critical player. Then select the player with the smallest weight for each coalition and minimize such players to get the smallest critical player 'ever'. Denote this player by w_i' . Let $S_1' \in W_1$ denote the coalition containing w_i' . If in N_2 an $S_2 \in W_2$ exists such that $\sum_{i \in N_2 \setminus S_2} w_j > w_i'$, then $(S_1' \setminus \{i\}) \cup N \in W$, but $S_1' \setminus \{i\} \notin W_1$. This proves $W^* \subset W$ and $|W^*| < |W|$ as claimed for $m \geq 2$ because we can apply the same reasoning for each partition of chamber h . ■

Comment 3.13 (i) While Thms. 3.10 and 3.12 are not particularly 'deep', they do have a fair amount of cutting power because they show that there is an inherent disadvantage to classification. If, for example, two jurisdictions have the same voting rules, e.g. Canada and the US, but different propensities to classify, the jurisdiction with the higher propensity will have a higher number of rejected reorganization plans.⁹

The effect, however, can be cancelled out by methods lying outside the pure structure of the rule. 'Cram down' is one such example. This is the case when a class of claimants fails to achieve a majority in favour of the reorganization plan (i.e. the plan fails in a component game) but the plan is ratified by the court because the members of the class are not worse off under the plan than under liquidation. In other words, cram down actually changes the decision rule.

(ii) Thms. 3.10 and 3.12 also limit the strategic incentive of the proposer of a plan to create classes as a way of getting around other rules such as 'equal treatment', which stipulates that each creditor of the same type must receive the same amount on their debt. In the absence of cram down, any attempt by the proposer of a plan to pair off creditors by, for example, their discount factors (i.e. level of patience) and therefore pay one creditor less than the other even though they are

⁹ This is in fact an empirical hypothesis that should in principle not be too difficult to test but one which is beyond the scope of this paper.

of the same broad type, could be counter productive because the plan will be more difficult to approve.

(iii) Finally, Thms 3.10 and 3.12 indicate that classification directly affects the trade-off between Type-I and Type-II errors (Rem. 3.4). That is, with each new partition the probability that a good reorganization plan will be rejected also increases. The upshot is that while classification may be considered as a means of protecting pre-bankruptcy entitlements because it ensures that all interests are represented in the final decision, it does come at a cost: plans that ought to be passed are not, which is evidently a welfare loss because it is increasingly likely that a firm will be liquidated when it is socially preferable to maintain it as a going concern.

3.2 Resistance under Coalition Formation

Despite the clarity and simplicity of Coleman's \mathcal{A} and Felsenthal and Machover's \mathcal{R} and of Thms. 3.10 and 3.12 they have an essential drawback: they assume that all coalitions are equally probable. This assumption has a very important implication for it essentially fits the description of the voting rules as game forms (Com. 2.6) and not necessarily as games in the usual sense because the utilities or payoff functions are not yet attached to the outcomes. While this may be a useful starting point – and has in particular normative appeal because of its Rawlsian veil of ignorance character – it is necessary to consider what happens if we introduce utility and coalition formation which will restrict the members of W that form. Doing so, however, will have consequences for both the effect of the majority quota and partition effect in that it is no longer clear that a higher q or increasing the number of component games decreases the probability that a proposal will be approved.

The theory that we will introduce is fairly elementary and requires us to consider the characteristic function representation of an SVG (Def. 2.3). Recall from Rem. 2.4, and Com. 2.5 that v describes an SVG as zero-sum game and where only the winning coalition obtains positive value and each member of the coalition obtains a positive payoff. If we make some additional assumptions that appear to fit the context of bargaining over corporate reorganization quite well, namely, that the voters are rational, have complete and perfect information, can make side-payments, and control membership of a winning coalition, then we can restrict our attention only to the set W^m . That is, only minimal winning coalitions will form. This is a very well-known proposition in political science known as the *size principle* (Riker 1962). In its most general form the size principle says:

Proposition 3.14 (Riker 1962, p. 32–33) *In n -person, zero-sum games with side-payments, participants create coalitions just as large as they believe will ensure winning and no larger.*

Riker demonstrates that there are no circumstances for an n -person game so defined wherein an incentive exists to form a coalition other than a MWC. His reasoning is as follows (without the technical details):

(i) Given that only winning coalitions have positive value, the zero-sum nature of the game implies that losing coalitions have negative value. Since a losing coalition has no positive value to distribute among its members, it would only form as a pretender to eventual winning status. The possibility that a losing coalition could eventually become winning provides a strong incentive for a winning coalition to pare off superfluous members before they and other disaffected members, to whom they cannot offer sufficient payoffs, defect.

(ii) If the excess members are not rejected, the winning coalition becomes vulnerable to offers from a losing coalition that could promise enough defectors greater rewards in a prospective MWC so as to actually constitute such a coalition. Because the members of a MWC have complete and perfect information, they know when they have enough members in the coalition to win.

Remark 3.15 (i) For our purposes, there are two points to note about the size principle. Firstly, rationality motivates the voters to obtain the benefits of being in a winning coalition. Secondly, since any win is the sole determinant of value, the ejection of superfluous members from a winning coalition means the same total amount of value can be divided among the fewer members of a MWC.

The size principle seems to fit the context of bargaining and voting on reorganization plans quite well. The reorganization plan is, as we noted in the introduction to this essay, about the division of the future value of the firm (or more generally, its capital structure). Thus any party that proposes a plan, be it the debtor or claimant, will want to maximize their own payoff, which means excluding surplus voters who will receive nothing above the amount they would receive under liquidation (or, if one abstracts from the liquidation floor, nothing at all).

Although it could be argued that assumptions of complete and perfect information is at odds with much of the corporate finance literature that highlights informational asymmetries between classes of claimholders, this is not a too serious objection. Firstly, the informational asymmetries will not lead to oversized coalitions because rational agents simply would not give anything above the liquidation floor to a 'dummy' claimant, i.e. a claimant which cannot change a coalition from winning to losing. The fact that this claimant might still vote in favour of the plan due to the informational asymmetries will not affect his or her payoff, unless, of course the asymmetries are so severe that the players are unable to determine what is a MWC (which seems unlikely). Such players simply agree with the outcome without gaining pecuniary benefit. One could say that such oversized coalitions are formed by 'luck' (Barry 1980a, 1980b, Holler and Packer 1983). Secondly, and as a corollary of the first reason, we can safely bet that informational asymmetries will only affect *which* of the set of MWCs will form. Given that we do

not know *a priori* the nature of the asymmetries we must assume that each MWC is equally likely.¹⁰

(ii) One of the major criticisms of the size principle is that MWCs are unstable and therefore coalitions tend to include excess members as ‘security’ against defections.¹¹ This is a very valid point for political games, i.e. when the winning coalition can divide the spoils of office only for a given period. In the context of reorganization plans, once the court has ratified the plan that is it: the cake is cut once and for all and there can be no collapse of a MWC.¹²

Observation 3.16 Although by Rem. 2.2 W^m completely determines W , it is mistaken to believe that if only MWCs form and we replace $|W|$ by $|W^m|$ in (3.1) and (3.2), Claim 3.6 and Thms. 3.10 and 3.12 hold. They do not. For WVGs, $|W^m|$ is only weakly decreasing in q : as q increases for a given assembly and distribution of weights, the number of MWCs may be the same or decrease. For $M_{n,(n/2)+1}$ it always decreases.

Example 3.17 (i) Consider,

$$[50; 50, 15, 15, 10, 10] \cap M_{n,(n/2)+1} \cdot |W^{*m}| = 6$$

$$[66; 50, 15, 15, 10, 10] \cap M_{n,(n/2)+1} \cdot |W^{*m}| = 6$$

(ii) Now consider,

$$[50; 35, 30, 20, 10, 5] \cap M_{n,(n/2)+1} \cdot |W^{*m}| = 7$$

$$[66; 35, 30, 20, 10, 5] \cap M_{n,(n/2)+1} \cdot |W^{*m}| = 4$$

Remark 3.18 The reason for the ambiguity of the effect of increasing q on $|W^{*m}|$ is not hard to explain. For WVGs (the first component of the games in Ex. 3.17), we see that as q increases one generally needs more voters to form a MWC. This would reduce $|W^{*m}|$. On the other hand, an increase in q may now make some voters more important, they are now in more MWCs than previously, which can increase $|W^{*m}|$. This in fact occurs for (i) if there is no numerosity requirement – $|W^{*m}|$ actually increases from 4 to 6 with an increase in the quota from simple to

¹⁰ This actually raises an interesting theoretical and empirical question about which types of coalitions do indeed form in corporate reorganizations, particularly if the choice of corporate capital structure is endogenous. But it should be noted that this is essentially a different question to that addressed in this paper because it concerns predicting coalition formation and not the *ex ante* assessment of voting rules. The two issues overlap, but are not necessarily the same. In fact, Kordana and Posner (1999: 204–205) suggest an interesting approach to this question based upon the discounted liquidation values of the creditors.

¹¹ See Laver and Schofield (1990, pp. 144ff) for a detailed discussion of this criticism. See also Frohlich (1975) for a bargaining argument that will result in coalitions containing excess members.

¹² Note that Kordana and Posner’s (1999) multi-party bargaining model implicitly presumes the size principle because they assume the proposer of a plan maximizes their expected value in the plan and can freeze other creditors who receive none of the surplus above the liquidation floor (or nothing at all if the liquidation floor is assumed away).

qualified majority by weight. Which effect is predominant depends upon N and the distribution of weights.

Claim 3.19 *If only MWCs form, then ceteris paribus, it can be the case the Canadian/US and German rules have equal resistance if \mathcal{R} is measured by $|W^{*m}|$.*

Proof From Rem. 3.18 ■

Remark 3.20 Claim 3.19 is somewhat counter-intuitive. What this implies is that for a given N and (w_1, \dots, w_n) there is a range of q below a threshold quota, \hat{q} , which will not decrease the probability that a proposal will be approved (where \hat{q} is the point where we obtain $M_{n,n}$, i.e. $|W| = |W^m| = 1$). That is, by increasing q we do not necessarily make it harder for one group of individuals to act against the aims and interests of another, if only MWCs form.

We now need to examine the partition effect under the size principle, viz. what is the effect of the number of component games on $|W^{*m}|$, and therefore on the resistance of the rule.

Observation 3.21 From the proof of Thm. 3.10 it follows that for $M_{n,(\eta/2)+1}$ a partition will reduce $|W^{*m}|$. For WVGs, however, the partition effect is ambiguous. Whether $|W^{*m}|$ increases or decreases depends upon (w_1, \dots, w_n) , q and where the partition falls.

Example 3.22 Consider the case of a set of 9 voters with weights (31,25,20,15,5,1,1,1,1). (For notational convenience we denote simple majority by weight as SM and qualified majority as QM.

(i)	SM	[50; 31, 25, 20, 15, 5, 1, 1, 1, 1]	$ W^m = 8$
	QM	[66; 31, 25, 20, 15, 5, 1, 1, 1, 1]	$ W^m = 12$
(ii)	SM	[48; 31, 25, 20, 15, 5] \times [2; 1, 1, 1, 1]	$ W^m = 20$
	QM	[64; 31, 25, 20, 15, 5] \times [2.6; 1, 1, 1, 1]	$ W^m = 16$
	SM	[38; 31, 25, 20] \times [12; 15, 5, 1, 1, 1, 1]	$ W^m = 3$
	QM	[50.6; 31, 25, 20] \times [16; 15, 5, 1, 1, 1, 1]	$ W^m = 14$
(iii)	SM	[38; 31, 25, 20] \times [10.5; 15, 1, 1] \times [1.5; 1, 1, 1]	$ W^m = 9$
	QM	[50.6; 31, 25, 20] \times [14; 15, 1, 1] \times [2; 1, 1, 1]	$ W^m = 6$
	SM	[20.5; 20, 15, 5, 1] \times [28; 31, 25] \times [1; 1, 1]	$ W^m = 3$
	QM	[27.3; 20, 15, 5, 1] \times [37.3; 31, 25] \times [1; 1, 1]	$ W^m = 1$

Comment 3.23 (i) Although Ex. 3.22 makes it abundantly clear that resistance can decrease with classification under the size principle, this effect will be off-set in the German and Canadian/US rules due to the numerosity requirement. This means that the $|W^{*m}|$ will either stay the same or fall.

(ii) Obs. 3.21 and Ex. 3.22 raise two important methodological issues. First, in the absence of the full specification of the WVG – i.e. we know both q and (w_1, \dots, w_n) – our ability to make general inferences regarding the effect of classification is rather limited when the size principle is at work. Second, and as a corollary of the first, given that the size principle produces results that are in sharp contrast to those when we assume all coalitions are equally likely, it is extremely hazardous to ignore coalition formation when analysing voting rules such as those in insolvency law.

The importance of this second point should not be overlooked. Recall from the introduction to this section that we said that the equiprobability assumption implies that we are not really dealing with a cooperative game but merely the set of coalitions that can win, i.e. the game form. In other words, the elements of coalition formation as assumed by Riker's size principal are exogenous to this set of subsets of the set of voters. Once these elements are introduced – which is essentially the introduction of utilities – then the 'rule' (or game form) becomes 'transformed' into a game. And what we infer about a game should not be taken to hold for the rule (game form).

4. *Distribution of Value*

At the outset of this essay we said that reorganization plans ultimately concern the distribution of value among those with pre-bankruptcy entitlements. The obvious question that arises is how do the rules differ in terms of the distribution of value. It should be noted, however, that the issue of concern here is purely *descriptive*, and not normative, although clearly it has normative interest. To be explicit, the idea in this section is *not* to test, for example, the functioning of rules governing equal treatment (each creditor in a class gets the same number of cents to the euro). This would be wholly mistaken because equal treatment does not entail equal distribution; it only entails that that equal claims receive equal amounts. What we do show, however, is that as number of classes increase the distribution of value tends towards equality, regardless of differences in the initial size of the claim. This has the strange implication that unequal treatment – placement of equal claims in different classes – can lead to more equal outcomes.

Our claim is easy to grasp: imagine each creditor is put in a separate class; then we would have a unanimity game at the 'top' (see Com. 3.9). If we assume, as Kordana and Posner (1999) do, that only members of a winning coalition receive value (i.e. we abstract away from the liquidation floor), and as there is only one such coalition, viz., $\{N\}$, in this game this means that each claimant gets something; which is obviously more equal than some claimants getting nothing.¹³

¹³ Note that if the payoff would be proportional to voting power as say measured by the

Given that this is an extreme case – the law generally would not permit such a classification (particularly as it can violate equal treatment) – we will on the whole have more than one winning coalition. Given this, the natural step is to use the winning coalition with the least number of members as our baseline. Phrased differently, we can compare the distribution of value in terms of the cardinality of the smallest possible winning coalition, i.e. the MWC with the least number of members. We denote such a coalition by S^l .

While the procedure of taking an S^l as the measuring stick slots nicely in Riker's size principle, a theory of coalition formation is not actually necessary here. The S^l is but a heuristic device: it allows us to assess how many creditors at *the very least* will get something from the plan (although not how much). Obviously, as the size of an S^l increases *more creditors gain some value*, and if the value is a fixed sum, then it will necessarily be distributed more equally as S^l grows in size.

In the previous section we compared both q and the number of chambers. Here we will concentrate only on the latter because for WVGs the size of S^l is dependent upon both (w_1, \dots, w_n) and q which makes generalizations very difficult. Ex. 4.1 shows that it is quite possible that the different majority quotas produce the same S^l or at least the same sized S^l :

Example 4.1 Consider an assembly with five voters with (40, 25, 15, 15, 5). Under simple majority by weight we obtain three MWCs each with three voters; and under qualified majority we also obtain three MWCs also with three voters. In both cases the MWCs are also S^l .

In contrast, however, we can make some general inferences about the effect of partitioning the creditors into disjoint classes. While this does not allow us to make general comparisons across the rules, we are in a position to make some general inferences about each of the rules given the propensity to create classes of creditors.¹⁴ Essentially this boils down to saying that, *ceteris paribus*, the higher the propensity to classify, the greater the distributional equality.

Lemma 4.2 *Let W be a WVG $[q; w_1, \dots, w_n]$ with $\sum_{i \in N} w_i / 2 < q < \sum_{i \in N} w_i$. If N is partitioned into m disjoint non-empty chambers playing a product game W^* with the same q in each chamber, then $|S^l| \leq |S^{*l}|$.*

classical measures of voting power such as the Shapley-Shubik (Shapley and Shubik 1954) or Banzhaf (1965) indices, then the payoffs would be equal under a unanimity game, regardless of the voting weights (claims) of each claimant.

¹⁴ The propensity to classify is determined among other things by, for example, whether classification is by legal entitlement only or by economic interests. The German rules permit only the former while the US rules also permit the latter.

Proof First, partition N into two disjoint non-empty subsets N_1 and N_2 playing the product game $W^* = W_1 \times W_2$. Next, let W_1 be $[q_1; w_1, \dots, w_l]$ and W_2 be $[q_2; w_{l+1}, \dots, w_m]$. From Thm. 3.12 each S^{*l} is also winning in W . Thus by definition of S^l it cannot be the case that $S^{*l} \subset S^l$ and therefore $|S^l| \leq |S^{*l}|$ holds. ■

Corollary 4.3 A necessary and sufficient condition for the strict inequality of Lemma 4.2 is that there exists an S^{*l} such that the surplus given by $sur(S^{*l}) =_{\text{def}} \sum_{i \in S^{*l}} w_i - q \sum_{i \in N} w_i$ fulfils $sur(S^{*l}) > \min_{i \in S^{*l}} w_i$ because in that case $\sum_{i \in S^{*l}} w_i - q \sum_{i \in N} w_i - \min_{i \in S^{*l}} w_i > q \sum_{i \in N} w_i$. Therefore at least one member in S^{*l} is redundant in S^l .

Lemma 4.4 Let W be a $M_{n, (n/2)+1}$, i.e. $q = 0.5$. If N is partitioned into m disjoint non-empty chambers playing a product game W^* with the same q in each chamber as in W and n_h is even in at least one chamber h , then $|S^l| < |S^{*l}|$.

Proof Partition N into two disjoint non-empty subsets N_1 and N_2 playing the product game $W^* = W_1 \times W_2$. Let each game apply the same quota $q = 0.5$. Use the surplus condition of Cor. 4.3 by representing $M_{n,k}$ by a WVG. Then $sur(S \in W^m) \leq 1$ by the property of an $M_{n,k}$. For $q = 0.5$ this implies $sur(S \in W^m) = \{0, 1\}$. Consider the case that n_1 and n_2 are even numbers. Then the threshold $q \cdot n_h = 0.5n_h$ is an integer and implies $sur(S_h \in W_h^m) = 1$. Therefore $sur(S) = sur(S_1 \in W_1^m) + (S_2 \in W_2^m) = 2$. Then $S \notin W^m$ and thus $|S^l| < |S^{*l}|$ for even n_h . Now consider n_2 as an odd number. Then $q \cdot n_2 + 0.5$ is an integer and $sur(S_2 \in W_2^m) = 0.5$. Thus $sur(S) = sur(S_1 \in W_1^m) + (S_2 \in W_2^m) = 1.5$ which again implies $S \notin W^m$ and by that $|S^l| < |S^{*l}|$ for the case that n_1 is a even and n_2 is an odd number as claimed for $m \geq 2$ because we can apply the same reasoning for each partition of chamber h . ■

Remark 4.5 A similar argument applies for $M_{n,k}$ with $n > k \geq (n/2) + 1$ in general. For each q we have to check in which cases the surplus $sur(S^{*l})$ exceeds 1. In these cases at least one player in S^{*l} becomes redundant when we move from W^* to W and thus $|S^l| < |S^{*l}|$.

We can now use Lemmas 4.2 and 4.4 to state the following proposition:

Proposition 4.6 Classification of a set of voters under the Canadian/US and German rules never reduces distributional equality (i.e. never makes the distribution of value more unequal).

Proof Immediate from Lemmas 4.2 and 4.4. ■

Remark 4.7 Lemma 4.2 says that for an WVG the size of S^l never decreases with a partitioning of the set of voters, while Lemma 4.4 says that S^l will increase if at least one chamber has an even number of voters. In other words, if a very weak

condition is met, then classification under the Canadian/US and German rules *always* increases distributional equality. Another way of putting Prop. 4.6 is to say that classification never harms pre-bankruptcy entitlements in the sense that it ensures that an increasing number of claimants get something (although this invariably means that some will get less), although it may have adverse implications for efficiency if it results in too many good plans being rejected.

Comment 4.8 (i) The implication of Prop. 4.6 is important because it more or less says that in the absence of cram-down procedures, classification works against the interests of the party proposing a reorganization plan (be it the debtor or creditor), because the cake will have to be shared among more creditors.¹⁵ It should be noted, however, that the propensity to classify is not a property of the decision-rules *simpliciter*, but the legal and strategic environment. That is, classification is not the decision rule, but rather defines it, because it defines the number of component games and therefore the distribution of value.

(ii) While we said that using a S^l is only a heuristic device, it can also be considered as a more restrictive version of the size principle, i.e. only a S^l will in fact form. The argument here is that if bargaining costs are high, the proposer of a reorganization plan (debtor or creditor) will have the incentive to put a MWC together with as few members as possible because negotiations will be easier and it will also be easier to keep the coalition together for the period until it is certified by the courts.¹⁶

(iii) From a conceptual standpoint, Prop. 4.6 indicates that the voting rules on reorganization plans do not simply reflect pre-bankruptcy entitlements over influencing whether or not a plan is accepted, but once coalition formation is taken into account, become allocation mechanisms. Phrased another way, voting rules do not merely allocate bargaining or voting power, but the distribution of payoffs.

5. Concluding Remarks

We wish to tie up this essay by setting our analysis and results in the context of the normative problem of institutional design. In their wide ranging analysis, Kordana and Posner (1999, pp. 202–208) question, for instance, the necessity of having a ‘count and account’ rule. They point out that under certain reasonable and mild assumptions, a simple majority rule (head count) would be sufficient and socially

¹⁵ Note that our result is not dependant on which party proposes the plan.

¹⁶ Leiserson (1968) actually introduced the notion of a *least member* winning coalition in an entirely different context. While one could argue that the incentive is in fact to form a *weight* minimal-MWC (Brams and Fishburn 1995), i.e. a MWC with the least weight, for the debtor, this might in fact involve more voters and the gains associated with reducing the payout maybe offset by the increased bargaining costs.

efficient. This is not the place to analyse their claims, but we wish to use this opportunity to show how our analysis can help answer this question of whether count and account is required.

If we recall Lemma 4.2 we see that for a WVG classification will never reduce distributional equality, i.e. it is never the case that a partition will reduce the size of the winning coalition with the smallest number of voters. Said otherwise, if we could find reasons to drop the numerosity requirement in the Canadian/US and German rules one of the consequences could be that although the resistance of the rule may fall, there could be less creditors getting something over and above the liquidation floor, i.e. more creditors getting the liquidation floor.

A further fact of major importance to the design of voting rules is that if only MWCs form the numerosity requirement in the Canadian/US and German rules has a very significant impact. As Ex. 3.22 demonstrates, the number of MWCs can actually increase with classification, making it easier to pass a plan. In the absence of the numerosity rule this would definitely give the party with the power to classify an incentive to gerrymander.

On a final methodological note our analysis picks up characteristics of the rules and voting games that a non-cooperative game theoretic approach cannot capture. And in this regard, one should be aware that a non-cooperative approach is very dependant upon the specification of preferences and payoffs which limits the generality of the results. As our examples demonstrate, when dealing with voting rules it is not clear that one can generalize from stylised examples. Possibly the main message of this essay is that future research on voting rules in insolvency law must recognize that coalition formation matters and this is a function not only of player preferences but of the formal structure of the decision rules.

References

- Aghion, P., O. Hart and J. Moore (1992) 'The Economics of Bankruptcy Reform', *Journal of Law, Economics, and Organization* 8: 523–46.
- Arrow, K.J. (1963) *Social Choice and Individual Values*. New York: John Wiley.
- Arrow, K.J. (1995) 'A Note on Freedom and Flexibility', in K. Basu, P.K. Pattanaik and K. Suzumura (eds) *Choice, Welfare and Development: A Festschrift in Honour of Amartya K. Sen*. Oxford: Oxford University Press.
- Aumann, R.J. (1959) 'Acceptable Points in General Cooperative n -Person Games', in A.W. Tucker and L.R. D (eds) *Contributions to the Theory of Games IV*. Princeton: Princeton University Press.
- Baird, D.G. and R.C. Picker (1991) 'A Simple Noncooperative Bargaining Model of Corporate Reorganizations', *Journal of Legal Studies* 20: 311–49.
- Balz, M. and H.-G. Landfermann (eds) (1999) *Die neuen Insolvenzgesetze*, Düsseldorf: IDW-Verlag.
- Banzhaf, J.F. (1965) 'Weighted Voting Doesn't Work: A Mathematical Analysis', *Rutgers Law Review* 19: 317–41.
- Barry, B. (1980a) 'Is it Better to be Powerful or Lucky? Part 1', *Political Studies* 28: 183–94.
- Barry, B. (1980b) 'Is it Better to be Powerful or Lucky? Part 2', *Political Studies* 28: 338–52.
- Bavetta, S. (1999) 'The Measurement of Specific Freedom', *Notizie di Politeia* 56: 107–13.
- Bavetta, S. (2004) 'Measuring Freedom of Choice: An Alternative View of a Recent Literature', *Social Choice and Welfare* 22: 29–48.
- Bebchuk, L.A. and H.F. Chang (1992) 'Bargaining and the Division of Value in Corporate Reorganization', *Journal of Law, Economics, and Organization* 8: 253–79.
- Benn, S.I. (1967) 'Power', in P. Edwards (ed.) *The Encyclopedia of Philosophy*, VI. London: MacMillan.
- Berlin, I. (1969) *Four Essays on Liberty*. Oxford: Oxford University Press.
- Bernheim, B.D., B. Peleg and M.D. Whinston (1987) 'Coalition-Proof Nash Equilibria I. Concepts', *Journal of Economic Theory* 42: 1–12.
- Bilal, S., P. Albuquerque and M.O. Hosli (2001) 'The Probability of Coalition Formation: Spatial Voting Power Indices', mimeo, European Centre for Development Policy Management, London; Ecole d'Ingénieurs, Geneva; University of Michigan, Ann Arbor.
- Braham, M. and F. Steffen (2002) 'Local Monotonicity of Voting Power: A Conceptual Analysis', mimeo, University of Hamburg.
- Brams, S.J. (1976) *Paradoxes in Politics*. New York: The Free Press.

- Brams, S.J. (1990) 'Negotiation Games'. London: Routledge.
- Brams, S.J. (1994) 'Voting Procedures', in R.J. Aumann and S. Hart (eds) *Handbook of Game Theory*. Amsterdam: Elsevier Science.
- Brams, S.J. and P.J. Affuso (1976) 'Power and Size: A New Paradox', *Theory and Decision* 7: 29–56.
- Brams, S.J., D.S. Felsenthal and Z. Maoz (1986) 'New Chairman Paradoxes', in A. Diekmann and P. Mitter (eds) *Paradoxical Effects of Social behavior: Essays in Honor of Anatol Rapoport*. Heidelberg: Physica-Verlag.
- Brams, S.J., D.S. Felsenthal and Z. Maoz (1988) 'Chairman Paradoxes Under Approval Voting', in G. Eberlein and H. Berghel (eds) *Theory and Decision: Essays in Honour of Werner Leinfellner*. Dordrecht: D. Reidel Publishing Company.
- Brams, S.J. and P.C. Fishburn (1983) *Approval Voting*. Stuttgart: Birkhäuser.
- Brams, S.J. and P.C. Fishburn (1995) 'When is Size a Liability? Bargaining Power in Minimal Winning Coalitions', *Journal of Theoretical Politics* 7: 301–16.
- Brams, S.J. and M. Lake (1978) 'Power and Satisfaction in a Representative Democracy', in P.C. Ordeshook (ed.) *Game Theory and Political Science*. New York: New York University Press.
- Brams, S.J. and F.C. Zagare (1977) 'Deception in Simple Voting Games', *Social Science Research* 6: 257–72.
- Brams, S.J. and F.C. Zagare (1981) 'Double Deception: Two Against One in Three-Person Games', *Theory and Decision* 13: 81–90.
- Carreras, F. and J. Freixas (2001) 'Weighted Games with Consensus: A Trading Characterization and a Power Analysis', mimeo, Polytechnic University Catalonia.
- Carter, I. (1992) 'Being Free When Being Forced to Choose', *Politics* 12: 38–9.
- Carter, I. (1999) *A Measure of Freedom*. Oxford: Oxford University Press.
- Carter, I. (2004) 'Choice, Freedom, and Freedom of Choice', *Social Choice and Welfare* 22: 61–81.
- Cohen, G.A. (1979) 'Capitalism, Freedom and the Proletariat', in A. Ryan (ed.) *The Idea of Freedom: Essays in Honour of Isaiah Berlin*. Oxford: Oxford University Press.
- Cohen, G.A. (1988) *History, Labour, and Freedom: Themes from Marx*. Oxford: Oxford University Press.
- Coleman, J.S. (1971) 'Control of Collectivities and the Power of a Collectivity to Act', in B. Lieberman (ed.) *Social Choice*. New York: Gordon and Breach.
- Colomer, J.M. and M.O. Hosli (2002) 'Decision-Making in the European Union: The Power of Political Parties', in P. Moser, G. Schneider and G. Kirchgässer (eds) *Decision Rules in the European Union*. Basingstoke: Macmillan.
- Conger, J.A. and R.N. Kanungo (1988) 'The Empowerment Process: Integrating Theory and Practice', *Academy of Management Review* 13: 471–82.
- Dahl, R.A. (1957) 'The Concept of Power', *Behavioral Science* 2: 201–15.
- De Sinopoli, F. (2000) 'Sophisticated Voting and Equilibrium Refinements under Plurality Rule', *Social Choice and Welfare* 17: 655–72.

- Deb, R. (1994) 'Waiver, Effectivity and Rights as Game Forms', *Economica* 61: 167–78.
- Deb, R. (2004) 'Rights as Alternative Game Forms', *Social Choice and Welfare* 22: 83–111.
- Deegan, J. and E.W. Packel (1978) 'A New Index for Voting Games', *International Journal of Game Theory* 7: 113–23.
- Dhillon, A. and B. Lockwood (2004) 'When are Plurality Rule Voting Games Dominance-Solvable', *Games and Economic Behavior* 46: 55–75.
- Dowding, K. (1991) *Rational Choice and Political Power*. Aldershot: Edward Elgar.
- Dowding, K. (1992) 'Choice: Its Increase and its Value', *British Journal of Political Science* 22: 301–14.
- Dowding, K. (1996) *Power*. Buckingham: Open University Press.
- Dowding, K. (2004) 'Social Choice and the Grammar of Rights and Freedoms', *Political Studies* 52: 144–61.
- Dowding, K. and M. van Hees (2003) 'The Construction of Rights', *American Political Science Review* 97: 281–93.
- Dworkin, G. (1970) 'Acting Freely', *Nous* 4: 367–83.
- Farquharson, R. (1969) *Theory of Voting*. New Haven: Yale University Press.
- Feddersen, T. (1993) 'Coalition-proof Nash Equilibria in a Model of Costly Voting under Plurality Rule', mimeo.
- Felsenthal, D.S. and M. Machover (1998) *The Measurement of Voting Power*. Cheltenham: Edward Elgar.
- Felsenthal, D.S. and M. Machover (2001) 'The Treaty of Nice and Qualified Majority Voting', *Social Choice and Welfare* 18: 431–64.
- Felsenthal, D.S., A. Rapoport and Z. Maoz (1988) 'Tacit Cooperation in Three-Alternative Non-cooperative Voting Games: A New Model of Sophisticated Behaviour Under the Plurality Procedure', *Electoral Studies* 7: 143–61.
- Fishburn, P.C. (1982) 'Monotonicity Paradoxes in the Theory of Elections', *Discrete Applied Mathematics* 4: 119–34.
- Fishburn, P.C. and W.V. Gehrlein (1980) 'The Paradox of Voting: Effects of Individual Indifference and Intransitivity', *Journal of Public Economics* 14: 83–94.
- Fleurbaey, M. and W. Gaertner (1996) 'Admissibility and Feasibility in Game Forms', *Analyse & Kritik* 18: 54–66.
- Fleurbaey, M. and M. van Hees (2000) 'On Rights in Game Forms', *Synthese* 123: 295–326.
- Franks, J.R., K.G. Nyborg and W.N. Torous (1996) 'A Comparison of US, UK, and German Insolvency Codes', *Financial Management* 25: 86–101.
- Frohlich, N. (1975) 'The Instability of Minimum Winning Coalitions', *American Political Science Review* 69: 943–6.
- Gaertner, W., P.K. Pattanaik and K. Suzumura (1992) 'Individual Rights Revisited', *Economica* 59: 161–77.
- Gaertner, W. and Y. Xu (2004) 'Procedural Choice', *Economic Theory* 24: 335–249.

- Gal-Or, E. and R. Amit (1998) 'Does Empowerment Lead to Higher Quality and Profitability?', *Journal of Economic Behavior and Organization* 36: 411–31.
- Gärdenfors, P. (1981) 'Rights, Games and Social Choice', *Noûs* 15: 341–56.
- Garrett, G. and G. Tsebelis (1996) 'An Institutional Critique of Intergovernmentalism', *International Organization* 50: 269–99.
- Garrett, G. and G. Tsebelis (1999a) 'More Reasons to Resist the Temptation of Power Indices in the European Union', *Journal of Theoretical Politics* 11: 331–8.
- Garrett, G. and G. Tsebelis (1999b) 'Why Resist the Temptation of Power Indices in the European Union?', *Journal of Theoretical Politics* 11: 291–308.
- Garrett, G. and G. Tsebelis (2001) 'Even More Reasons to Resist the Temptation of Power Indices in the European Union', *Journal of Theoretical Politics* 13: 99–105.
- Gehrlein, W.V. and P.C. Fishburn (1976) 'Condorcet's Paradox and Anonymous Preference Profiles', *Public Choice* 26: 1–18.
- Gibbard, A. (1973) 'Manipulation of Voting Schemes: A General Result', *Econometrica* 41: 587–601.
- Goldman, A.I. (1970) *A Theory of Human Action*. Englewood Cliffs: Prentice-Hall.
- Goldman, A.I. (1972) 'Toward a Theory of Social Power', *Philosophical Studies* 23: 221–68.
- Goldman, A.I. (1974) 'On the Measurement of Power', *Journal of Philosophy* 71: 231–52.
- Harré, R. (1970) 'Powers', *British Journal of the Philosophy of Science* 21: 81–101.
- Harsanyi, J.C. (1962a) 'Measurement of Social Power in n -Person Reciprocal Power Situations', *Behavioral Science* 7: 81–91.
- Harsanyi, J.C. (1962b) 'Measurement of Social Power, Opportunity Costs, and the Theory of Two-Person Bargaining Games', *Behavioral Science* 7: 67–80.
- Harsanyi, J.C. and R. Selten (1988) *A General Theory of Equilibrium Selection in Games*. Cambridge, MA: MIT Press.
- Hart, J. (1976) 'Three Approaches to the Measurement of Power in International Relations', *International Organization* 30: 293–305.
- Haunsperger, D.B. and D. Melville (1996) 'Voting Power When Using Preference Ballots', *Social Choice and Welfare* 13: 457–65.
- Henry, S.M. (ed.) (1999) *The Portable Bankruptcy Code and Rules: 1999*, Chicago: American Bar Association.
- Hoede, C. and R.R. Bakker (1982) 'A Theory of Decisional Power', *Journal of Mathematical Sociology* 8: 309–22.
- Hohfeld, W.N. (1923/1978) *Fundamental Legal Conceptions as Applied in Judicial and Legal Reason*. Westport: Greenwood Press.
- Holler, M.J. (1982) 'Forming Coalitions and Measuring Voting Power', *Political Studies* 30: 262–71.
- Holler, M.J. (1998) 'Two Stories, One Power Index', *Journal of Theoretical Politics* 10: 179–90.

- Holler, M.J. and E.W. Packel (1983) 'Power, Luck and the Right Index', *Zeitschrift für Nationalökonomie* 43: 21–9.
- Holler, M.J. and M. Widgrén (1999) 'The Value of a Coalition is Power', *Homo Oeconomicus* 15: 497–511.
- Hosli, M.O. (1997) 'Voting Strength in the European Parliament: The Influence of National and Partisan Actors', *European Journal of Political Research* 31: 351–66.
- Hosli, M.O. (2002) 'Preferences and Power in the European Union', *Homo Oeconomicus* 19: 311–26.
- Hume, D. (1739 [1951]) *A Treatise of Human Nature*. Oxford: Oxford University Press.
- Instruments, H.s.S. (ed.) (1991) *Insolvency Rules 1986*, London: Butterworths.
- Johnston, R. (1978) 'On the Measurement of Power: Some Reactions to Laver', *Environment and Planning* 10: 569–77.
- Kordana, K.A. and E.A. Posner (1999) 'A Positive Theory of Chapter 11', *New York University Law Review* 74: 161–234.
- Kramer, M.H. (2003a) 'On Being Free Without Having Any Choices', *Journal of Theoretical Politics* 15: 445–71.
- Kramer, M.H. (2003b) *The Quality of Freedom*. Oxford: Oxford University Press.
- Kristjánsson, K. (1996) *Social Freedom*. Cambridge: Cambridge University Press.
- Lane, J.-E. and R. Maeland (1995) 'Voting Power Under the EU Constitution', *Journal of Theoretical Politics* 7: 223–30.
- Laruelle, A. and F. Valenciano (2004a) 'A Critical Reappraisal of Some Voting Power Paradoxes', *Public Choice*.
- Laruelle, A. and F. Valenciano (2004b) 'Inequality in Voting Power', *Social Choice and Welfare* 22: 413–31.
- Laver, M. and N. Schofield (1990) *Multiparty Government: The Politics of Coalition in Europe*. Oxford: Oxford University Press.
- Leiserson, M. (1968) 'Factions and Coalitions in One-Party Japan: An Interpretation Based on the Theory of Games', *American Political Science Review* 62: 770–87.
- Lepelley, D. and M. Martin (2001) 'Condorcet's Paradox for Weak Preference Orderings', *European Journal of Political Economy* 17: 163–77.
- Levínský, R. and P. Silársky (2001) 'Global Monotonicity of Values of Cooperative Games: An Argument Supporting the Explanatory Power of Shapley's Approach', in M.J. Holler and G. Owen (eds) *Power Indices and Coalition Formation*. Dordrecht: Kluwer.
- Lindner, I. (2003) *Power Measures in Large Weighted Voting Games: Asymptotic Properties and Numerical Methods*, mimeo: University of Hamburg.
- List, C. (2004) 'The Impossibility of a Paretian Republican', *Economics and Philosophy* 20: 1–23.
- LoPucki, L.M. and G.G. Triantis (1994) 'A Systems Approach to Comparing U.S. and Canadian Reorganization of Financially Distressed Companies', *Harvard International Law Journal* 35: 267–343.

- MacCallum, G.C. (1967) 'Negative and Positive Freedom', *Philosophical Review* 76: 312–34.
- March, J.G. (1955) 'An Introduction to the Theory and Measurement of Power', *American Political Science Review* 49: 431–51.
- McKelvey, R.D. and R.G. Niemi (1978) 'A Multistage Game Representation of Sophisticated Voting for Binary Procedures', *Journal of Economic Theory* 18: 1–22.
- Messner, M. and M.K. Polborn (2002) 'Robust Political Equilibria under Plurality and Runoff Rule', mimeo, University of Western Ontario, Canada.
- Miller, D. (1983) 'Constraints on Freedom', *Ethics* 94: 66–86.
- Miller, N.R. (1982) 'Power in Game Forms', in M.J. Holler (ed.) *Power, Voting, and Voting Power*. Würzburg: Physica Verlag.
- Morriss, P. (1987/2002) *Power: A Philosophical Analysis*. Manchester: Manchester University Press.
- Moulin, H. (1983) *The Strategy of Social Choice*. Amsterdam: North-Holland.
- Nagel, J. (1975) *The Descriptive Analysis of Power*. New Haven: Yale University Press.
- Napel, S. and M. Widgrén (2001) 'Inferior Players in Simple Games', *International Journal of Game Theory* 30: 209–20.
- Napel, S. and M. Widgrén (2002) 'The Power of a Spatially Inferior Player', *Homo Oeconomicus* 19: 327–43.
- Napel, S. and M. Widgrén (2004) 'Power Measurement as Sensitivity Analysis – A Unified Approach', *Journal of Theoretical Politics* 16: 517–538.
- Niemi, R.G., B.H. Bjurulf and G. Blewis (1983) 'The Power of the Chairman', *Public Choice* 40: 293–305.
- Niemi, R.G. and A.Q. Frank (1982) 'Sophisticated Voting under the Plurality Procedure', in P.C. Ordeshook and K.A. Shepsle (eds) *Political Equilibrium*. Boston: Kluwer.
- Norberg, S.F. (1998) 'Debtor Incentives, Agency Costs, and Voting Theory in Chapter 11', *Kansas Law Review* 46: 507–50.
- Nozick, R. (1974) *Anarchy, State, and Utopia*. New York: Basic Books.
- Nurmi, H. (1999) *Voting Paradoxes and How to Deal with Them*. Berlin: Springer.
- Oppenheim, F.E. (1960) 'Degrees of Power and Freedom', *American Political Science Review* 54: 437–46.
- Oppenheim, F.E. (1961) *Dimensions of Freedom*. New York: St. Martins Press.
- Oppenheim, F.E. (1981) *Political Concepts: A Reconstruction*. Chicago: University of Chicago Press.
- Ordeshook, P.C. (1992) *A Political Theory Primer*. London: Routledge.
- Pattanaik, P. and Y. Xu (1990) 'On Ranking Opportunity Sets in Terms of Freedom of Choice', *Recherches Economiques de Louvain* 56: 383–90.
- Pattanaik, P.K. and K. Suzumura (1996) 'Individual Rights and Social Evaluation: A Conceptual Framework', *Oxford Economic Papers* 48: 194–212.
- Pattanaik, P.K. and Y. Xu (1998) 'On Preference and Freedom', *Theory and Decision* 44: 173–98.

- Peleg, B. (1984) *Game Theoretic Analysis of Voting in Committees*. Cambridge: Cambridge University Press.
- Peleg, B. (1992) 'Voting by Count and Account', in R. Selten (ed.) *Rational Interaction: Essays in Honor of John C Harsanyi*. Berlin: Springer.
- Peleg, B. (1998) 'Effectivity Functions, Game Forms, Games and Rights', *Social Choice and Welfare* 15: 67–80.
- Penrose, L.S. (1946) 'The Elementary Statistics of Majority Voting', *Journal of the Royal Statistical Society* 109: 53–7.
- Pettit, P. (1996) 'Freedom as Antipower', *Ethics* 106: 576–604.
- Pettit, P. (1997) *Republicanism: A Theory of Freedom and Government*. Oxford: Oxford University Press.
- Pfeffer, J. (1992) *Managing with Power: Politics and Influence in Organizations*. Boston MA: Harvard Business School Press.
- Rae, D.W. (1969) 'Decision-Rules and Individual Values in Constitutional Choice', *American Political Science Review* 63: 40–63.
- Rescher, N. (1975) *A Theory of Possibility*. Oxford: Oxford University Press.
- Riker, W.H. (1962) *The Theory of Political Coalitions*. New Haven: Yale University Press.
- Rosenbaum, E. (2000) 'On Measuring Freedom', *Journal of Theoretical Politics* 12: 205–27.
- Roth, A.E. (1995) 'Bargaining Experiments', in J.H. Kagel and A.E. Roth (eds) *The Handbook of Experimental Economics*. Princeton: Princeton University Press.
- Rusinowska, A. and H. de Swart (2002) 'On Some Properties of the Hoede-Bakker Index', Mimeo, Tilburg University.
- Rusinowska, A. and H. de Swart (2003) 'Generalizing and Modifying the Hoede-Bakker Index', Mimeo, Tilburg University.
- Ryle, G. (1949) *The Concept of Mind*. New York: Barnes and Nobel.
- Samuelson, P.A. (1947) *Foundations of Economic Analysis*. Cambridge, MA: Harvard University Press.
- Selten, R. (1975) 'Re-examination of the Perfectness Concept for Equilibrium', *International Journal of Game Theory* 4: 25–55.
- Sen, A.K. (1970a) *Collective Choice and Social Welfare*. Edinburgh: Oliver and Boyd.
- Sen, A.K. (1970b) 'The Impossibility of a Paretian Liberal', *Journal of Political Economy* 78: 152–7.
- Sen, A.K. (1979) 'Personal Utilities and Public Judgements: or What's Wrong with Welfare Economics', *Economic Journal* 89: 537–58.
- Sen, A.K. (1984) *Resources, Values and Development*. Cambridge, MA: Harvard University Press.
- Shapley, L.S. (1953) 'A Value for n-Person Games', in H. Kuhn and A.W. Tucker (eds) *Contributions to the Theory of Games*. Princeton: Princeton University Press.
- Shapley, L.S. (1962) 'Simple Games: An Outline of the Descriptive Theory', *Behavioral Science* 7: 59–66.

- Shapley, L.S. and M. Shubik (1954) 'A Method for Evaluating the Distribution of Power in a Committee System', *American Political Science Review* 48: 787–92.
- Simon, H.A. (1953) 'Notes on the Observation and Measurement of Power', *Journal of Politics* 15: 500–16.
- Spreitzer, G.M. (1995) 'Psychological Empowerment in the Workplace: Dimensions, Measurement, and Validation', *Academy of Management Journal* 38: 1442–65.
- Spreitzer, G.M. (1996) 'Social Structural Characteristics of Psychological Empowerment', *Academy of Management Journal* 39: 483–504.
- Steiner, H. (1983) 'How Free: Computing Personal Liberty', in A.P. Griffiths (ed.) *Of Liberty*. Cambridge: Cambridge University Press.
- Steiner, H. (1994) *An Essay on Rights*. Oxford: Blackwell.
- Steiner, H. (2001) 'Freedom and Bivalence', in I. Carter and M. Ricciardi (eds) *Freedom, Power and Political Morality*. Basingstoke: Palgrave.
- Steunenberg, B., D. Schmidtchen and C. Koboldt (1999) 'Strategic Power in the European Union: Evaluating the Distribution of Power in Policy Games', *Journal of Theoretical Politics* 11: 339–66.
- Straffin, P.D., M.D. Davis and S.J. Brams (1982) 'Power and Satisfaction in an Ideologically Divided Body', in M.J. Holler (ed.) *Power, Voting, and Voting Power*. Würzburg: Physica Verlag.
- Sugden, R. (1978) 'Social Choice and Individual Liberty', in M.J. Artis and A.R. Nobay (eds) *Contemporary Economic Analysis*. London: Croom Helm.
- Sugden, R. (1985) 'Liberty, Preference, and Choice', *Economics and Philosophy* 1: 213–29.
- Sugden, R. (1998) 'The Metric of Opportunity', *Economics and Philosophy* 14: 307–37.
- Taylor, A.D. (1995) *Mathematics and Politics*. New York: Springer.
- Taylor, A.D. and W.S. Zwicker (1999) *Simple Games: Desirability Relations, Trading, Pseudoweightings*. Princeton: Princeton University Press.
- Taylor, C. (1979) 'What's Wrong with Negative Liberty', in A. Ryan (ed.) *The Idea of Freedom: Essays in Honour of Isaiah Berlin*. Oxford: Oxford University Press.
- Tsebelis, G. and G. Garrett (1996) 'Agenda Setting, Power Indices, and Decision Making in the European Union', *International Review of Law and Economics* 16: 345–61.
- Tsetlin, I., M. Regenwetter and B. Grofman (2003) 'The Impartial Culture Maximizes the Probability of Majority Cycles', *Social Choice and Welfare* 21: 387–98.
- van Deemen, A.M.A. (1997) *Coalition Formation and Social Choice*. Dordrecht: Kluwer.
- van Deemen, A.M.A. (1999) 'The Probability of the Paradox of Voting for Weak Preference Orderings', *Social Choice and Welfare* 16: 171–82.
- van Hees, M. (1995) *Rights and Decisions*. Dordrecht: Kluwer.
- van Hees, M. (1998) 'On the Analysis of Negative Freedom', *Theory and Decision* 45: 175–97.
- van Hees, M. (2000) *Legal Reductionism and Freedom*. Dordrecht: Kluwer.
- van Hees, M. and M. Wissenburg (1999) 'Freedom and Opportunity', *Political Studies* 48: 67–82.

- Vannucci, S. (2002) 'Effectivity Functions, Opportunity Rankings, and Generalized Game Desirability Relations', *Homo Oeconomicus* 19: 451–67.
- von Neumann, J. and O. Morgenstern (1944) *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.
- White, M.J. (1989) 'The Corporate Bankruptcy Decision', *Journal of Economic Perspectives* 3: 129–51.
- Wrong, D.H. (1979) *Power: Its Forms, Bases and Uses*. Oxford: Basil Blackwell.
- Zagare, F.C. (1979) 'The Geneva Conference of 1954', *International Studies Quarterly* 23: 390–411.