UNIVERZITA KARLOVA V PRAZE

FAKULTA SOCIÁLNÍCH VĚD

Institut politologických studií

Michal Parízek

The Organizational-Cybernetic Approach to International Institutions

Disertační práce

Praha 2013

Autor práce: Michal Parízek, M.Sc.

Školitel: Doc. PhDr. Jan Karlas, M.A., Ph.D.

Rok obhajoby: 2013

Bibliografický záznam

Abstrakt

Jedním z klíčových úkolů institucionálního výzkumu v mezinárodních vztazích by mělo být zkoumání toho, jak dosáhnout pomocí změn designu mezinárodních institucí jejich lepšího fungování. Institucionální design je přitom třeba chápat jako jediný faktor, který má na fungování institucí přímý kauzální vliv, a který lze zároveň vědomě manipulovat. Cílem této disertační práce je posoudit, do jaké míry lze vůbec tohoto úkolu dosáhnout a, pokud ano, za jakých podmínek. Za tímto účelem představuji v disertaci nový konceptuálně-teoretický rámec pro studium mezinárodních institucí postavený na poznatcích z oboru organizační kybernetiky.

Na základě tohoto rámce definuji koncept kapacity přenosu informací jako vhodné závislé proměnné pro výzkum vlivu designu institucí na jejich fungování. Následně představuji tzv. model životaschopného systému, tedy organizačně-kybernetický model, který na obecné úrovni identifikuje klíčové informační kanály, které jakékoliv schéma vládnutí (governance) musí mít k dispozici. Platnost mého přístupu ilustruji ve dvou empirických kapitolách týkajících se, zaprvé, fungování Rady Evropské unie po východním rozšíření a, zadruhé, personální politiky několika významných mezivládních organizací.

S pomocí zde představovaného konceptuálně-teoretického rámce jsme schopni problém designu a fungování mezinárodních institucí zachytit způsobem, který je hluboce zakořeněn v existujících teoriích mezinárodních vztahů, ale který do nich zároveň přináší četné nové poznatky a analytické nástroje.

Abstract

One of the key tasks of the institutionalist research in international relations should be to explore how the functioning of international institutions can be enhanced through alterations of their design, where design is understood as the only factor that can have a direct causal effect on the institutions' functioning and that, at the same time, is at least in principle amenable to conscious manipulation. The task of this dissertation is

to elaborate on whether this can be at all done and, if yes, how exactly and under what conditions. To deal with these problems, I present what is best labelled as the organizational-cybernetic approach, a theoretical and conceptual framework based on insights from the field of organizational cybernetics.

Building on the organizational-cybernetic framework, I outline the concept of information transmission capacity as the right dependent variable for the study of how design impacts on functioning of institutions. Subsequently I present the viable system model, an organizational-cybernetic model that identifies on a general level all the key information channels necessary for viability of any governance scheme. Using this approach, we can proceed in the direction of systematic investigation of the effects of design on international institutions' capacity. I illustrate validity of the presented theoretical and conceptual framework in two empirical chapters, one on decision-making in the Council of the European Union after the eastern enlargement, and one on the politics of staffing in several major inter-governmental organizations.

With the help of the conceptual and theoretical framework I develop in the dissertation, we are able to approach the problem of international institutions capacity and design from a perspective that on the one hand is deeply rooted in the existing theories of international relations, but that on the other hand brings in a number of new insights and analytical tools from outside of the discipline.

Klíčová slova

Mezinárodní instituce, institucionální design, institucionální kapacita, organizační kybernetika

Keywords

International institutions, institutional design, institutional capacity, organizational cybernetics

Rozsah práce / length

71 353 slov / words

1. Prohlašuji, že jsem předkládanou práci zpracoval samostatně a použil jen uvedené prameny a literaturu.

2. Prohlašuji, že práce nebyla využita k získání jiného titulu.

3. Souhlasím s tím, aby práce byla zpřístupněna pro studijní a výzkumné účely.

V Praze, 27. dubna 2013.

Michal Parízek

Acknowledgements

Contents

Chapter 1

Introduction

If there is any response to most global challenges of today's, it lies in well functioning global governance bodies, and in particular in international institutions the states create to address them. Seen from a slightly normatively laden perspective, one of the key tasks of the institutionalist research in international relations (IR) should be to explore how these bodies can be made work better. We should be able to provide insights into how functioning of the international institutions can be enhanced through alterations of their design, where design is understood as the only factor that can have a direct causal effect on their functioning and at the same time is at least in principle amenable to conscious manipulation.

The task of this dissertation is to elaborate on whether this can be at all done and, if yes, how exactly and under what conditions. I answer the former question of *whether* in the affirmative, but with a number of important qualification that significantly limit the scope of the design tools that it is worth considering and of the tasks we may have the ambition to achieve through alterations in design. For the latter question of *how exactly*, I present a detailed discussion in the course of several chapters of the dissertation where I argue for

adoption of what is best labelled as the organizational-cybernetic approach, a theoretical and conceptual framework for the study of international institutions' design and functioning based on the insights from the field of organizational cybernetics.

With the help of this framework, I argue, we are able to approach the problem of design for better functioning of international institutions from a perspective that on the one hand is deeply rooted in the existing IR theories, but on the other hand brings in a number of new tools and insights from outside of the discipline. As a matter of fact, I will argue that the combination of some of the key insights made within the discipline of IR with those brought in from the field of organizational cybernetics forms a synthetic framework, one able to productively accommodate the interests of the institutionalist research paradigm in the design and functioning of international institutions but also the more sceptical views on the role of international institutions held especially by realists and neorealists. In other words, the framework enables us to deal in detail with how international institutions work and could be made work better, while taking very seriously the power- and interest-driven nature of international politics.

This theoretical argumentation ultimately has a simple purpose: to generate a solid framework that motivates interesting empirical questions and from which testable hypotheses about these questions can be derived. There are two key set of conditions for the possibility of success of any such new theoretical and conceptual framework, and these are also the criteria for judging it *a priori*. First, the framework should be sufficiently generally applicable to be interesting, while maintaining internal coherence. Second, it needs to secure that empirical analyses based on the framework can be conducted, it must be empirical-research-friendly. Achieving that *both* these criteria are met is, obviously, not an easy task, as they often go against each other. The task of parts I and II of the dissertation is to convince the reader

that the first criterion has been met. Part III presents two empirical applications of the framework and thus tries to demonstrate that also the second condition for its success is present.

If the argumentation I present in the dissertation is correct, a new way of how we conduct the analysis of international institutions' design opens up, and the list of specific design features we should focus on in empirical research changes. For example, in many existing international organizations, discussions are being held about the alterations of such design features as the voting procedures or the rules for the distribution of votes across the member states. While obviously important from the perspective of the actors themselves, within the framework I develop here such matters appear as secondary at best. In fact, if implemented, design alterations focusing on such factors may easily produce unintended consequences that will work for the detriment of the organizations as such. They may create costs an order of magnitude higher than the immediate benefits they may bring. In other words, if the arguments I develop in the dissertation are right, important implications for how international institutions ought to be designed follow.

In the rest of this introductory chapter I discuss in more detail the research problem I deal with in the dissertation, I present in as-concise-as-possible a manner the argumentation I develop in the chapters to follow, and I outline the structure of the dissertation and how the individual chapters fit together to form a coherent whole.

1.1 The research problem: delimitation and its relation to the existing research

The underlying problem of the entire dissertation can be summarized in a single question: *how do we design better international institutions?* More specifically, how can international institutions be designed in a way that makes them better perform their key function, namely enabling cooperation among states? In principle, this question boils down to three operational questions that I directly answer in the dissertation, and that jointly provide the answer to the underlying question. These are

1. What does it mean – in causal terms – that an institution performs its function?

2. Given the answer to the previous question, how do we determine whether an institution does have the ability to perform its functions?

3. What design adjustments are likely to increase the institutions' ability to perform its functions?

My target in the dissertation, then, is to develop a set of general theoretical and conceptual answers to these questions, and show how these answers can guide empirical research. The entire endeavour of the dissertation is obviously substantively motivated by the presupposition according to which numerous existing international institutions actually do not work as they should. In such areas as international trade liberalization, international environmental protection, or international financial regulation, what we observe today are often deadlocked negotiations and a principle inability of the key actors to agree on mutually beneficial cooperative schemes. Similarly, there is a number of international institutions that fail to deliver on their tasks, be it in the area of security (the United Nations Security

Council) or for example in economic and social development (World Bank, United Nations Development Programme). The lack of truly cooperative outcomes in international politics puts the suitability of many aspects of the institutional structure of contemporary global governance into question. The desired global targets are notoriously not being met, and the international institutions supposed to contribute to them necessarily take part of the blame. The problem of how they can be better (re-)designed, and thus how to improve the prospects of mutually beneficial cooperation, lies at the core of numerous current lay as well as academic debates.

To address the problem, I build on a number of existing literatures. Their choice also delimits more precisely the approach I take to addressing this problem, as well as the assumptions I work with. These literatures can be broadly categorized into three groups. First, I build heavily on the existing theoretical literature on the international institutions, in particular on the debates between neorealist and neoliberals in early 1990s (e.g. Baldwin, 1993) and on the more recent literature on the theory of international governance (Reinicke, 1998; Rosenau, 2000). While the former provides important insights into the nature of international politics, and the constraints it imposes on the functioning of international institutions, the latter points out the enormous complexity of the problem the actors face. I take these two strands of large-scale theorizing as my points of departure, and my work is guided by an effort to synthesize their key insights, even though they are sometimes portrayed as representing opposite views (cf. Held & McGrew, 2002, part III).

Second, I make extensive use of the organizational-cybernetic literature (Beer, 1979), building on the more general earlier information, cybernetic, and complex systems theory literature (Shannon, 1948; Ashby, 1956; Gell-Mann, 2002, respectively). Although sometimes of an earlier date, the major insights from these works have so far not been systematically

considered in the IR research – an omission I attempt to remedy, and one that in fact received attention in some of the most recent IR theorizing as well (Albert *et al.*, 2010). I use this literature to develop the synthesis between the two major IR literatures on international cooperation mentioned above. I bring this literature into my discussion because it fruitfully transcends the contradictions between different IR-theoretical perspectives, and thus enables us to study empirically important phenomena within a framework that takes both the more power- and interest-oriented neorealist and neoliberal theories *and* the complexity-oriented governance approach seriously.

Third, I naturally engage thoroughly the existing literatures that deal directly with the problem of design of international institutions or with their functioning. The research on design entails first and foremost the rational design literature (Koremenos *et al.*, 2001), i.e. literature that explains why states design institutions the way they do. Prominent within this field is the literature on delegation of power to international organizations (Hawkins *et al.*, 2006). On the other side of the design ⇒ functioning relationship, the research on functioning of institutions then revolves around several concepts that are supposed to capture what it actually means that institutions function (cf. Gutner & Thompson, 2010). I combine this literature with that from the information theory and cybernetics to develop a concept of institutional capacity that, I argue, tells us how well international institutions do what we want them to do better than these existing alternatives.

Each of the two major theoretical perspectives on international institutions in the current rationalist IR – the more standard neoliberal and neorealist power- and interest-oriented account on the one hand, and the complexity- and problem-solving-oriented governance approach on the other hand, provide essential insights into important empirical and theoretical problems. Precisely because each has a valid point, none can actually work alone, disregard-

ing the other. A synthetic approach is needed.

1.2 The argument: the organizational-cybernetic framework for the design and capacity of international institutions

How can we make the institutions work? What design changes can we propose? The key argument I will keep returning to throughout the dissertation is that we cannot do much. More precisely, we cannot hope to cover with design a large range of issues, there are many goals we cannot achieve. What we can achieve, and what we need to focus on, is that a limited set of key functions the institutions need to perform are actually performed. I try to show that the real problem with design does not necessarily lie in the design as such, but rather first and foremost in a careful diagnosis of those key functions. Once we know where the possible shortages in their provision are, the design solutions may not be as difficult.

I develop my argumentation on the basis of two broad classes of assumptions and theoretical approaches that stem from them. First, I take very seriously the power- and interest-driven nature of international politics, and the necessary constraints these realities impose on the role international institutions can play. Building on the results of the debate between neorealists and neoliberals (the neo-neo debate) I re-iterate the neoliberal argument that 1) cooperation is possible and 2) that international institutions can help the states achieve it[1].

More specifically, I argue that the concrete mechanisms through which international

---

[1]Following Keohane (Keohane, 1984, p. 51-52), I understand that "intergovernmental cooperation takes place when the policies actually followed by one government are regarded by its partners as facilitating realization of their own objectives, as the result of a process of policy coordination".

institutions enhance cooperation ultimately rest in their ability to secure enhanced transmission of information. In other words, if we take the actors' powers and interests as the key determinants of international outcomes, the way in which institutions causally impact on cooperation lies in better transmission of the information the actors need for cooperation. I later elaborate on why information transmission is so important for cooperation and I try to demonstrate that, within a strict rationalist framework, enhanced information transmission is the most dominant way in which international institutions can improve the prospects for cooperation. To be sure, an argument that focuses only on information transmission is very reductionist, it excludes from the analysis all other factors. Yet, such a minimalist position has the advantage that it takes very seriously the power- and interest-driven view of international politics. It is synthetic in the sense that it is consistent with both neorealist and neoliberal positions[2]. The fact that such an approach builds on the common ground of these two key rationalist theories, and is in this sense very conservative, arguably more than justifies the narrow focus on information transmission. Anything beyond this focus could easily be attacked for inconsistency with some of the major theoretical findings in the discipline.

Second, beyond power and interests as important factors for cooperation I take very seriously also the extremely complex nature of international cooperation, that is the fact that there are always numerous relevant actors involved, that the links between them are both numerous and multi-dimensional, and that the actors face in their decisions high levels of uncertainty. As a result, if we are interested in the design of international institutions, we cannot *simply* reconstruct the situation as the actors perceive it, and create (design) new

---

[2]To be sure, there are still scholars who claim to be realists or neorealists, such as Mearsheimer, who probably would not share this view (Mearsheimer, 1994). It has been demonstrated, though, that such a position suffers from significant inconsistencies Keohane & Martin, 1995.

institutions that will be able to alter the actors' incentives and consequently behaviour in the desired direction. The institutions are political systems and their systemic nature means that they are too complex animals for that. Instead, we need to limit our analysis to some small set of general functions the institutions *need* to perform as political systems, if they are to be at all viable, and try to devise simple institutional solutions for these.

Putting these two classes of assumptions and theoretical propositions together, I arrive at what is best labelled as a cybernetic approach to international institutions, where the adjective cybernetic refers to the multi-disciplinary field of cybernetics, the science of communication and control. Cybernetics, and its subfield I focus most, the organizational cybernetics, is a science dealing with the problem of information flows within complex regulatory systems. Hence, it is by its very nature and by its key interests very suitable for our purposes. Use of cybernetic approaches in politics is not new (Deutsch, 1963) and it has received renewed attention in recent years (cf. Albert *et al.*, 2010). Yet, the important insights from the field of organizational cybernetics, developed primarily by British cybernetician Stafford Beer (Beer, 1972), have so far not been introduced to the analysis of international institutions.

The most important contribution of organizational cybernetics is the so-called viable system model (VSM) developed by Beer, a framework that identifies formally all the key information channels in any governance system, as well as the amount of information these channels need to be able to transmit (Beer, 1972, 1979). Following the original information theory formulations (Shannon, 1948), I define a new concept of *capacity* as the maximum amount of political information any institution (or its component) is able to transmit. In other words, the concept of institutional capacity captures how much information a given institutional framework is at all able to transmit among the cooperation actors and, con-

versely, what are its limits in enhancing the prospects for cooperation. I argue that it is this concept of capacity that serves best the purpose of analysis of how design of institutions impacts on their functioning. Capacity is the right dependent variable for the kind of causal analysis I argue for in the dissertation.

The key causal relationship the dissertation addresses, then, can be summarized as

$$\text{institutional design} \Rightarrow \text{capacity},$$

where capacity is conceptualized (so far informally) as the maximum amount of information an institution of a particular design can transmit. For a given set of actors' preferences and power relations, design of the institution determines its capacity, as I define it (in chapter 5). The notion of design refers to the set of explicit formal features of which the institutional arrangements consists[3], the rules defining the interaction (North, 1990; Goodin, 1998; cf. Koremenos *et al.*, 2001, 762, Karlas, 2008).

In some limited sense the fact that the design of institutions determines their capacity may be seen as the overall hypothesis of the dissertation, but it is not a truly empirically testable proposition. This is because, for a given set of preferences and power relations, capacity is conceptualized in a way that makes the statement true by definition. I do not present a single hypothesis in the strict sense. What the framework developed within this

---

[3]Note that by *formal* I do not necessarily mean written. A rule can be considered formal if it is agreed in some standard setup that secures high level of shared understanding about both its content and its binding nature. A design feature may be stated verbally and still be considered formal, although in the context of international institutions formality will in practice usually imply written form. The definition of design then obviously considers only explicit rules, as implicit rules cannot be formal. The notion of design will then apply to a broad range of features, from those studied by Koremenos et al (Koremenos *et al.*, 2001), to individual components of legal treaties, to explicit formal agreements by actors, e.g. on how they will deal with problems of specific kind.

dissertation enables us, however, is to develop actually empirically testable hypotheses about how specific design features impact on the institutions' capacity. Two illustrations of such empirical applications are presented in chapters 8 and 9 where I do have specific hypotheses and where I also test them. From the point of view of the dissertation as a whole, these are illustrative applications of the framework rather than its comprehensive tests.

I should note that when I refer to enhanced prospects for cooperation I mean it exactly, that means I do not argue that in any specific case concrete design changes will lead to, or prevent, cooperation. There is no direct link between the design of international institutions and cooperation. What design can in principle do is to provide favourable conditions for cooperation. It can provide conditions that will be favourable for cooperation under varying constellations of power and interest across the actors. Yet whether cooperation actually does or does not take place is within any rationalist framework ultimately always given precisely by the actors' powers and interests, that jointly give rise to their utility calculations, and only then by the institutional framework within which they interact.

If we wanted to understand also the next step from the institutions' capacity to the actual levels of cooperation, we would need to include in the analysis the distributions of power and interest prevailing in each specific empirical situation of our interest. The relationship would then be correspondingly more complex:

institutional design $\Rightarrow$ institutional capacity

$$\left. \text{distribution of interests} \right\} \Rightarrow \text{cooperation.}$$

distribution of powers

Given the theoretical orientation of the dissertation I do not present here any discussion of my methodological approach. In the empirical chapters, I use game theoretical and statistical

14

analysis, and I discuss possible methodological issues in more detail there. In general, I adopt a rationalist approach broadly defined, so I assume that actors in international politics have complete and transitive preferences over important outcomes (see e.g. Shepsle, 2010, ch. 2). Yet, I by no means assume that they act under the conditions of certainty and that they possess complete information; on the contrary, in many situations the actors face extreme information shortages (or overloads) and information asymmetries. The world is complex, and the degree of the actors' ability to effectively collect and process the relevant information dramatically influences the outcomes of their interaction (Rasmusen, 2006). As a matter of fact, this problem of how the actors work with information lies at the very core of my approach, as it does in much of the contemporary rationalist research (in IR, the most prominent being that by Fearon, 1995) and as it does in the organizational cybernetic framework that I adopt. This broad rationalist framework therefore also delimits the scope conditions of my argumentation.

To summarize, I present a broadly applicable theoretical framework for the study of how design impacts on functioning of the international institutions. The framework is synthetic in the sense that it takes seriously the key insights from approaches that are usually not considered jointly, the neo-neo power- and interest-oriented framework and the problem- and complexity-oriented governance approach. I arrive at the synthesis by embedding these two approaches within a unified framework, that is partly newly constructed and partly imported from the organizational cybernetics.

1.3 Development of the argument and the structure of the dissertation

A lot of what the dissertation is about has already been outlined, but it will be helpful at this point to briefly discuss how the dissertation is structured and how the individual chapters and parts link together. The dissertation is composed of three parts where each provides a partly independent argumentation but that all together create a coherent whole.

Part I outlines the key building blocks of the framework I develop in the dissertation. In chapter 2 I present the main findings of the neo-neo debate in IR and argue that institutions increase the prospects of international cooperation by enhancing the transmission of information among the actors.

Chapter 3 present the complexity- and problem-oriented governance approach to international institutions. It argues that there are severe constraints that the complexity of the matters imposes on our ability to predict how particular design changes impact on functioning of the institutions. Consequently, we need to adopt a more systemic approach and focus only on a limited set of vital functions the international institutions as political systems need to perform. We should not try to fine-tune every aspect of the institutions' functioning.

Building on these two arguments I introduce in chapter 4 the cybernetic approach to the study of institutions' design. Cybernetics is the science of communication and control in complex regulatory systems, and as such it provides the solution to the demands raised by the theoretical argument of chapters 2 and 3.

In Part II, I move into a somewhat more practical discussion of specifically how – given the theoretical argument presented in the previous chapters – we can actually approach the design problem. In chapter 5 I outline the concept of capacity as the most appropriate

concept for assessment of how well international institutions perform their tasks, or more precisely for how much they perform their function of enhanced information transmission. Institutions with high capacity can improve the prospects for cooperation; those with low capacity cannot.

This is followed in chapter 6 by a presentation of the viable system model (VSM) I import from the field of organizational cybernetics. VSM is a deductive framework that identifies all the key information channels that *any* viable governing system (system that is supposed to regulate behaviour of some actors) needs to posses. If we were able to secure that in an institution all the information channels identified in VSM had sufficient capacity, the institution is correctly designed and has the ability to secure transmission of all the information the actors need for cooperation.

In other words, having defined in chapter 5 the dependent variable of capacity, in chapter 6 I elaborate on the viable system model that describes what levels this dependent variable in general needs to have for the institutions to be able to perform their functions. If the values of capacity are lower than those identified in the model, we know that the design of the institution – as the determinant of its capacity – is inappropriate.

In chapter 7 I discuss in some detail what all the theoretical debate up to that point actually implies for what design tools we should expect to be effective in increasing the institutions capacity. I provide a classification of these promising design tools, that can then orient concrete empirical investigations.

Overall, in parts I and II I provide answers to the three major questions identified above, in a slightly more general wording:

1. What does it mean – in causal terms – that an institution performs its function? *The answer is the concept of capacity.*

17

2. How do we determine the minimum level of capacity that is necessary for the institution to be viable? *The answer lies in the viable system model I adopt from the organizational cybernetics literature.*

3. What design adjustments are likely to increase the institutions' capacity? *The answer has a form of a classification of design tools presented in chapter 7.*

In part III, I leave the theoretical and conceptual debates and present two empirical chapters that illustrate how the framework I develop can be used. These chapters are not intended to cover the entire scope of applicability of my arguments. In fact, given the very general nature of my argumentation, that would not be possible. Rather, the purpose of the chapters is to apply the framework directly to two very specific problems, that is to an analysis of concrete challenges concrete institutions face.

Specifically, in chapter 8, I discuss how the Council of the European Union could have survived the Eastern enlargement by ten new member states and the correspondingly increased level of complexity the Council negotiations entail. There are more actors now in the Council, with more diverse interests, and I study how it has dealt with this new situation. I argue that the Council maintained for itself a sufficient level of decision-making capacity by actually delegating part of its agenda downwards to the Committee of permanent representatives, Coreper. I substantiate the argument with extensive quantitative evidence.

Chapter 9 present an argument that the way some of the major international organizations dealing with socio-economic matters – the International Monetary Fund, the World Health Organizations, the United Nations Development Programme, and UNICEF – are staffed reflects the need of these IOs for access to information about their client countries. Using the notion of soft information (information that is not easily quantified, standardized,

and transmitted) I argue that it is important for the IOs to have in their staff a sufficient number of citizens of the developing countries. This, however, goes against the interest of the powerful developed countries in maintaining control over the IOs by having them staffed by their own citizens. I elaborate empirically on how these two antagonistic forces interact in the IOs.

Finally, by the way of a conclusion, I summarize in chapter 10 the main findings of the dissertation, provide some reflections on both the dissertation as its stands and on the possible avenues of future research, and return back to the broader context of research on international institutions and to more general implications of my findings.

Table 1.1 summarizes the structure of the dissertation in a non-linear manner, indicating the links between the components of the argumentation I develop. While in principle each of the chapters provides to some extent a self-standing argument, but all the parts as well as the individual chapters form together a single coherent whole and should be understood as such.

	Assumptions, observations of the nature of the reality	My core arguments		Empirical application of the framework
		Conceptual argumentation	Theoretical argumentation	
Part I	Chapters 2 & 3		Chapter 4	
Part II		Chapters 5 & 6	Chapters 7	
Part III				Chapters 8 & 9
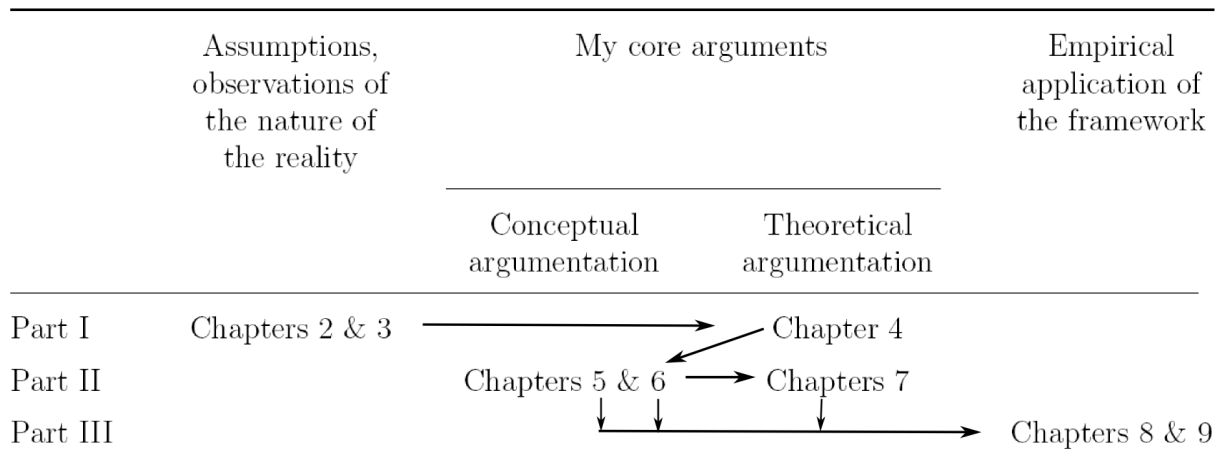
Figure 1.1: Structure of the dissertation: a non-linear representation

Part I

International Institutions

Chapter 2

Power, Interests, and the Role of International Institutions

Are institutions only derivatives of the states' power and interests, mere epiphenomena without causal effects on international politics, or do they actually matter? If they do, in what precise way? This chapter reviews the key arguments that speak to this question, in particular the so-called *neo-neo* debate between neoliberal institutionalists and neorealists about the possibility of cooperation. By doing so, it sets the grounds for the theoretical arguments I present in the later chapters.

In particular, I argue in this chapter that – while there are various views on what role institutions play in international politics – ultimately within the rationalist framework what institutions can secure is (only) enhanced transmission of information among the cooperating actors. This is a synthetic position that builds on the results of the neo-neo debate, and in a way re-iterates the earlier findings made within the regime theory.

The chapter is composed of three parts: I first present what has at least since E.H. Carr's

Twenty Years Crisis (1939) been the default in IR, the realist power-based position towards the role of international institutions. Second, I review the counter-argument raised by regime theorists, in particular by Keohane (1982), and re-iterate their position according to which institutions can help the actors achieve cooperative outcomes. Taking the regime-theory works one step further, I argue that the ways in which the institutions help cooperation can all be plausibly reduced to the function of enhanced information transmission (see section 2.2). Third, I review the later stages of the neo-neo debate in which neorealists raised arguments that were supposed to undermine the regime theory findings, but in which these findings proved – I would say conclusively – to be valid even under (neo)realist assumptions.

The chapter can be read as a critical review of the neo-neo debate, trying to establish its key results and their relevance for the study of design of international institutions. The most direct purpose of the chapter, however, is to formulate the first building block of my later argumentation.

2.1 The skeptics: realist and neorealist view on international institutions

The most significant defining feature of international politics is its anarchical nature. The system consists of sovereign states of varying power capabilities; because these are in their behaviour driven by the will to dominate – inherent in human nature –, whatever specific goals they may wish to pursue, their primary concern has to be power. In international politics, what matters is not intentions (whether good or bad) but power determined primarily by actors' military strength. This basic realist proposition was formulated comprehensively by Morgenthau (Morgenthau *et al.*, 1948) and in a more rudimentary version already by

Carr (1939). Despite intuitive plausibility of the insights presented in these works, in this classical realist theory the key explanation for persistently violent international outcomes rests not only in international politics as such, but also in the problematic *assumption* of men's evil lust for power, or *animus dominandi*, inherent in the human nature (Morgenthau *et al.*, 1948; Waltz, 1954).

In his *Theory of International Politics* (1979), Kenneth Waltz attempted to remedy these problems by getting rid of the 'unscientific' assumption of evil human nature and presented a comprehensive but parsimonious argumentation for why the nature of international politics *has to be* violent, whatever we actually think about human beings. Waltz moves all the explanatory power of the realist theory to the international 'systemic' level: the reason why states fight wars does not lie in human nature, but in the *anarchic structure* of international politics. Classical realism is superseded by Waltz's realism of structure (or neorealism).

The starting point of Waltz's analysis lies in the distinction between systemic and reductionist theories of international politics. Most existing IR theories are reductionist in their nature, that is they find explanation for international outcomes solely on the national or sub-national level (Singer, 1961), either in the first image of human nature, or in the second image of the national political systems (Waltz, 1954). In this view, we can understand the whole of the international system simply by investigating the properties of the units forming it, in practice primarily by investigating states' interests and power capabilities. For Waltz, however, this is fundamentally insufficient, since what matters is not only the units as such, but also the structure of the environment in which they operate, that is the way their interaction is organized by the nature of the *system as a whole.*

> If the organization of units affects their behaviour and their interactions, then one cannot predict outcomes or understand them merely by knowing the char-

acteristics, purposes, and interactions of the system's units. [...] In international

politics, systems-level forces seem to be at work (Waltz, 1979, p. 39).

Waltz does not claim that it is the system-level factors that ultimately determines the actual behaviour of the units; what it does however, is that it constraints their behaviour in the same way competitive market constraints behaviour of firms. Either the states do conform to the 'third-image' systemic forces, or they perish. It may matter what states want in their actual behaviour, but in general "[i]t is not possible to understand world politics simply by looking inside of states" (Waltz, 1979, p. 65).

The defining feature of the international system is anarchy, understood as a lack of government or of a central enforcing authority. Ultimately, as Waltz shows, it is this lack of central enforcing authority that causes the violent outcomes characteristic for international politics. "Nationally, relations of authority are established. Internationally, only relations of strength result" (Waltz, 1979, p. 112). Given the general scarcity of resources, international politics is a realm of potentially unrestricted violence, and the primary rational target of states thus needs to lie in securing their ability to defend themselves from the threats posed to them by the other states.

Under these conditions, very little space for stable cooperation among states is left. Much of international interaction takes on the form of a prisoner's dilemma, where states may gain by cooperating but they may individually gain even more by defecting from cooperation, provided that others continue to cooperate. Because states know that others have strong incentives to defect from cooperation, and because there is no central authority to prevent them from doing so, cooperation is unlikely to emerge. This logic is most painful in security matters, where the states face the security dilemma (Herz, 1950): to secure its defence, each state has to build military capabilities, but by this it in turn decreases security of others.

Arms-races emerge, and general security is threatened, without anybody actually intending to threaten anyone. As summarized by Jervis, "[g]iven this gloomy picture, the obvious question is, why are we not all dead" (Jervis, 1978, p. 170).

The key point to be taken from this Waltzian analysis is that domestic intentions of the states on their own are not sufficient to secure cooperation, i.e. states cannot simply decide to cooperate – the system always gives them and their potential counter-parts the incentives to defect from the cooperative scheme and take advantage of the others. In this interpretation, when we do observe international institutions, this is because the powerful states devised them to further their own interests, often at the expense of the weaker ones. The institutions do not have any causal effects on the international outcomes, because they only mirror the prevailing power distribution. Institutions and their activities are endogenous to the power distribution, and hence do not and cannot independently contribute to cooperation.

2.2 Institutions and international cooperation: the regime theory

The picture painted by Waltz is gloomy but the simplicity and logical coherence with which he builds his theory provide it with a significant degree of plausibility. And yet, a deeper scrutiny shows that some of the building blocks of the neorealist theory are not completely solid.

2.2.1 Problems of the realist and neo-realist theory

Broadly speaking, three problems of the neo-realist theory can be identified, the third one being the most relevant for our discussion. First, the concept of power central to the realist

thinking is by no means unproblematic. In the realist tradition power is understood primarily as military capability, that is ability to make others do what one wants with use of, or threat of , violence. However, it is not obvious how military power as such necessarily increases ability of a state to shape outcomes in other areas than in security. Overwhelming military power on its own does not grant a state any substantial advantage, as translating the military predominance into actual influence may in many issue-areas be difficult (Keohane & Nye, 1977). Power may, but also may not, be fungible. As a result, even if cooperation may be difficult or even impossible to achieve in security matters (but see Jervis, 1982), it may well flourish in other areas. In the end, the actual impact of cooperation in these areas on the existing balance of power may be marginal and the ability of the actors to take decisive advantage of one another is very limited (Lipson, 1984).

Second, even under the conditions of anarchy, cooperation may in fact emerge among rational egoists if what they play is a repeated rather than a one-off game. In a series of highly influential texts, Robert Axelrod (Axelrod, 1981, 1984) demonstrated analytically and with the use of computer simulations that even in the hostile situation of PD structure rational egoists may find it more profitable to cooperate than to defect, if the game is played indefinite number of times. In particular, the 'nice' strategy of tit-for-tat – that is to cooperate in the first round and then do whatever the counterpart did in the previous round – proves consistently to yield the highest payoffs. The strategy of defection that is dominant in a one-off PD ceases to be unconditionally dominant if there is a sufficiently high probability of playing the game again, that is if the game is played under the 'shadow of the future'. In this situation, the optimal strategy for an actor becomes conditional on the actual strategy adopted by the opponent. The key result of the Axelrod's analysis is simple and stunning: cooperation is possible even in the world as depicted by neorealists.

Third, in the early 1980s an entire new field of study emerged, dealing with what has been labelled by Ruggie (Ruggie, 1982) as international regimes, that is "the principles, norms, rules, and decision-making procedures around which actor expectations converge in a given issue-area" (Krasner, 1983, p. 1). According to this literature, outcomes in international politics are determined primarily by interests of the rationally acting states, but international regimes may enter the picture where the partially conflicting interests of states could lead them to Pareto-inferior outcomes if pursued unwittingly. In these situations, regimes serve as 'intervening variables', enabling cooperation where it would not be possible without them by providing a stable framework of the interaction.

On the most basic level, regimes emerge to address problems of cooperation in situations of *common interest* and *common aversion* (Stein, 1982). In many situations, states' interests are mostly harmonious, that is if each state pursues its individually rational strategy a collectively optimal outcome results (as in the ideal-typical competitive market). In many other situations, however, this is not the case, and in order to achieve collectively optimal outcomes states have to mutually adjust their actions, i.e. they have to engage in joint decision making (Stein, 1982, p. 304). One instance of such a situation is the dilemma of common interests captured e.g. by the prisoner's dilemma game. Here, to reach the single collectively optimal outcome of mutual cooperation, states have to *collaborate* and devise mechanisms, usually formalized, for effective punishment of unilateral defection. The other instance of a situation requiring joint decision making is the dilemma of common aversion where states have to *coordinate* their actions in order to avoid a specific Pareto-inferior outcome. This situation, depicted usually either by a game of coordination or as a battle of sexes, does not – at least in principle – require strong mechanisms enforcing cooperative

behaviour[1].

2.2.2   The functional theory of regimes

These three insights, however essential, do not amount to a full-fledged theory of international regimes. This was developed by Robert Keohane in his *After Hegemony* (1984). Keohane adopts the same starting point as Waltz: states are unitary, rational, and egoist actors, pursuing their goals under the conditions of anarchy. Similarly to Stein, Keohane points to the fact that anarchic structure of international politics impedes cooperation, since states can never be sure that their counterparts would stick to the agreed cooperation. As a result, "even where common interests exist, cooperation often fails" (Keohane, 1984, p. 6).

One way this unhappy situation can be overcome is through order imposed on international politics by the hegemonic power, as is proposed by the *theory of hegemonic stability* (Kindleberger, 1986). In this view, a hegemon is by the virtue of its power able to create international institutions and by maintaining them to secure orderly conduct of international affairs in a given issue-area. Whether the costs of maintenance are in the end borne by the hegemon, or passed to the weaker states, the key point is that it is the hegemonic power that secures operation of the regime. But should not we then, as pointed out by Keohane, observe a decline and dissolution of the existing regimes, as the U.S. hegemony was declining throughout 1970s? How is it possible that international regimes endure, even when conditions that gave rise to them change? Keohane argues that regimes persist because they serve the states certain functions.

I will not review the functional theory here in its entirety (see e.g. Hasenclever *et al.*,

---

[1]Although this does not mean that achieving cooperation in the first place is unproblematic. Especially in the battle of sexes game, the actors may actually find it extremely difficult to agree on which of the two equilibria they want to reach (see also Krasner, 1991).

1997, ch. 3), but I will briefly refer to the presumed effects of the regimes the theory identifies. What I will try to show is that all these effects are ultimately concerned with information transmission. I start with the very concept of international regime, lying at the core of institutional analysis in international relations. As mentioned above, in his consensual definition, Krasner conceptualizes regimes as the (explicit or implicit) "principles, norms, rules, and decision-making procedures along which actors' expectations converge in a given issue-area" Krasner, 1982, p. 185. In this definition, the four components in the first part stipulate the specific form regimes take, i.e. they are built of principles (general share understandings), norms (general shared conceptions of appropriate behaviour), rules (specific behavioral prescriptions) and decision-making procedures. The second part – referring to convergence of the actors' expectations – then deals with the causal mechanism through which the regimes matter, i.e. regimes matter by having the actors build shared expectations about various parameters of the cooperation scheme. This is a relatively subtle, but very important point: the mechanism through which institutions affect the actors' behaviour lies in changes of their beliefs, and in particular beliefs about what the consequences of their different actions would be. This mechanism can take in different ways.

Firstly and most importantly, regimes (and institutions in general) enable the cooperating actors to build reputations and thus develop more credibility vis-à-vis their partners. The notion of reputation lies at the core of cooperation theory (Axelrod, 1984) as whether cooperation is or is not possible depends on what record of cooperation and defection the actors carry (Baldwin, 1993). Because institutions are (by definition) relatively more stable than one-off agreements, they provide a more constant framework for interaction, which in itself opens space for reputation building.

Less obviously, though, reputation-building is enabled and encouraged also by issue-

linkages and side-payments. In the neoliberal institutionalist literature, issue-linkages and side-payments are usually though to have causal effects because they enable overcoming of negotiation deadlocks (Keohane, 1982; Martin, 1992). This is certainly correct. But the way they do so actually also has an information basis. The ability of the institutions to encourage cooperation does not lie in issue-linkages as such – after all the states can easily devise trade-offs without institutions. What the institutions are good for, however, is that they explicitly and formally stipulate the relationship between behaviour of actors in one area and the likely responses of their counterparts in other areas. They inform the actors about the likely consequences for them in the other areas, for the case in which they decide to defect on some of their commitments. The actors could create such linkages on their own, without the institutions, but existence of the institutions enables them to make these links clear to everyone and to make obvious to the potential defectors what the consequences of their possible defection would be. The causal mechanism through which institutions help issue-linkages and side-payments is through making of the link between the areas explicit and easily interpretable, not in the link as such.

Secondly, institutions reduce uncertainties faced by the actors duo to the unforeseeable behaviour of their cooperation counterparts. They do so through specification and formalization of the rules of behaviour and potentially also through legalization of their commitments (Abbott *et al.*, 2000). In the non-hierarchical world, these processes of formalization and legalization *per se* do not constrain the actors, but they increase the chances of shared understanding of the agreed upon desirable behavioral patterns. In other words, formalizing agreements results in the actors' better ability to interpret whether their own behaviour will be perceived as conforming to the agreed rules, and whether their counterparts' behaviour does. Legalization can then be seen as yet another step towards clarification of the agreed

rules, as it embeds them in a coherent, relatively broadly shared framework for their interpretation, and it directly establishes procedures (the dispute settlement mechanisms) for impartial assessment of compliance (Raustiala & Slaughter, 2002). Again, then, the causal mechanism through which institutions influence international political outcomes lies in enhanced transmission of information among the actors, as they will now be more able to assess compliance of their own as well as others' behaviour with the commitments.

Thirdly, institutions reduce what is usually referred to as the transaction costs of cooperation, or the costs of organizing it. In principle, institutions are stable bargaining fora, and hence the costs of establishing relations among the actors and of deciding on the bargaining procedures only have to be paid once (Keohane, 1984). As a result, once these fora are established, cooperation becomes much cheaper for the actors. As in the previous cases, the actual causal mechanism through which institutions perform this functions is based on information-transmission, as what the institutions really do is that they focus and fix the actors' expectations about how the cooperative arrangements will be set. In devising cooperative schemes, the problem does not lie in some material costs – flights and hotels of the diplomats - but in the need to discuss how the cooperative scheme will be run. Such negotiations often involve distributive bargaining and are in this sense political; they are potentially highly inefficient and politically costly (Fearon, 1998). What existence of institutions as stable cooperation fora enables is setting of these debates and costly processes once and for all (within some time scope), because once the bargaining is done, and if power distribution among the actors does not change, the actors can relatively easily adopt the agreed upon scheme also in the new cooperation area. The form of the agreement from the previous negotiations reflects the distribution of the states' underlying capabilities and power positions, and hence it is suitable also for the new arrangements. The actors' expectations

about what the outcome of the potential negotiation about the new scheme would look like therefore converge on the already existing model.

Clearly, the actors do not have to make use of this information. It is essentially a description of a cooperation blueprint that proved successful in the past and if it does not conform to the actors' interests anymore they probably will not adopt the blueprint. This is why existing institutional arrangements are neither completely stable, nor blindly transferred to new issue areas. But the existing setups offer useful institutional focal points of likely convergence.

So I argue that what is usually labelled as transaction costs are, ultimately, information costs, or more specifically that the causal mechanisms through which the institutions help the actors overcome these transaction costs are based on enhanced information transmission. In Keohane's original interpretation, enhanced information-transmission is one aspect of reduction of transaction costs, albeit a prominent one (Keohane, 1982, p. 338, fn. 28). Later on, he explicitly acknowledges that "information is at the core of [his] positive theory of international regimes" (Keohane, 2006, p. 77). The argument I present here is clearly consistent with the newer conceptualization; all transaction costs are information costs, not vice versa.

To be sure, whether we want to understand transaction costs as information costs depends on what our analysis is about – it is a matter of interpretation and in many contexts not adopting the information-based interpretation will be practically more suitable. Yet, if we are building a unified framework for the study of the institutions effects, being specific about the actual causal mechanism through which all these transaction costs operate – and this mechanism is information-based – is essential. Hence, while different interpretation of the relationship between transaction costs and information costs are possible, for the purpose of

institutional analysis I put forward here it is important to understand that what is normally referred to broadly as transaction costs has in fact an informational causal basis.

Last but not least, international institutions may not only enable easier transmission of information among the actors, they may also actively collect, process, and disseminate it. The specific institutions that have the potential to do so are international organizations, with the necessary administrative bodies. The primary function of these institutions – or the causal mechanism through which they improve the prospects for cooperation – is again transmission of information, most prominently about the actual behaviour of the states. The IOs enable cooperation thanks to their monitoring capabilities. In more bargaining contexts, in which there is as yet no behaviour to monitor, only some vaguely defined negotiations forum, IOs may enable cooperation by lowering the price of information on who wants what in the cooperation scheme, because the IO can collect the information 1) impartially and therefore more credibly, and 2) centrally, hence saving the costs of each member finding out about each other member's interests individually.

In sum, institutions do not constrain actors' choices by the virtue of their existence, but 1) by the virtue of making it for the actors easier to compel the others that cooperation will actually be profitable for them (they enable signalling of commitments and of reputation building), 2) by providing institutional focal points for new cooperation schemes (by showing blueprints of successful schemes), and 3) by lowering the price of information on interests and behaviour (through centralized monitoring and information collection).

We should note that the information-based notion of institutions is by no means unique to IR. The functional theory of regimes itself was developed explicitly on the basis of earlier works on transaction costs economics (Coase, 1960; Williamson, 1979). Quite specifically, the notion of institutions lying at the basis of cooperation theory in IR figures prominently

also in economic institutionalist works, as exemplified by Greif's analysis of institutions of long-distance trade (Greif, 1993). On a more general level, the core of economic analyses of markets, as opposed to command economies, is built on the notion of markets as mechanisms (institutional frameworks) that are more efficient in their key information transmission function (Hayek, 1945; Hurwicz, 1972; cf. Plott, 2000; Myerson, 2009).

In IR, as well as in economics, the reasoning behind the concept of institutions as mechanisms of enhanced information transmission is the same. If states wish to cooperate to reap gains that would otherwise be unattainable, but do not have a central authority to enforce agreements, they can still devise institutional mechanisms that provide them with easier access to the information they need. Because states play in international politics repeated rather than one-off games, their behaviour in the previous rounds of the game determines how trustworthy as cooperation partners they are in the current round. By reciprocating cooperative behaviour, states build their reputations as partners that can be trusted. International institutions encourage this reputation building by providing lasting and stable frameworks of interaction, and in general by lowering the price of information. For successful cooperation, whether economic or political, actors need information, and institutions can make it cheaper.

This notion of international institutions goes somewhat beyond the original formulation of Keohane's, in that he does not explicitly ground all the functions of the institutions in enhanced information transmission. Yet, my interpretation is perfectly consistent with the original accounts, it only attempts to go as close to the actual causal mechanisms through which the institutions help the actors deal with such generally defined concepts as transaction costs or issue-linkages. This section then really formulates one of the key theoretical arguments on which I build later. In the remainder of this chapter I will go back to the orig-

inal neo-neo debate and review the arguments that neorealists raised against the Keohane's analysis, and that might potentially threaten my interpretation of the institutions' role as well. I show that they do not.

2.3 The neo-neo synthesis: cooperation is possible, and institutions help it

Neoliberals claim to disprove neorealists' unconditional argument about impossibility of sustained cooperation by presenting compelling arguments for why 'cooperation under anarchy' (Oye, 1986) is indeed possible, even when one fully accepts the harsh neorealist assumptions about the nature of states and of the international system. The neorealist counterargument, presented most clearly by Joseph Grieco (Grieco, 1988; Grieco *et al.*, 1993), is simple: neoliberals are almost completely missing the point.

> [N]eoliberal institutionalism misconstrues the realist analysis of international anarchy and therefore it misunderstands the realist analysis of the impact of anarchy on the preferences and actions of states (Grieco, 1988, p. 487).

First of all, neoliberals see states as atomistic utility-maximizers, caring only about how well they fare. For realists, this view fundamentally misses the key structural determinacy of international anarchy where states cannot consider only their own welfare, but have to care primarily and constantly about their ultimate survival. Because their survival directly depends on their power standing in relation to other states (which may potentially threaten them), they have to care about how well they fare *in relation to them* in the first place: "states are positional, not atomistic, in character" (Grieco, 1988, p. 499). In other words, for states

under the conditions of anarchy the primary concern is the *relative* gains; *absolute* gains are only secondary. For liberals, states are *rational egoists*; for realists, they are *defensive positionalists* (Grieco, 1993, p. 303).

This observation has direct implications for the possibility of cooperation. If the primary concern of states is not to increase their absolute gains, but to secure their position vis-à-vis other states, negotiating particular agreements will be in practice much more difficult, since states not only have to find ways how to reap the additional gains lying on the table, but also how to distribute them equally. Related to this, the overall positive-sum situations on which the neoliberal analyses rely so heavily turn into zero-sum situations at the margin as during negotiations over the distribution of the benefits one's gain is another's loss.

Secondly and consequently from the previous point, neoliberals may well be giving good reasons for why the prisoner's dilemma structure of the situation of international cooperation may not completely impede it, but this does not tell us much about why neorealism should be flawed (as neoliberals suggest). The PD game, on which most neoliberals concentrate, in fact captures international politics as the neoliberals see it, and not as the neorealist do. For neoliberals, the world of international politics is full of unrealized gains from cooperation, it is a world of opportunity to increase actors' welfare that fails only because there is no central authority to prevent states from cheating. In their world, ability to overcome the problem of cheating is by and large sufficient for achieving cooperation.

This, however, is not how realists understand the situation. For them, to the contrary, the world is one of constant struggle over allocation of gains in which each state's ultimate goal is to secure its position in the power competition. The problem for states is not to reach or at least get closer to some distant hypothetical Pareto-frontier, they already 'live on it' (Krasner, 1991). Rather, the core of international politics lies in a struggle over which

of the many particular equilibria on the Pareto-frontier, with the respective redistributive consequences, will be chosen; international politics is about power competition, not about efficiency-enhancing cooperation. Although both neorealist and neoliberal research programs are in principle compatible with this view, it is only the former that possesses the analytical tools to fully appreciate this fact (Grieco, 1993, pp. 321-2; Krasner, 1991, p. 362; Powell, 1991). Representing the situation states face as one of a prisoner's dilemma – as neoliberals do – is thus not appropriate; a far more realistic representation is provided by the battle of sexes game (Krasner, 1991), or even by the deadlock game where actors actually prefer defection to cooperation (Jervis, 1988). This, obviously, is a fundamental critique, one potentially rendering much of the Keohane's argumentation irrelevant.

Neoliberal response came from Duncan Snidal, who showed analytically that as the number of actors in the system increases, the degree to which relative gains considerations matter for their decisions goes very quickly down (Snidal, 1991). In other words, Grieco's argumentation is not really valid, as under the conditions as defined by neorealists themselves the states will only care about relative gains, as opposed to absolute ones, marginally.

To arrive at this conclusion, Snidal develops a solid complex analytical framework capturing the key aspects of the neo-neo debate, and makes several non-trivial sub-arguments worth reviewing. Firstly, the real problem for international cooperation is the possibility of cheating, and not the potential gaps between the payoffs individual states receive from cooperation since, if there at all are some gaps, states can simply alter the terms of the agreement to remedy them[2]. As a result, the entire analysis of cooperation among rational egoists can focus on the problem of cheating exclusively.

---

[2]It should be noted, though, that this argument was later opposed strongly by Grieco according to whom, by assuming that finding the appropriate balance is possible and easy, Snidal avoids the central realist consideration instead of actually facing it (Grieco, 1993, p. 319). Although Grieco's position on this probably deserves some merit, I believe Snidal defends his view much more successfully (Snidal, 1991, p.

Secondly, to focus and simplify the discussion further, he demonstrates analytically that as concerns of states for relative gains increase, all the games that are most relevant for analysis of cooperation among two states, even the harmony game where both actors strictly prefer mutual cooperation, turn into a prisoner's dilemma (Snidal, 1991, pp. 705-10). So the key problems of cooperation are indeed reducible to the analysis of the prisoner's dilemma game structure.

The major point of Snidal's analysis has to do with how the entire problem of cooperation changes when we move from the simplifying bilateral to a more realistic multilateral setting. The major Snidal's finding is that the impact of relative gains considerations on possibility of cooperation drops off very quickly as the number of actors in the system increases. In a multilateral setting, for cooperation to be impeded by relative gains considerations, states 1) have to care almost exclusively about *relative* gains (and mostly disregard absolute gains), and 2) they have to behave paranoically in the sense of not considering at all the *probabilities* of being attacked by particular states but rather treating all states as constantly extremely threatening (Snidal, 1991, p. 719; Keohane, 1993, p. 282). Unless both these conditions hold to a high extent, cooperation is possible. The entire neorealist argumentation, then, is conditional, and in fact fully applies only in very specific and indeed extremely unrealistic circumstances.

Although Snidal's analysis is certainly impressive, it by no means 'wins' the debate for neoliberals. Besides the fact that it does not disprove realism but only makes its claims conditional, the model it is based on is in one crucial aspect fundamentally problematic. Paradoxically, this problem was not 'discovered' and discussed by neorealists, but by Robert Powell, himself significantly more inclined towards the neoliberal research program. In his

---

703; Grieco *et al.*, 1993, p. 738).

article on *Absolute and Relative Gains in International Relations Theory* (Powell, 1991), Powell sets himself an ambitious goal to develop a unified framework of analysis capturing both neorealist and neoliberal perspectives on international politics. As a starting point he adopts the neoliberal view that states' ultimate motivation is their welfare, that is that they seek primarily absolute gains. At the same time, he fully appreciates the fact that states cannot care only about absolute gains, because they are seriously threatened in their pursuit of wealth by other states. Broadly speaking, states are trying to maximize their wealth, but under serious constraints imposed on them by the fact of the potentially violent international anarchy (Powell, 1991, p. 1304).

The key to Powell's analysis lies in his realization that modelling this situation as a repeated game – as all neoliberals do – is ultimately inappropriate. The key problem for states is not maximization of gains, but survival. In other words, states' immediate concern is not how much they earn in the long run, but whether they survive the next few rounds of the game. Precisely this states' concern with survival, though, is something the models based on repeated games cannot capture, because they assume that the actors enter each new round of the game 'alive' and with unaltered relative capabilities. The most fundamental of the realist insights is thus in the repeated-games framework completely missing. To address this shortcoming, Powell develops a two-stage model in which being a sucker in the first round decreases a state's capabilities to such degree that it is overwhelmed by the opponent in the second round . In the logic of the model (Powell, 1991, pp. 1311-16), in the first round the states decide whether or not to cooperate on certain economic issues, and in the second round they decide whether or not to fight a war. Receiving the sucker's payoff in the first round puts a state to such a technological disadvantage that it actually loses the war in the second round. A state that cooperates unilaterally in the first round, may perish in the

second.

Now, can cooperation ever emerge even in these conditions? Powell shows that indeed achieving stable cooperation is highly problematic in this setting, more problematic than the standard neoliberal analyses suggest, but also that it clearly is possible if the costs of fighting a war are high. Even in this extreme setting where being a sucker once is enough for an actor to perish, "cooperation reemerges once the use of force is not an issue" (Powell, 1991, p. 1313) because the costs of fighting would be too high.

Powell's argumentation has been welcomed by Grieco who appreciates its serious effort to understand the realist view fully. At the same time, Grieco maintains that Powell is still missing some of the important consideration states may have with respect to cooperation, and thus it still is not accurate enough (Grieco, 1993, pp. 312-13; Grieco *et al.*, 1993). Overall, I would contend that the counterarguments Grieco present are rather weak, if not marginal, and that the synthetic framework Powell develops captures the realist standpoint accurately. In fact, it seems that at this later stage of the debate Grieco is becoming slightly selective and incoherent in how he applies his own arguments, even to the point where – as claimed by Snidal Grieco *et al.*, 1993, p. 741 – he chooses arguments simply according to whether or not they lead to the results he likes. I believe that Grieco has at disposal – mostly thanks to Powell's sophisticated analysis of the issue – some good arguments for refuting the results Snidal arrives at. However, the evasive 'yes, but still...' type of argumentation he engages in relation to Powell's framework (Grieco, 1993, pp. 314-16) strongly tends towards what I believe could in the end earn neorealism a label of the lakatosian 'degenerative' research programme.

Since early 1990s, further developments have taken place on the issue of cooperation under anarchy. Mulford and Berejikian, for instance, claim to have develop a unified realist-liberal

framework on the basis of experimentally better substantiated assumptions about actors'
behaviour (2002). Halás has argued that the way the problem of relative and absolute gains
was conceptualized was inappropriate, and has proposed an alternative specification (Halás,
2009). Mosher (2003) employs arguably empirically more realistic assumptions to challenge
Snidal's analysis of the role number of actors in the system plays for possibility of cooperation.
On the other side, Glaser (Glaser, 1994) presented an influential argument according to
which in fact the neorealist position itself, if taken into its own logical implications, predicts
cooperation instead of adversarial behaviour under many circumstance.

A completely new research approach towards the problem of cooperation has been formu-
lated by Moravcsik (Moravcsik, 1997), based on disaggregation of the states and explicit con-
siderations of the domestic politics as the sources of international politics (Putnam, 1988).
Especially in the area of international political economy such approach continues to earn
considerable support.

The key result of the neo-neo debate, however, is that mutually beneficial cooperation
is in international politics possible, the structural anarchy pressures notwithstanding. In
other words, the insight that institutions can effectively help the actors achieve cooperative
outcomes holds. This notion of institutions sounds very functionalist – institutions perform
the information transmission function. Yet, it should be clear that it takes power relations
among states as seriously as possible. If the major critique of the functionalist theory of
regimes has been that regimes do not help because the key problem of cooperation are the
problems of distribution of gains (Krasner, 1991; Grieco, 1993), the theory I adopt here says
that precisely because of these power considerations the only contribution institutions can
have to cooperation consists of the enhanced information transmission. In particular, when
it comes to distribution of gains among the cooperating actors, institutions enhance the

prospects of cooperation by formalizing and clarifying the links between issue areas that in the end enable the actors the agree. Precisely in this lies the power of the Keohane's original functional theory: institutions do not constrain the actors directly, they inform them.

If we want to take states in the international politics seriously, we know that institutions can help them achieve cooperation only through better transmission of information. To be sure, the fact that all mechanisms through which institutions impact on the actors' behaviour are ultimately based on enhanced information transmission should not be interpreted as meaning that all institutional analyses should build on such a notion of institutions. Quite the contrary, for many specific purposes treating institutions as rules that actually put on the actors' behaviour direct constraints may be useful – it is an analytical convenience how deep into the *institutions* ⇒ *behaviour* mechanism one wants (and needs) to go. But when devising a general framework for the study of how institutional design impacts on their functioning, we are in need of a unified concept of the functions institutions perform. Information-transmission is the single underlying function, all others – such as imposition of constraints on the actors' behaviour – are derived from it.

Also, what I am trying to accomplish here is to disentangle the causal effect of design on functioning of the institutions and hence we need to get as close to the immediate causal mechanism through which institutions function as possible (George & Bennett, 2005). When we want to understand how design impacts on functioning of institutions, we are effectively conducting a causal inquiry. We are trying to establish a causal relationship between two factors. I tried to show in this chapter that this in fact entails focusing on a single key effect that institutions have – or may have, to a varying degree. I argued, or rather re-iterated the now several decades old argument, that in a non-hierarchical world of international politics ultimately the *only* way in which institutions matter is through enhanced transmission of

information among the cooperating actors. Clearly, institutions are created because their creators believe they will bring about the desired changes in the actors' behaviour. But this is *not* what institutions do, or not directly. What they do is that they make available and cheaper information, the availability of which *induces* the actors to behave in the desired ways. The effect of institutions on behaviour is indirect; they provide information relevant for behaviour of the actors, and based on this information, together with other factors, the actors chose their actions. Institutions do not constrain, they inform.

My entire argumentation builds on this exclusive focus on information transmission. It is reductionist in that I exclude all other potentially relevant effects international institutions have. As discussed by constructivists, international institutions do not only constraint actors, they may also endogenously shape their preferences, they may socialize actors into certain roles (Johnston, 2001). Similarly, I do not consider the fact that interests of the interacting actors may change over time simply due to learning and long-term evolutionary pressures. I do not consider these factors, and I explicitly take the conservative approach in which we cannot assume institutions to have these effects, however desirable they often may be. The disadvantage is a theoretical reduction.[3]The major advantage is that adopting this minimalist position makes the entire theoretical scheme defensible against the fundamental realists' arguments, i.e. it takes the power- and interests-driven nature of international politics seriously. I try to demonstrate in the text that the costs paid in exclusion of certain aspects of institutions' functioning are more than outweighed by the analytical leverage we obtain when we build on the information-transmission perspective what I will refer to as the organizational-cybernetic framework.

---

[3]One could argue, however, that all theories are based on enormous reductionism. Good theories are in this sense distinguished from bad ones according to how the reduction is performed, not how much reduction there is.

Chapter 3

The Problem of Complexity

In the previous chapter I outlined the mainstream power- and interest-oriented approach to the role of international institutions. This approach is mostly driven by descriptions of the actors' interaction, where the actors understand what their and the counterparts' available strategies and payoffs are. They may not know this at the beginning, i.e. they may need to build the knowledge over time, but in principle at some point they are able to understand the structure of the game they are playing and what actions are likely to lead to what outcomes[1].

In this chapter, I turn to a completely different type to theorizing about international institutions, one that could be best labelled as the governance approach. According to this governance perspective, the characterizing feature of most problems in international politics is their overwhelming complexity and the multiplicity of actors that are relevant for solution of these problems. Actors make their choices under the conditions of extreme uncertainty, they can only rarely predict what outcomes their actions will lead to and what the unintended consequences will be, and their ability comprehend the situations in their entirety is limited

---

[1]Technically speaking, they have a *complete* information about the game, even though they may not have perfect information McCarty & Meirowitz, 2007.

(Rosenau, 2000).

The implication of this theoretical approach to the study of international institutions' design is fairly clear: the political systems formed by the institutions and the actors are so complex, that we cannot hope to determine a priori what would, in any situation, happen with the outcomes if we changed the institutional framework. In other words, due to the enormous complexity of the matters we know close to nothing whether any specific design alteration would improve or hinder the prospects of cooperation.

The chapter consists of four sections. I first outline the problem of complexity, in both empirical and theoretical terms. In the next step, I discuss its implication for the study of political system of international institutions and the problems complexity brings to the notion of design. Third, I re-iterate the implications identified in the previous section with use of the more mainstream approaches, showing that the problems of design I discuss are not completely foreign to the neorealist and institutionalist positions either. Lastly, I discuss how these problems can be productively dealt with.

3.1   International politics is complex

International politics is complex. Most problems of international cooperation involve un-countable relevant actors, and the numbers tend to grow over time (Rosenau, 2000). The actors pursue all different kinds of interests (e.g. Allison, 1969), at different levels (Putnam, 1988; Moravcsik, 1997). The ways in which the actors are inter-connected are both numerous and multi-dimensional (Keohane & Nye, 1977). Actors' abilities to comprehend events as they occur and react to them are limited at best (Jervis, 1976), up to a point where they are uncertain not only about others' interests and about how their actions might translate into

outcomes, but also about what their own interests are likely to be in the future. To some extent, the actors need to make decisions from behind a veil of uncertainty (Young, 1989, p. 361; Young, 1991). As if this was not enough for anyone's ability to comprehend what is going one, "many damn things happen at one time" (Rosenau, 1997). In practice we hardly ever are able to deal with certain problems without at the same time directly or indirectly influencing many other areas, and obtaining a number of unintended consequences.

We know that things are complex. In fact, though, they are much more complex than we are usually willing to admit or even able to imagine. A system of just four actors where any actor can either be or not be connected to another one has 64 possible states. If the direction of the connections matters, so for instance if one element is dominant in the relationship, there are already $2^{12} = 4096$ possible states. If we move to (still only!) five actors, there are 1024 possible states of the system if the directions of the relationships do not matter, and 2^{20} – over a million – if the directions do matter. The number of possible states is given by the formula

$$variety_{non-directional} = 2^{\frac{n(n-1)}{2}} \tag{3.1}$$

for non-directional connections and

$$variety_{directional} = 2^{n(n-1)} \tag{3.2}$$

for system where direction of the connections actually matters (cf. Beer, 1979, p. 37).

These numbers apply to system where each pair of actors can be connected in only one way. What if two different kinds of connections are possible, e.g. a no-connection (0), a weak connection (1) and a strong one? In this case, already of system of only 3 actors where the

direction of the connections matters has 729 possible states; with 5 actors the number rises to three and half billion. With somewhat more actors, the sheer number of possibilities rises beyond comprehension, as illustrated in figure 3.1, where on the horizontal axes numbers of actors (N) and numbers of possible levels of their connections (s) are depicted, and where on the vertical axis the number of possible states of the system is showed, on a logarithmic scale (i.e. every unit increase raises the value 10 times, so 3 signifies 100times more than 1, and 6 signifies million times more than one).
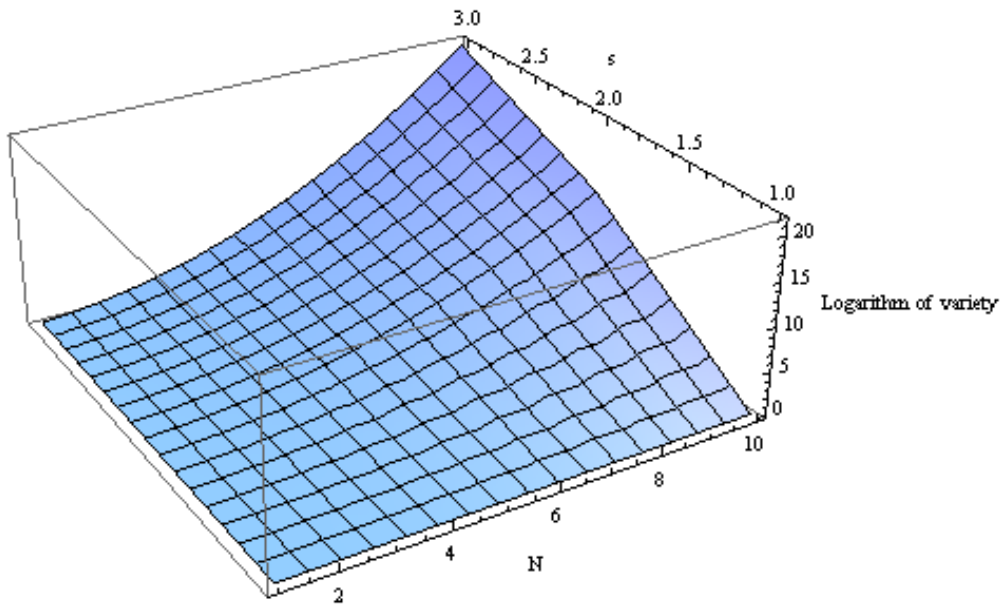
Figure 3.1: Complexity increase with number of actors and types connections

In his textbook on international politics, Bueno de Mesquita uses for illustration of the use of spatial models of politics and of game theory the (real) case of the Iranian nuclear programme (Bueno de Mesquita, 2009, chs. 2 and 3). The data on interests and power resources of the actors he uses were obtained from expert interviews of specialists on the Middle-Eastern politics and according to Bueno de Mesquita, these experts identified more than 80 different actors (individual or collective) that are directly relevant for the situation.

47

Yet, understandably for education purposes, he reduces the description of the situation to only 6-7 actors, and even these are for the actual analysis clustered into only 4 groups. To be sure, we are dealing with a textbook analysis rather than with a proper elaborate research text, but when one considers with how many actors most actual research texts deal with, reaching above these numbers would be very rare (indeed Bueno de Mesquita would most certainly count among those theorists who treat complexity of situation as a modelling challenge rather than as an annoyance that can be assumed away, as exemplified e.g. in Bueno de Mesquita *et al.*, 2003).

To take a historically more distant but perhaps even more interesting example, it is worth considering the historical accounts of the origins of the World War 1. Van Evera shows that one of the key reasons for why the War started was the overwhelming complexity of the relationships among the actors:

> Blunders and miscalculations played an important role in causing the war. For instance, the Germans launched the July crisis in the mistaken expectation that Britain and Belgium would not resist their offensive, and that Italy, Sweden, Rumania, and Japan would fight on the side of the Central Powers. The British, partly because they were unaware that Russian mobilization meant war, failed to move forcefully to prevent it. The Russians began mobilizing in the false hope that Germany would acquiesce to their partial mobilization, and that their general mobilization could be concealed from Germany. Russian, German, and French officials also exaggerated one another's preliminary mobilization measures, which spurred all three to take further measures, fuelling the spiral of mobilizations.

In part these blunders reflected the confusion sown by the simultaneous action of eight states. The dispersion of power scattered each state's focus of attention. As a result, each was likely to misconstrue the actions, interests, and perceptions of the others. Large numbers increased the costs of information and raised the risk of error (Evera, 1985, p. 111).

A rather similar interpretation can be taken from the memoirs of Lloyd George, the then Chancellor of the Exchequer and later the British prime-minister, who observed that "Not even the astutest and most far-seeing statesman foresaw in the early summer of 1914 that the autumn would find the nations of the world interlocked in the most terrible conflict that had ever been witnessed in the history of mankind" (George, 1938, p. 32). Note that this (for most people overwhelming) complexity involves only six to eight actors and only a relatively small set of available actions (ultimately the choice between mobilize and wait). Somewhat on the same note, the pre-war German *Kaiser* Wilhelm II. notes in his memoirs that Bismarck "was able to juggle with five balls of which at least two were in the air", which the Kaiser himself certainly "was not capable of", and neither was Bismarck's successor in the office Count Caprivi, himself no novice to foreign affairs (Wilhelm II, 1922, p. 9). Politics, whether high or low, *is* complex.

In fact, since the *Kaiser's* times, the actual complexity of the situations the actors need to deal with has risen significantly. As summarized in figure 3.2, various important indicators for complexity of world politics have been rising steeply over the recent decades. The number of states has almost tripled in the last five or six decades. The World population has risen from about 3 billion to about 7 within the same period. An only in the last twenty years, the stock of foreign direct investment (as measure of significance of multinational economic activity, and primarily of the multinational corporations (MNCs)) has increased 10 times,

from two billiard (10^{12}) USD to twenty.
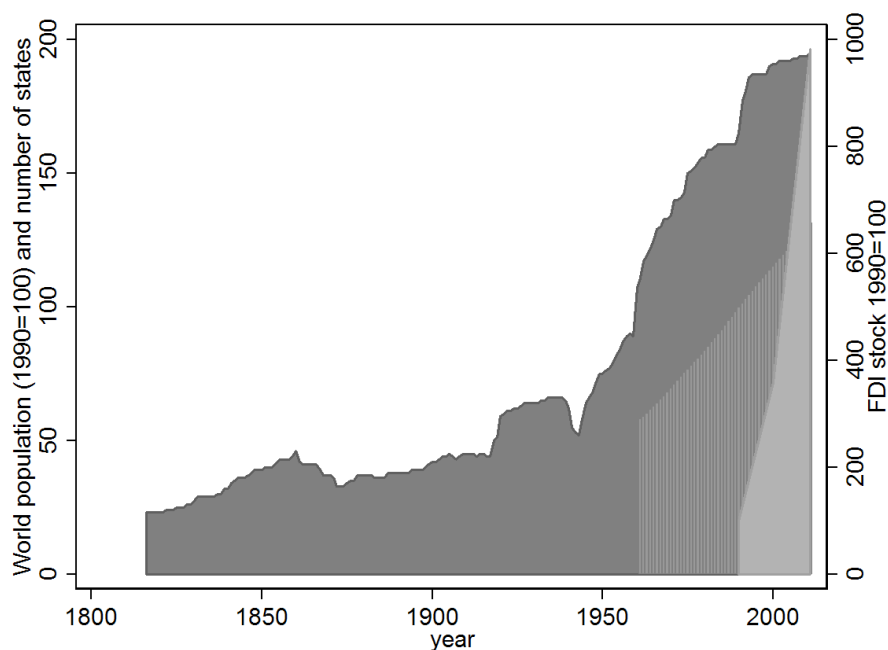
Figure 3.2: World trends: number of states, population size, and FDI stocks

In a space of a fixed size, an increase in density of its population leads to an increase in the amount of interaction between the actors. Politics was always complex, but the level of complexity today is much higher than ever in the history. And it keeps rising (Alberts & Czerwinski, 1997; Rosenau & Durfee, 2000; Held & McGrew, 2003).

When the task we set for ourselves is to try to design institutions that would enhance the prospects of international cooperation, we cannot avoid *embracing* this complexity, *counting* with it, and *coping* with it. We cannot assume the reality's complexity away, and we cannot simplify the situation in a way that does not respect the variety of the situations we deal with. If we do so, we fool ourselves, and the complexity will take us by surprise, to our detriment (Beer, 1979, p. 96), as I discuss in more detail in chapters 4.

In his now classical study, Scharpf points out very correctly the constraints imposed

on political research by the complexity of the issues we deal with. He argues that if we want to understand political phenomena, and especially if we want to generate the kind of understanding that can be actually used in practical political affairs, we cannot impose on them simplicity analytically, and casually assume complexity out. We need to deal with the complexity. In Sharpf's view, the only way in which complexity of the political phenomena we study is reduced is through existence of institutions that structure the situations and provide some regularities to it, i.e. it is the actual existing institutions that help us do political research and arrive at approximately meaningful statements about the situations we study, it is not the assumptions (Scharpf, 1997).

However true this may be, the problem is that when we are inquiring into the possibilities of better institutional design, this help by definition disappears. Subsequently, the necessary but unwelcome implication of taking the world's complexity seriously, i.e. of trying to paint at least a somewhat accurate picture of the problems we study, is that we should also *dramatically* increase complexity of the models we use to capture them. And given the serious constraints of our cognitive abilities, whether we perform qualitative, quantitative, or formal research, our ability to devise such models and use them for prediction of outcomes of institutional re-design will face significant *technical* limits.

3.2 Complexity hinders meaningful design efforts

There is also a deeper *theoretical* problem implied, however. If we believe that politics take place within a complex system of interactions, we need to be ready to live with a number of interesting but disturbing phenomena that are inherent to the nature of complex systems, and that unfortunately make the task of institutional design rather demanding. Jervis (Jervis,

1998) discusses these phenomena with plenty of illustrations, so I will present them only in a brief summary.

In complex systems, effects follow from their causes indirectly, via various mechanisms, often with time-lags. Ecosystems and climate system are among the most popular examples[2]. Another feature is that within complex systems effects often do not magnify additively, but in interactive ways (cf. Taagepera, 2008, ch. 5). The use of interaction models in statistics (Brambor *et al.*, 2006) or configurational approaches in qualitative research (Ragin, 1987, 2000) shows increasing appreciation of these phenomena in political research (cf. Hoffmann & Riley, 2002). Yet, *prediction* is a task on a different level of difficulty than observation and explanation (cf. Taagepera, 2008; Schneider *et al.*, 2011), and the odds that we will be able to predict existence of an interaction effect between such factors as various design features are depressingly low.

This is further reinforced by the fact that in complex systems relations between variables are only rarely linear or close to linear; non-linearity is the norm, linearity the exception (Jervis, 1998). Extreme cases of non-linearity are described in the chaos theory according to which in complex systems un-measurably small differences in initial conditions of the systems lead to very significant differences in outcomes. This is referred to as the so-called "butterfly effect", following the hypothetical example in which a small displacement of air mass at one place on the Earth can lead to creation of a tornado at some other distant place (e.g. Gleick, 1988). However unlikely, this mostly metaphorical illustration captures a real phenomenon in which marginal differences in initial conditions can result in enormous differences in outcomes, as the non-linearity of relationships between the factors lead to that

---

[2]In ecology, there is an entire journal devoted just to the issues of complexity of the ecological systems (*Ecological Complexity*). As far as the climate is concerned, the incredibly complexities of inter-connection among the relevant factors have been studied for years and need no emphasis (see e.g. Rind, 1999).

the differences propagate upwards in the orders of magnitude. In effect, then, the causal relations through which the effects of these small differences materialize are intractable and behaviour of complex systems then becomes unpredictable.

This notion directly precludes use of, or reliance on, the comparative method in the design research, as the underlying conditions within which the design should be implemented are never the same. We know from comparative politics how important it is that local conditions are taken into account, and the notion that one size does *not* fit all has been mainstreamed (Fukuyama, 2004; World Bank, 1997). In international politics, the same problem applies as well: with small differences in such underlying conditions as distributions of power and interests, the same institutional settings may easily lead to very different outcomes.

The last key problem for successful analysis of how design will impact on functioning of institutions is the so-called emergence, a phenomenon where in complex systems effects develop that are fundamentally irreducible to the features of the individual component units. Perhaps the most famous early example of an emergent phenomenon was provided by Thomas Schelling in his analysis of racial segregation in the U.S. cities. Schelling showed that under a set of very plausible and extremely simple assumptions about how individuals and families choose the neighbourhoods they want to live in, racial segregation into white and black communities naturally arises. The key surprising result of Schelling's was, that for segregation to emerge, only very moderate levels of preference for neighbours of the same race are sufficient, i.e. only very mildly 'racist' individuals will tend to create, over time, highly segregated societies, even if no one individual preferred or could foresee such an outcome (Schelling, 1969).

The insights from the research on emergent phenomena of this kind (see Holland, 1999; Gell-Mann, 2002) are making their way into social sciences also today (Sawyer, 2005; Miller

& Page, 2010), and the same applies to international relations (Cederman, 1997; Albert *et al.*, 2010). Structural theories such as those of Waltz (Waltz, 1979) or Wallerstein (Wallerstein, 1979) point at certain very important emergent phenomena, and the power of these theories (as well as of other structural theories) lies precisely in that they fully appreciate these phenomena[3]. What the research on emergence shows is that presence of phenomena not derivable from properties of the component units is inherent to complex systems – it is a rule, not an exception. In the context of the design research, emergence in effect means that we cannot hope to be able to find out how individual design features impact on functioning of an institution; what causes the effects (or the outcomes we observe) are the institutional setups as wholes, not the sums of their parts reducible to these parts.

Put together, all these factors mean that our ability to predict how adjustments in design will impact on behavioral outcomes and hence on the prospects of cooperation is extremely limited. Quite simply, the sheer unpredictability and enormous complexity of considerations we need to take into account when designing the institution renders the notion of optimizing design hugely problematic.

3.3   Informal politics in institutionalized cooperation

To be very sceptical about the possibilities of modelling behavioural outcomes through institutional re-design, we do not need to invoke the notion of complexity. We may arrive at similar conclusions also from a more mainstream rationalist perspective, in particular when we consider explicitly the problem of design endogeneity and of the relationship between

---

[3]Although they usually achieve this only at the expense of the neglect of the actors' agency. The newer approaches are in many ways able to overcome this problem, and model the structural phenomena as emerging from interaction of individually rational actors.

formal design of institutions and the reality of informal politics that take place outside or at the margin of these formal institutions.

First, there is the key argument that states only commit to cooperation to the extent (or depth) to which it conforms to their interests, i.e. that they design formal institutional mechanisms only when they are sure they will not be constrained by them (Downs *et al.*, 1996). In this sense, actors' compliance with the institutions' prescriptions is automatic, the actors would not act differently even if the institution was not there (Raustiala & Slaughter, 2002). The literature most directly relevant for this view is that focusing on the rational design of institutions, claiming that state devise institutions rationally to reflect their own interests (Koremenos *et al.*, 2001). A particularly important example is the research on various escape clauses and flexibility measures built in to the institutions to provide a space to manoeuvre for the actors if their interests diverge (usually only temporarily) from the agreed upon rules (Rosendorff, 2005). In sum, the design of the institutions, or the way they are set up, is to a considerable extent endogenous to their interest and power constellations.

Second, even if we admit that the particular cooperative arrangements (as formalized in the agreements) may become restrictive for the actors (and I happily acknowledge this possibility), it is not obvious why the powerful actors could not simply come together and devise new informal agreements at the margin or outside of the formal institutional framework. We may want to change the design of institutions in certain ways, to achieve the desired outcomes; but by doing so, we run the risk of inducing the powerful actors to try to avoid the constraining rules by acting outside of the formal rules, if it suits their interests. Indeed, there is a lot of evidence that this is precisely what actors often do (e.g. Kleine, 2010; Reh *et al.*, 2013; Héritier, 1999).

In international politics, this danger is especially pronounced when there is a variety of

institutional fora within which the specific problems may be solved, as is increasingly the case (Drezner, 2009; Raustiala & Victor, 2004). When we change the rules in an effort to induce certain types of behaviour, we may induce a range of consequences we possibly do not like and certainly did not intend.

This problem actually gets further dimension when we consider what I would refer to simply as a paradox of institutionalization, namely that the more institutionalized a certain cooperative arrangement among actors is, the more cooperative outcome we may get (e.g. Smith, 2003), but the less will the particular design of formal institutions matter and the more the cooperative arrangement will be based on informal practices. In other words, because the actors know each other better and because they have a (probably positive) common cooperative experience, they do not have to care about the particular formal rules that are supposed to constrain their behaviour in the first place. In fact, a lot of evidence shows that even in contexts much more formalized than that of any IOs, informal politics are omnipresent. In the case of the EU, notorious for its excessive legalization, research shows that the formal rules do not really matter too much in actual practice, that political outcomes are mostly driven by power and interests (e.g. Héritier, 1999). Or consider the issue of log-rolling, pork-bareling, and even vote-buying in the U.S. Congress, the legal sovereign of the country with arguably very strong rule of law (World Bank, 1997); again, interests are notoriously known to prevail over formal rules.

In other words, more successful international cooperation does not mean less power- and interest-politics. The conflicts may not be so violent, and hence the ultimate target of international cooperation would be fulfilled (Fearon, 1995), but this should not obscure the point that power will matter in determination of political outcomes as much in cooperation with formal institutional arrangements as it does everywhere else in (not only) international

politics.

It is interesting to see what this approach tells us about the problem of dysfunctionality of the existing international institutional arrangements, for instance about such matters as the often deadlocked negotiations in the United Nations Security Council. I would argue that these problems call for a very careful interpretation as to the role played by the formal institutions. If we observe, for example, a failure of the Security Council to prevent mass atrocities in certain civil conflicts, the reason does not need to be seen in its formal design. If the key power-holders in the Council shared the view that intervention was desirable, they could decide on one and implement it, no matter what the formal decision-making rules say. Yet, the very purpose of the veto of the major powers in the Council is to ensure that this may not happen as long as at least one of the major powers does not share the perspective that an intervention should take place. The problem of the Security Council, in this sense, does not lie in its formal arrangements, but simply in the fact that there is a significant disagreement among the key power-holders about what policies to adopt.

More generally, inadequate formal rules do not in this perspective necessarily impair cooperation. As long as the concerned states share interest in the cooperative outcomes they should be able to find the ways to achieve them, no matter what the formal procedures, often inherited form the past, dictate. The formal arrangements may, more or less successfully, induce cooperation, but only by the virtue of ensuring, more or less successfully, enhanced transmission of information. It is not the formal rules per se, that prescribe or proscribe specific behaviour, or that identify appropriate decision-making procedures, that determine the success of cooperation. What matters is the possible impact of these formal rules on the flows of information.

Taking all these findings about the informal nature of international politics seriously, how

can we learn about what institutions work and what do not? We may simply empirically observe that in some areas the cooperative schemes operate smoothly while in others they do not. But, given the discussion above, it is hard to infer from such observations what role design of the institutions played in this, as opposed to e.g. simple convergence of interests or pressure by powerful states. In other words, our ability to learn comparatively what institutional design works and what does not is limited at best.

In sum, we cannot hope to fine-tune the design of international institutions. By trying to model all the individual aspects of the institutions' functioning we may do more harm than good; however clever the models on which we build our design prescriptions might be, they will always capture only a small part of the problem's complexity.

3.4   Focus only on the vital system functions

Given the gloomy prospect for design identified above, can we do anything? My answer is that the best we can hope for is a robust performance of a small set of key underlying functions, needed for the system's resilience and stability (cf. Janssen & Ostrom, 2006; Thompson, 2010). We should focus only on few big underlying aspects of the system's functioning that we know are vital to its survival and elementary stability (Wilson, 2002). This is as much as we can possibly achieve. We need modesty instead of ambition in our efforts to model outcomes through changes in institutions.

If we do focus only on some limited set of vital aspects of the institution's functioning, the otherwise extremely demanding task of institutional analysis becomes somewhat less daunting. First, we obviously lower the complexity of our task because we lower the number of elements we need to trace. We restrict our attention to a limited amount a long-term

features of the system – its internal conditions – and simply exclude from our attention all other specific rules the system might be composed. Anything that does not serve the purpose of elementary viability of the system will be left off the analysis.

Second, if we focus only on the key functions, we somewhat increase our chances to learn important insights comparatively. With all the possible specific design features we observe in the existing institutions, we do not know whether their infringements really matter, as the actors may have designed them for whatever specific purposes they had, perhaps purposes tangent or even opposite to the actual collective purpose of the institution (Downs *et al.*, 1996). So knowing whether existing design does or does not make the actors obey these contingent rules does not matter that much for the analysis. If we want to study institution's design and their performance comparatively, what we need to study instead is whether the key systemic functions that *any* institution must perform are performed. Only performance of these functions, and how institutional design impacts on them, is comparable across institutions, across issue-areas, and over time. The actors cannot decide whether or not they want them to be performed, these functions cannot be by-passed.

Third, even more importantly, focusing on the actors' compliance with the various specific (contingent) rules in the existing institutional setups may be counterproductive. It is easily possible that the actors in fact by-pass the rules, or simply disobey them, precisely in an effort to improve their cooperation, as the specific rules may at that point be constraining and inefficient. Héritier shows convincingly how by-pass of the formal rules in the EU at all enabled propagation of European integration over the years in a situation in which, within the formal institutional framework, divergences of the actors' interests would have lead to a deadlock (Héritier, 1999). Enforcing compliance with these rules, perhaps through strengthening of punishments for infringements, would than harm the system rather than

59

help it. Again, by focusing on the key vital functions we know this cannot happen, and therefore the otherwise very important problem of unintended consequences loses part of its relevance.

So if we limit our ambition to securing of performance of only the vital functions, our task is somewhat easier than the previous general arguments would imply. It should become clear in the course of the text, that the problem of design may be further simplified when we are able to first precisely diagnose where the particular problems we want to address lie. If we focus on how the design helps with performance of only the key vital system functions, and if we are able to diagnose carefully where the shortcomings in their performance lie, devising design solutions may turn out simpler.

In political and social theory, it has not been rare that authors tried to create models that would describe what these key vital functions are. For instance, Almond's list of key system functions developed within his structural-functional analysis is an attempt at a model that identifies the necessary conditions for viability of a political system (e.g. Almond *et al.*, 2003). These conditions can be formulated as performance of the following functions: interest articulation, aggregation, decision-making, rule-implementation, adjudication, political communication and socialization. The earlier Easton's works presenting political systems from the input-output view, and focusing on the communication loops, go in the same direction.

Another alternative could be the Parson's general AGIL scheme (Parsons, 1951), where he presents societies as systems that need – to secure viability and stability – to maintain performance of four key functions. The first is adaptation, or the ability of the society to interact with and adapt to the changes in its environment. The second is goal attainment, the ability to define and pursue important societal goals. The third is integration – maintenance of elementary coherence of the key societal values. The fourth is latency, maintenance of the

elementary institutions serving for transfers of beliefs and values. A related framework, but on a more general level, has been developed by Miller in his monumental Living Systems (Miller, 1978). Miller outlines a general scheme that is supposed to capture the flows of inputs and outputs in all living system, from individual cells, across organs, organism, to societal groups, states, and the international system.

Each of these schemes provides important insights, but they also suffer from significant problems: Parsons' and Miller's works are too general and only loosely related to the study of political systems. Almond's scheme offers an inductive list of functions rather than a unified deductive framework (cf. Finer, 1970). Most importantly, neither of the schemes offers a direct focus on information transmission within the systems, and hence none is as theoretically suitable for my purposes.

If we combine the systemic perspective I advocate in this chapter with the focus on information flows called for in the previous chapter, we are driven to an area that has not received much attention from political scientists, the cybernetic theory, or the science of information flows within systems and of design of systems for regulation. One sub-field of cybernetics, the organizational-cybernetics, offers tools that precisely cater for our needs, i.e. that focus on information flows within complex organizational systems. The so-called viable system model, developed by Stafford Beer, is a model that does the 'job' of joining within a unified scheme the focus on information with the system-needs perspective. It is this model that I will use in the next chapters to solve the problem of the design of international institutions elaborated in this and the previous chapter.

According to Rosenau and Durfee, trying just *to understand* world politics, with all their infinite complexities, is "sheer craziness" (Rosenau & Durfee, 2000, p. 1). Yet, we obviously need to try and do it, as the alternatives to getting at least some understanding

are far worse. But when we want to go beyond understanding and try to model world affairs through changes in design features of the existing institutions – as I assume the actors in world politics do –, we better take the demands imposed on our task by this crazy complexity very seriously. In the next chapter I start outlining the theoretical approach within which we can achieve this difficult task.

Chapter 4

Cybernetic Approach to International Institutions

In this chapter I outline a framework which deals, on the theoretical level, with the problems of institutional analysis (and of design) discussed in the previous chapters. I argued that 1) we need to focus the analysis on the information-transmission problem, and that 2) we need to approach the design analysis from a systemic perspective and only focus on the vital functions of the cooperation systems.

These two arguments, when put together, lead directly to the field of cybernetics, the science that has precisely the problems of information transmission within systems of regulation and control at its very core. One of its subfield, organizational cybernetics, has developed a range of conceptual and theoretical tools with the help of which the problems of information transmission in international cooperation can be fruitfully addressed. In this chapter I discuss the relationship between cybernetics and political institutional analysis, present some basic principles of cybernetics, and outline the core of the cybernetic analysis

of governance systems. By this I finish laying down the theoretical underpinnings for the more practical questions of the study of design and functioning of international institutions I present in part II of the dissertation.

4.1 Information in systems: the realm of cybernetics and the realm of politics

By now, it should be clear that the problem of information transmission is central to international politics and to politics in general, and that – given that politics take place within political systems – how information flows through these systems is essential for their viability. Yet, it turns out that cybernetics offer for this kind of analysis at least as useful tools as are those political science as such has currently at its disposal. I start this chapter by discussing briefly how the two fields – cybernetics and politics – relate to each other.

4.1.1 The general problem of control

Cybernetics is a science of control and communication, as defined by one of its two founding fathers, Norbert Wiener (Wiener, 1948). The science has emerged out of the need to address analogous general problems persistent in a number of otherwise unrelated disciplines, from engineering to physiology. The underlying notion of cybernetics is that the problem of control is essentially a problem of communication, or information transmission (cf. Wiener, 1948; Ashby, 1956).

There is a number of key concepts around which cybernetics are built, among the most popular being the notions of feedback loops and equilibrium. Feedback loops are mechanisms through which a variable P that is affected by another variables R, has an effect back on

variable R, as in

$$R \rightarrow P \wedge P \rightarrow R.$$

If the two relations run in the same direction, the feedback is positive; if the direction are opposite, the feedback is negative. Systems with built in positive feedback are inherently unstable, as even minor increases in levels of certain variables may multiply and drive the system into ever stronger oscillation. A suitable current example of a positive feedback in social matters could be the relationship between the state of the economy and the people's expectations about future economic development. Poor state of the economy triggers poor expectations, this leads to lower consumption and investments levels, and this in turn undermines economic performance of the state; and the cycle begins again.

A negative feedback has exactly the opposite effects, i.e. it tends to keep the system in or close to a stable state, equilibrium. The notion of equilibrium state of some system (whole) can be defined in terms of negative feedback among its parts: "the whole is at a state of equilibrium if and only if each part is at a state of equilibrium in the conditions provided by the other part" (Ashby, 1956, p. 83). In principle, most existing systems are most of the times in equilibrium since, as noted by Weinberg, one of the most important insights of general systemic analysis is the vastly under-appreciated fact that most of the times *nothing of any significance happens* (Weinberg & Weinberg, 1980, p. 3). Most systems do have strong negative feedback mechanisms, otherwise they would disintegrate under even weak disturbances from the environment.

The standard example of early self-regulating systems with a negative feedback loop is

the mechanical centrifugal governor used since the 17th century and perfected for steam engines by James Watt. The purpose of the governor, depicted in figure 4.1, is to maintain a close to constant speed of an engine, regardless of the demanded power (as the demanded power varies e.g. due to the weight of the train load or due to the slope of the rails). The governor consists of a simple negative feedback loop. When the speed of the engine – transmitted to the governor through rotation of the vertical axis in the figure – rises, the increased centrifugal force pushes the ball weights further from the axis. This movement is mechanically transmitted (here via a leverage) to the valve that decreases the amount of fuel incoming to the engine. In the opposite direction, at low speeds the gravity force pushes the ball weights down, and this causes the leverage to open the valve, as depicted.
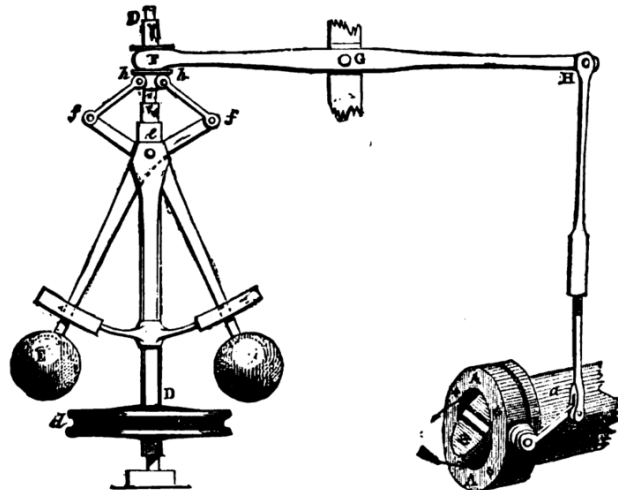
FIG. 4.—*Governor and Throttle-Valve.*

Figure 4.1: Centrifugal governor

The steam engine governor may seem unrelated to the problem of political institutions but it illustrates the underlying commonality of goals of political institutions as mechanisms regulating interaction among the political actors and of mechanical regulatory devices – in both cases the key to functioning of the system lies in its ability to detect the levels of

66

certain variables of interest (in this case speed of the engine), and through some mechanism devise an appropriate reaction that would keep the system as a whole in some desired range around its equilibrium. As I will discuss later on, this is precisely what political systems do, be it in terms of maintaining stable levels of support from citizens, securing of a balance of powers among states, or – as will be our primary interest – maintaining at least the minimal necessary levels of inputs into the international cooperation schemes.

The problems of cybernetics are in principle co-terminus with the notion of governance. Both terms stem from the same Greek origin, in particular from the Greek word for *steersman*, and the label of cybernetics was intentionally devised by Wiener and his colleagues in this way (Wiener, 1948, p. 11). It is worth mentioning that despite its only relatively recent popularity in political research, the term governance has a truly venerable pedigree, dating back to Plato and Aristotle.

In spite of the possibly exotic notion of a relationship between cybernetics and politics, there is a close correspondence between the two disciplines, given precisely by the general interest in how functioning of systems – mechanical, social, or political – can be regulated in some productive way, or how the systems can be governed. The important thing is that this correspondence does not stop on the ontological level, with observation of a vaguely defined common interest in 1) flows of information 2) within systems. Numerous works discussing the problems of governance and politics stop at this abstract level and only mention, in passim, that governance and cybernetics actually stem from the same origin (Rosenau, 1995). My key target in this work is to show that the correspondence indeed does go much deeper and in particular that it can be fruitfully exploited for political institutional analysis. This is what I try to show later in this chapter and in chapter 6.

To start with, making use of insights from cybernetics is nothing new in politics, as a

comprehensive cybernetic framework has been introduced into political analysis already by Karl Deutsch in his Nerves of the Government (Deutsch, 1963). In Deutsch's view, political systems can be conceptualized as "self-modifying communications network[s]" or as "learning nets" Deutsch, 1963, p. 80. The key function of any political system, then, is to secure flows of the necessary information among the relevant actors. Unless that can be secured, the system as such is threatened. Also based on cybernetics, he develops the notion of political power as an ability to afford *not to learn*, a concept revolutionary at that point in its relational nature (Albert & Cederman, 2010, p. 6). Although the cybernetic model of Deutsch's has certainly been overshadowed by his research on *security communities*, we should note that this transactionalist theory is also in principle based on the same underlying notion of communication as the key to politics (Deutsch *et al.*, 1957).

The cybernetic approach was adopted also by social theorists, most notably by Luhmann. For Luhman, societies *are* systems of communication, i.e. there is nothing more to social systems than communication among their members (Luhmann, 1995; cf. Seidl & Schoeneborn, 2010). This Luhman's view corresponds with the turn to the so-called second order cybernetics in 1970, where the attention of the cybernetic research moved from the study of mechanical systems to the study of societal systems (cf. Clarke, 2011)[1]. A small field of the so-called socio-cybernetics, branch of social theory interested in mechanisms through which societies maintain coherence and reproduce themselves, exists within sociology (cf. Geyer & Zouwen, 2001). The cybernetic approach has in fact been developed also into a specific decision-making theory, on a similar level of generality as for example the rational choice theory, but with the specific focus on the problem of uncertainty and learning (Steinbruner, 2002). The core of the cybernetic theory of decision-making lies in the acknowledgment of the enormous complexity the decision-makers face, and the essential impact of how they are

able to process the input information on what decisions they in the end adopt.

All these fields share a common interest in the problem of control and regulation mechanisms, where the essential function of the system is monitoring of its internal processes as well as of its environment, and formulation and implementation of appropriate responses.

4.1.2 Cybernetics and related disciplines

It may be worth clarifying at this point the relationship between cybernetics and some related disciplines, notably the general systems theory. In principle, cybernetics and the general systems theory developed approximately at the same time (late 1940s and early 1950s) and adopted closely related systemic approaches. Yet, while the general systems theory was developed mostly by biologists – such as Ludwig von Bertalanffy (Bertalanffy, 1968) – creation of cybernetics was primarily driven by the demands of engineering[2]. In some way, cybernetics – despite its extremely ambitious project – could be seen as a subdiscipline of the general systems theory, as cybernetics are only concerned with systems of control (or regulation), while the general systems theory does not limit itself in this way (Weinberg, 1975). In other perspective, though, the two disciplines are essentially dealing with the same problems and differ mostly in from where they come to them. In political research, cybernetics inspired first and foremost Deutsch, the general systems theory was built upon e.g. by Easton and Almond.

---

[1]To be more precise, the move to second-order cybernetics signifies an increasing attention of cyberneticians not only to mechanical systems observed from outside, but also to the fact that the observers are always themselves necessarily parts of the investigated scheme. In other words, the observer of the system – a cybernetician – is also a cybernetic system itself, that interacts with the object of its investigation. Such an approach directly enables the cybernetic study of social systems.

[2]Concretely, for example, Wiener started dealing with what would now be considered as cybernetic problems during the Second World War, when devising mechanisms for control of the anti-aircraft defence. In this case, the problem was how to devise a mechanism that would predict the position of the airplane in some future moment, and adjust the rotations of the defence gun to the actual (changing) speed and trajectory of the airplane.

Complexity theory, and in particular the study of complex adaptive systems (CASs), can be seen as the today's heir of cybernetics. The multidisciplinary research on CASs, the bulk of which is carried at the Santa Fe Institute in the United States, focuses primarily on the questions of how complex adaptive systems – systems with the ability to learn from the environment – collect and process information, and in particular on what general problems this activity entails (Gell-Mann, 2002).[3]

Given that both cybernetics and the general systems theory enjoyed in the 1950s and 60s a considerable degree of popularity, an obvious question arises as to why it might be worth returning to them – in this case to cybernetics – despite the general decline in attention paid to them. Is not the entire field of cybernetics, similarly to the other multidisciplinary approaches such as the general systems theory or chaos theory, just a bubble that after a while bursts without really leaving in the research any tangible traces (cf. Rosser, 1999)?

In principle, the approach I propose here is agnostic as to the further potential of the multidisciplinary approaches such as cybernetics for uncovering new general laws. Clearly, part of the decline of the interest in cybernetics has been driven by the fact that in some disciplines the concepts from cybernetics have been applied much more easily and successfully than in others – artificial intelligence being the prime example – and hence that the core of cybernetic research is being done back within the confines of individual disciplines. This can, obviously, reflect the institutional aspects of how science is done at least as much as it reflects the potential of the discipline of cybernetics as such. Furthermore, the CASs research is in both ontological terms (focus on information) and in terms of the ambition at maximum generality a clear descendant discipline of cybernetics; the fact that it distinguishes itself from cybernetics or the general systems theory may well be given by the different roots from

---

[3]In the Czech Republic, the research on CASs is conducted for example at the Center for Theoretical Study, the joint center of the Charles University and of the Academy of Sciences.

which it grows – in this case mostly physics and biology.

More importantly, the question of the potential of cybernetics as of a general multidisciplinary approach is not really in any way important for the work I present here. I build on some recognized aspects of the political reality of international relations – in particular their enormous complexity and the key role of information for cooperation among states – and seek what tools various disciplines offer for the study of systems of such characteristics. As cybernetics were developed precisely to address the problems of information flows in systems, they can offer the entire conceptual apparatus suitable for such a research.

Beyond the conceptual apparatus, however, cybernetics and its related disciplines – such as its today's heir the theory of CASs and of emergence – provide general insights into the nature of systems, i.e. into how any system of certain characteristics behaves, no matter whether it is biological, mechanical, or social. To be sure, these insights may have higher or lower predictive power for the particular research tasks we are interested in, as the effects described by these insights may be made effectively overruled by other important factors, but this is the case with any theory. To give an example, if we believe that international politics are complex, than the insights from complexity theory *by definition* do apply to the phenomena we study (Rosenau, 1997). In other words, we may be more or less optimistic as to the potential of the kind of multidisciplinary enterprises as the general systems theory, cybernetics, or complex adaptive systems, but the insights already made within these fields are valid and social science needs to take them into account, as it needs to take into account the constraints imposed on social reality by the physical laws (Wendt, 2006). In sum, I will build here on what has already been proven and what has potentially enormous consequences for political research, even though it is usually neglected by the mainstream discipline.

As a matter of fact, though, even in recent years we have seen in IR also signs of a revival of

71

interest in the truly general theories of systems, exemplified by the edited volume by Albert, Cedermand and Wendt (Albert *et al.*, 2010). In this volume, a number of perspectives is offered on how the study of systems in IR, as opposed to the constituent units (usually the states), is necessary for explanations of some of the most important phenomena. Cederman, for instance, argues that it is in principle impossible to understand the revolutionary changes international politics have been undergoing without adopting a systemic perspective, and in particular insights from the multidisciplinary research on complex systems (Cederman, 2010).

A similar shift back to social systems theories – but on a higher level – can be seen in the already mentioned area of the complex adaptive systems. Miller and Page, in a recent book, review a whole range of works that actually adopt the CAS perspective and try to explain important social phenomena on the basis of its conceptual and methodological apparatus (Miller & Page, 2010). One should also not forget that one of the key insights in IR of the 1980s, namely the Axelrod's analyses of cooperation, is in principle built on the models similar in kind to those used in CASs, i.e. computational models in which we study the aggregate outcomes of a highly iterated interaction of individual actors (Axelrod, 1981).

4.1.3   Systemic and individual approaches

The discussion so far may have already suggested one important point that needs to be clarified as regards the ontological assumptions of my approach, and in particular the relationship between actors and the systems. I have started in chapter 2 with a very mainstream rationalist account of international politics, and in particular with the neoliberal institutionalist and neorealist schools. Later on, however, I have increasingly argued for a more systemic approach, one focusing on the vital functions of the political systems of international regimes we study. Are these approaches compatible?

First of all, I already indicated that my approach does not need to side either with structure or agents – I do not ascribe any powers to the system as such through which it could enforce its own logic over the will of the actors. But the system does have its needs, i.e. it only operates well if the actors behave in certain ways. In other words, I assume that the system has its own logic of functioning and that if this logic is disrespected by the actors, the system will cease to perform the functions the actors expect it to perform and ultimately that it will collapse. If the actors want to maintain the system – e.g. a political system of cooperation in a given international issue area – they must behave in the ways the system requires. Hence, similarly to the kind of Waltzian analysis, I assume that the system induces – more or less effectively – the actors to behave in certain ways. If the actors do understand how the system operates, they can improve their lots by following its logic. If they do not understand, the system on its own does not force them, but the outcomes of the actors' actions will be suboptimal for them. Rational actors should be able to realize the pressures generated by the system (the structure) and adjust their behaviour to it. In this sense, the system – the structure – is not external to the actors, it is constituted by them but, as any system, it supersedes in its functioning the sum of the behaviour of the actors (cf. Waltz, 1979; Wendt, 1999). There is no ontological boarder between the system and the actors, they are mutually constituted.

Hence, if I treat international institutions – or specifically international regimes – as vital systems I am not referring to any their innate ability to secure their survival. I am referring to these systems as being composed of individual rational actors who, because (and if) it is in their interest, act in ways that are expected to maintain viability of the system. It is always the individual actors who chose strategies, and through their interaction the system emerges and persists or perishes. But if it is in their interest that the system is viable, so for

example that an international regime for cooperation in a specific area persists, they will have incentives to invest resources into maintaining its viability. There is an enormous difference between such a agency-driven perspective on the system and its role, and the original, e.g. physiological concepts of systems in which every subsystem's role is determined in a top-down manner by the overarching system. I do not claim that the need for cooperation will force the states to act in certain ways; they may well decide to disregard the needs of the cooperation schemes as they are agents in their own right, not merely parts of the larger system. The systemic cybernetics-based perspective I elaborate on enables us to identify the key specific functions the international cooperation schemes need to perform, for cooperation to be viable. It does not mean they will necessarily be performed, it only means that when the actors disregard them and do not act so as to ensure their performance, cooperation will not be maintained. In this sense, the systemic perspective provides the reference point, or how the actors should behave, and not necessarily how they behave in the actual empirical reality.

Second, I should clarify that the starting point of my analysis is always a set of individuals that rationally pursue their interests under the conditions of *uncertainty* and for whom one of the key tasks is to make sure that they can navigate through this uncertainty[4]. This means, they need to be able to collect the relevant information, process it, update their beliefs, and act accordingly. The level of uncertainty the actors face varies. Uncertainty may concern the other actors' preferences, the states of the world and the causal connections operating within it, or uncertainty about the already realized behaviour of others (Koremenos *et al.*, 2001, p. 773). Uncertainty may go even one step further, as indicated in chapter 3: actors may be uncertain about what they themselves will want in the future. More precisely, what institutional arrangements are likely to bring them in the future higher payoffs, hence the

74

notion of veil of uncertainty (Young, 1989). The principle problem the actors face is how to reduce this uncertainty. Again, such a view is perfectly consistent with the mainstream rationalist accounts, even though in many contexts much of the literature tends to make stronger assumptions about what all information the actors hold. For my purposes, the level of uncertainty – given by the available information – is a variable or a parameter, depending on the purpose of the analysis.

Third and perhaps most importantly, I would argue that there is no contradiction between my individualistic point of departure – the rationalist mainstream – and the more systemic approach I move towards from chapter 3 onwards also when it comes to the problem of general versus partial equilibrium perspective. Lake and Powell discuss this problem convincingly: the mainstream rationalist strategic perspective usually assumes the partial equilibrium perspective, i.e. that not everything is significantly related to everything else, and that it is often worth putting aside some feedback loops within the given system, for the sake of clarity of the models we build (Lake & Powell, 1999, p. 17). The general equilibrium perspective, on the other hand, does not assume this, it tries to incorporate all the relevant actors, with all their links and interactions, into the model. These two types of analysis usually lead to very different forms of outcomes – highly simplified 'cut to the bone' formal models versus relatively richer and more complex, often narrative descriptions.

I would like to stress, and I build on the fact, that these are not two distinct positions, but only two ends of a continuum on which most empirical studies are somewhere in between the extremes. I do believe it is often useful to simplify the situations we study and focus only on the key actors; yet if we empirically know that there are many important actors, and that

---

[4]Technically speaking, the actors act under the conditions of risk and not uncertainty, as the former implies that they are able to at least establish with some confidence the probabilities of outcomes, while the latter would imply that they perceive possible outcomes as completely random.

they do interact in complex ways, we have no other option but to take one or several steps in the more systemic direction. We will certainly lose on clarity of representation by introducing the complexities and the feedback loops into the analysis, but we will maintain elementary level of correspondence between our models and the reality we try to capture with them (see Albert *et al.*, 2010, Introduction). In fact Lake and Powell themselves argue in their text – the flagship of the mainstream strategic rationalist perspective – that "[a]n important goal of future research [...] is to deepen our understanding by incorporating the significant feedback channels which our analyses may have initially excluded" (Lake & Powell, 1999, p. 18). So my analysis departs from the very same strategic analytical approach characteristic for the rationalist mainstream, but by taking seriously the complexity of the environments many of the existing studies bracket it tries to paint a more faithful picture of the reality. To be sure, this is mostly a meta-theoretical question, but clearly an important, as so much of research in IR in the last two decades has been driven by such meta-theoretical problems (Fearon & Wendt, 2002).

There is one definitional note that it might be worth making for the sake of clarity of the further discussion: how we define the system we are interested in in our analysis, or more precisely on what level we locate it, is purely an analytical convenience. When we are interested in the design and functioning of international regimes and IOs, it is these regimes and IOs that constitute the systems of interest. Yet, the states who act within these systems are themselves systems, on a lower level, embedded in the system of the regime. This is nothing new: each state forms a political system, in fact it is a political system. A number of such systems, when engaging in a stable cooperation in international politics, together form a higher level system. From the other side, a national political system contains in itself a number of regional or local political systems. By definition, each system is composed of

other systems on a lower level, as regimes are composed of the member states. This recursion can go high and low, and it is only the specific purpose of the analysis that determines what the system of our interest will be, and hence what will be the systems on the lower level that actually constitute it.[5]

In general, this section should have shown that approaching the problem of design from the cybernetic perspective should not raise any suspicions. Cybernetics is a general multi-disciplinary science, and the political realities are ruled by cybernetic principles no less than any other aspects of the World. Cybernetics are well suited for providing positive solutions to the problems I identified in the previous two chapters, problems that might otherwise lead us to an altogether dismissal of the entire research programme on institutional design. In the next two sections I lay down the building blocks on which these positive solutions could be found; in the successive chapters I try to show that the solutions indeed do exists, and what they are.

4.2 Information theory, cybernetics, and the problem of control

The key problem of cybernetics is that of information transmission within the controlling system (the regulator) so it is not surprising that the ancestor of modern cybernetics lies in the information theory, a branch of applied mathematics pioneered by Wiener himself and especially by the British mathematician Claude Shannon. In his foundational *The Mathe-*

---

[5]I use the language of some systems being embedded in others when I refer to schemes where some clear hierarchy can be discerned, e.g. where the relationship between an international institution and its many member states is concerned. When we speak about systems on the same level (e.g. about a group of states), the terminology of their functional overlaps is more appropriate.

matical Theory of Communication (1949), Shannon developed a comprehensive framework for dealing with the principle problem of actually how to measure information flows and, in the next step, how to assess sufficiency of the existing information channels for transmission of the necessary amount of information. It should be clear by now, that these questions lie also at the very core of my interest in this text. The framework is built around two inter-related concepts: information entropy, or simply entropy, and capacity.

Entropy, as defined by Shannon (Shannon & Weaver, 1949, p. 48) and alternatively by Norbert Wiener (Wiener, 1948), is a measure of the information content of a signal or a message. Given a particular set of available signals, for example some alphabet, it is possible to calculate the information content of a transmitted signal. Such a measure is usually given in bits (abbreviation for binary digits), where one bit conveys information about a single binary choice (0 or 1, yes or no, etc.). That means, information is understood as a signal, or selection of a message, that identifies, for the receiver, a single specific outcome from a set of possible outcomes, i.e. that tells which specific element of a set of possibilities was materialized[6]. A letter in a text is one such signal, that has been chosen by the sender from a set of approximately 25 possible signals, the alphabet.

Formally, entropy is defined by Shannon as

$$H = -K \sum_i p_i \log p_i, \tag{4.1}$$

---

[6]Information can have various physical manifestations, but on the most general level information is a stable pattern identifiable in the transmission medium. The pattern can be carried by varying electrical current in a wire, as in telephone, by radio waves of various length, as in radio broadcast, TV broadcast, or mobile communication, or e.g. by interrupted masses of smoke from the Indian fires. The notion of information as of a pattern, irrespective of the material that carries it, corresponds also to the etymological origin of the word *information*, coming from Latin *to give form*.

where K is a constant giving the unit in which entropy H is measured, e.g. in bits or nuts, and p_i is the probability that signal i is used, out of all the signals available in the alphabet. The sum sign then sums the entropies of the individual signals into a single value consisting of i signals (e.g. letters). The negative sign in the equation may be confusing, but it is given by the mathematical property of the logarithmic function. Since each of the probabilities p_i will lie between 0 and 1, and since logarithm of a number between 0 and 1 is always negative, each component of the sum and therefore the sum as well will be negative. The negative sign in front of the sum then secures that the equation 'goes' in the right dimension; the higher the information content of a signal or message, the higher its entropy[7].

Shannon observed that the information content of a particular signal, or its entropy, is determined by the size of the available alphabet and the probabilities with which the signals are used. Quite intuitively, he found that the more different signals are available in the alphabet (e.g. letters) the higher the information content of each of them, as knowing the chosen signal (letter) conveys information about a choice from a larger set of possibilities, and thus reduces our uncertainty more. For example, the maximum possible information content of a single letter selected from an alphabet consisting of 32 letters is 5 bits, as getting to know the selected letter would take 5 binary choices ($2^5 = 32$ or, perhaps better, the other way round $\log_2(32) = 5$.) With an alphabet of 16 letters, the maximum achievable information content of a single letter is 4 bits ($\log_2(16) = 4$.)

---

[7]The term entropy has been chosen by both Shannon and Wiener due to the structural equivalence between the quantity they discovered to be the right measure of information content and the notion of entropy in physics, in particular in the second law of thermodynamics. According to this law, informally, in closed systems the key physical and chemical properties of its parts, such as temperature and pressure, tend to equate over time. The situation in which all the key properties have the same values across all the parts is the state of maximum entropy. The second law dictates that closed systems tend towards this state. So, for example, we can pour two masses of water into a container that is then isolated from its environment. If one of the masses is hot and the other cold, over time all segments of the water volume will converge on the same temperature, somewhere between 'cold' and 'hot'.

Shannon also found that the information content of a signal is highest when the probabilities of use of each of the signals available in the alphabet are equal. If the probability that some signal occurs is (on average) higher than that of the others, such as that of letter e in the English language, the information content contained in that signal is lower than the maximum available. So, for example, in the alphabet of eight letters a, b, c, d, e, f, g, h, if all the letters have equal probability of occurrence ($p = 0.125$), a single letter message conveys the information of exactly 3 bits (as it takes 3 bits to identify a choice from eight possible outcomes). If however, the probability of, say, letter f is higher ($p_f = 0.3$), and that of all the other letters correspondingly lower ($p_{-f} = 0.1$), the information content of a single letter message is on average only approximately 2.85 bits. Since in no natural language the probabilities of occurrence of all its letters are equal, they contain large redundancies. Shannon estimated that the redundancy of the English language is about 50%, i.e. that each letter in an English text contains only about half of the information it could contain in theory, if all the letters appeared in English with the same frequency. Parts of this redundancy serve a good purpose as it prevents misunderstandings due to small random errors in transmission, but nevertheless it is there.

We should stress one very important feature of Shannon's analysis, that is already implied in the note about the 50% redundancy of the English language. In his information theory, information and meaning are two strictly distinct concepts. A sequence of letters 'XQZKWOWB' has as much information content as the sequence 'POLITICS'.[8] This on the one hand allows for the most general use of the concepts – information of any type, in any format, in natural languages as well as in artificial codes can be captured – but its use for social analysis is obviously limited as what is usually more important than the amount of bits in a particular message is the amount of meaningful information, its credibility, simply

its content.

Given entropy as the measure of information content, Shannon defines the concept we will be really interested in, capacity of a channel. Capacity is conceptualized as the amount of information a particular channel is able to transmit in a given period of time. I will devote an entire chapter to the concept of capacity later in the text (chapter 5), here it is sufficient to say that capacity of a channel gives the maximum information content that can be transmitted through it in a given unit of time.

Now relating back to the cybernetic framework, when we know that the problem of control or governance entails that of transmission of information (e.g. instructions), it does not surprise that the concept of capacity needs to apply in cybernetics as well, with only a slightly broadened meaning. In cybernetics the primary concern is not transmission of information *per se*, but of information as a means to regulation, or perhaps *the* means. The task of regulation consists in that the governing system changes its own state, in an effort to induce change in the governed system. To take the example from Ashby's, the traffic lights as the governing system regulates traffic by changing their own state from Red to Red+Orange, to Green, and through Orange back to Red. It is through the change in the state of the regulator that the subject of regulation can be (with more or less success) induced to change its own state. The regulatory capacity of a governing system is then defined by, and limited to, the number of states the governing system itself can adopt, in the given period of time. Ashby defines this quantity as variety ($\mathcal{V}$). In the above case, the variety of the traffic lights is $\mathcal{V}$(Red; Red+Orange; Green; Orange) $= 4$.

Although the terms used are somewhat different from those in the information theory,

---

[8]In fact it most probably has higher content as the letters used in the former sequence are less usual in the English language than those in the word POLITICS, and hence have statistically speaking higher information content.

the substantive meaning of the concept of capacity is in both fields equivalent, as is the mathematical delimitation (Ashby, 1956). In both cases capacity is given by the variety of the system, that means by how many different states the system can be in in the given period, and how many times it can change its own state to another one (from 0 to 1, from Red to Green, etc.).

In our analysis we are really concerned only with information and not regulation as such, because in politics there is no one-to-one correspondence between instructions (rules) and the regulatory effect on behaviour, as discussed in the previous chapters. So in the world where the only thing that can really be directly influenced by the design of the institutions is how effectively they transmit the instructions, and not how the actors in the end behave, the notion of capacity really needs to boil down to how much information can be transmitted in the given time unit through the given communication channel. Traffic lights, for example, can transmit to the drivers instructions for four different types of behaviour, but they cannot force them to follow these instructions. This is obviously very different e.g. from the engineering context, where – at least in most situations – the engineers can assume that if the system is well designed the component parts actually will follow the instructions, as they do not have the will to disobey them.

The last, but certainly the most important insight from cybernetics we need to review to be able to derive the arguments that come later concerns the so-called Law of Requisite Variety (LoRV), or Ashby's law. With a significant degree of simplification, LoRV implies that any regulatory system must have at least as high variety as the system it is supposed to regulate. That means, a regulator needs to be able to take on at least as many different states, within a given time unit, as the regulated system. If it cannot do so, it may happen that the regulated system moves or is pushed into some undesirable state, perhaps into

a state which threatens its very existence, and the regulator will not be able to respond to it. We may recall the notion of equilibrium discussed above; if a system is to be able to maintain its internal equilibrium, and thus secure viability, it needs to have sufficient variety to respond to all possible disturbances coming from outside (or from it own internal functioning, for that matter). It has been shown that this principle implies that "[]every good regulator of a system must be a model of that system", or what is referred to as the good regulator theorem (Conant & Ashby, 1970). The LoRV is considered a natural law in that it is a principle (logical) feature of any systems and any regulatory situation, its effects cannot be avoided (Ashby, 1956, p. 208).

In the context of the notion of capacity, LoRV dictates that a system that is to be viable needs to possess sufficient information transmission capacity so as to be able, among other things, to transmit to its parts instruction at a sufficient speed. In the traffic lights example above, the standard traffic lights do not have sufficient variety to inform the drivers that they should slow down – the limitations of their design do not allow them do transmit this instruction, they do not have the capacity to do so. In the context of international institutional analysis, LoRV will dictate that the information channels forming the institutions need to have sufficient capacity to transmit information about the complexities of the problems that the institutions are supposed to regulate. In other words, LoRV and the cybernetics built on it will help us determine whether the institutions we have or we plan to design will be at all able to fulfil their functions. As noted by Ashby more than 50 years ago, "[c]ybernetics offers the hope of providing effective methods for the study, and control, of systems that are intrinsically extremely complex" (Ashby, 1956, pp. 5-6).

4.3 Cybernetics of organization: introducing the viable system model

Following the previous brief exposition of the key concepts of information theory and cybernetics, in this section I present the building blocks of the so-called viable system model (VSM), a model developed by the British cybernetician Stafford Beer that specifies all the information channels in any viable regulatory system. The specifics of the model will be elaborated upon in chapter 6. I should stress that I will restrict myself to presentation of only those components of the model that are essential for my very purpose of outlining the key channels every regulatory system, and hence any international political system or an international regime, needs. VSM as such is much richer and provides a whole range of other important insights. A somewhat more detailed view can be obtained from several of the article-long introductions to VSM (Schwaninger, 2006; Yolles, 2003, 2005; Malik, 2011) or at best from Beer's monographs, most prominently *The Heart of Enterprise* (1979) or *Brain of the Firm* (1972; see also Beer, 1984).

The entire VSM is derived from the fundamental principle of cybernetics introduced above, the Law of Requisite Variety (LoRV). The very purpose of the VSM is to propose a set of general requirements, or guidelines, for how regulatory schemes need to function, so that they accommodate the implications of the LoRV in a productive way. VSM gives a set of necessary conditions for viability of any system, in the face of pressures given by the LoRV and the need of the system to deal with the complexity of its environment.

On the most general level, the way VSM is derived from LoRV is surprisingly simple (cf. Beer, 1979). The starting point of the analysis is the inherent complexity of the world or of the environment in which the systems we study operate. These may be e.g. firms on the

84

market, departments within large organizations, governments in domestic politics, or states in international politics. Because of the world's complexity, decision-making (governance) needs to be distributed across potentially numerous levels of the governing scheme. As each level, and in fact any decision-making body, can process only a limited amount of information, and hence effectively cope with only a limited amount of the complexity, all decisions that can be done at the lower level need to be done there. In other words, no information that is not of direct concern to the higher levels should actually be transmitted there, as otherwise the higher levels become overloaded or, more usually and certainly worse, they simply make decisions without first processing all the relevant information. We should note here that precisely this – i.e. making the decisions without the necessary information – is the situation in which the LoRV propagates itself in a destructive manner. In other words, the LoRV is a law, it applies no matter how the system is designed, and if it is designed poorly, so that the decision-makers somewhere in the system need to make decisions without having the capacity to process all the relevant information, the LoRV will manifest itself through the ignorance of their decisions and through the potentially harmful consequences of the ignorant decisions (Beer, 1979, p. 96).

The key to successful design then lies in creation of mechanisms through which the enormous complexity of the world is gradually reduced to only the most essential pieces of information that pass to the topmost levels in the regulatory system. Parts of the complexity – the variety of the environment – are absorbed at each level of the system, and only those pieces of information that are essential for the decision-makers on the higher levels are passed on. to give a concrete example, a minister does not need to know all the details of each of his official's work – it is the task of the various management levels to ensure that the officials do what the collective purpose requires and they only pass to the minister the information

that is actually relevant to the decisions that the minister should do. So the creation of many level in the system (the so-called recursion, as I discuss in chapter 6) is one means that enables it to cope with all the complexity.

The second means is what is referred to by Beer as variety amplifiers and attenuators. These are mechanisms that effectively translate the information, after being processed at one level, into a different code, usually more compact, that is more suitable for the higher levels. Again to give an example, a social policy official acquires in the terrain work a large number of detailed information on the state of a certain problem in a given area – say drug addiction prevalence in a city district. All this information may be highly relevant for the decisions this official has in his or her competence. For the purposes of the higher level management, though, the enormous amount of information may need to be reduced perhaps to a single number or a single index summarizing a whole range of factors. Since the higher level needs to process inputs from a number of lower level officials, it is important that it receives only the key information it needs, and that the information comes in a concise form, e.g. in a form of an index or some specific aggregate statistic.

The design problem lies in that if the number or the index reflects all the information that is necessary for the decision-makers higher in the hierarchy, the system will work well (in this case it is appropriate that the index is used, and its particular construction is good). A well designed index, or a summary statistic, is an appropriate variety attenuator, that translates the enormous variety of the empirical reality (observed by the lower level official) into a much more concise form of certain summary statistic (suitable for more aggregate decisions). If the index does not capture all the information that is necessary for the higher levels, and yet the index is the only information it has, the higher level does not have the requisite variety. Due to its ignorance, the higher level will make poor decisions.

So, given these restrictions imposed on governance by LoRV, how do we set up a scheme of information channels in such a way that those who decide have all the information they need? How do we design the communication network that will enable, in the international political context, the actors that cooperate in pursuit of some shared goal to transmit among themselves all the information they need?

It turns out that to model such a scheme we need a set of five different types of Systems. In figure 6.4 I present the graphical summary of the model, as adopted from Beer (1972, p. 130). The entire scheme appears, at first, as a complicated tangle of connections, but it is carefully deductively derived according to the principles imposed on any governance mechanism by the LoRV. There are several Systems One (A,B,C,D) with their managements (1A, 1B,1C,1D) that jointly form the overall System. For example, there are four states A, B, C, and D, with their governments, 1A, 1B, 1C, and 1D, forming together an international regime (the cooperation scheme, composed of Systems 2-5).

Each of the five Systems needs to perform certain specific functions for the overall System to be viable, and I elaborate on these in chapter 6. At this point, however, I would like to briefly discuss two key aspects of the model that cannot be apparent from the discussion so far, but will become clear in the later chapters. They deal with the two tasks the model will do for our analysis of design in the future chapters, i.e. why it is at all useful to have such a model, in the context of the discussion so far.

First, what the scheme really does is that it depicts in a summarizing way all the five systems (again, for illustration, several Systems One – A, B, C, and D – are included) as well as the required connections between them. The value added of the scheme, then, is that it explicates the list of necessary Systems, each with its function. Unless a cooperation scheme possesses the tools to perform the functions performed in the model by each of the
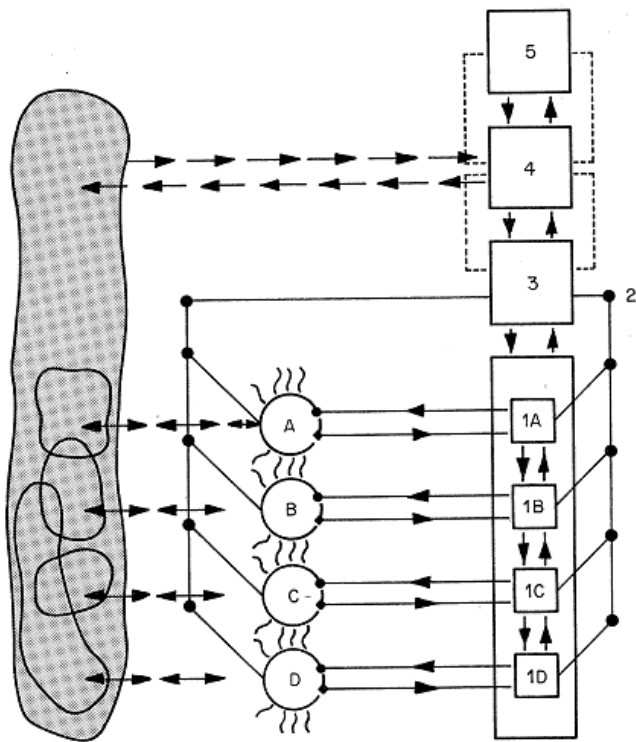
Figure 4.2: Viable System Model scheme

five Systems, it will not be viable. The decision-makers will not be able to process all the information about their internal and external environment they need for qualified decisions.

Second, VSM allows us to directly specify the required capacity of each of these channels. This is something that the figure does not indicate well as all the connection look alike, but the model itself does identify the differences, as we will see later. With VSM, what we have is a general scheme for the study of the IOs' capacity from the information-based organizational cybernetics perspective. The model gives us a guidance for assessment of whether all the key information channels among the actors work.

I would like to stress how powerful a statement this is: *the model gives us a guidance for assessment of whether all the key information channels among the actors work.* We have learned in the previous chapter that what institutions can do for the actors' cooperation is better transmission of information. VSM gives us the tool to diagnose whether the institutions do it. When we want to design international institutions in a way that makes them viable, here we have the blueprint.

In summary, in this first part of the dissertation my task was to outline a range of key theoretical problems related to the question of successful institutional design. In the next part I translate this theoretical debate into more concrete insights into how we can conceptualize and measure capacity of international institutions (chapter 5), what particular channels, based on VSM, we should look at (chapter 6), and what all this implies for the actual design tools we should focus on in the empirical research (chapter 7).

Part II

The Organizational-Cybernetic

Approach

Chapter 5

Institutional Capacity: the Concept

The first target of a systematic analysis of effects of design on functioning of institutions needs to be the definition of the dependent variable. We need to have a concept that captures how well the institutions perform the functions we want them to perform. I outlined above the argument that ultimately the only thing we may want the institutions to do is to help transmit the important information that would otherwise be for the actors too costly or impossible to obtain. In this section I develop the concept of the institutions' ability to deliver on their tasks in detail.

5.1 Defining the dependent variable of institutions' functioning: the existing alternatives

Several good candidates for the right dependent variable can be found in the IR literature. In the earlier literature the fairly general notion of regime *strength* was often used (Ruggie, 1982), although the specific meaning of strength seems to vary somewhat across its users (cf.

Keohane, 1989, p. 15 and Hasenclever *et al.*, 1997, p. 2). Since early 1990s the more specific concept of regime *effectiveness* has been developed, and considerable amount of research has been guided by it (Underdal, 1992; Helm & Sprinz, 2000; Young, 2001; Hovi *et al.*, 2003; Breitmeier *et al.*, 2006). Effectiveness can be understood as ability of a regime to induce changes in the behaviour of the relevant actors or in the physical state of the world.

Recently the research agenda has been turning to the relatively comprehensive notion of international organizations' performance, understood most often also as ability of the institutions to contribute to solution of the given problems, but specifically addressing the problems of functioning of international organizations (Gutner & Thompson, 2010). The approaches taken in this strand of literature reflect this focus on the organizational nature of IOs, ranging from constructivist analysis of bureaucratic culture (Barnett & Finnemore, 2004) to public choice oriented works on IOs as budget maximizers (Vaubel, 1986) and as purposeful autonomous agents able to diverge in their actions from the goals of their principals (Reinalda & Verbeek, 1998; Hawkins *et al.*, 2006).

A very important concept on which I build further in the text is that of institutions' *robustness*, where we consider as robust those institutions that are persistent under varying constellations of power and interests (Shepsle, 2006). Perhaps the closest alternative in the mainstream IR literature would be the notion of institutional stability (Rosendorff, 2005; Thompson, 2010) or, in a more general sense, that of regime resilience and sustainability, particularly frequently used in the environmental regimes literature (Janssen & Ostrom, 2006; Ostrom, 2009).

Last but not least, for a whole range of works the key dependent variable defining institutions' functioning has been compliance, i.e. the degree to which the actors obey the rules to which they committed themselves (cf. Raustiala & Slaughter, 2002). Again, a number

of both conceptual and theoretical studies (Chayes & Chayes, 1993; Downs *et al.*, 1996), as well as empirical applications (Mitchell, 1994; Zürn & Joerges, 2005), have been guided by this concept.

Each of these concepts attempts, in one way or another, to capture the degree to which institutions bring to the primary actors' interaction some desirable effects. None of them is, however, on its own suitable for the type of analysis that I pursue, that is analysis aimed at understanding of the *causal effects of the institutions' design on their ability to enhance the prospects for cooperation*, taking into account the limitations imposed by the power- and interest-driven nature of international politics.

For that purpose I propose a concept which I label as *institutional capacity*, referring to the maximum ability of the institutions to transmit information among the actors. This meaning directly reflects how the notion of capacity is used in information theory and especially cybernetics, i.e. the field that deals with the problems of information transmission in systems of regulation and governance[1]. I start with developing the concept of capacity as it is used in these fields first, and in the next steps I introduce into the analysis the issues of power and interests.

---

[1]Although I build the concept of institutional capacity with reference to the cybernetics and information-theory literature, it has a direct counterpart at the core of the literature on comparative politics of transition and development where it usually refers to some broadly understood ability of the institutions to shape the political process, including by ensuring effective implementation and enforcement of law (e.g. World Bank, 1997; Fukuyama, 2004).

5.2 Information theory and cybernetics: capacity as ability to transmit a message

If our primary concern is transmission of information, arguably the best starting point for our inquiry is the information theory itself, or the field that directly builds on it, cybernetics. The first and foremost interest of the information theory is the speed and reliability of information transmission. The typical question an information theorist would ask is: How much information can be transmitted through a particular communication channel in a given time? The answer to this question is then stated in terms of the channel capacity, were capacity conceptualized as the amount of information a particular channel is able to transmit in a given period of time.

I already mentioned that the measure of information content – a bit – is completely general, it has nothing to do with the meaning of the message. This on the one hand allows for the most general use of the concepts – information of any type, in any format, in natural languages as well as in artificial codes can be captured – but its use for social analysis is obviously limited as what is usually more important than the amount of bits in a particular message is the amount of meaningful information, its credibility, simply its content. As a result, I propose that at this stage we can measure the transmitted information content generically by the number of *messages* (M), where by a message I mean any meaningful piece of politically relevant information that actors in political setting may want to communicate to others.

The specific content of message as a unit of information will depend on the context, but when we deal with functioning of international institutions messages may provide information about such issues as others' current degrees of compliance with the agreed rules, the content

and intensity of their interests, or e.g. on their reputations from past cooperation. To give an example, when we study compliance of member states with rulings of the World Trade Organization, a comprehensive overview of particular country's applied tariffs and quotas may constitute a message, transmitted by the WTO monitoring facility to the other member states. And so may the decision of the Dispute settlement body communicated back to the states.

In a more subtle sense, for example, an established successful institutional arrangement constitutes a message as well, as it informs the actors about a successful blueprint for co-operation and thus structures (focuses) their expectations in that it makes convergence to some cooperation scheme more likely than on others. In general, any pattern of behaviour can turn a message, if the relevant actors are aware of the pattern and are incorporating existence of the pattern into their calculations. Any political information is a message to the extent to which it reduces the actors' uncertainty about their counterparts' future behaviour.

The notion of message is fully generic, as is that of bit, but we require that a message has for each of the communicating actors some meaning, that it conveys some politically relevant information. It will always depend on the context, how such a unit of information content is specifically operationalized. Similarly, we will leave the time measure unspecified, we will only refer to it as unit time t and because it is a unit time we will in fact mostly disregard it in the future analysis (the unit can be e.g. one year, one election cycle, one intergovernmental conference).

Using this conceptualization, we can provisionally define capacity of an institution as

$$C_T = \frac{M}{t}, \tag{5.1}$$

95

i.e. as the number of messages it can transmit in a unit of time. The subindex T in C_T stands for *technical*, to distinguish it from a more generally applicable concept of capacity I introduce later.

An IO secretariat, for example, has presumably higher technical capacity if it consists of 1000 employees than if only of 500 employees, as the 1000 can handle more information, send more regulatory signals, comprehend and respond to more inputs from the social reality – it can transmit more messages. The technical capacity of a committee for relationship between employers and trade unions is higher if the committee meets monthly than if it meets yearly, simply because the higher amount of meetings allows for more participants expressing their views, it allows for more elaboration of the problems – again, higher transmission of information among the actors. The information capacity of an international agreement, to give an example, clearly cannot surpass the number of statements or instructions it gives (measured e.g. in paragraphs, provided that each §on average specifies one instruction, i.e. sends one message), so in general the longer and more elaborate the agreement, the more complexity in the social reality it can at least potentially regulate in that it informs the actors about how they should behave and how they can expect the others to behave. A very short agreement leaves unspecified and unregulated larger share of possible events, than a more elaborate one.

This technical notion of capacity is too simplistic for political analysis as there is no space in it for what we normally conceptualize as the main drivers of political outcomes: power and interests.

5.3 Power, interests, and information transmission

In politics actors often have interests in certain information not being transmitted. They may use the resources they have at disposal to prevent information transmission, even if the particular institution supposed to secure its transmission might otherwise have the technical capacity to do so. That means, capacity of an institution is directly dependent also on the underlying constellations of power and interest. If information is to be transmitted, holders of the majority of power resources must be supportive of the transmission. The more power lies behind the actors who are against transmission, the less likely is the information to be successfully transmitted.

The key to the success of transmission of a message then lies in how likely the message is to 'meet' with support, as opposed to opposition. Do the powerful actors want it transmitted, or do they want to block it? For each message, each actor decides whether to support or to oppose its transmission, and how strongly to do so. The probability that a message is transmitted via the channel is then formally given by S in

$$S = \frac{support}{support + opposition},\tag{5.2}$$

where *support* and *opposition* stand for the sums of supportive and opposed power across all the actors (an both take any non-negative value) and where $S \in \langle 0; 1 \rangle$. This value describes the power and interest constellation for any particular individual message.

Together across many individual messages, the levels of *support* and *opposition* add up to what can be understood as long-term support for the institution as such. The higher the support for transmission of the individual messages, on average, the more support commands the institution responsible for their transmission itself (cf. Easton, 1975). We should

stress one important thing, however: this long-term support does not belong to the particular institutional scheme or design *per se*, but rather to its performance of the message transmission function. The level of long-term support of the institution is in this sense given by how much the actors, on average, want the messages handled by that institution to be transmitted; it is not given by some intrinsic value of the particular institutional framework. It is important to note this because an institution may be appreciated by some actors even if it does not, or perhaps precisely if it does not, perform its functions, as some actors may prefer a poorly functioning institution to a well functioning one. For my analysis what is important, hence, is to what extent the actors support the institutions' cause or functional purpose, information transmission as such, and not its specific arrangement.

Such quantity of long-term support can be intuitively likened to what is referred to in institutional analysis as *robustness*. An institution is said to be robust "if after no history of experience would any decisive coalition [of actors] wish to implement some alteration of the arrangement" (Shepsle, 1989, p. 142) or if it survives the test of history. If we once again focus not on the specific design arrangements (as Shepsle does), but on the function of the institution, we may adjust the notion of robustness to denote *the degree of long-term support the decisive amount of actors provides to the institution to secure continued information transmission.*

This means that I do not subscribe to the formal-institutional perspective where robustness is a feature of the particular institutional framework. In this formal institutionalist approach, an institution is robust when its specific arrangements correspond to the interests of the key actors. For example, an institution may be seen as being robust when its decision-making rule – as one of its key components – adequately puts into favourable position the most powerful actors. The argument I present does not deal with such formal

institutional arrangements, it is not about whether the key actors like the specific current institutional form. Instead, the notion of robustness should refer to whether the key actors support the institutions' information-transmission purpose, i.e. whether they actually want the institution to help them achieve the outlined cooperative outcomes.

Consequently, an institution is said to be robust, when a sufficiently large group of actors supports its information transmission function, even after a change in the actors' preferences or the distribution of their power. For example, some states may decide that it is not anymore in their interest to report to an international institution on its behaviour. The institution is said to be robust if the overall level of the support it enjoys is still high enough, so as to secure that this information is transmitted. The state that planned to stop reporting may be, for example, compelled to continue to report by the other states who value the cooperation enough to either threaten it or to provide it with sufficient side-payments. The higher the margin of support over that of opposition, the more robust the institution is. Or in other words, the more support the institutions' information transmission function enjoys, relative to the opposition it faces, the more can the distribution of the actors' preferences and powers change without putting the cooperative scheme in jeopardy. If the margin of support over opposition is small, even a small change in preferences of the actors may result in that the institution cases to have sufficient support to perform its function.

In a strict sense, robustness is not a feature of the specific institutional setting, it is a feature of the distribution of the key actors' preferences, with respect to the institutions. Put simply, do the actors want the institution to continue transmitting information, even if their power and interest constellations change? If yes, the institution can be – in the long-term – considered robust.

This notion of robustness also directly relates to the problem of credibility. Robustness

of the institution – again understood as support for its information transmission function – is what will determine credibility of the messages it transmits. Take the example of a simple bilateral bargaining over a fixed cake. The bargaining problem can be represented as essentially one of the need to find out which of the actors is willing to forego the payoff for a longer time, and hence keep bargaining more toughly and patiently (Muthoo, 2000; Fearon, 1998). From the perspective of information transmission, none of the actors supports transmission of their private information about how patient they will be, as such information would give the counter-part a potentially decisive advantage. Robustness of such a simple bilateral bargaining scheme is thus very low as the key actor – the owner of the private information, the sender – does not have an interest in the information actually being transmitted.

Hence, in bilateral bargaining, even if both the actors may well have the technical capacity to exchange information of their interests and willingness to continue bargaining, this technical capacity is very strongly undermined by the generally very low robustness of such an institutional setup. The fact that the actors do not support transmission of information about their own interests (their intensity) will result in that both in any immediate interaction, and in the longer term from the perspective of the institutional setup as such, capacity of simple bilateral bargaining to transmit the important information will be low.

This notion of robustness can then capture the problem of information credibility. Each of the actors individually has incentives to pretend to be willing to bargain longer and sustain the no-agreement status quo for longer time (and thus to induce an earlier concession from the opponent). This is an important matter since credibility of information transmission constitutes one of the important areas of political research, indeed the entire field of signalling games, among others, deals precisely with information credibility (McCarty & Meirowitz, 2007, ch. 8). In my framework, credibility of the message is given by the ratio of support

and opposition for transmission of the particular peace of information.

In the language of spatial analysis, the concept of robustness can be directly derived from that of political core of a meta-institutional (constitutional) game. As discussed e.g. by Tsebelis (Tsebelis, 2002), political core is the area of stability in the political space where it is — due to the existing constellation of actors' interests and power — virtually impossible to arrive at a change. To make use of the existing term I will refer to the notion of long term support developed here as to robustness, to stress the formal correspondence with the term used in comparative politics. But we should always keep in mind that we are not speaking about robustness of a specific institutional arrangement, but of the support the key actors provide to the institutions' function, i.e. transmission of information.

Formally, in the specific context of information transmission the concept of robustness R is equivalent with the degree to which the actors want, on average, the information handled by the institution to be transmitted. This desired level may reflect either the interest in that only part of a message is actually transmitted in any specific case (for example 20% is not transmitted), or simply a long-term frequency of support for transmission as opposed to opposition (so in 80% of comparable cases the actors support transmission, in 20% they opposes it). I denote this desired amount of transmission with T^*. That means,

$$R = T^*. \tag{5.3}$$

The correspondence between robustness and the desired amount of transmission is not arbitrary, in fact it is derived from a simple spatial model in which the actors decide whether to support or oppose the transmission, and with what amount of power to do so. This underlying micro-foundation of the model is therefore given by the actors' calculations when

they compare their payoffs in a situation in which the institution performs its function (transmission=1) and in which it does not (transmission=0).

If we assume a linear utility function, than the level of their support will be given by the distance of their ideal point T^* from the *reversion* point of zero transmission. Similarly, the level of their opposition is given by the distance of their ideal point from the maximum transmission level, where the institution functions perfectly[2]. Formally,

$$support = supp = |T^* - 0| \text{ and} \tag{5.4}$$

$$opposition = opp = |T^* - 1|. \tag{5.5}$$

Consider figure 5.1 that depicts this decision-making act on whether or not to support the information transmission. The further the ideal point of the actor (ideal share of transmitted information) is from the situation in which the institution fails to function and hence no information is transmitted (value 0), the more supportive of the institution the actor actually must be. On the other hand, the further is his ideal point from the situation in which the institution performs its function fully (full transmission, value 1), the more opposed to the institutions the actor is. This is arguably the simplest way how to capture the decision made by the individual actors.

Consequently, the correspondence in equation 5.3 needs to be seen just as a special case of a general spatial model in which the state's support for a proposal is given by the relative position of the reversion point, the proposal, and its ideal point (cf. figure 5.1). Note again

---

[2]The assumptions of linear utility functions simplifies the formal expression considerably, but the logic applies to any definition of utility functions, provided that they are monotonically decreasing with distance from the ideal point.

support

0   T*   1

utility decresing
with distance
from T*

opposition
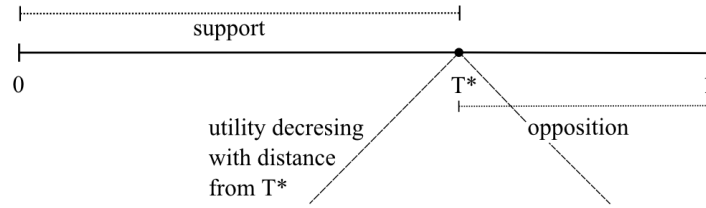
Figure 5.1: Robustness – spatial definition

that in this case we are not speaking about a specific proposal but about support for the institutions, where the reversion point corresponds to prevention of information transmission (value 0) and the proposal corresponds to the full, unblocked information transmission (value 1). The formula of the type presented in equation 5.2, i.e.

$$S = \frac{support}{support + opposition},\tag{5.6}$$

then translates into

$$R = \frac{|T^* - RP|}{|T^* - RP| + |T^* - Proposal|}\tag{5.7}$$
$$= \frac{|T^* - 0|}{|T^* - 0| + |T^* - 1|}.\tag{5.8}$$

Since for positive numbers $|T^* - 1| = |1 - T^*|$ this reduces to

$$R = T^*.\tag{5.9}$$

The reader might have noted that this calculation does not really take into account the

103

relative power of the actor, or his ability to actually influence through his decision whether the information will or will not be transmitted. Even when we do take this into account, however, the results do not change. To see this, we can introduce into the utility calculation also the power of the actor to cause the information transmission to fail, if he decides to withdraw support and instead to oppose it; the standard measure of this power is the Banzhaf (or also Penrose-Banzhaf) index (Banzhaf, 1964; see. Plechanovová, 2004, ch. 5). Banzhaf index is defined as the number of coalitions a particular actor i can change from winning to losing, over the sum of such values across all the actors. If an actor is powerful, he is able to turn a higher share of coalitions from wining to losing; similarly, a weak actor can only turn a very small share of all the possible coalitions into losing ones. Formally, hence, Banzhaf index is defined as

$$\beta_i = \frac{\Psi_i}{\sum_{i=1}^{n} \Psi_i}, \tag{5.10}$$

where β_i gives the power index of actor i, Ψ_i gives the number of coalitions actor i can turn from wining to losing, and $\sum_{i=1}^{n} \Psi_i$ gives the sum of values Ψ_i over all the n actors. In our context, turning a wining coalition into a losing one means causing the information transmission to fail. If we incorporate this factor into the actors' decisions, we get

$$supp' = \beta_i |T^* - 0| \text{ and} \tag{5.11}$$

$$opp' = \beta_i |1 - T^*|. \tag{5.12}$$

104

Consequently, robustness of the institution form the perspective of actor i is calculated as

$$R_i = \frac{\beta_i|T^* - 0|}{\beta_i|T^* - 0| + \beta_i|1 - T^*|} \tag{5.13}$$

$$= \frac{\beta_i T^*}{\beta_i} \tag{5.14}$$

$$= T^*. \tag{5.15}$$

Even though the power-considerations will, on average (note that the power index gives an average – *a priori* – power value, not power in a given individual case), not alter the individual actors' decisions on whether to support or oppose information transmission, it will obviously have an important effect on the aggregate outcome. That means, if more powerful actors decide to oppose transmission, the result will be different than if weaker actors decide to do so.

To include in the formula this possibility that powers are across the actors distributed unequally, we can simply weight the individual robustnesses R_i by the same Banzhaf power index. Since $\sum \beta_i = 1$, the aggregate value across all the actors is simply

$$R = \sum \beta_i T_i^*. \tag{5.16}$$

Whether plain (as in equation 5.3) or weighted (as in equation 5.16), the notion of robustness has two important features: it takes on values between 0 and 1, and it is a monotonically increasing concave function of support (it rises more steeply when support is low than when it is high, relative to opposition). The first feature is very convenient as

it provides robustness with the ability to serve as the moderator of technical capacity; the mathematical definition thus corresponds to the theoretical and conceptual requirements of robustness being the moderating factor for technical capacity. The second feature is equally suitable from the theoretical perspective in that it corresponds to the general principle of diminishing marginal returns, or in economics the law of diminishing marginal returns. Figure 5.2 illustrates these dynamics in a plot where the horizontal axes represent a range of values for support and opposition, and on the vertical axis provides the resulting values of robustness[3].
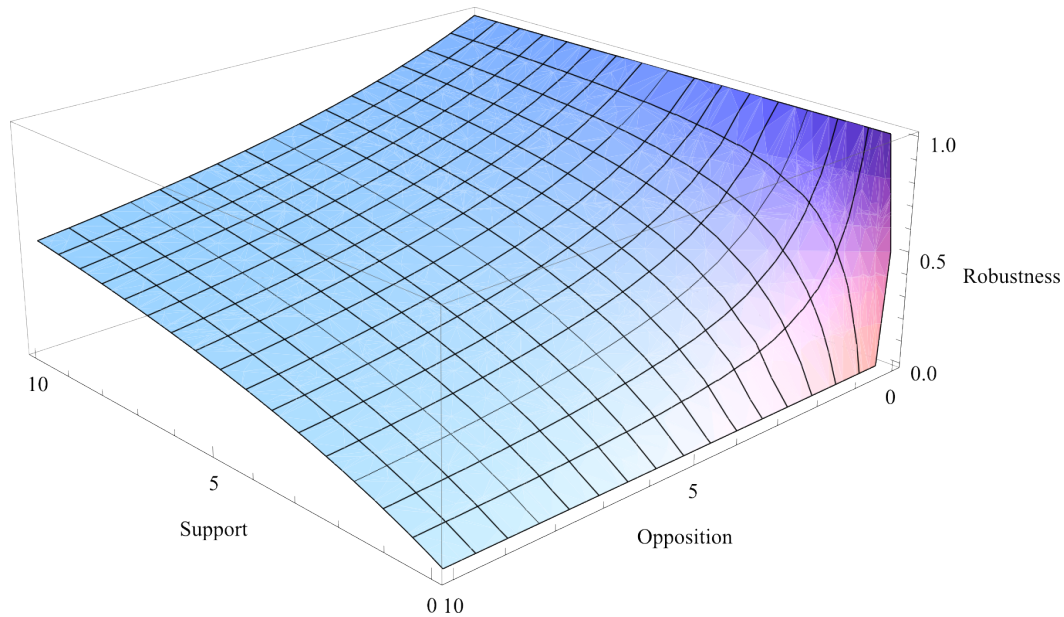
Figure 5.2: Robustness as given by support and opposition

It is worth mentioning that such notion of robustness is common in the IR literature. Downs et al. (Downs *et al.*, 1996), for instance, build on it their critique of the compliance literature as represented by Chayes and Chayes (Chayes & Chayes, 1993). For Chayes and

---

[3]Given that *supp* $\in \langle 0; \infty \rangle$ and *opp* $\in \langle 0; \infty \rangle$ the plot obviously cannot depict all possible combinations. Since robustness is defined by the relationship between the two variables, though, the plot correctly represents the shape of the robustness function, i.e. the pattern does not change however high the actual values of *supp* and *opp* are.

Chayes and others, the fact that states comply with their obligations is a sign of the role of international law. Downs et al. rightly point at that this would only be the case if the agreements were biting in the first place, i.e. if they were designed so as to actually constrain the actors. Using the notion of robustness, the same argument would read: compliance only demonstrates causal effect of those institutions that are robust, that is stable under varying power and interest constellations. Similarly, the notion of robustness is directly introduced into the analysis of regimes by Mitchell, who argues that for assessment of the causal effect of the institutions it is important to include in the analysis the difficulty of the task, or how much a unit change in the outcome variable of interest 'bites' (Mitchell, 2006). Finally, Rittberger's and Zürn's analysis of East-West regimes hints at the importance of robustness as they show that to demonstrate the regimes' causal role we need to show that they are resilient (or robust) even if the overall relations deteriorate (Rittberger & Zürn, 1990).

All these findings point in one direction: for the institutions' performance of their information transmission function it is essential to what extent they are able to secure long-term support of sufficiently powerful actors for their purpose of information transmission. The more of the long-term support they enjoy, the more information they will also be able to transmit.

5.4   Institutional capacity

It is possible now to complement the preliminary definition of technical capacity (C_T) with the notion of robustness. Capacity is then defined as the number of messages an institution

transmits in a given time, moderated by the institutions' robustness. Specifically

$$C = \frac{MR}{t}, \tag{5.17}$$

where $C \in \langle 0; \infty)$. This definition has several nice features. First, it conforms to the micro-foundation outlined above, in that the capacity directly reflects the probability that a particular message meets with support or opposition from the relevant actors. For example, the resulting capacity is half of the technical capacity if the supportive and opposed powers are equally strong. Second, it maintains the non-linearity given by definition of the robustness function, i.e. it also conforms to the law of diminishing marginal returns.

The definition also has a very strong resemblance to how capacity is defined in information theory in the more general case in which noise is allowed. We discussed above the simplified situation of a noiseless channel. Shannon, however, describes also the more general end empirically realistic situation of a noisy channel (Shannon & Weaver, 1949, pp.), and there the capacity is also defined as some technically available amount of channels for transmission (e.g. bandwidth in radio transmission) moderated by the power of the signal relative to the power of the noise. In our case, noise corresponds to opposition to transmission. The stronger the signal, given by support for transmission, in relation to the noise, given by opposition to it, the more information will make it through the available channels.

The concept of capacity as presented here establishes the upper bound of what an institution can, given its design, be expected to accomplish. In other words the variable of capacity captures the extent to which the institution can contribute to the prospects of cooperation. This value is given by the institution's design and by the support of the actors for its cause, information transmission. Having such an estimate is obviously for many purposes useful on

108

its own.

The real value, however, comes when the level of capacity can be related to another reference value. First, the value of capacity can be meaningfully compared with the actual transmission level (T^{actual}). A value defined as ($e = \frac{T^{actual}}{C}, e \in \langle 0; 1 \rangle$) will then identify efficiency of the institution, i.e. how much the institution makes use of its available resources (support it receives from actors, and its design). Such a measure will directly reflect how well the institution is managed, its *internal functioning*. I should stress here that $C \neq T^{actual}$, i.e. that capacity cannot be a priori equated with the actual transmission levels; these are two distinct concepts. In any institution, some efficiency gap is likely to occur, the size of which will be given by how good the management of the institution is.

Yet, it may be perfectly plausible for empirical political research in many contexts to assess capacity of an institution essentially by considering the actual levels of information transmission. To be sure, for some specific purposes we may want to distinguish the two also empirically, e.g. when quality of internal management of the institutions becomes itself a matter for political analysis. In many other contexts, though, assessing capacity through the proxy of actual transmission is reasonable, especially if we are interested in variation over time (or across states) and we do not have any reason to expect that the quality of internal management changes across the compared cases or over time. In many contexts, furthermore, also theoretically the value of actual transmission and capacity will converge; in simple bargaining fora, for example, there is little role management can play, and hence actual information transmission is likely to be very close to the capacity levels[4].

Perhaps even more importantly from the perspective of political analysis, capacity of an

---

[4]Although clearly one could consider such factors as quality of the mediators that might be employed in the forum. If this was an important matter for the specific political analysis, then the values of capacity and actual information transmission would clearly have to be distinguished.

institution can be compared with the amount of information transmission that is actually *needed* for cooperation among the actors to work, that is with the required amount of transmission. If the capacity of an institution is lower than the required volume, it is immediately clear that the design of the institution is fundamentally inappropriate. By comparing capacity with the required levels of information transmission, we are addressing how well the institution has been designed, i.e. we are considering its relation to the *external political conditions* that led to its form.

Consider, as an example, the capacity of the World Trade Organization as of a bargaining forum for the Doha multilateral negotiations talks. In the 2-level setting of the WTO negotiations (as of other international negotiations), the capacity (and consequently also the actual information transmission) is lower than the technical capacity because the negotiators have incentives not to transmit correctly to their negotiation counterparts the information on their true domestic political constraints, i.e. on what are the true boundaries beyond which they cannot go in the negotiations (Schelling, 1960; Putnam, 1988). The diplomats in the Doha Round may have a reasonably high technical capacity as communication channels, because by now they have spent ten years discussing who wants what and trying to find out the possible zones of agreements. Yet, the key domestic interest groups potentially negatively affected by the negotiation outcomes clearly do not support the transmission of the information on what concessions each state can make. The technical capacity of the negotiators will then, provided that the interest groups have power in their respective domestic political systems, be seriously undermined by the strong opposition of these actors on the domestic level.

In other words, there are many powerful actors who do not want the messages sent by the governments to be transmitted. Consequently, even though the diplomats may have

the technical capacity to transmit to their counterparts the information about what their governments want, they certainly do not have the actual capacity to do so. The support and opposition levels will thus be determined not only by the interests of the negotiators, but also of the key actors who determine what the negotiators themselves can and cannot do. Since the information about each state's interests are, in general, necessary for success of the negotiations, capacity of the bargaining forum is clearly lower than is required. That means that the design of the bargaining forum is inappropriate.

As indicated, given the existing design that determines capacity, there may still be variance in actual information transmission, i.e. in how much of the existing capacity is actually used. This may be influenced by such matters as personal relations between the negotiators, the abilities of the Director General as a mediator, and the like. This degree of use of the available resources signifies efficiency of the internal WTO functioning, i.e. how efficiently it makes use of the input (design and relations among the actors) for production of the output (information transmission).

For the purpose of further clarification, the notion of capacity presented here can be compared with that of effectiveness as developed in the regimes literature, and in particular with the conceptually most developed piece, the Oslo-Potsdam solution of the problem of regime effectiveness (cf. Helm & Sprinz, 2000). In the Oslo-Potsdam approach, effectiveness is measured as the ratio between the value added of the institution and the total size of the desired change; this corresponds to its causal effect. Formally effectiveness is given by the distance in a spatial model between the actual performance of the regime and the no-regime counterfactual, standardized by the distance from the no-regime counterfactual to the collective optimum. The top part of figure 5.3 shows the definition of effectiveness in the OPSE.

$$E = (AP-NR)/(CO-NR)$$

NR No regime
counterfactual

AP = Actual level
of transmission=T^{actual}

CO = Collective optimum,
required level of transmission

$$e = T^{actual}/C \qquad D = C/CO$$

0

C = Capacity

e = Efficiency
E = Effectiveness, Oslo-Potsdam
D = Design appropriateness
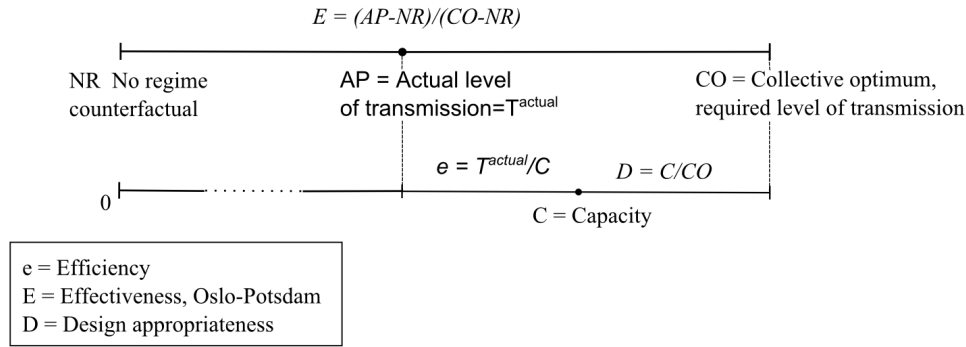
Figure 5.3: Capacity in relation to the Oslo-Potsdam concept of effectiveness

The bottom part than indicates how the concept of capacity relates to that of effectiveness. In particular, it indicates that capacity lies between the actual performance in the regime and the optimal state. In the Oslo-Potsdam approach, these two points are defined in terms of some desirable physical properties of the system (e.g. thickness of the ozone layer), in my analysis these are naturally defined in terms of information transmission. So instead of actual performance AP in OPSE I refer to the actual level of information transmission. Similarly, I do not deal with some notion of optimal performance of the regime, but instead with the notion of required information transmission, i.e. the level of transmission that is necessary for the system to be viable. One conceptual advantage of the notion of capacity is that the level of information transmission has a natural zero point (where no information is transmitted), so we do not need to estimate any counter-factual.

That means, the notion of capacity does not capture the direct causal effect of the institution, as OPSE tries to achieve. The concept of capacity is not concerned with existing causal effect of the institutions. Instead it offers a theoretically deeply rooted notion of the institutions' ability to provide space for successful cooperation.

In general, an important conceptual strength of the notion of capacity lies in its ability

to capture both the more technical problems of information-transmission in cooperation and the underlying power and interest relations concerning the institutions' functioning. To be sure, these power and interest distributions do not capture the substantive interests of the actors in specific issue areas. I mentioned in chapter 1 that to understand whether or not cooperation in the end occurs we need to consider what substantive interests the cooperating actors have, and what is the power distribution among them. This is something the concept of capacity does not capture. But what it does encompass is whether and how strongly the actors support the institution as such, and this is – for the purpose of analysis of the institution's functioning – the key.

5.5   Operationalization and measurement

Having defined the concept of capacity as the appropriate dependent variable for the kind of institutional endeavour I discuss in the dissertation, two difficult questions need to be addressed for the organizational-cybernetic framework to be complete. First, we need at least some broad understanding of how we can operationalize the concept of capacity and how we can approach empirically the problem of information measurement. Second, we need to find out what capacity do the political systems of international institutions need, so what is the required amount of information transmission to which we will then be able to relate the institution's capacity and see whether it is sufficient. I leave the latter task for chapter 6, where I discuss it in detail. The former task – operationalization – will be briefly discussed in this section.

The qualifier *briefly* in the last sentence is in the right place: the problem is that it is enormously difficult to provide a general discussion of how information transmission and

capacity can be measured, simply because the operationalization procedures will be highly context dependent. There are many different forms of information and each can be operationalized and measured in different ways; chapter 7 that discusses various appropriate tools for the design of international institutions will review a range of various aspects of institutions' functioning, and these will indicate also the various forms information in international institutions can take.

Yet, an elementary classification of methods for assessment of information flows and capacity is certainly desirable. I propose to divide the approaches to operationalization according to two criteria: 1) qualitative and quantitative, and 2) indirect measurement (based on assumption or observable implications) vs. direct measurement.

The qualitative and quantitative divide is obviously to some extent artificial, as qualitative information can be quantified, and quantitative information can provide qualitative insights (e.g. in showing broad patterns of relationships). Nevertheless, distinctions can be made in terms of the form in which the data is collected from the primary sources. It is essential to understand that approaching the measurement problem from the qualitative perspective is by no means inferior. This is implied by the need for information in social and political context to have meaning, i.e. rather often the major research challenge will lie in identification of whether information of a particular meaning can be transmitted, rather than simply how much of information is transmitted. A quantitative approach is only possible when the problem of meaning can be plausibly abstracted from.

According to the second criterion, we can distinguish methods that are directly accessing information flows from those that do it only indirectly, via some (plausible) assumption or via implications of the information being or not being transmitted. Table 5.1 summarizes the scheme and several examples that I discuss below.

	Qualitative	Quantitative
Direct	Costly signals in public statements	Amount of communication – e-mails, web-sites visited. Experiments.
Indirect	Change in responsibility allocation	Numbers of assigned staff

Table 5.1: Approaches to operationalization of capacity: examples

First, an example of directly observable information that can be measured qualitatively are certain public statements that the relevant actors make in certain occasions. The reasoning behind such an approach stems from the branch of game theory focused on communication, the so-called signaling games (see e.g. Myerson, 1991, ch. 6). According to the analyses in this field, in negotiations in which each actor has a private information about how much he or she values the possible outcome the actors may try to settle on certain outcome by making public statements that bind them to their preferred outcome, i.e. statements that would impose on them costs, were they to back off from them later. Fearon, for example, has showed that precisely this effort of international negotiations may lead to escalation of conflicts. What the negotiators do is that they bind themselves to their preferred outcomes by publicly claiming that other outcomes are unacceptable to them; the domestic public records these statements, and – presumably – would punish the negotiators later, were they to give up on these positions (Fearon, 1994). Making use of this need of the statements to be public (as otherwise they cannot be costly, they are mere cheap talk), we can assess the capacity of a negotiation forum by studying the statements the negotiators make, and the degree to which they are binding. Recently, such an approach has been applied to the case of the negotiations in the World Trade Organization (Parízek, 2012b).

Second, there is a whole range of ways how to measure certain information flows directly and quantitatively. The most obvious data susceptible for such an approach is data monitoring online activity. There are, for example, tools that provide measurement of access of web-sites from specific locations, the most prominent and a highly functional one being the Google Analytics[5].

A completely different way how to observe directly quantitatively flows of information can be based on experimental research. Experiments have been conducted on functioning of the beliefs-updating process, i.e. on how the actors are able to learn about their counterparts preferences, and how much they are able to incorporate into their conceptions of the counterparts the newly obtained information (e.g. Huck & Weizsäcker, 2002). Similarly, there have been experiments on how quality of estimates of others' positions relates to both individual and collective success of the negotiations (Thompson, 1991).

Third, it is possible to assess quality of information flows indirectly. An example of how this can be done in a qualitative way is provided in chapter 8, where I present an argument about how sufficient capacity of the Council of ministers in the EU could have been maintained after the Eastern enlargement. The argument there is that the enlargement, by dramatically increasing variety of interests in the Council, increases the amount of complexity it needs to deal with, and hence demands higher capacity. The key role of the Council as of a bargaining forum is to have the actors communicate to each other their interests and devise an agreement that is compatible with their communicated positions. The capacity assessment is therefore built on the *assumption* that with more actors with more diverse interests, this task will – on average – be more difficult and that it will require from the Council more (decision-making) capacity.

---

[5]Accessible at http://www.google.com/analytics.

Fourth and last, an example of a quantitative indirect measurement is presented in chapter 9. There I argue that the operational multilateral organizations are in their activities dependent on soft information about the broad social, cultural, and institutional patterns in the countries with which they cooperate. Such soft information on local contexts is notoriously difficult to quantify and transmit. I assume in the analysis, therefore, that to possess this information the IOs need some minimum amount of staff from the countries, and I show that this logic is justified by the internal documents of the IOs themselves. I then run a quantitative analysis based on the number of staff from each of these countries in the IOs, arguing that indeed the IOs devise mechanisms to secure that they do have (relatively more) staff from the countries with which they cooperate.

Overall, the operationalization problem is severe in the analysis of information flows, but it is not principally more difficult to measure information flows than it is to measure other concepts with which political research and IR have to deal with.

In summary, with the concept of capacity we have the appropriate dependent variable for the kind of institutional analysis in which we are interested in how design of institutions impacts on their functioning. Capacity captures how much the institutions are able to fulfil their information transmission function, i.e. how an institution of a particular design can contribute to the prospects of cooperation. In the next chapter I present the scheme that helps us identified how high capacity of an international institution needs to be, if the prospects of cooperation are to be high.

Chapter 6

Institutional Capacity of Political Systems: the Viable System Model

Having outlined the concept of capacity of an information channel in the previous chapter, it is time we move to the key task of identifying what capacity the individual channels that form the political system of a regime need. This entails two tasks. First, we need to identify what are the key channels about which we, so to speak, should care. If I laid down the argument in chapter 3 that we need to focus on the vital system functions, and we know that these will consist in transmission of information, what are the specific channels that secure these vital transmission? Second, we need to establish what the capacity of these channels ought to be. These tasks are accomplished in sections 6.1 and 6.2, respectively.

6.1   System components and their functions

The very purpose of the VSM is to propose a set of general requirements, or guidelines, for how regulatory schemes need to function, so that they accommodate the implications of the

law or requisite variety, or LoRV, in a productive way (see chapter 4). Because the world is inherently complex, decision-making (governance) needs to be distributed across several (potentially many) levels. As each level can process only a limited amount of information, all decisions that can be done at the lower levels need to be done there, otherwise the higher levels become overloaded or, more usually and certainly worse, they simply make decisions without first processing all the relevant information. Given the restrictions imposed on governance by LoRV, how do we set up a scheme of information channels in such a way that those who decide have all the information they need? How do we design the communication network that will enable the actors that cooperate in pursuit of some shared goal to transmit among themselves all the information they need?

It turns out, and it has been demonstrated deductively by Beer (Beer, 1979, chs. 8-10) that to model such a scheme we need a set of five different types of Systems, where the first type (System One) identifies a principle organizational unit, simply one of the (corporate) actors who are coming together to achieve some shared goal through joint governance. Systems Two through Five then form the governing System that is necessary if a collect of two or more Systems One (of individual actors) is to be viable as a coherent whole, i.e. when all the fundamental actors are to form together a viable system. Figure 6.4 later in this chapter summarizes the entire scheme, and we will build it step-by-step in this chapter.

The entire scheme we will discuss is recursive in that the governance System as a whole is itself a System One in a collectivity one level higher in recursion. For example, departments of a ministry can be understood as Systems One, where the ministry as a whole is treated as the overall governance (viable) system of interest. One level of recursion higher, the ministry as a whole forms a System One in another (higher) viable system of interest, the government (other ministries are other Systems One). On an even higher level of recursion,

the government may be understood as one System One in higher viable system of interest, an international regime. For the study of IOs, then, we would usually consider as the most interesting Systems One the member states. The IOs as such, or the entire regimes, would be treated as the viable systems of interest, and we would study how information flows within the regimes or IOs. Bearing in mind this recursive nature of any viable system, i.e. of any system that is supposed to be able to cope with the variety of its environment, we can proceed to describe how one level of such a recursive scheme should look like. AS mentioned, such a scheme will be composed of five systems, and I will describe each in turn in some detail and thus describe the entire scheme.

The basis of the entire scheme is a set of Systems One. As a viable system, each System One has its own identity and purposes (or interests), it conducts certain activity, and it has its own governance mechanisms (in Beer's original treatment *the management*). The System also has its environment, from which it takes its inputs, and where its outputs go. Figure 6.1 depicts these components, with the environment on the left side, the unit as such or its operational unit in the middle (the circle A), and its governance (or management) on the right side (the square 1A). In the graph the three components have been separated to indicate the links between them but in reality the System is embedded in its environment, and its management in the System itself.
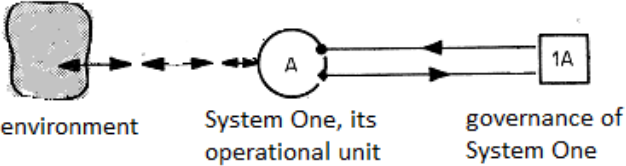
Figure 6.1: System One: its environment, its operational unit, its management (governance)

If we again consider the example of international regimes as the viable systems of our

interest, each state is treated as a System One. Each state has its own governance mechanism (the government) its operational units as such (the entire political system of the state), and its environment (the social, economic, and other systems within the state, and the international environment, cf. Almond *et al.*, 2003). One key implication of LoRV for this model of one state is that if its governance mechanisms are to be functional, they need to be able to accommodate all the relevant variety (complexity) in the political system of the state, and the political system in turn needs to be able to accommodate all the variety of its environment, e.g. of its social and economic systems, and of its international environment. The channels linking the System as such with its environment and with its governance will also need to have sufficient capacity to transmit the information needed to capture this variety (Ashby, 1956). When several states come together to create a viable system of some international cooperation scheme, the scheme cannot succeed if the political representations of the states – the negotiators, the foreign affairs ministers, the presidents – themselves do not have the capacity to regulate their domestic affairs. This is not a surprising conclusion, but it illustrates the kind of reasoning inherent in the VSM.

Given a set of Systems One, the purpose of System Two is to serve them as a means to prevent destructive oscillations in their activities, or to make sure that their activities are in some basic technical sense coordinated. In the words of IR theory, System Two is a scheme of *international cooperation* as understood in the regime theory, i.e. a conscious effort of egoistic states to coordinate policies in a Pareto-efficient way (Keohane, 1984). System Two is built consciously by Systems One when they pursue their individual interests; it is essentially a communication and coordination device serving them (each individually) to avoid destructive interference in each other's interests. This System is depicted in the right part of figure 6.2.
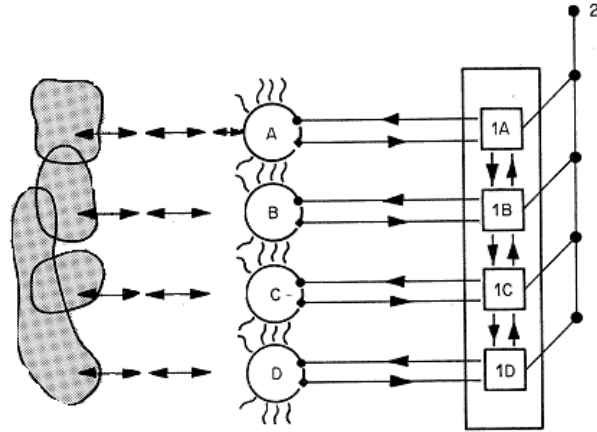
Figure 6.2: System Two with four Systems One and their environments and governance(s)

This changes with introduction of System Three, which already incorporates the notion of a system that lies above the original units, i.e. a system that has its own purposes, that acts as a coherent whole, a metasystem. The task of System Three is to make Systems One coordinate their efforts *in the way that is most profitable to the System as a whole.* In the case of international cooperation, System Three secures that states align their policies in such ways that increase the long-term stability of their relationships, that increase trust among the actors, and in general that improve the internal environment within which the states (Systems One) cooperate.

The principle difference between System Two and System Three is that while the former only *serves* Systems One to coordinate among themselves, and in this sense requires explicit consent from each of them, System Three is not concerned with position of any of Systems One but only with viability of the corporate cooperative arrangement as a whole. System Three can impose sanctions or otherwise harm any System One, if it improves the internal functioning and hence viability of the overall system as such. In policy terms, when moving from System Two to System Three we are crossing the border between regulatory and

distributive (and redistributive) policies. The distinction closely corresponds to the notion of Type-I and Type-II governance identified by Hooghe and Marks, where by Type-I they refer to integrated systems with redistributive capacity and by Type-2 to un-integrated task-specific systems Hooghe & Marks, 2003. In figure 6.3 System Three is added to the scheme; it is connected to the governance systems of the individuals System One, to the System Two in the right, and directly to the operational units of all Systems One.
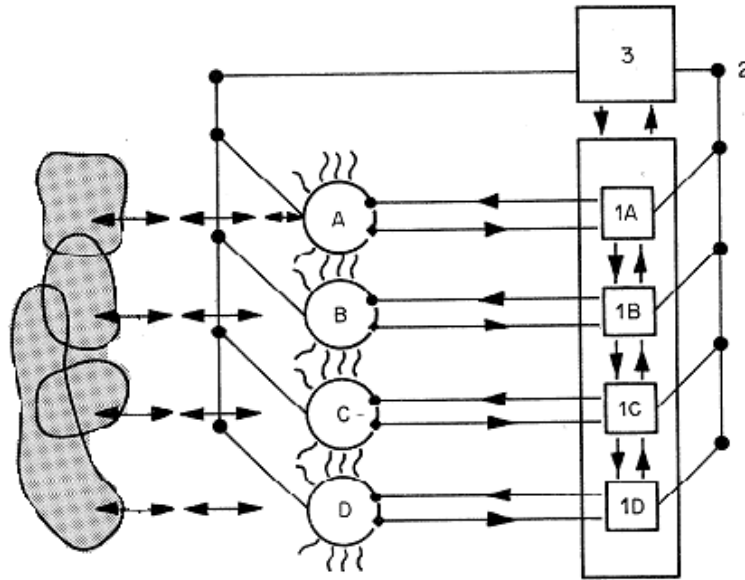
Figure 6.3: System Three, with its links to System Two as well as the Systems One

The overall governing system, if it is to be viable, needs to be able to secure not only its internal functioning (for this System Three is responsible), but it also needs to make sure it is able to adapt to its changing external environment. It was mentioned in the exposition about Systems One that each has its environment. The environment of the overall system, however, amounts to more than just a sum of the environments faced by the individual Systems one. The governance system as a whole has its environment, and it needs to deal with it. For example, in international finance each state (System One) faces certain conditions in its

own financial matters, such as its access to credit. But the environment faced by the global financial regulation (the viable system of our interest) clearly surpasses the sum of the individual environments of the member states; on the global level, entirely new dynamics that ought to be regulated emerge (cf. Andrews, 1994). The purpose of System Four in the overall governance scheme, then, is to secure its interaction with this total environment.

Finally, the task of System Five is to weigh the demands of Systems Three and Four and thus to secure viability *vis-á-vis* the external environment while maintaining coherent internal function of the entire scheme. From a somewhat different perspective, System Five maintains the identity of the overall scheme, whereby it provides coherence into the needs and purposes of all the Systems One (as transmitted to it via Systems Two and Three) on the one hand, and the demands of the System Four on the other hand.

The entire scheme is summarized in figure 6.4, adopted from Beer (1972, p. 130). On the left, we see the individual environments as well as the total environment. In the central part of the scheme several Systems One are depicted as circles A-D, with their individual governance mechanisms (1A-1D). System Two is depicted by the communication lines to the right from Systems One, leading to System Three. System Three is directly connected both to the governance schemes of all the Systems One and to the Systems One as such (the left loop from System Three directly to the circles A-D). Systems Four and Five are connected with Systems One through System Three.

I mentioned previously that we want the VSM to do two things for us: 1) explicitly list all the key functions, and 2) tell us what capacity the individual components of the scheme need. The first part has been done now, the value added of the scheme, then, is that it explicates the list of Systems, each with its function; unless a cooperation scheme possesses tools to perform the functions performed in the model by each of the five Systems, it will
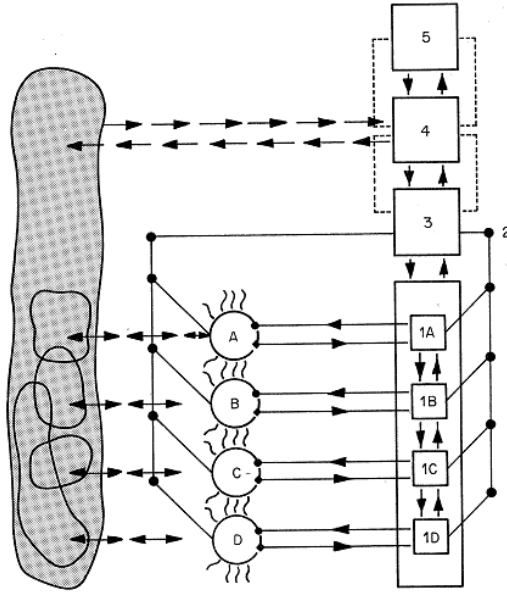
Figure 6.4: Viable System Model scheme

not be viable, because the decision-makers will not be able to process all the information about their internal and external environment they need for qualified decisions. This is a result derived deductively by Beer – the *sine qua non* conditions of viability. The second part – assessment of the requisite variety – is presented in the next section.

6.2 Achieving the requisite variety

The VSM allows us to directly specify the required capacity of each of these channels. This is something that figure 6.4 does not indicate as all the connections look for representational purposes alike, but VSM does identify the differences. In particular, the key to assessment of the required level of capacity lies in complexity of the environment within which the System operates; as I discussed in chapter 4, this complexity is measured as the variety of the environment, or the number of possible states in which it can be. Therefore, the higher the variety of the environment within which the System we study operates, the higher will

need to be the capacity of the channels that form the System as such.

Given this exogenously given value, we can estimate what capacity each component of our System will need to possess. If a System One has highly complex environment, then its own governance (e.g. 1A or 1B in figure 6.4) needs to have high capacity to be able to respond to it; the governance needs to match the variety of the environment. Similarly, we know that the communication channels among the several Systems One need to match variety of the overlap of their functioning, i.e. the activity through which they directly influence each other's functioning. Capacity of the channel between System A and System B, for example, needs to be at least as high as the variety of their relationship – otherwise it will not be able to govern all the phenomena that emerge from their interaction.

But we know more than that. Each System One is itself a viable system, and each viable system is – recursively – a System One of a viable system on a higher level. That means that the channel capacities need to match all over the model, from the Systems One to System Three. If they did not, capacity on the higher level of recursion would be lacking. Capacity of System Four needs to match the complexity of the overall environment, and System Five needs sufficient capacity to moderate the relationship between System Three and Four. In sum, we know, what the capacity of all the components of the System needs to be, for any given complexity of the environments.

One of the very interesting insights directly deducible from VSM concerns the relationship of varieties of Systems One and System Three. We know that System Three should not do more than what is essential to maintain cohesion of the system as a whole, as doing more would preclude functioning of the System One management, it would constrain its management ability and impose an unnecessary level of top-down dominance. So the vertical command line in figure 6.4 needs to be relatively subtle, or as subtle as possible, without

126

undermining coherence of the overall System as a whole. At the same time, though, if the entire System is recursive and needs to match the complexities on all the levels of recursion, than System Three needs to have the same variety as all the Systems One.

For the capacities of Systems One and System Three to match it is clear that another channel of communication between Systems One and System three needs to exist, because System Two only deals with the interaction of all Systems One and the vertical command line carries relatively little variety.

Beer argues that such a communication channel is, in many organizations, represented by such mechanisms as the audit. These are mechanisms that monitor directly functioning of the operational units of Systems One and if they indicate problems, these are reported to the management in System Three. Given that audit works with randomly chosen operational units, i.e. it always investigates samples of activities of the units, it does not interfere with the management as such. But it has the ability to detect irregularities, and thus it is able to secure that the units operate according to some general rules, that are superior to the actual management activity.

The EU structural funds management might be a good example. The Commission does not rely on the governments themselves in controlling the expenditures from its budget, but it conducts its own investigations on samples of the projects (European Court of Auditors, 2012). If the frequency and seriousness of irregularities in a sample reaches some threshold, the Commission stops inflow of the funds. This is an activity that does not intervene directly into the domestic politics, but it is an intervention with extremely high variety directly into operation of the programmes.

In an even less oppressive way for the Systems One, the direct channel between operational units of Systems One and System Three is there also to secure that even those aspects

of interaction of the units that for any reasons do not go through the managements are controlled for. So the states, for example, may only have established cooperation in certain area, but suddenly new patterns of behaviour in the operational units (within the states) emerge that are not accounted for. It is essential that the System Three monitors them and, if needed, initiates a new regulation. It is important that we understand that this monitoring function is primarily also a service function, in that it saves the individual System One managements the need to monitor their operational units constantly also from the perspective of the System Three as such. So this channel secures that also the variety is accounted for that, for any reasons, would not be dealt with (and reported) by the management.

Put together, these three channels need to have as much capacity as there is variety in the environments faced by all the Systems One. So specifically, System Two transmits only the part of information that is directly concerned with the mechanical interaction between the Systems One, the oscillation. The direct vertical managerial channel transmits the information (variety) that concerns instructions dealing with strategic goal-orientation of the overall system. The monitoring channel constantly checks if any of the operational units deviates in its behaviour from the range that is acceptable for the system as a whole; if it does, the pertinent System One is informed and instructed to deal with the problem.

This is obviously a very abstract discussion, but it in any specific example it translates into concrete requisite varieties of the channels, i.e. to the levels of capacity the channels will need to posses. Any environment has certain variety, it generates some level of complexity. Within each System One (e.g. a state), the political system and the government needs to be able to cope with this complexity. If it does not have the variety, the environment (economy, social system) will lie beyond the reach of its control. Once we move to the international level, we know that the international regime will need to have the capacity to deal mechanically

with interaction of the states' activities. The more the states interact, the more variety their interaction will produce. This variety then needs to be matched by the capacity of the international regimes. In the next step, if the states wish to form together a viable system of cooperation, they will need mechanisms that will not only cater for their individual needs, but also of the needs of their joint cooperative scheme. These mechanisms will then need to have the capacity to set regulate interaction of the states not only in the egoist Pareto-efficient manner, but – if needed – also to intervene in their relations in a redistributive manner, i.e. to make sure that the states follow a jointly defined common interest, not only the smallest common denominator of their interests. If Systems Two and Three have the necessary capacity, smooth internal functioning of the viable system is possible.

As far as System Four is concerned, it first and foremost needs to have the capacity to reflect all the complexity of the overall environment and possible future environments. Yet, this is not sufficient since it needs to absorb the complexity of the environment *and* the complexity of the overall system as such, at least in some highly simplified form. A strategic office of a ministry of foreign affairs, for example, must have the capacity to process information from the external environment it deals with – of international politics – but also of the very political system of which it is the System Four. Otherwise it could not provide the strategic advice on how the system should behave.

The role of System Four is thus rather delicate and, as observed by Beer, System Four is usually dramatically under-appreciated; Systems tend to concentrate on maintenance of immediate internal stability at the expense of long-term strategic planning (Beer, 1979). When one considers the state of most IOs, one is led to the conclusion that the problem with weak System Four is prevalent there as well. In general, IOs have very little capacities to devise long-term strategies for dealing with the problems they are supposed to address,

beyond immediate interests of the member states.

Finally, the necessary capacity of System Five is relatively low, given that its purpose is not to decide on a number of problems, but only to maintain the balance between System Three and Four. That means, the demands and varieties of Systems Three and Four should naturally be in approximate balance, and System Five's intervention will only be necessary if a major difference between the two occurs. In formal terms, System five functions as the closure of the entire system as a whole, it is the mechanism that closes the information flows and makes the decision on the balance between Systems Three and Four. Put differently, System Five maintains identity of the system as a whole (this is a concept of identity developed already by Deutsch, Deutsch, 1963). Most of the executive decision-making, however, should be done in System Three that maintains the everyday functioning of the coordination scheme among all the Systems One and thus secures immediate internal viability of the System.

In sum, the five systems together are able to cope – if designed according to the requirements identified in VSM – with the essential problem of extreme complexity of the System's environment on the one hand, and the need for order and relative simplicity within the System.

Let me illustrate the argumentation with an example of the problem of overfishing and of the allocation of the fishing grounds to different states. A successful international scheme to regulate fishing entails a number of actors, both state and non-state, and a whole range of tasks the regulatory body needs to perform. First of all, the task of System Two will be to collect information on all these actors and on their interests, and to devise agreements that will coordinate who, and when, has the right to fish at all the grounds. This will ensure that conflicts among the states (Systems One) are minimized. The System Two will therefore

need sufficient capacity to capture all the relevant aspects of the problem, where the size of their variety will be given primarily by the actual degree to which ships from different states tend to make use of the same fishing grounds, i.e. the more overlap there is, the higher the space for conflicts and the higher the need for regulation.

The purpose of System Three, in this example, will be to ensure allocation and re-allocation of incorrectly appropriated resources and in general solution of conflicts among the states. While for System Two the function was to secure elementary coordination, the function of System Three is to ensure smooth solution of the allocation conflicts. Its capacity, therefore will be given, on the most general level, by the incentives of the states to disrespect the agreements. In other words, whenever any actor disregards the agreed allocation, System Three needs to step in and ensure that this disrespect does not endanger the entire cooperation scheme. System Three achieves this through two channels. First, System Three intervenes through the direct vertical channel to the individual states' governments (the governance mechanisms of the Systems One), where it can instruct the governments to adjust their policies so as to ensure compliance with their obligations. Second, System Three needs also mechanisms for direct monitoring of the sub-state actors, notably of the fishing companies, via the audit channel identified in the top-left part of figure 6.3. The combined capacity of these two channels, together with that of System Two, needs to match the variety of all the key actors in the fishing industry and most importantly of the size of their potential conflict.

The function of System Four in the governance scheme will be to ensure that long-term trends, for example in the population of fish, are monitored and that the needs of the entire maritime ecosystem are duly respected. This means that System Four needs to have the capacity to collect and process information about all the relevant phenomena that lie

131

outside of the perspective of the individual states, but that are nevertheless important for sustainability of the resource as a whole. This way the 'tragedy of the commons' in which the resources is depleted due to the individual states' interests is prevented (Hardin, 1968).

The purpose of System five, institutionally embodied usually in the highest governing body of the international regime, will be to weight the interests of System Three in smooth and conflict-less allocation of the fishing grounds on the one hand, and the long-term sustainability of the resource on the other hand. Its requisite capacity will then be given by the degree of divergence between these two goals. If, for instance, the oceans vast and the fish population very rich, the long-term sustainability demands of System four will be minimal, and so System Five could also have only low capacity. If, however, the conflict between the immediate interests of all the key actors and the long-term sustainability of the resource is intense, System Five will need very high capacity. If all these capacity requirements are met, the institutional framework ensures that a successful cooperation and governance of the problem can be achieved.

Perhaps a short clarification is due at this point of the relationship between the notions of complexity, variety, requisite variety, and necessary capacity. First, as discussed previously, the notions of complexity and variety are for the purposes of our analysis coterminous – they both refer to the number of different states within which specific system or its environment can be in a given unit of time. In general, the more interaction there are among actors in a system, the more variety (complexity) they produce. The key insight of cybernetics and of the law of requisite variety is that any governance system always needs to match the variety of its environment; it needs to have the requisite variety. Requisite variety than identifies the level of variety any governance system needs to have, to be able to regulate. Since in international politics governance systems are composed of independent states pursuing their interests, we

132

cannot discuss ability of the governance scheme to regulate (as this would require a single global authority), we can only discuss its ability to make sure that all the information that is necessary for regulation is available to the actors. International institutions do not regulate the affairs directly, they do not constrain; what they can do is to transmit the information that the actors that possess the power resources – usually the states – need to be able to regulate. Capacity captures this ability.

6.3   Design diagnosis

Given the discussion above, it is probably clear that in principle all institutional pathologies that are relevant from the organizational-cybernetic perspective have to do with insufficient capacity of some of the important channels.

One important and convenient implication of the discussion above about the necessary levels of capacity of the individual components of the System is that for many purposes we do not really need to *calculate* the necessary levels precisely. Instead, we need to be able to assess the variety (complexity) of the environment, and from this we will immediately see what the capacity of the System as a whole needs to be. The scheme just presented then tells us how this capacity should be distributed across the channels from which the System is formed.

The diagnosis and design task then comprises 1) the study of what share of complexity of the environment the current arrangements do not match and 2) creation of mechanisms to balance a potential lack. The system is functioning well if there is a stable balance between the inflows of variety into the system and the outflows from the system to the environment. What we need is to monitor the balance between inflows and outflows, and intervene when

the inflows of variety supersede those of the outflows produced by the system (Beer, 1979, p. 288). So the task of design really lies in analysis of what parts of the complexity of the environment are relevant, and identification of those parts, that are not matched by the capacity of the regulatory system.

With VSM, we have a general scheme for the study of the international regimes' capacity from the information-based organizational cybernetics perspective. The model gives us a guidance for assessment of whether all the key information channels among the actors work. In the field of organizational cybernetics, works can be found that provide lists of specific criteria and requirements for such an assessment, making the entire VSM closer to being readily operational. For empirical research, however, adjustments will be needed for primarily *political* rather than managerial or organizational analysis.

Overall, together with the concept of capacity, VSM provides a unified framework for analysis of international institutions' functioning. The framework – by focusing exclusively on the problem of information transmission – on the one hand takes power and interests of actors in international politics very seriously, but on the other hand creates space for a positive institutional research agenda. The next chapter presents the key implications of the general framework developed here for the specific problem of institutional design.

Chapter 7

Increasing the Capacity of International Institutions

I argued in chapter 3 that designing institutions is difficult as the situations we deal with are incredibly complex, and hence that we need to focus our efforts only on a limited set of key functions that are vital for the cooperative scheme. I also argued that the sole way in which institutions directly contribute to cooperation is through enhanced information transmission, and hence that we should focus our design efforts on how the institutional setup contributes to this (chapter 2). In chapters 5 and 6 I presented a conceptual framework within which such an analysis can be performed – I identified sufficient capacity of the vital information channels as the ultimate target of our design efforts, and I also presented VSM, a model with which we can identify what those vital channels are and what capacity is considered sufficient.

Now it is time we return explicitly to the very initial question: what can we learn and do about (better) institutional design? How much, if anything, can be done to improve an

international institution's functioning, by conscious changes in its design? In the light of the previous chapters, this question reads: How can we, through changes in the design, increase the capacity of all the information channels that are vital for the cooperative systems?

There is a whole range of design tools that are potentially useful for this task. I propose that, with some simplification, they can be divided into eight classes, as summarized in table 7.1. First, the tools can be either *political* in their nature, or *technical.* The tools of the former class address robustness of the institution, those of the latter class address its technical capacity, as identified in equation

$$C = \frac{MR}{t},$$

elaborated upon in chapter 5. The political tools can be further divided into the tools oriented at prevention of abuse of power by the actors against the institution, and the tools oriented at giving the actors incentives for sharing information. The technical ones can be also divided, into the tools oriented at more efficient coding of the transmitted information, and the tools aiming at higher capacity of the transmission channels. Each of these four classes of tools can be further divided along whether they deal with international institutions as structures (rules, norms) or with institutions possessing genuine agency (IOs). That means, different tools might be used to increase capacity of passive structures and of institutions that can be treated as active agents, with own interests and some actual physical existence.

Altogether, we obtain eight logically possible classes of design tools, each of which is briefly discussed in the relevant section indicated in table 7.1. Clearly some of the tools appear more empirically plausible than others, or more politico-analytically interesting. Yet

	Political		Technical	
	Interests	Power	Coding	Channel
Passive, structure-based (rules and norms)	Section 7.1.1	7.1.2	7.2.1	7.2.2
Active, agency-based (IOs)	7.1.1	7.1.2	7.2.1	7.2.2

Table 7.1: Design tools classes: overview

we should understand that even the use of the supposedly more technical tools may have serious consequences for functioning of the political system as a whole, and therefore their use or lack of use has important political implications. This gets further importance in that, as I argue in section 7.3, the technical tools can be to a considerable extent treated as substitutes for the political ones, and not necessarily only as their complements.

One thing should be stressed with regard to the target of this chapter. It can by no means be considered to provide an exhaustive list of potentially useful design principles. There are many more. What this chapter attempts is to classify the types of tools that we can plausibly expect to perform the function of enhanced information transmission, and by the way of this classification to illustrate the variety of tools that we have at disposal and especially the variety of specific institutional targets we may address with these tool. The classification is absolutely minimalist in that it considers only the major axes along which we may want to divide the design tools. More nuanced classifications can then be arrived at either theoretically, or through empirical research.

7.1 Political tools for design

The *political* design tools are those that are aimed at increase in the institution's robustness, that means at increase in the level of support for transmission of information the institution is supposed to transmit, relative to the opposition. Informally speaking, the purpose of the *political* design tools is to make sure that the institution enjoys sufficient political backing and that its information-transmission function cannot be impaired by individual actors who, because of their specific targets, have reasons to disturb the information flows relevant for the political system.

It is possible to divide the political design tools into two groups: those aiming at alterations of the actors' incentives for information-sharing and revelation, and those aiming at the institutions' ability to resist pressures from the powerful actors.

7.1.1 Interest-based tools: altering incentives of the actors

Rules and norms   We start the debate on suitable interest-based design tools by considering the case of institutions as rules, or passive structures. One key way how to ensure, through design of the rules, more support for information transmission is through alterations of the incentives (and hence of utility calculations) of the cooperating actors. This approach is based on the strategies the actors adopt in hiding their private information, and on that when we take these strategies into account in designing the institutions, we may be actually able to induce the actors to reveal the information because under the given design it is in their interest to do so. The literature most directly addressing these problems is that on mechanism design Hurwicz, 1973; Myerson & Satterthwaite, 1983; Holmstrom & Myerson, 1983, concerned precisely with development of institutional schemes (mechanisms) in which

the actors find it in their interest to truthfully reveal their private information even if, under alternative designs, they would clearly prefer to keep it private.

For example, Myerson has argued that the specific way many political regimes have been set up reflects the need of the founders – or in general of any leader or group of leaders – to commit credibly to certain positions and thus induce right at the outset support from the important power groups (Myerson, 2008). Specifically, Myerson argues that a successful regime needs to posses mechanisms through which the key supporters of the leader can communicate effectively to each other their grievances towards the leader, i.e. his lack of reward to them, and if they decide so, to remove him from power. Various courts in which these supporters – e.g. nobility, military captains – can regularly meet and discuss the leader's treatment of them individually provide such functions. In other words, it is the clever design of a powerful court of supporters that enables any founding leader (a king) to reveal his private information about his intentions to his supporters credibly; by setting up some court that can be strip him of power, the ruler demonstrates to the key power players his intention to care about their interests.

Important insights into how design induces actors to reveal private information may be obtained also from signalling games and screening games research (Akerlof, 1970; Stiglitz, 2000). Stiglitz, for example, argues that one of the key functions of institutionalized educational systems in many states is to provide the opportunity for the future employers to 'screen' the applicants, i.e. that the mechanisms of standardized testing and grading serve as important information-transmission mechanisms, whereby an information about the applicants' capabilities, that would otherwise be private to them in the application process, can be collected by the employers from the applicants' past formal records (Stiglitz, 1975). Similarly, Akerlof discusses how such institutions as guarantees (based on credible commit-

139

ment to quality) or brand-names (based on reputations) can help participants in market interactions overcome the information asymmetry problems (Akerlof, 1970).

The strategic information transmission literature also shows interesting insights. Gilligan and Krehbiel have argued, for example, that some restrictions on self-selection into legislative committees (i.e. ability of high demanders to get a position in the committee) increase information efficiency of the legislative process. That means that preventing policy outliers from getting positions in committees lead to better information of the plenary legislatives about the discussed issues. Similarly, they argue that the closed rules in which neither committee members nor the plenary can make amendments to the existing proposals conveys more information than rules with the possibility of ammendments (Gilligan & Krehbiel, 1989). In IR, the major insights of the signaling games approach have become part of the mainstream: in particular, signals of intentions and interests, to be credible, need to be sufficiently costly for the senders, were they to back off from them (Fearon, 1994). Morrow, for example, argues that the specific information asymmetry conditions during wars with respect to treatment of the prisoners of war has led the states to granting of monitoring authority to state or even non-state (Red Cross) third parties. Such a mechanism, despite its many limitations, provides a relatively high level of information transmission even under the extremely hostile conditions (Morrow, 2001). Hence, the states that are willing to comply with the monitoring scheme (Red Cross) show credibly their willingness to comply also with their commitment to treatment of the prisoners of war. In general, though, clearly more specific and empirically tested studies of actual information-transmission enhancing features of international institutions are clearly needed.

All these insights are no doubt valuable, and indeed they have earned several of the authors just mentioned Nobel Prizes in Economics. Yet, the potential prescriptive power

they bring to the study of international institutions' design may be more limited than we might hope.

First, many of these studies, especially the formally more sophisticated ones, rely in their results on strong solution concepts of the games they model; in other words, their results are only valid to the extent to which we believe the actors to actually play the games as the modellers assume (for instance to perform Bayesian belief updating). Since in the information-oriented games the solution concepts tend to be particularly demanding for the actual actors, this may be a more important problem than usually (cf. Rasmusen, 2006). For instance, in the signalling-games-based analyses of committee compositions and the quality of their signals mentioned above, competing models with starkly different predictions have been developed. The differences of these predictions are driven only by the notion of equilibrium behaviour that was used in the analyses, i.e. both studies are game theoretical, they both deal with the same actors, they just use a different notion of how the actors are likely to act under the same conditions[1](Krishna & Morgan, 2001; Krehbiel, 2001).

Second, many of the specific models make strong assumptions about the structure of the interaction, in particular about the sequences of the moves. This, however, does not seem warranted in the political analysis in which – as we discussed in chapter 2 – informal politics and by-pass of formal rules are usual rather than exceptional.

Third, there have so far not been that many works in the literature that would actually yield results from which one could directly derive some specific principles for design of international institutions. The strand of literature concerned with prescriptive design analysis – the mechanism design literature – has found it very difficult to come to positive solutions

---

[1]Which is obviously very common in game theory. Even the simplest game, such as the battle of the sexes, may be solved differently, depending on what solution concept we choose (the famous Nash equilibrium being only one possibility among many).

in other than very specific environments, such as those of auctions. Translation of these findings into political contexts have so far not very successfully followed (cf. McCarty & Meirowitz, 2007, ch. 11).

In sum, unless we are confident about actual empirical applicability of the solution concepts of the games, and unless we may have some reasonable confidence that the structure imposed on the actors by the designer can be actually enforced in the international political context, the value of the research relevant for the indirect (strategy oriented) design tools may lie primarily in the general insights about the incentives of the actors for certain behaviour.

International organizations   If we consider now, instead of rules and norms, international organizations, the simplest way in which they can induce the actors to reveal their private information is through explanation and persuasion. In some cases, the IOs may be able to obtain from the cooperating actors the necessary information if they are able to explain to them and compel them that this information is needed for the cooperative system to function. The design tools that may enhance the IO's ability to do so are also simple – it is enough if the IO has sufficient resources available to prepare compelling analyses of why the information is necessary. We can expect, though, that the practical scope of use of such tool will be rather limited as the cooperating actors may in general be well aware of the information needs of cooperation, and when they decide to conceal some information they likely do it for some specific purpose (not out of ignorance).

7.1.2 Power-based design tools: creating new powerful actors

An alternative way to increase robustness of an institution – alter the support/opposition relation – is through a change in the power distribution underlying the cooperation system.

Rules and norms   Starting with the structural (passive) institutions, clearly these can only be considered to dispose with power resources if we adopt the notion of structural power, i.e. power exercised not by individual intentional actors, but by the very structures the actors inherit (Barnett & Duvall, 2005). This is most natural for the critically oriented Gramscian approaches that stress the role of ideology in politics, and the way in which embededdness of the actors in the cultural and social environment prevents them from realizing their genuine interests (Lukes, 2005; Cox, 1983).

In the context of international rules and norms, those norms that promote transparency, openness, and information-sharing among the cooperating actors clearly contribute to smooth information-transmission. It can be argued, for example, that the Wilson's insistence on public and open diplomacy after the first world war was a case of an attempt to create a strong norm, albeit perhaps not very successful one, targeting information-transmission behaviour of the actors (cf. Sofer, 1988; Nye, 2008). Usually we do not perceive such norms as transparency and openness as coercive, but it is probably fair to say that the diplomatic environment today is more open than that one hundred years ago, and that this openness is by most actors *taken for granted.* In this sense, the conscious creation and definition of the norm after 1918, and its later development, have effectively constrained the actors' choices and contributed to smoother information flows in cooperative systems.

In a rationalist framework this structural notion of power is on its own problematic as in principle the actors should be able to assess their own interests and should not let themselves

be compelled into believing in something that is tangent or even opposite to their interests. In some sense, this can be understood as the core assumption of the rational choice theory, together with preference transitivity (e.g. Shepsle, 2010). In a broader rationalist framework, though, the impact of the structural power of rules' and norms' design on behaviour could be relatively easily conceptualized as a bias in the actors' belief updating mechanisms, i.e. as an imperfection in their ability to perceive accurately their environment and learn from it.

International organizations   The much more obvious way in which alterations in power distributions enhance the information flows is through effective creation of new powerful actors in whose interest it is to have the information transmitted. This consist first and foremost in delegation of more authority and resources to the bodies that are supposed to collect the information, the international organizations (Hawkins *et al.*, 2006; Pollack, 1997). Individually, states may try to conceal some information and hence prevent effective monitoring of their behaviour. Alternatively, they may use their power resources (e.g. budgetary contributions) to influence the IOs in their monitoring of other states. Either way, it is essential for the bodies supposed to collect, process, and disseminate information to possess the resources necessary for elimination of these pressures. In other words, the bodies need to have power, and if we expect them to be exposed to pressures from other actors, they may simply need to be endowed with more resources and to be able to face it.

The standard autonomy-enhancing mechanisms – such as independence of budgets and in internal matters – may be the right tools (Reinalda & Verbeek, 1998). So, for example, the European Commission and the European Court of Justice have been shown to possess significant monitoring capabilities that have been delegated to them and that facilitate coop-

eration among the EU member states (Pollack, 1997; Tallberg, 2002). Similarly, strong and independent dispute settlement arrangements in the WTO have been expected to secure that the rulings are, as compared to GATT, less subject to pressures from the powerful members (Busch & Reinhardt, 2003; Smith, 2004). In my perspective, this corresponds to a better performance of the information transmission function, in which the institution is better able to transmit to the cooperating actors the messages about the behaviour of all the counterparts. It has been also argued that the U.S. were able to influence the conditions under which the IMF provided loans to the US' allies (Momani, 2004 – autonomy and independence of IOs from such pressures is necessary if they are to fulfil their information-transmission function even against the interests of the powerful actors.

Over the last several decades, we have witnessed increases in authority of IOs (Zürn *et al.*, 2012). Interestingly, Hooghe and Marks show directly that concerns with information processing, precisely in the sense discussed here, impact on the territorial composition of states, i.e. whether they are unitary or federal (Hooghe & Marks, 2012). It seems that empirically the use of delegation is among the most popular design mechanisms through which information-transmission among the actors is enhanced.

Delegation creates problems of agency costs, as well as of accountability. Clearly highly relevant for these purposes is the literature on the principal-agent relations, although this time rather from the opposite perspective than above, i.e. from the view of how autonomous international bureaucracies obtain significant information advantages over the principal actors, the states, and thus effectively preclude monitoring of their own behaviour (Nielson & Tierney, 2003). As a result, serious concerns can be raised as to their potentially pathological behaviour (Vaubel, 1986; Barnett & Finnemore, 2004).

Delegation of authority and power resources to some independent agents may be very

difficult also for political reasons (cf. Zürn *et al.*, 2012; Hooghe & Marks, 2003). Yet, the amount of resources that needs to be delegated in order to secure the basic task of information transmission may be surprisingly low. This is because of the specific nature of information – already a demonstrated unwillingness of actors to let themselves be monitored or to transmit the information that they should transmit implies harm to their reputation, it signals that they do not cooperate. In this sense, in many circumstances the information-collecting body has a large advantage already at the outset of the interaction. In some situations, though, crude bargaining power given e.g. by financial resources of the institution might be necessary – only with this power may the IOs be able to induce the actors to reveal the necessary information.

Table 7.2 summarizes all the political tools. In general, the problems identified with our current knowledge of mechanism design are rather daunting. Prescriptively, then, we can clearly have much more confidence in the solutions based on delegation of authority and resource to bodies in whose direct imminent interest it is to have the pertinent messages transmitted.

7.1.3   Going down the levels

Beyond the interest-oriented and power-oriented tools there is a class of design tools that are also oriented at robustness of the institutions but that are of completely different type. So far all the debate, with the exception of a short mention of the liberal approach to international institutions in chapter 2, was exclusively concerned with cooperation on the inter-state level, where for the purpose of the analysis the states represented by their governments are considered as the key actors, and it is the governments' decisions that determine the cooperative outcomes. The analysis has been conducted solely in the Waltzian third image (Waltz, 1954;

	Interests	Power
Passive, structure-based (rules)	Inducing private information revelation through tools from mechanism design theory	Norms of transparency, information-sharing
Examples	Sequencing of moves, rules structuring the interaction	Promotion of ideology of openness, embedding rules in broader ideological framework of transparency
Active, agency-based (IOs)	Campaigns for information-sharing (explanation and persuasion)	Power and autonomy of the information-collecting bodies (IOs)
Examples	Provisions of legal and material resources to IOs for such campaigns.	Delegation of authority and resources to the IOs

Table 7.2: Political design tools

cf. Singer, 1961). Yet, clearly international institutions do have effects also on the lower levels, and in this sense design tools focused on these effects may also be relevant.

Moving down from the international level to the level of domestic politics, the institutions may be able to induce transmission of vital cooperation information through involvement in the domestic politics of the relevant states; that means, they may attempt to alter the behaviour of the states by a bottom-up pressure on the governments. A prime example is the human rights regime where, as elaborated by Simmons (Simmons, 2009), the international organizations have been able to modify the governments' behaviour through empowerment of the domestic civil society groups.

Consequently, it is possible to endow, through appropriate design of their mandates, the institutions with th ability to enter the domestic political realm of the countries via empowerment of the interest groups that, for example, push for more transparent behaviour

of the governments in the pertinent areas. Support by the World Bank and World Trade Organization and other organizations for non-governmental projects aimed at transparent public procurement is an obvious example, constituting an important part of the Doha Round of multilateral trade negotiations within the WTO (Parízek, 2012b). Another example is the already mentioned human rights regime where the enhanced ability of domestic non-governmental organizations to monitor violations of human rights creates on the governments additional pressure (Simmons, 2009). These examples, however, also show the practical limits of such an approach, as the willingness of the governments to subject themselves to the potentially increased domestic pressures, due to activities of international institutions, is very limited. Contemporary Russian Federation, for example, has been systematically preventing international institutions from supporting its domestic democratically oriented non-governmental organizations (BBC, 2012). Design changes increasing the ability of the institutions to penetrate domestic politics of the states are unlikely to prove effective, unless the states themselves want to make use of the institutions to signal their genuine intentions.

On the individual level, in the first image, better transfer of information needed for co-operation may be secured through one specific tool: creation of individual legal obligation of certain actors to actually transmit the information they are supposed to transmit. The motivating example of such an approach may be the international criminal law, where individuals, even if they act on behalf of their governments, can be sued (cf. Popovski, 2000). Clearly endowing the institutions with such possibilities will always face stark opposition from at least some states, as exemplified by the approach of the U.S. towards the International Criminal Court. The depth of cooperation argument is at play here – states do not like taking commitments they realistically expect they might want to break (Downs *et al.*, 1996). Yet, one can imagine that in areas of overwhelming concern for the majority of states,

such as nuclear security, states might be willing to subject their own public officials to international legal sanctions in case they do not properly report to the international institutions. This is obviously only a hypothetical example and to my best knowledge there are no such arrangements in the present international law, but it clearly constitutes a logically possible alternative way through which design of institutions can enhance their capacity – it can give the key actors in the process strong personal incentives to transmit the information, even against the potential pressure of their own governments.

The obvious limits on the design tools dealing with the lower levels are given by that they are effectively avoiding the international political contest; the desired effects of the institutions' design are imported into either domestic politics or to the individual legal realm. In both cases, these tools can only work if the concerned governments provide at least passive support, i.e. the design of international institutions is unlikely to be modified in the suggested directions if some of the states fear that their interests would be hurt. Against the will of the powerful states, such strategies are likely to work only if the institutions, especially international organizations, actually do posses the resources to implement such strategies autonomously.

7.2 Technical design features

Politically and analytically the simplest, but perhaps the most practically relevant class of design solutions is concerned with the more technical aspects of design for information transmission, and relates to the technical capacity of an institution. In this perspective, institutions may fail their information-transmission tasks even when the cooperating actors in principle want them to function, because they do not have the technical means to perform

their functions. The range of technical tools to address lacks of capacity can be classified –
based on the information theory – into two groups: coding-based, and channel capacity-based
(again, see the elaboration in Shannon & Weaver, 1949).

7.2.1 Technical design features concerned with coding

Each message, if it is to be transmitted, needs to be coded in some way, that means it needs to
be formulated by the senders in a form that makes it transmissible and then comprehensible
for the receivers. How the message is coded has a significant effect on whether, how, quickly,
and how reliably the information in the message can be transmitted[2]. As in the previous
section, I start the discussion of the appropriate design tools with those addressing capacity
of rules and norms, i.e. with how the capacity of the structures that shape the actors'
interaction can be increased.

Rules and norms Rules and norms are statements about appropriate behaviour, i.e.
messages stipulating what the relevant actors are expected to do or not to do, and possibly
also what sanctions they can expect in case they do or do not follow the prescriptions. The
key for successful cooperation, then, lies in how successfully these messages are transmitted.
One way to improve the prospects for transmission, i.e. to increase the capacity of an
institution as of a communication channel, is to formalize the messages, that means to
state them in such terms in which their meaning will become more readily apparent, and
less subject to widely divergent interpretations. Formalization may, for example, consist
in detailed definitions of the key previously vague concepts around which the cooperative

---

[2]Note that I do not refer here to the problem of language and shared meanings, as this problem would
lie one level deeper than the rest of my analysis. I am concerned with specific way how information relevant
to international political outcomes can be coded, leaving the language and its intricacies outside (constant).

scheme is oriented. A specific type of formalization may be also quantification of the key terms of cooperation, i.e. creation of clearly stated and easily interpreted measures of what does and what does not constitute compliance (for instance specification of the parameters of permissible military equipment in de-militarization talks).[3]

The instances of formalization of messages in international politics are numerous, the most obvious example being the creation of various quantitative measures of compliance and indexes of performance. So, for example, the currently hotly debated measures for slowing down of the climate change have to a considerable extent revolved around establishment of the quantitative target for increase in the temperature level of no more than 2°C as compared to the pre-industrial level (UNFCCC, 2009).

A special case of a design change of rules resulting potentially in increased information-transmission capacity is legalization, which not only translates the messages (e.g. conditions of some agreement) into more formal terms, but which also embeds them into a whole legal-interpretative system. This system not only possesses an elaborate specialized language, but also numerous mechanisms that evolved in it to determine the precise meaning of prescriptions and the degree to which actors' behaviour reflects these (Abbott *et al.*, 2000).

Embedding the rules in a legal system is potentially a highly efficient complexity reducing mechanism, as it enables reduction of the entire complexity of the actors' behaviour to the binary statement comply/not-comply, or to a statement about the degree of non-compliance. The case of the WTO dispute settlement mechanism is very instructive in these terms, as it directly requires that the amount of damage caused to the cooperating partners by measures

---

[3]Note that in the Introductory chapter I define the concept of design as the set of explicit formal features of which the institutional arrangements consists, i.e. the very notion of design already entails some level of formalization. Yet, this level can vary significantly, from short verbal statements about shared understandings to highly specific elaborations on the terms of cooperation. In this sense, formalization is a specific design feature that may vary in its amount.

adopted by one of the actors is expressed quantitatively. In the next step, then, the affected parties to the agreement have the right to adopt measures themselves that will bring them unilaterally benefits in the amount corresponding to the damage they suffered (cf. Desmedt, 2001). In other words, the elaborate dispute settlement mechanism is based on the notion that all the complexity of compliance and non-compliance can be reduced to the simple quantitative statements of how much damage was made and, in effect, how much needs to be provided in compensations.

It should be clear that formalization and legalization are not thought to constrain the actors' behaviour by their inherent normative appeal, but by the virtue of the fact that they enable easier and better quality de-coding and interpretation of the information about what constitutes appropriate behaviour, and of conformity of the actual actors' behaviour with the standards.

International organizations A very different way in which the information coding procedures may be improved is through centralized, harmonized, or simply more effective training of the staff whose work it is to transmit the information, be it of the national staff responsible for the transmission on the domestic level (i.e. of national bureaucrats) or on the level of the particular IOs.

Programmes for training of public officials collecting statistics within individual states may be an example. Especially in the context of the recent problems with reliability of national statistics in the Eurozone, it is clear how essential harmonization of the procedures through which the key data are collected is. Similarly, especially for smaller developing countries it may be important that they receive technical training for their international law specialists. In the WTO, again, the issue of lack of technical ability of some of the countries

to deal with the legal agenda has been identified as a major obstacle for cooperation (Busch & Reinhardt, 2003). As trivial as it sound, one example from this class of design tools may lie in improved and more extensive translating services provided for the cooperating parties (EU being the prime example in heavy investment in these services).

Although these mechanisms aimed at improved coding may seem as not being of primary political importance, we should be aware of their direct relation to the VSM and the Law of Requisite Variety from which it is derived. Unless the information that needs to be transmitted among the actors is coded in such a way that actually allows the transmission and easy decoding, the entire cooperation system may not be viable. Furthermore, as all governance schemes, the international cooperation systems face the daunting problem of complexity. If they are to be viable, the information they receive and on which they base the decisions needs to be as concentrated and efficiently coded as possible, otherwise the imminent threat of information overload may materialize. It is therefore essential for principle reasons that complexity the governance systems deal with is effectively reduced into (short) pregnant comprehensible statements that provide all the necessary information as concisely as possible.

7.2.2 Technical design features 2: channel capacity

The logical alternative to the problem of coding is the increase in the information transmission capacity of channels as such, i.e. strengthening of the very mechanisms that transmit the coded information. Again, the design tools addressing this technical problem seem unproblematic from the political perspective but the consequences of their absence are detrimental for cooperation.

Rules and norms   In the case of rules and norms, their technical information transmission capacity is directly given by their length. In particular, provided that on average each paragraph of an international treaty contains specification of appropriate behaviour of one aspect of the cooperative scheme, the capacity of the agreement is given by the number of articles or paragraphs. The more extensive the agreement (a set of rules), the larger share of social reality it covers, i.e. it transmits information about appropriate behaviour in a larger number of more varied situations. To be sure, this only applies on average, but *ceteris paribus*, the more elaborate the agreements, the higher their technical capacity.

International organizations   In international organizations, the most obvious way in which their capacity to transmit information can be technically increased is through the increase in the amount of disposable work-hours of the personnel, i.e. effectively through increases in the number of staff assigned to the information transmission related tasks, such as collection, processing, and dissemination of the available data.

In other contexts, it may be possible to increase the capacity of the decision-making bodies through increases in the number of meetings among the key actors. If we understand the decision-making process as an effort to find mutually acceptable solutions to specific problem, giving the actors more time to exchange views and find out who wants what may be essential, however trivial this may sound. A nice example of such a problem is the decision-making in the Council of the EU after the eastern enlargement. As documented by Best and Settembri (Best & Settembri, 2008), there the ministers simply due to the rising time-demands of the negotiations increasingly opt for sending their deputies or other staff to the meetings.

In some cases, the solution might actually lie in creation of new bodies within the existing

154

ones. If we take again the case of the EU, I will argue in chapter 8 that precisely because of the lack of time-resources of the ministers in the Council the subordinated administrative body, the Committee of Permanent Representatives, has been gradually taking more of its work. This means that effectively some pieces of agenda that have been previously decided by the ministers have now been transferred for decisions to a different body which has more capacities for it.

A somewhat different but related example may concern staffing of IOs aimed at increased variety across professional fields, so that the secretariats are able to comprehend higher variety of the social reality they deal with (e.g. not only economic, but also social and political factors) or across paradigms (conceptual frameworks within which the actors situate their interpretations of events and statements). A special case of this problem is the distribution of staff of IOs across the member states, in particular the large excess representation of the OECD countries, in comparison especially to the African countries. From an information-transmission perspective, such an imbalance is highly undesirable as it reduces the amount of soft context information the IO may need for its activities in the developing countries. I present an empirical analysis of this phenomenon in chapter 9.

Table 7.3 provides a summary of all the technical design tools mentioned in this chapter. As in the case of political design tools, the list is by no means exhaustive, but it provides a clear indication of the kind of design features that we may expect – on the basis of the theoretical argumentation presented in the previous chapters – to have a positive effect on the prospects for cooperation.

	Coding	Channels
Passive, structure-based (rules and norms)	Clarification and easier interpretation of commitments	Scope of coverage of social reality with rules
Examples	Formalization, legalization.	Increased length of agreements
Active, agency-based (IOs)	Improved ability of IOs to encode and decode information efficiently	Increases in number of available work-hours
Examples	Training of the personnel in statistics, in law, in information gathering and processing techniques	Hiring more personnel

Table 7.3: Technical design tools

7.3 Technical and political tools: complements and substitutes

A whole range of design tools has been briefly discussed in the previous sections – tools of both political and technical nature, i.e. dealing with robustness of the institutions or with their technical capacity. Application of any of the design tools is on its own likely to increase the capacity of the institution. Yet, different tools or approaches may provide variable improvement in different situations. In this last section of this chapter I present a general scheme for decisions on when the technical tools be more appropriate than the political ones, and vice versa.

7.3.1   Technical and political design tools as substitutes

On the first sight, the capacity equation from which a considerable part of this discussion on design is directly or indirectly derived implies that robustness and technical capacity are complements, i.e. that to achieve the desired target of high capacity we need both robustness and technical capacity. Yet, the relationship between the two components is to some extent asymmetrical. In many conceivable situations, increases in technical capacity may well provide effective substitution for the lack of robustness. This has two reasons.

First, already from the fact that $R \in \langle 0; 1 \rangle$ while $M \in \langle 0; \infty \rangle$ it is apparent that the space for increases in capacity is much higher in the technical domain. The political domain serves more as a moderating force, where very low robustness may prevent the institution from functioning, but even very high robustness does not ensure high capacity (while even with moderate robustness, high technical capacity may ensure a high enough capacity). To be sure, the technical capacity itself will, at some point, face the law of diminishing returns, i.e. adding new technical capacity will be increasingly costly, with the given available tools.

Second, robustness tends to have steep positive effect on capacity only at low levels – at higher levels its contribution diminishes very quickly. This is given by the concave shape of the robustness function, with steep rise close to the minimum but very flat rise later on. Consequently, it is plausible to assume that in many situations the power- and interested-oriented design tools, if we are unable to make use of them or if they fail, may be substituted with the technical ones. In other words, for most politically relevant information the institutions may be able to obtain and transmit it through simple deployment of a larger amount of technical capacity, even if the degree of the institution's robustness is lower.

This corresponds to the earlier observation that the interest-oriented design tools may be difficult to use, i.e. that because of the complexity of the situations we deal with, we may

simply not be able to come up with clever-enough design solutions. In that case, we may want to substitute strategic sophistication (which we necessarily lack) with either power-oriented tools, such as delegation of authority to IOs, or with simple technical tools.

There are clearly situations in which even high technical capacity will not be sufficient for transmission of the key information, as the relevant actors may be able to keep the information private and effectively undetectable. In such situations, the power- and interest-oriented design mechanisms will be necessary, and robustness *will* be the key.

It can be argued, though, that such situations may be surprisingly rare. To be sure, information e.g. about preferences of autocratic regimes may be genuinely private in that even the close collaborators of the rulers may not have access to it. Yet, this is likely to be a rare case (and one in which any standard political tools would fail to induce information revelations as well). Within complex organizations, such as states, even the most secret information must flow among the individual actors; the information is being transmitted somewhere. Usually, when robustness surpasses some relatively low levels, lack of further robustness can be substituted with more technical capacity.

This is very suitable also practically. We certainly cannot hope to be able to model interests and power relations among the actors and, as discussed, our ability to devise strategically sophisticated mechanisms that would give the actors the right incentives for information sharing is more than limited. As a result, increasing the technical capacity may prove to be the only practically available option for improved institutions' functioning. Corresponding to this, the two empirical chapters (8 and 9) deal with the technical aspects of the institutions' capacity. In the following section, I briefly describe the general principles which may guide the choice of the appropriate approach, i.e. the choice between technical and political tools.

7.3.2 Substitution rules: formal expression

Which tools should be used in individual situations can be formally describes with a system of inequalities based on the marginal benefits of use the use of the individual tools. For each of the components of the capacity equation

$$
C = \frac{MR}{t}
$$

$$
= \frac{M\left[\frac{supp}{supp+opp}\right]}{t},
$$

the marginal effects of increases in technical capacity (E_{tech}), support (E_{supp}), and opposition (E_{opp}) are given by the following partial derivatives (note that, without any effect on the analysis, we can exclude the unit time t):

$$
E_{tech} = \frac{\partial C}{\partial M} = \frac{supp}{supp + opp}. \tag{7.1}
$$

$$
E_{supp} = \frac{\partial C}{\partial supp} = M\left[(supp + opp)^{-1} + supp(-1)(supp + opp)^{-2}\right] \tag{7.2}
$$

$$
= M\left[\frac{(supp + opp) - supp}{(supp + opp)^2}\right] \tag{7.3}
$$

$$
= M\left[\frac{opp}{(supp + opp)^2}\right]. \tag{7.4}
$$

$$
E_{opp} = \frac{\partial C}{\partial opp} = M\left[supp(-1)(supp + opp)^{-2}\right] \tag{7.5}
$$

$$
= -M\left[\frac{supp}{(supp + opp)^2}\right]. \tag{7.6}
$$

These equations then identify the conditions under which tools addressing technical ca-

pacity or support and opposition, respectively, should be applied. In particular, tools addressing technical capacity should be applied if

$$\frac{\partial C}{\partial M} \geq \frac{\partial C}{\partial supp} \text{ and} \tag{7.7}$$

$$\frac{\partial C}{\partial M} \geq \frac{\partial C}{\partial opp}, \text{ i.e. when} \tag{7.8}$$

$$\frac{supp}{supp + opp} \geq M \left[\frac{opp}{(supp + opp)^2} \right] \text{ and} \tag{7.9}$$

$$\frac{supp}{supp + opp} \geq M \left[\frac{supp}{(supp + opp)^2} \right]. \tag{7.10}$$

Note that in the last equation the direction was switched into positive, so the design tools will not be used to increase opposition, but to decrease it. Inequality 7.9 holds when

$$supp \geq M \left[\frac{opp}{(supp + opp)} \right], \text{ i.e.} \tag{7.11}$$

$$\frac{supp(supp + opp)}{opp} \geq M \tag{7.12}$$

$$\frac{supp^2}{opp} + supp \geq M, \tag{7.13}$$

and inequality 7.10 holds when

$$supp \geq M \left[\frac{supp}{(supp + opp)} \right], \text{ i.e.} \tag{7.14}$$

$$(supp + opp) \geq M. \tag{7.15}$$

Conditions 7.13 and 7.15 then jointly identify the situation in which technical, as opposed to political design tools should be used. If they do not both hold, political tools will have, *ceteris paribus*, higher positive effect on the institutions' capacity.

Differencing further between the problems of low support or too high opposition, as long as

$$supp \geq \pm\sqrt{opp}, \qquad (7.16)$$

condition 7.13 is more restrictive than 7.15 and hence addressing lack of support yields higher benefit than addressing high opposition. Otherwise the reverse is true.

Figure 7.1 illustrates the relationship between the use of the components graphically. Since we can only plot two independent variables, we hold in this case the level of *opposition* constant at 0.5, and show how changes in support (horizontal axis) and technical capacity (vertical axis) impact on the overall capacity, captured by the rising contours (the lightest colours, in the tor-right corner, show the highest capacity). It holds (and is given by the equation above) that at any point the tool should be used that secures steeper increase of the function, i.e. steeper rise towards the lighter areas in the graph.

Last, considering the problem of costs of the individual available tools, if unit costs of the tools addressing technical capacity, support, and opposition, are $p_{tech}, p_{supp}, p_{opp}$, respectively, the choice of the tools should reflect the marginal effects of these tools and their price ratios.

Hence, the tools will be used optimally when conditions
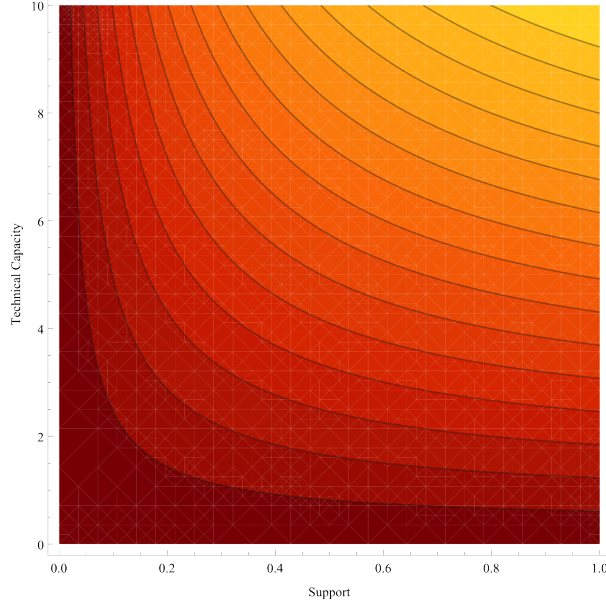
Figure 7.1: Capacity as given by support and technical capacity (with opposition held constant at 0.5)

$$\frac{E_{tech}}{E_{supp}} = \frac{p_{tech}}{p_{supp}}, \tag{7.17}$$

$$\frac{E_{tech}}{E_{opp}} = \frac{p_{tech}}{p_{opp}}, \text{ and} \tag{7.18}$$

$$\frac{E_{opp}}{E_{supp}} = \frac{p_{opp}}{p_{supp}} \tag{7.19}$$

all hold. If, for example, $\frac{E_{tech}}{E_{supp}} > \frac{p_{tech}}{p_{supp}}$, re-allocating part of the resources from political tools (in this case tools addressing the problem of low support) to the technical tools would increase the achieved capacity, under the same budget constraint.

7.4 Institutional design revisited

In general, what I present in this chapter is a notion of design that aims at dealing with complexity. With the tools I describe here, we can expect positive increases in capacity. The tools focus on the vital functions of the system, as discussed in chapter 3.

We should understand that the notion of information may in many contexts be relatively subtle. Take the example of the 50% majority voting rule. From the perspective developed here, it does not matter that much if the rule is based on a 49% or 51% share – after all with an even number of deputies the simple majority amounts to 50.5 anyway. What matters, though, is that the 50% thershold symbolizes the majority principle, the underlying decision-making principle of democratic competition. The specific form (49-50-51%) does not necessarily matter for constraining the political actors, as in many ways political struggles depends on power and interests rather than on formal rules. But form of the institutions – their design – does matter for how effectively information is transmitted within a system. So a rule that says "try to find what most deputies think" is much less specific and more open to diverging interpretations that a norm "count the votes and the side that has 50%+ wins". The latter design has a much higher ability to transmit the message about the underlying principle of decision-making, it endows the rule with much higher capacity.

Again, this is a relatively subtle problem, because we are used to thinking about institutions as rules. This may present little problem many research contexts, where rules simply imply compliance. But in politics this is not always so. Rules do not imply compliance as compliance is given by both rules and interests of the actors. When we want to understand the causal effects of the design of the rules on the prospect for cooperation (or prospects for compliance), we need to focus on how the design of the rules affects their ability to transmit

effectively the messages they should transmit, i.e. messages about what behaviour is appropriate and what is not. So what we need to focus on when considering appropriateness of the design of institutions is how it impacts on their capacity. If we focus our resources on capacity of the channels that transmit the information that is essential for the system as a whole, we are increasing the chances that the system we deal with will be viable, i.e. we are increasing the prospects of cooperation.

The purpose of all these proposed design steps is to make sure that more information transmission capacity is available for the key channels forming the cooperation system. Some of the proposed steps may appear uninteresting from the perspective of political analysis, and not enough sophisticated for dealing with such a subtle issue as transmission of information. In principle, my argument is that given all the complexities and unpredictability, we may opt for the most direct power-based solutions where we endow those actors with interest in the transmission with more power, or for the simple and obvious technical solutions aimed at increased technical capacity of the information channels. This means, however, that the real design problem may not so much lie in more clever or sophisticated design but instead in the analysis of functioning of the international institutions, i.e. in careful diagnosis of their deficiencies. The concept of capacity and the VSM discussed in the previous sections are the tools for such diagnosis. Once we know what information is missing and what information flows are insufficient, finding out the institutional remedy may not be as difficult. The tools for remedy are simple but potentially powerful, if we are able to find out where the information-problem lies.

Part III

Empirical Applications

Chapter 8

Eastern Enlargement and the Capacity of the Council of the European Union

The ability of the European Union (EU) to take decisions was expected to be strongly affected by the 2004/2007 Eastern enlargement of ten-plus-two new member states[1]. It was understood that the dramatic increase in the number of member states, amplified by their structurally distinct political and socio-economic systems and considerable cultural differences, might paralyse EU decision-making processes. The institution expected to be hit most strongly by these effects was the Council of the EU (henceforth the Council), due to its notable intergovernmental character.

Several years after the enlargement, I try to use the framework developed in this thesis to understand why this threat did not materialize, i.e. what factors helped the Council 'survive'

---

[1] This chapter to some extent draws on a previous text by M. Parízek, M. O. Hosli, and B. Plechanovová, *Avoiding Paralysis: The Eastern Enlargement and the Council of the European Union* (currently under review in a peer-reviewed journal). My contribution to the co-authored text was of more than 50%. The present chapter takes part of the empirical material of the previous text as a basis and considers it from the perspective of the theoretical argument developed in the dissertation. A substantial part of the data that form the basis of the empirical analysis were collected within the GAČR 407/09/1747 project.

the enlargement and prevented it from becoming paralysed. How did the Council preserve its capacity to make decisions in an efficient and timely manner, in spite of the dramatic increase in EU membership?

The argument I put forward is that member states' representatives have delegated an additional part of the overall agenda from ministerial meetings to the Committee of Permanent Representatives (Coreper) and the Council working groups (WGs). As a result of this change, or with the help of it, the Council as a whole is still capable of operating smoothly, despite the current membership of twenty-seven states. While this argument is fairly straightforward, I show that the delegation mechanism may in practice face important obstacles. In order to explore the dynamics underlying this process in more analytical detail, I conceptualize decision-making in the Council after the Eastern enlargement and the threat of its paralysis as a situation of common-pool resource (CPR) extraction. With the help of a simple formal model based on this conceptualization, I show how the two-level design of Council decision-making based on the interaction of the ministerial level and the Coreper, and in particular the nature of the Coreper as a decision-making body, helped the Council cope with the effects of the Eastern enlargement.

I test the 'delegation-hypothesis' on the basis of quantitative data on all 3811 proposals transmitted to the Council by the Commission in years 2000 through 2007. The evidence capturing the amount of A-points and B-points on the Council agenda and the trends in their ratio allows us to assess the central argument, and further also to discuss the more general question of the relative involvement of ministers and permanent representatives in the actual Council decision-making.

The purpose of this chapter, hence, is to illustrate how an appropriate design of the main EU decision-making body helped it secure sufficient capacity to cope with the increased

complexity brought about by the Eastern enlargement. The key design factors that enabled it are 1) availability of Coreper as of an alternative decision-making forum, and 2) its very nature, i.e. its high stability and its encompassing scope of activity. These design principles, when combined, led the fact that the Council could have maintained sufficient capacity to come to decisions even after the enlargement.

The chapter is structured as follows: The next section reviews the previous research relating to expected impacts of the Eastern enlargement on the EU political system and specifically on decision-making in the Council. Section three develops the conceptualization of the problem as one of CPR extraction and formulates the central delegation-hypothesis. Section four presents a formal model which captures the interplay between the ministerial meetings and Coreper in an easily discernible two-by-two game; this game then sheds light on how the specific design of Coreper helped the Council survive the enlargement. Section five assesses the argument on the basis of quantitative empirical evidence.

8.1 The threat of Council Paralysis: Theoretical and Empirical Expectations

The Eastern enlargement has almost doubled the number of actors in Council decision-making and has brought a higher diversity of EU member states' interests. In cybernetic terms, it has significantly increased the requisite variety of the EU decision-making system, i.e. the complexity of the realities the system needs to deal with. Within a given institutional framework these changes can be expected to render EU decision-making more complicated in practice (cf. Hertz & Leuffen, 2011).

In the EU political system, more member states effectively implies more veto players

(Tsebelis, 2002) as a substantial share of decisions is in the Council still adopted by unanimity, which holds notably for some of the most politically contentious domains such as taxation and most aspects within the area of common foreign and security policy. Besides, the tendency within the Council is to seek as much consensus as possible so a wide support for a given proposal is in practice usually strived to (only about 19 percent of the legislative acts are contested by one or more ministers, see Plechanovová, 2011b). Even if the qualified majority voting (QMV) strictly applied in practice, the political core of stability is very likely to expand with the Eastern enlargement (Tsebelis & Yataganas, 2002; Tsebelis, 2008). Hosli and Machover (Hosli & Machover, 2004) analytically show that, if we assume independent coalition culture where there is no restriction in terms of assumed preference orderings of the members forming a coalition, with rising membership the share of winning coalitions in the Council in total decreases (with the effect being stronger under the unanimity rule as compared to QMV). All these general theoretical findings point at the expectation that an increase in EU membership is likely to lead to increased Council inertia: the Eastern enlargement has brought many more occasions on which one or several member states could find the emerging common position unacceptable and effectively veto, or at least substantially complicate, its adoption.

Analogous expectations are arrived at also in the more empirically oriented research on the EU political space. Some authors find Council politics to be largely determined by the traditional left-right ideological positioning of the member states' governments (Mattila, 2004; Hagemann & Hoyland, 2008). Others contend that the best predictor of actors' positions in the Council is their geographic location (Mattila & Lane, 2001; Thomson *et al.*, 2004; Zimmer *et al.*, 2005). Yet others argue that contestation in the Council is less given by a single general determinant but will vary depending on the particular policy issues at

stake (e.g. Hayes-Renshaw *et al.*, 2006). Various statistical techniques (Hagemann, 2007; Plechanovová, 2011a) and various data types (see König *et al.*, 2006) have been employed to assess the respective claims empirically but whichever turns out to be the closest to the empirical reality we can be confident that the new member states will occupy positions at a distance from those of the older EU members or, at minimum, that their positions will not be completely absorbed (Tsebelis & Yataganas, 2002). Indeed, recent empirical findings strongly support this view: research on dimensionality of Council politics shows that new salient divisions among the actors have emerged in the period after the enlargement and that the new MSs differ in their voting behaviour from the older ones (Mattila, 2009; Plechanovová, 2011b).

Consequently, as result of what Zielonka describes as an 'enormous injection of economic, political, legal and cultural diversity' brought about by the enlargement (Zielonka, 2007, p. 188), we can expect an increased pressure on the Council, as well as on the EU political process more generally (Hertz & Leuffen, 2011; König & Bräuninger, 2004).[2]

Technically speaking, higher number of actors as such does not necessarily lead to problems as increased membership 'expand[s] the possibilities for tradeoffs among the members' and thus possibly enhances cooperation (Koremenos *et al.*, 2001, pp. 784-85). In this manner, König and Junge (König & Junge, 2009) explain the fact that no Council policy-gridlock seems to occur after the enlargement on the basis of extensive log-rolling within the specific ministerial meetings and in Coreper. Although certainly plausible, this insight only holds when one disregards the potentially significant transaction costs inherent in conduct of the negotiations, where by transaction costs I mean the costs incurred by the actors due to their

---

[2]It should be noted that Golub (Golub, 2007) comes to different results, arguing that more states in the Council may actually speed the legislative process up. However, this argument presupposes that the newcomers' preferences do not diverge significantly from those of the old members (Golub, 2007, p. 169).

need to communicate to each other their interests, including saliences they attach to the given issues, and to devise – on the basis of this information – proposals that will be acceptable for a sufficient number of actors. In other words, we need to take into account that 'bargaining over side-payments and package deals increases the duration of the decision-making process' (Schulz & König, 2000, p. 656) as figuring out the right balance of powers and interests involve potentially significant amounts of communication and information processing by the actors.

So while König and Junge (König & Junge, 2009) do offer an answer to the problem of how the Council manages, after the enlargement, to avoid policy-gridlock and reach decisions, they do not address the question of how it manages to reach them within a reasonable amount of time, i.e. whether it has the necessary capacity to cope with its agenda in a timely manner. The purpose of this study is to fill this gap and thus complement the existing analyses. In this light, and in the context of the theoretical and conceptual debates presented in this dissertation, I propose to conceptualize the dangers faced by the Council as those of system overload (Deutsch, 1963), in particular overload of what is in the viable system model labelled as System Three, the system that secures internal functioning of the EU. This threat of overload results from the increased variety brought to EU politics by the Eastern enlargement, and hence from the possibly insufficient capacity of the decision-making system.

8.2 Council Decision-Making Capacity as a Common-Pool Resource

A Council that operates efficiently is essential to EU policy-making and to the ability of member states to put their interests onto the EU agenda. In the current era, a considerable share of political decisions, whether national or European, involves the EU level in one way or another (Hix, 2005, pp. 3-4). It is clear, then, that even though actors obviously during the negotiations in the Council aim to achieve decisions as close to their nationally defined ideal points as possible, the 'keep-the-Council-working' imperative is, explicitly or implicitly, also present in their calculations (Lewis, 2010, p. 655). Indeed, these considerations have been on the table in the Council for many years (cf. Council Doc. 7105/03 and Official Journal of the EU, L325 (2009), art. 20) and in the Council preparatory bodies this perspective seems to be present even more strongly (Egeberg, 1999; Egeberg *et al.*, 2003; Hayes-Renshaw *et al.*, 2006; Lewis, 2005).

The degree to which the actors care about functioning of the EU decision-making system then gives the measure of their support for the system. On the other hand, their insistence on their national interests, possibly at the expense of the system's viability, corresponds to their opposition. If, for example, on average they cared about their individual national interests equally as they did for functioning of the system, the robustness of the system would be 0.5. The Eastern enlargement implies more diversity among actors and potentially more challenging process of consensus-building; at the same time, beyond doubt, it is in all actors' interest to keep the Council operational. How are these antagonistic trends reconciled in practice?

In order to explore the effects of the increased membership of the Council on its func-

172

tioning a more formal representation may be helpful. It is plausible to define the variety the Council needs to deal with ($\mathcal{V}$) as depending on the number I of actors i, the distance of their ideal points x_{ij}^* to the median position of member state ideal points x_{MEDj}^*, that means the spread of their positions, the number of dimensions (issues) dealt with J, and the salience actors attribute to these issues σ_{ij}. Formally,

$$\mathcal{V} = \sum_{i=1}^{I} \sum_{j=1}^{J} \|x_{ij}^* - x_{MEDj}^*\| \sigma_{ij}. \tag{8.1}$$

This definition, albeit to some extent arbitrary, captures the intuitive notion that the variety the Council needs to cope with, and hence the difficulty of labouring out the compromises, is given by how many actors there are, how many and how important issues they need to address, and how divergent (spread) their interests on these issues are. Due to the increase in the number of actors and also in the number of issues and the actors' divergences on them, it is clear that the variety has increased after the eastern enlargement ($\mathcal{V}_{EU15} < \mathcal{V}_{EU27}$).

Subsequently, the key question is, whether this increase has been matched by a corresponding increase in the Council's capacity. Council is a decision-making body and hence, as indicated, its capacity is best conceptualized as the ability to get all the important interests communicated among the actors and devise, in some limited time, a widely acceptable solution. In other words, the job of the Council consists in securing transmission of the information on who wants what in the negotiations, and how much – this is the essence of multilateral bargaining (cf. Muthoo, 1999; Fearon, 1998). Following the capacity equation

$$C = \frac{M \frac{supp}{supp+oppo}}{t}, \tag{8.2}$$

from chapter 5, we know that it is given by two key components: the technical capacity, and robustness. We have seen that the degree to which states care about their specific interests as opposed to viability of the system gives the measure of robustness. There is little evidence that the states would be, after the enlargement, any more willing to sacrifice their national interests for the sake of the EU as such. In fact, if anything there seems to be a growing opposition towards the EU, and less and less support, as exemplified by the gradual shift from permissive consensus towards a constraining dissensus (Hooghe & Marks, 2009). Hence, robustness of the EU decision-making mechanisms is constant at best, if it has actually not been decreasing. If the complexity the system needs to deal with has increased (and hence so has its requisite variety), this increase had to be matched by technical capacity. In the rest of the text I try to demonstrate how this has been achieved.

To handle this problem in a way that permits analytical discussion, I propose to conceptualize the situation as one of a common-pool resource extraction, as defined and extensively elaborated upon by Ostrom and her collaborators (Ostrom, 1990, 2005; Ostrom *et al.*, 1994), where the time spent on the negotiations is the scarce (common-pool) resource. The key characteristic of a CPR is that it is very costly to exclude potential consumers from its use, which is a feature CPRs share with public goods. At the same time, it holds for CPRs that use of the resource by one actor lowers its availability to others, which is a typical feature of private goods (Ostrom, 2005, p. 80). This combination creates a vicious circle in which all actors have the incentives to extract the valuable scarce resource without limitations, but this leads to congestion, overuse, and eventually destruction of the resource. Since no actor has incentives to limit her extraction of the resource unilaterally, the CPR is depleted. This is the 'tragedy of the commons' as described early on by Garrett Hardin (Hardin, 1968).

This logic also applies to Council decision-making and the negotiations connected to it.

The time available for the Council negotiations – here conceptualized as a CPR – is fixed as the total amount of time ministers can devote to negotiations in Brussels is limited (cf. Best & Settembri, 2008). As long as the variety of interests the ministers need to deal with is lower than the capacity, the Council functions well and is able to arrive at decisions in a timely way. Once, however, the requisite variety exceeds the capacity, the system becomes overloaded and the Council is unable to handle its entire agenda. Each member state seeks to promote its preferences in the Council negotiations and to put on the table its own agenda, but because the interests of member states often contradict each other there is a need to offer side-payments, to devise package deals and to conduct lengthy negotiations (Hertz & Leuffen, 2011; König & Junge, 2009); by this the CPR of Council time is extracted. A situation that is structurally equivalent to the one of CPR extraction occurs, a situation that leads to increased overload and eventually to system paralysis.

Given this gloomy prediction, how can the Council deal with such a threat? The dynamics captured by equation 8.1 imply that if member states wish to preserve Council decision-making they have to react by lowering the number of issues discussed in the ministerial meetings. Because the number of actors involved is fixed (as is within the rationalist framework the distances of their ideal points and the saliency they attach to the issues), reducing variety to sustainable levels effectively means reducing the number of topics that ministers themselves have to decide on, i.e. having part of the agenda decided by someone else.

Building on the existing accounts of Council decision-making, and in particular of the role of the Council preparatory bodies (e.g. Häge, 2007, 2008; Panke, 2011), I hypothesize that the most natural locus for this delegation of decision-making from the ministerial meetings is the administrative level of Coreper and WGs. In the Council decision-making process,

175

for each individual point on the agenda, representatives of member states decide whether it can be solved by experts and bureaucrats in WGs and Coreper (as an A-point) or whether they want it to be decided on the level of the ministerial meeting (as a B-point). In fact, on average about 80 per cent of all the points on agenda are decided in Coreper, only about 20 per cent are substantively discussed and decided by ministers (see table 5 further in the text). Consequently, my explanation of how the Council 'survived' the enlargement is that in order to provide ministers with enough space for deciding the key conflictive points on the agenda an informal mechanism has developed within the Council through which more of the relatively less salient issues are after the enlargement 'delegated', or more precisely 'left', to bureaucrats.[3] That means, the ministers in the Council can only deal with that much variety; the rest needs to be absorbed by the lower level. Due to the Coreper's specific design features I discuss in the next section, it is able to absorb this excess variety that would otherwise cause the Council decision-making process to be blocked.

8.3 Decision-making within the Common-Pool Resource: A Two-by-Two Game Exploration

The delegation hypothesis outlined above is in general terms very straightforward. Yet, I show in this section with the help of a simple model derived from the conceptualization of the situation as one of CPR-extraction that the delegation mechanism may in practice face some difficulties and that its functioning may be impeded by individually rational but collectively detrimental behaviour. Further, I elucidate how the two-level design of the Council,

---

[3] Note that the delegation procedure I discuss in principle takes place on the intra-state rather than the supranational European level (Pollack, 2003).

consisting of the ministerial meetings and of the Coreper, and in particular the institutional environment of the Coreper as such, can effectively help deal with these difficulties.

Let us first consider some problems the Council – as a bargaining forum – faces. For simplicity, we can represent the bargaining situation as one of two actors, one standing for an individual member state or a group of member states as a collective (subscript i) and the other standing for *all other* member states as a collective (subscript $-i$)[4]. The state representations are led by the ministers and the permanent representatives. In general, we can say that the default payoff from any bargaining is some kind of equilibrium that corresponds to the power and interests distribution among the actors. This can be for example the Nash bargaining solution (Nash, 1950) or, even more simply, the so-called compromise model (Achen, 2006). Either way, we can understand these *basic payoffs* (denoted with $\pi_{i,-i}$) the actors will get from the negotiations as the payoffs actors receive from pure intergovernmental bargaining, with the outcomes of the contestation primarily determined by the distribution of states' interests and relative power capabilities. This is what states can, in general, expect to gain, given their overall power standing.

Yet, we know that in bargaining not only pure interests and power relations matter, but also the actual strategies the actors adopt do. In other words, the interests can be expressed and defended by the actors in various ways, leading to different outcomes. In particular, the key strategic decision by the actors is on how much they will be willing to compromise or, on the contrary, how toughly they will negotiate. We can distinguish two basic situations: actors either *Press* their positions and are not willing to make concessions, or they do *Not press*, i.e. they are willing to make concessions. These are the two available strategies (S) the states' representations can adopt.

---

[4]What matters for the analysis is that the groups need not be of comparable strength; whatever their relative power capabilities, the game yields the same results.

		Actor $-i$	
		Press	Not press
Actor i	Press	$\pi_i - \beta$; $\pi_{-i} - \beta$	$\pi_i + \delta$; $\pi_{-i} - \delta$
	Not Press	$\pi_i - \delta$; $\pi_{-i} + \delta$	π_i ; π_{-i}

Table 8.1: General payoff matrix

If one of the actors negotiates toughly, while the other is willing to compromise, the result will favour the hard bargainer. Formally, if one actor (for example i) chooses *to press* her position and unconditionally insists that her interests are accounted for, while the other ($-i$) shows willingness to compromise and plays *not press*, i achieves a bargaining advantage over $-i$ and the overall payoff of the *pressing* actor i is increased by a premium denoted with $\delta(\delta > 0)$, to the level of $\pi_i + \delta$. The overall payoff of the actor $-i$ that unilaterally does *not press* accordingly decreases by the same amount to $\pi_{-i} - \delta$.

What is important, if both actors choose *to press* their positions ($S_i = S_{-i} = P$), the negotiations get harder and more lengthy. In such a case, the negotiations on the particular matter extract the limited CPR of decision-making time of the ministerial meeting and impose on actors additional transaction costs of extensive ministerial bargaining $\beta(\beta > 0)$. In general I assume that the bargaining advantage component is more important than the bargaining costs, i.e. $\delta > \beta$. The setup so far is summarized in table 8.1.

If we plug into the general formulas the relations between the various factors, in particular the assumption that negotiation advantages are more important than the bargaining costs ($\delta > \beta$), the general payoff matrix 8.1 directly translates into the preference orderings indicated in table 8.4, where 4 indicates the highest payoff and 1 the lowest. (Note that the orderings is simply given by comparison of sizes of the different components of the actors'

178

		Actor -i	
		Press	Not press
Actor i	Press	**2 ; 2**	4 ; 1
	Not Press	1 ; 4	3 ; 3

Table 8.2: The general logic of Council decision-making

utility, so – for example – $\pi_i - \beta > \pi_i - \delta$, because $\delta > \beta$.)

What emerges is the standard Prisoner's dilemma game with one Nash equilibrium (indicated in bold font), the sub-optimal (Press ; Press) equilibrium. In other words, the actors have individual incentives to adopt more aggressive and non-concessional bargaining positions, even though it brings on them the extra costs of bargaining they could avoid in the Pareto-efficient (Not press ; Not press) outcome. This is the essence of the Council problem – unilaterally the actors have incentives to adopt aggressive negotiation strategies that are collectively sub-optimal, leading to system overload.

I indicated that the solution to the problem, the way how the Council can maintain sufficient capacity, lies in delegation of part of the decision-making load to Coreper. Doing so significantly increases the amount of time the decision-making body has available for work as, first, there are two decision-making bodies instead of one and, second, for the permanent representatives the Council decision-making is their only job, which obviously is not the case for the ministers. As a result, having the option of delegation to Coreper represents for the Council a very significant increase in its technical capacity. This is nothing new. Such delegation, after all, takes place in most decision-making bodies – the less important agenda is decided by the subordinate units (in our case Coreper).

The real problem, however, is that Coreper itself is subject to the very same logic that impedes capacity of the ministerial meetings, i.e. the permanent representatives have the

same incentives to play *Press* and thus achieve bargaining advantages over the others. If they all do so, Coreper will not be able to arrive at an agreement and it will have to pass the decisions to the ministers. So even though delegation is an obvious way how to alleviate the pressure on the ministerial meetings, it is not at all obvious whether the delegation mechanism can at all work.

It turns out that the mechanism indeed may work, provided that Coreper is designed in a specific way. To understand in what specific way, let us model the whole situation explicitly, so that we can discuss the actual dynamics in more detail.

The model outlined in table 8.4 can serve as a good basis, but we need to extend it a little. First, we need to consider explicitly the possibility for the ministers to delegate part of their decision making load – the variety they need to process – to the Coreper. We can model this possibility by introducing the rule that if at least one of the actors chooses in the negotiations *Not to press*, the issue can be decided already in Coreper, as an A-point. In other words, as the points on agenda move up from the Working groups where they are pre-negotiated, unless all (both) the actors in Coreper *Press* their positions, a decision can be made already there. In such a case, the 'delegation' takes place or, more precisely, the delegation mechanism built into the system is activated[5]. Since the ministers lose in this situation control over the outcomes, some low costs of monitoring $\mu, \mu \in [0, 1]$ are incurred by them. As a result, each game will represent not only the result of the negotiations but first and foremost the interplay between Coreper and the ministerial level; it will determine on what level the decision is actually made.

---

[5]Note that in practice we cannot really speak about delegation, as the sequence of moves is reversed: the ministers do not formally delegate issues, rather delegation takes place if the issue is decided in Coreper without the ministers. I stick to the term delegation because of its wide use and because, obviously, Coreper can only decide in various matters because the ministers have delegated their authority to it in the first place. So in principle the term is appropriate.

		Actor -i	
		Press	Not press
Actor i	Press	$\pi_i - \beta$; $\pi_{-i} - \beta$	$\pi_i + \sigma_i\delta - \sigma_i\mu$; $\pi_{-i} - \sigma_{-i}\delta - \sigma_{-i}\mu$
	Not Press	$\pi_i - \sigma_i\delta - \sigma_i\mu$; $\pi_{-i} + \sigma_{-i}\delta - \sigma_{-i}\mu$	$\pi_i - \sigma_i\mu$; $\pi_{-i} - \sigma_{-i}\mu$

Table 8.3: General payoff matrix: full model

Second, we need to introduce into the model the notion of saliency, i.e. the degree to which the issue matters for the ministers. High saliency implies that the issue is politically important for the ministers. Technically, saliency (denoted with $\sigma, \sigma \in [0;1]$), can be best thought of as effectively moderating the impact of the varying components of the actors' payoff functions. In particular, it determines to what extent the premium δ is obtained and to what extent the costs of monitoring μ are paid by the ministers. The higher the saliency of an issue for an actor, i.e. the more is at stake for her, the more important the bargaining advantage δ is but also the higher are the costs of monitoring. This somewhat more extensive setup of the game is summarized in table 8.3.

As indicated, I assume that after the Eastern enlargement the costs of bargaining β rise as these are determined by the number and interests of the actors, while the costs of monitoring μ remain constant (these reflect domestic political-bureaucratic dynamics, unaffected by the enlargement). At some point, or for a certain part of the agenda, the costs of bargaining with as many as 27 member states approach and exceed those of monitoring. Neither monitoring costs, nor bargaining costs, however, are within this setup considered as significant as the bargaining advantages actors can obtain from unilateral pressing.[6]Given this setup, what can we learn about the possibility of successful delegation and hence of maintenance of sufficiently high decision-making capacity of the Council? Can it match the requisite variety

of the decision-making in the now-larger EU?

In the following tables (8.4 through 8.6), the equilibrium outcomes of the game are identified for two situations: the one where the issue discussed is for both actors of equal saliency ($\sigma_i = \sigma_{-i}$) and the one where it is salient asymmetrically ($\sigma_i \neq \sigma_{-i}$). With the help of these two models, we will be able to elucidate the incentive structure of the actors and hence deduce what design features Coreper needs to have.

Technically, as in the simpler model above, the preference orderings in the tables are directly derived from table 8.3 by setting σ to the respective values and comparing desirability of the individual outcomes for the actors on the basis of the relative weights of the additional payoff factors as described above (i.e. $\delta > \beta, \mu,$ and $\beta > \sigma\mu$). Equations 8.3 through 8.6 present the critical values of σ at which the strategies of the actors switch, i.e. these values indicate under what conditions the actors will play *Press* or *Not press*. At these values the actors are indifferent between the payoffs they receive from playing the respective strategies, that means they are simple algebraic expressions of the relationship between the factors that determine the actors' utilities.

Let us first consider the situation where the issue is salient for both actors equally. As can be seen from table 8.4, in a situation *before the Eastern enlargement* both actors' chose *to press* their positions, i.e. to oppose compromise proposals and in effect to pass decision-making to the ministerial level. Since the costs of bargaining on the ministerial level are relatively low, this outcome (P;P) is Pareto-efficient.

After the enlargement, however, as the costs of bargaining with 27 member states rise, the originally optimal result (P;P) becomes at least for some issues suboptimal and it is

---

[6]For the argument to hold it is sufficient to assume that $\delta > \beta - \mu$ but it is reasonable to expect that the bargaining advantage will be for the actors in practice much more important, i.e. that $\delta > \beta$. The condition is therefore even less restrictive.

		Actor -i	
		Press	Not press
Actor i	Press	**3 ; 3**	4 ; 1
	Not Press	1 ; 4	2; 2

Table 8.4: Issue fully salient for both actors, situation before the enlargement

collectively better if both actors do *Not press* and these issues are delegated to Coreper[7]. While optimally all issues of saliency lower than

$$\sigma_{i,-i} < \frac{\beta}{\mu} \tag{8.3}$$

should be delegated[8], due to the prisoner's dilemma-like payoff structure this will not be possible. No actor wants to provide the others with the bargaining advantage of δ and hence no actor has incentives to relax her pressure unilaterally. Consequently, only the issues with saliency of

$$\sigma_{i,-i} < \frac{\beta}{\delta + \mu} \tag{8.4}$$

will in fact pass the delegation threshold. All the issues with salience lying between these values, even though in principle of saliency sufficiently low for delegation to take place, will in fact be *pressed* on in the Coreper negotiations and therefore will have to be decided by ministers. The Pareto-inferior outcome (*Press ; Press*) will constitute a single Nash equilibrium of the game (table 8.5). An opportunity to further ease the pressure on Council

---

[7]For the argument to be relevant the bargaining costs need to surpass, at least for some issues, the threshold of $\beta > \sigma\mu$; given that the issues are distributed in their saliency over $\sigma \in [0,1]$ and given the assumption noted above that after the enlargement $\beta > \mu$, this is not constraining.

[8]Again,the value for $\sigma_{i,-i}$ is derived algebraically from table 8.3.

		Actor $-i$	
		Press	Not press
Actor i	Press	**2 ; 2**	4 ; 1
	Not Press	1 ; 4	3 ; 3

Table 8.5: Issue of equal saliency $\beta/(\delta + \mu) < \sigma_{i,-i} < \beta/\mu$ for both actors, situation after the enlargement

will be missed.

This suboptimal and potentially threatening outcome can only be avoided if the actors conduct their negotiations within a highly iterated setting. As showed theoretically by Axelrod (Axelrod, 1981) and empirically by Ostrom (Ostrom, 1990), if the actors are to overcome the perverse incentives of the prisoner's dilemma game they need to play the game with the perspective of 'infinite' repetition where the future iterations of the game give them the incentives to behave cooperatively in the present rounds. In our case, the actors need to believe that it is in their interest to forego the bargaining advantages δ for the sake of preservation of the ability to arrive at decisions within Coreper. In a highly iterative setting, the actors will be able to build the necessary reputations for cooperative behaviour, and hence avoid the need to pass the decisions to the Council. If Coreper is a highly iterative setting, the delegation mechanism can work.

This, however, is only one part of the solution. Let us now consider a situation of asymmetric saliency, where the issue discussed is highly salient for actor i and relatively non-salient for actor $-i$, in particular where

		Actor $-i$	
		Press	Not press
Actor i	Press	2 ; 1	**4 ; 2**
	Not Press	1 ; 4	3 ; 3

Table 8.6: Issue of asymmetric saliency $\sigma_i > \beta/(\delta + \mu)$ and $\sigma_{-i} < \beta/(\delta + \mu)$, situation after the enlargement

$$\sigma_i > \frac{\beta}{\delta + \mu} \text{and} \tag{8.5}$$

$$0 < \sigma_{-i} < \frac{\beta}{\delta + \mu}. \tag{8.6}$$

As captured in table 8.6, in this case i always prefers *to press* and *-i not to press*. The issue can thus be decided in Coreper, but with significant distributive consequences; whereas i achieves her preferred policy and the maximum payoff, $-i$, who is willing to compromise, receives the next-to-worst payoff.

This is hardly a surprising outcome given that for the first actor the issue is highly salient and for the second it is mostly insignificant. Yet, clearly actors may actually gain leverage over others in the negotiations by strategically revealing higher than true saliency for any single issue on agenda. By pretending that the issue discussed is for her of saliency higher than as indicated in equation 8.6 an actor may achieve more favourable policy outcome since the opponent would not have incentives to counter her pressure, as this would result in passing the issue to an already too busy ministerial meeting. Because both actors face the same incentive *to press* on individual legislative proposals and thus steer the distributive outcomes on their side, collectively a less-than-optimal amount of delegation will take place.

185

We see from the model that to deal with this threat, the decision-making body needs to have mechanisms that prevent the actors from strategically revealing higher than true saliencies of the issues and thus inducing uneven distribution of gains. This can be secured through a combination of two design principles. First, the body needs to be permanent, as discussed above. If the decision-makers negotiate within the same setting repeatedly, they can easily monitor whether some actors claim high saliencies for the discussed issues excessively, more than the others. They are also much more likely to develop rich knowledge of each other's domestic political settings and hence the ability to assess, on their own, the true state of the affairs in the individual negotiating states.

Second, and equally important, the body needs to cover in its authority as broad a range of issues as possible. Given that states must not, on average, claim higher saliencies than others, it is essential that all the possible systematic differences the states attach to the individual areas can be accounted for in the negotiations. If the rule of non-excessive claims of saliency was invoked within the individual issue areas, states would effectively lose the ability to devise mutually beneficial deals based on issue-linkages, as these are – by definition – enabled by linking of areas that are of unequal saliency for the different actors. Hence, the model shows that the decision-making needs to both be permanent and have a wide scope of coverage.

What this simplified model enables us to see is that even though bureaucratic delegation may for the Council present the most natural means for preservation of its decision-making efficiency and prevention of paralysis – as indeed my hypothesis in this paper puts it – in practice the process will face severe difficulties as individually the actors have incentives to act against it. In particular, efficiency of the delegation mechanism is impeded because unilateral pressing provides an actor with a bargaining advantage and because actors can

gain leverage over others by strategically misrepresenting their preferences (by revealing higher than true saliency). To secure sufficient capacity, the Council not only needs to have a body to delegate part of its decision load to, but this body also needs to be permanent and it needs to have a wide sectoral scope.

A closer look at the situation shows that indeed Coreper has been designed by the actors in such a way that effectively enables it to overcome these difficulties. Thanks to its permanent nature Coreper is uniquely equipped for dealing with the problem of actors' incentives for unilateral pressing and the PD-like payoff structure. To be sure, ministers do meet relatively regularly in the Council and in this sense their interaction is to some extent repeated. Yet, in practice the density of their interaction and its continuity are nowhere near those of the permanent representatives who meet both formally and informally every week and who tend to stay in the office over longer periods of time. The very *permanent* nature of Coreper enables it to deal with the incentives characteristic for the PD-like payoff structure of the single-shot game much more successfully than the ministerial sessions ever could.

The denser interaction among the permanent representatives effectively means that the problem of interest misrepresentation captured in table 8.6 is likely to be significantly lower in Coreper than it is among the ministers. The scenario in which delegation may be impeded because of individual actors' incentives to reveal higher than true saliency of the discussed issues rests on the assumption of high informational asymmetry among the actors in the Council with regard to the actual political saliency they attach to the individual points on agenda. As documented by qualitative inquiries, however, in Coreper this assumption clearly does not hold (cf. Lewis, 2010). Furthermore, because Coreper oversees the entire agenda it is much easier for the permanent representatives to control each other's eventual excessive extraction of the CPR of the Council time while not preventing the possibility of

187

issue-linkages. For ministers within the sector specific areas, it is much more difficult to keep track of how much the others tend to press their positions across areas and hence they may need to secure that no single actor claims high saliency excessively within their own areas. Coreper is not constrained in this way, so reaching of decisions is easier there.

Hence, both the key problems identified earlier as impeding the delegation mechanism are to some extent remedied by the fact that Coreper is an environment in general much more suitable for successful dealing with the problems of CPR-extraction-like nature. Consequently, once issues are discussed in Coreper, they can be decided there and do not have to be passed to the ministerial meetings. As identified by Ostrom, orderly use of CPRs requires, among other factors, high stability of populations (so that reputations can build up in repeated interactions) and extensive shared normative frameworks (Ostrom, 1990, pp. 88-89). Clearly, Coreper meets these criteria to a significantly higher extent than the Council *per se*, and therefore, again, the fact that it is available as an alternative decision-making locus makes reaching of the mutually beneficial (*Not press ; Not press*) outcome easier.

It should also not be forgotten that the permanent representatives have a strong private interest in securing that Council decision-making operates smoothly and in this sense they have incentives to create conditions for the delegation mechanism to be successful. As Lewis points out, Coreper is characterized by a shared sense of 'duty to "avoid Council"' (Lewis, 2010, p. 655), peer-pressure on individual permanent representatives to reach an agreement within Coreper is present. Alternatively, one could argue that the permanent representatives have a clear interest in delegation in the sense of bureaucratic politics (Niskanen, 1971). A rich qualitative illustration of this fact is provided by Lewis (Lewis, 2000) who shows how a highly political and sensitive topic of the 1994 Local Elections Directive was kept away from discussion at the ministerial level and – de facto as well as formally – decided by Coreper,

precisely because of the effort by permanent representatives to make sure that the process is not blocked in the politically and ideologically necessarily more loaded ministerial debate.

Interestingly, besides this simple 'delegating-the-decisions-down' mechanism there seems to exist yet another form of lowering Council workload and thus alleviating the pressure on it. In this form, decision-making is formally kept on the ministerial level, but *de facto* it moves downwards to Coreper and the WGs. In this case of the so-called *false B-points,* all member state representatives in principle agree on the solution found within the preparatory bodies and are aware of the fact that other solutions would be difficult to reach in the ministerial meeting anyway. However, for one or more ministers it would be politically dangerous in the domestic arena not to be seen to oppose the particular solution, and thus her space for pro-compromise behaviour in the Council could be dramatically reduced. To avoid this problem, the issue is symbolically 'discussed' on the ministerial level as a (*false*) B-point, the particular minister interested in the issue is given an opportunity to voice her objections (these can be recorded in domestic politics), and the whole process is not in fact significantly affected. This mechanism, described in other contexts by scholars (e.g. Hayes-Renshaw & Wallace, 2006; Lewis, 2003) as well as practitioners[9]can be understood as a particularly sophisticated form of workload reduction: while the symbolic opposition by the minister recorded in domestic media enlarges her domestic win-set (Putnam, 1988) as she can claim to have tried to defend the national interest, the decision-making process is not blocked or substantially retarded.

Clearly, delegation to Coreper was not the sole way in which the Council tried to cope with the Enlargement. In fact, the ministerial meetings themselves have been continuously under reform aimed at increased technical capacity of the meetings, in particular at more

---

[9]Author's interview with a member of the Council Secretariat, June 2009.

efficient use of the available time. A number of documents – rules of procedure and codes of conduct – has been adopted the purpose of which was to secure smoother functioning of the Council (Council of the European Union, 2009, Annex V). Yet, as I have hypothesized and as I show empirically in the next section, this apparently was not sufficient and delegation to Coreper needed to take place.

To summarize the key argument, what we see are two sets of key design principles that enable the Council to maintain sufficiently high capacity even when the requisite variety rose with the Eastern enlargement. First, the Council has an alternative body, one that has the authority to make the decisions. This is the simplest technical design feature – creation of an additional channel – but an essential one.

Second, the key to success lies in how is the body designed, namely that it is permanent and that it has very broad sectoral scope. Both these factors are essential in overcoming the information asymmetry problems that otherwise impede cooperation in less permanent and more narrowly focused bodies. Related to this, given that the representatives are permanent and that – to a considerable extent – their professional success depends on their ability to keep as much of agenda as possible off the ministers' tables, we have actors who have strong interest in reaching of compromise decisions. These design features address the problem of support and opposition. Compared to ministers, for the permanent representatives the weights they attach to individual negotiated issues as compared to viability of the Council decision-making system are relatively more favourable for the EU. Hence, robustness of Coreper is higher than that of ministerial meetings, and so is consequently its capacity.

Overall, we see that there are important problems inherent in delegation of decision-making in the Council; it is not at all obvious that delegation and relieving of the pressure brought on the Council by the enlargement will take place, the structure of the actors' in-

190

terests may impede the delegation mechanisms' functioning. However, because of the very stable nature of Coreper, its cross-sectoral reach, and interests of the permanent representatives, it can deal with these problems much more effectively than a body with a lower density of interaction, less permanent design, or narrower scope of authority could.

8.4 Empirical Insights into Attempts to Avoid Council Paralysis

To investigate whether the hypothesized effects as described above have materialized in practice I resort in this section to a quantitative empirical analysis. The dataset I base our analysis on is derived from PreLex, a database monitoring decision-making processes among EU institutions (see König *et al.*, 2006). The dataset contains information on Council decisions on *all proposals* transmitted to the Council by the Commission between January 2000 and December 2007.[10]This accounts altogether for 3811 cases with approximately 450 to 500 per year; of these 3432 (above 90%) have at least once entered the Council agenda (see table 8.7 for more details). The dataset is composed of 464 directives, 1559 regulations, 1768 decisions, and 19 framework decisions, it includes both legislative and implementation acts. In terms of crude data, it contains first and foremost information on interactions between ministers on the one hand and bureaucrats on the other (Coreper and WGs). This is accounted for by two core variables, one indicating how many times an individual proposal was put on the Council agenda as an A-point, and the other indicating how many times it was categorized as a B-point. The dataset also includes basic identification of each proposal and information on the date of its transmission from the Commission, and of the final decision

by the Council.

The indicator I use to assess the 'delegation-hypothesis' is the ratio of the total amount of B-points over the sum of total amount of A-points and B-points (i.e. B/(A+B)) for each individual case or, in case of the graphical presentation, in a given month. This indicator reveals whether over a longer time phase any move of decision-making activity between the two levels can be traced empirically.

If my hypothesis was *not* correct I would expect a move from the bureaucratic level of working groups and Coreper where actors find it more difficult to reach a consensual position upwards to the ministerial level, i.e. I would expect a relative increase in the amount of B-points. In this intuitive approach, bureaucrats on their own are unable to derive a decision because of the increased variety of member states' interests. Accordingly, they are forced to pass the points upwards for a genuine discussion by the ministers and more issues thus enter the Council agenda as B-points. The value of the indicator B/(A+B) would then *be increasing* over long period. This is what in general could be expected to result from an increase in the number of actors and heterogeneity of their interests.

If my hypothesis was correct, I should expect a different development. Because ministers' time and general capacities are fixed I expect the amount of B-points to remain stable and all the additional Council workload brought about by the enlargement to be absorbed by the bureaucratic levels of Coreper and WGs, and decided as A-points. As a result, I actually expect the ratio B/(A+B) *to decrease* over time.

The empirical evidence of a total of 3811 proposals over a time phase of 96 months provides some fairly strong support for the hypothesis. In figure 8.1 the data are shown in the form of a simple plot, with the months of transmission of the proposals by the European

[10]I use the terms 'proposal' to refer to directives, regulations, decisions, and framework decisions, that is to all types of Commission proposals the Council actually decides upon.

Commission shown on the horizontal axis, and the values of the indicator B/(A+B) aggregated into months on the vertical axis. The data show a fairly clear decreasing trend in the number of B-points relative to A-points over time, capturing a move of decision-making over some of the agenda from ministers to Coreper.[11]To allow for easier interpretation, line is fitted through the data on the basis of the ordinary least squares (OLS) method, although – strictly speaking – this analysis is purely descriptive and the coefficient reported in table 8.8 (Model 4) only provides supplementary information.
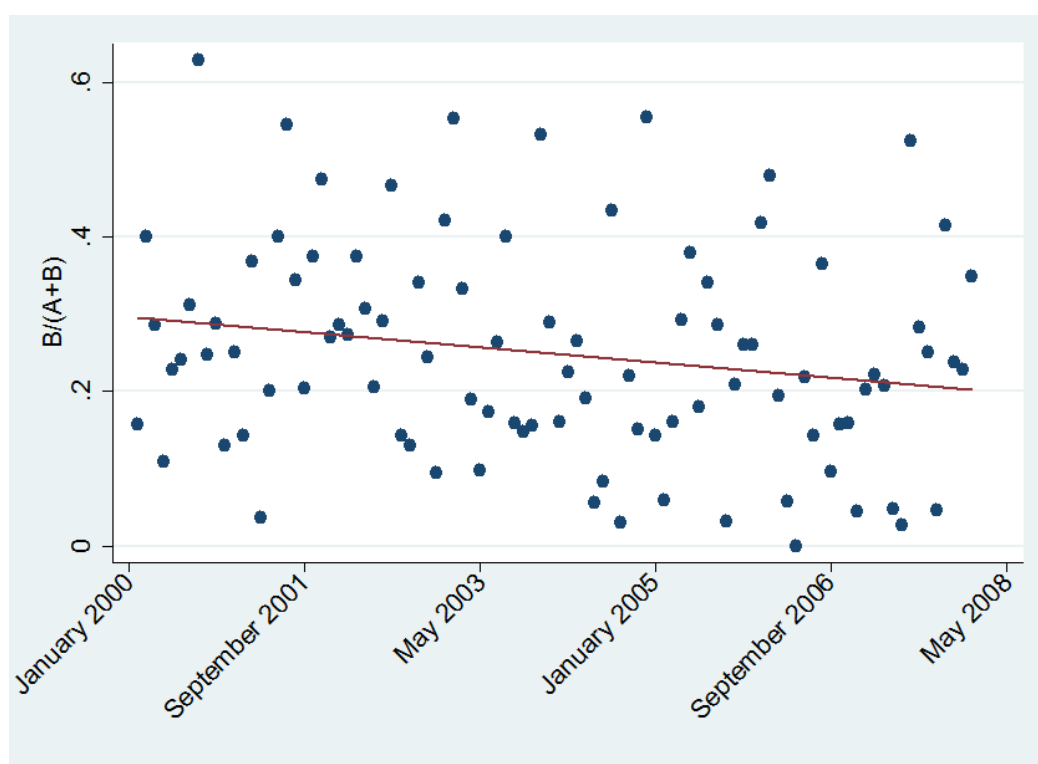
Figure 8.1: Relative position of ministers and Coreper (monthly aggregates displayed)

To assess presence and strength of the trend more rigorously I ran a series of regression analyses (again, OLS method) with a dummy variable for enlargement as the key predictor and with three control variables. First, according to the core logic of the model I would

---

[11]The variance of these aggregates also appears to somewhat increase over time.

expect that value of the B/(A+B) indicator is primarily determined by political saliency of the issue discussed, in particular the more salient issues should imply more B-points and therefore higher value of the indicator. I use the amount of time a particular issue 'takes' as a proxy for its saliency - more politically contentious issues should be more difficult to solve and therefore take more time. This proxy is not perfect as clearly in general duration of the decision-making process is itself to some extent endogenous to functioning of the Council. For this particular analysis, though, this does not cause problems as there is neither reason nor evidence in the data to expect that this system-level endogeneity would manifest itself differently across individual proposals.

Second, I expect that value of the indicator may be influenced by other systemic factors since the actors in both Commission and Council may strategically react to broader developments in EU politics by putting on or withdrawing from the table different points according to the actual political ambiance (see e.g. the sign of a peak in decision-making activity just before the enlargement, in years 2003 and 2004, in table 8.7). In particular, I include a variable capturing what might be called a learning period, i.e. a period of one year after all major institutional changes. This variable takes maximum values in May 1999 (entry into force of the Amsterdam Treaty), February 2003 (Nice Treaty), and May 2004 (Enlargement), and then always decreases towards zero over the period of twelve months. Symmetrically I include a variable capturing the periods of crises which takes maximum values in April 1999 (after Santer Commission resignation), July 2001 (first Irish veto), and June 2005 (Dutch and in July also French vetos); again, the value of the variable always decreases towards zero over twelve months. Quite simply, I expect that in the learning periods actors should be more cautious and that relatively less salient agenda would be put on the table (therefore the value of the dependent variable should decrease) and, to the contrary, that in periods

194

after the major crises the true size of conflict among actors will be revealed (and therefore the value of the dependent variable may increase).

Table 8.8 presents result of these regressions. Models 1 and 2 are based on analysis of all the individual proposals, with model 1 including only the key predictor and the proxy for saliency of the individual issues and model 2 adding the other two controls. Model 3 replicates model 2 on the data aggregated into months (96 months between January 2000 and December 2007) and model 4 gives numeric values defining the plot in figure 8.1.

Although the significance levels vary across the individual models, I obtain a very clear and stable picture in which all variables point in the predicted direction and both the enlargement dummy and the proxy for issue saliency are consistently significant at least at the $p=0.05$ level. Both the other control variables are also statistically significant but comparison of overall predictive values (R-squared) of models 1 and 2 shows that their contribution is very limited.

Beyond the analyses covering the entire dataset presented above I can assess our hypothesis empirically by investigating separately a subset of the relatively more contentious or salient issues, i.e. by investigating only those proposals which we would generally expect to be decided by ministers. If our argument is correct we would expect that a group of such cases is likely to emerge that after the enlargement will be decided in Coreper even though before the enlargement it would have been passed for decisions to ministers. While it is obviously difficult to assess such observable implication directly, as it involves counterfactual reasoning, I attempted to do so in quantitative terms. I selected from all the proposals those that are likely to belong among the most salient ones, specifically those that lasted more than two years in duration (altogether 311 cases, approx. 9% of all proposals).

As demonstrated in Figure 8.2, the hypothesized relationship holds for this subset as

195

well and, as a matter of fact, the trend proves to be even stronger than in the case of the entire population of proposals (model 5 in table 6 gives the regression coefficient of the fitted line). What is as important as the quantitative result is the qualitative insight we obtain from graphical inspection of the data. In particular, after the enlargement we observe and increased incidence of cases that do not at all appear on the ministerial meeting agenda, i.e. months in which not a single of the selected highly salient cases 'makes it' on the ministerial table. This observation provides additional direct evidence for my argument and lends the delegation hypothesis strong support.[12]
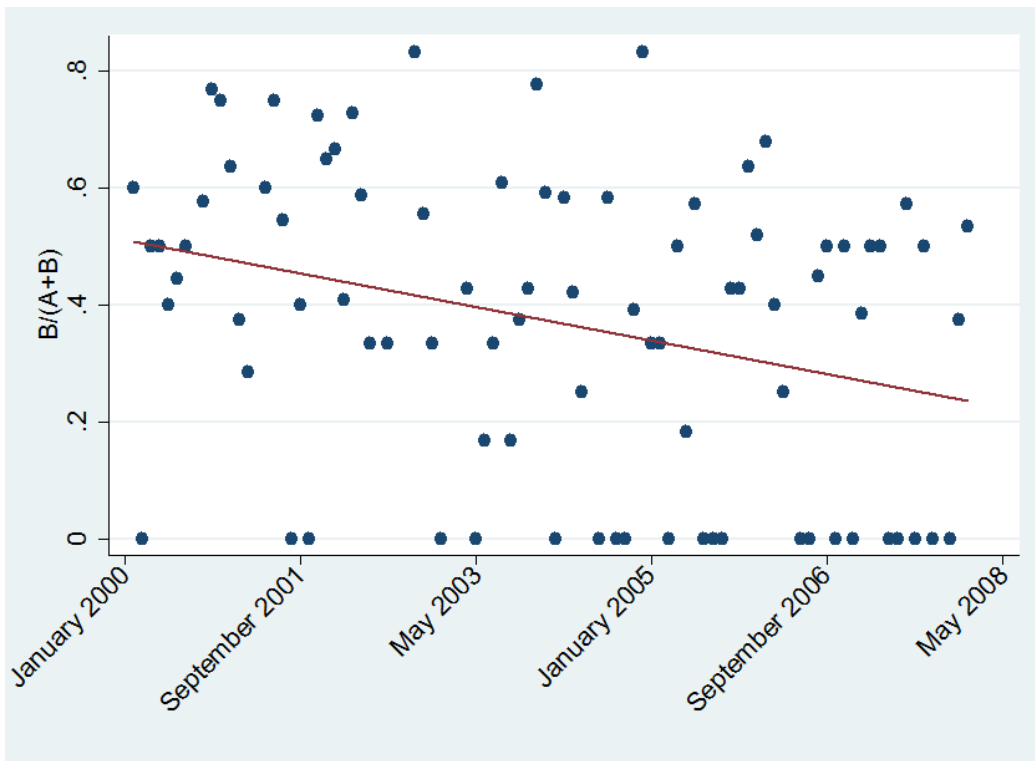
Figure 8.2: Relative position of ministers and Coreper for proposals longer than two years (monthly aggregates displayed)

In addition to the central analysis of this paper, the collected data provide insights into

[12]Note that this result holds irrespective of where exactly we draw the line for qualifying into the subset of the most salient issues; any criterion aimed at identification of some 5-20% of most salient cases yields essentially the same results.

the issue of relative involvement of ministers and bureaucrats in Council decision-making, i.e. into 'Who decides in the Council of the European Union' (Häge, 2008). The data presented in table 5 above reveal that out of the total of 3432 cases (for which the data is complete) 724 have at least once entered the Council as a B-point, i.e. 19 percent of the proposals have been at least once dealt with by ministers. This figure falls very close to the earlier often-quoted estimate by Hayes-Renshaw and Wallace (Hayes-Renshaw & Wallace, 2006, p. 79) who report the insider's estimate of the share of issues decided *without* ministers to be approximately 85 percent. Häge's estimate (Häge, 2008) differs from mine significantly (he argues that it is actually as much as 48 percent of all Community agenda that *is discussed* by ministers, and 35 percent that *is actually decided* by them) but the datasets used differ significantly and, perhaps most importantly, he considers only the legislative activity while my analysis includes all proposals on Council agenda, including the implementation acts.

8.5 Conclusions

Why does the Council of the EU continue to operate smoothly, in spite of the feared 2004/2007 enlargement? How can an international decision-making body of its type 'survive' an eighty per cent increase in its membership? In my analysis I put forward and test one possible explanation: the Council was able to cope with the enlargement by delegating an additional part of its agenda down to Coreper and therefore by reducing the amount of variety it needs to deal with. Because the total amount of time the ministers can spend in Brussels is mostly fixed and highly constrained, the most natural solution to the threat of paralysis of the Council is a shift of a part of the relatively less politically salient agenda to the bureaucrats.

The quantitative empirical analysis covering the entire Council decision-making load in years 2000-2007 shows that indeed the hypothesized trend can be clearly discerned in the data: the amount of work done by the ministers is stable, the additional decision-making load is absorbed by the Coreper. Over time, then, the share of decision-making done by the Coreper relative to that done by the ministers themselves is increasing.

Although the delegation hypothesis appears very intuitive, I demonstrate – by conceptualizing the delegation mechanism as a problem of CPR-extraction – that the delegation mechanism may in fact face severe difficulties. With the help of a simple model based on this conceptualization I subsequently show how the setup of the Council-Coreper interaction and the institutional environment of the Coreper helped overcome these difficulties and why the Council could have indeed maintained a sufficient level of capacity also after the enlargement.

My findings, however, point at some potentially dangerous trends in Council politics, namely at a still increasing need of the political representations to rely in European decision-making on the bureaucracy. I do not necessarily observe the power shifting from the member states towards the supranational administration, since clearly the move of decision-making in the Council from ministers to Coreper is a case of innovation that does not lead to greater leverage of the European Commission or other European bodies (cf. Tsebelis, 2008; Tsebelis & Yataganas, 2002). Nevertheless, a move of part of the Council agenda from at least in theory politically accountable ministers to the permanent representatives in the Coreper certainly does not make the decision-making processes more open. There may be an unfortunate trade-off in the Council between its operability and capacity on the one hand, and its congruence with the principles of transparency and political accountability; this trade-off may be due to the enlargement more severe than it was before.

Overall, I believe the pressure generated by the 2004/2007 enlargement on the Council's

capacities should not be underestimated now, several years afterwards, only because patterns of decision-making in the Council do not appear to have changed at first look; they may have changed, although perhaps only informally, in ways that are not easy to trace.

Table 8.7: Decisions and relative involvement of ministers and Coreper

	2000	2001	2002	2003	2004	2005	2006	2007	Overall
Proposals	508	453	439	482	520	442	489	478	3811
(out of which at least once on agenda)	(465)	(412)	(399)	(419)	(475)	(379)	(448)	(438)	(3432)
Mean duration of the decision-making process (in days)	244	266	247	278	249	250	269	247	261
A-points	485	435	410	459	479	374	464	422	3528
B-points	166	212	161	204	140	145	112	144	1284
Pure A-points so far	347	307	309	322	398	291	383	351	2708
Acts at least once on agenda as B-points so far	118	105	90	97	77	88	65	84	724
Share of cases at least once on agenda as B-points	0.25	0.25	0.23	0.23	0.16	0.23	0.15	0.19	0.21

Table 8.8: Decisions and relative involvement of ministers and Coreper

	Model 1 (N=3340)	Model 2 (N=3339)	Model 3 (N=96)	Model 4 (N=96)	Model 5 (N=86)
Enlargement	-0.0473***	-0.0485***	-0.0632*	-	-
	(0.0089)	(0.0090)	(0.0281)		
Political saliency	0.0003***	0.0003***	0.0000'	-	-
	(0.0000)	(0.0000)	(0.0000)		
Learning period	-	-0.0104*	-0.0153	-	-
		(0.0047)	(0.0148)		
Crisis period	-	0.0084'	0.0215	-	-
		(0.0053)	(0.0134)		
Month	-	-	-	-0.0010'	-0.0029**
				(0.0005)	(0.0010)
Constant	0.0711***	0.0735***	0.2252***	0.2962**	1.8855***
	(0.0071)	(0.0086)	(0.0415)	(0.0280)	(0.5048)
R-squared	0.1141	0.1168	0.1203	0.0388	0.0972

Chapter 9

International Organizations' Strive for Information: the Politics of Secretariat Staffing

The previous chapter on European politics concerned first and foremost the decision-making of the system, in particular the System 3 function of the viable system model. In this chapter, I illustrate usefulness of the entire scheme developed in the dissertation on what is perhaps the most fundamental functionality of any political system. I go well beyond the decision-making phase and study how several important IOs deal with the problem of how to actually obtain reliable information on the conditions in the countries with which they cooperate. In other words, I study how the IOs secure sufficient capacity in System One in their everyday business.

From a political research perspective, this issue may seem uninteresting, as it is apparently a technical problem for the IOs how to secure this. This is a matter for management science,

not politics. Yet, as I will try to show, in IOs the technical needs of management often collide with the political needs of control by the member states, i.e. the internal functioning of the system collides with the requirements from the external environment. In the case discussed here, the powerful member states have a strong interest in staffing the IOs with 'their' people, or with people who share with them the underlying view of the problems and possible solutions. At the same time, though, for many IOs the primary client countries are the developing ones, precisely the ones that are most severely under-represented in their secretariats. This creates possible legitimacy problems, but from my perspective it creates first and foremost an information transmission problem, as much of the key information the IOs need is of soft nature – it is hardly standardized, quantified, it is hardly mediated. This information is most reliably obtained by the IOs if they actually do have staff from the countries, i.e. staff familiar with the local culture, society, and in general institutional framework. Without this staff, the IOs don't have the capacity to deal with the variety implied by the specifics of these countries.

Consequently, in this chapter I show two things: First, the powerful states try to design the IOs' staffing mechanisms in a way that secure their positions. As a result, there is an enormous level of inequality in states' representation in the secretariats. Yet, second, in their need for information and localized expertise the IOs counter-act this design principle and over-represent these countries, *relative* to their power positions.

The chapter consists of three sections. First, I outline a theoretical argument describing the tension between the external demand for control of the IO and the internal demand for representation of the developing countries in the staff. Second, I present the empirical evidence showing how drastically – in the large picture – the developing countries are under-represented relative to the developed ones. Third, I show that this state of affairs is in

practice counter-acted by the IOs, and that the client countries of the IOs are represented relatively more than would correspond to their power standing or to the actual rules for staffing the IOs follow. The chapter consists of three sections. First, I outline a theoretical argument describing the tension between the external demand for control of the IO and the internal demand for representation of the developing countries in the staff. Second, I present the empirical evidence showing how drastically – in the large picture – the developing countries are under-represented relative to the developed ones. Third, I show that this state of affairs is in practice counter-acted by the IOs, and that the client countries of the IOs are represented relatively more than would correspond to their power standing or to the actual rules for staffing the IOs follow.

9.1   Staffing of international organizations and the theory of delegation

Each IO, as any organization, needs to secure first and foremost its survival, that means its budget (Niskanen, 1971). In IOs, the ability to secure their budget is given by their ability to secure support from the powerful member states, the main budget contributors. As a result, any IO needs to be able to perform its tasks a way that it convenes to the interests of the power states. This, in a nutshell, is the accountability problem as elaborated by Keohane: most today's IOs are highly accountable, but to secure survival they need to be accountable first and foremost to those who pay them, their principals (Keohane, 2003).

If we carry the principal-agent analysis further, we are able to identify on the basis of the existing literature the results to which this in practice leads. Since by the virtue of being experts in their areas the IOs obtain a significant information advantage over the states that

create and finance them, it is essential for the states – if they want to keep the ability to control the IOs – to secure that the IOs share to the maximum possible extent the states' own interests (Pollack, 1997).

According to the principal-agent theory, one of the most important reasons for which states create IOs is that they can become experts in their areas as they can collect information centrally (Hawkins *et al.*, 2006). So states create IOs to make them experts and for them to have an information advantage over their counterparts. Yet, this necessarily means that at some point the IOs are increasingly evade the control of the powerful states themselves; slippage or shirking can emerge, the IOs start acting in ways the principals that created them actually do not like (Conceição Heldt, 2010).

As a result, if the powerful states want to prevent the IOs from acting against their interests, they need to secure that they share their interests, i.e. that what the IOs as actors want is aligned with the goals with which the states endowed them as agents (Urpelainen, 2012). To be sure, the states may decide that they in fact need agents whose interests deviate from those of the individual member states. The clearest example is the delegation of the legislative initiative in the EU to the Commission, the long-term interests of which clearly deviate, potentially significantly, from the short-term interests of the member states (Stone-Sweet & Sandholtz, 1997). But in general, it is always in the states' interest to find (or design) such principals that will pursue similar interests to their own.

This, however, has significant implications for how the IOs' secretariats need to be staffed. The secretariats are not completely impartial bodies without the ability to make politically salient decisions. On the contrary, many IOs' administrations enjoy considerable degrees of autonomy and are allocated high (and rising) authority within their respective fields of activity (Zürn *et al.*, 2012).

Consequently, it is very important for the powerful states – the principals – to make sure that the administrations share their views on different problems, or quite simply that they pursue the states' interests. As a result, the states will have strong incentives to make sure that it is their citizens who are on the IOs' staff. The most prominent examples of this are the managing directors of the World Bank and the International Monetary Fund, the former always coming from the US and the latter from what is today the EU (in particular from Western Europe).

Hence, it is in a clear interest of the key states to make sure that they are well represented in the IOs' secretariats, and since they are those who provide the IOs with the funds, the IOs have a clear incentive to prioritize these countries in their personnel policy. Such a mutually beneficial arrangement can lead to a long-term viability of the system (the IO) in face of its environment (especially the powerful states' interests).

This, however, comes at a cost paid in terms of internal functioning of the IOs. In particular, skewing the staff distribution in the IOs towards the powerful actors – the principles – necessarily implies that the weaker states will be under-represented. Given, however, that it is precisely these weaker states that are often the primary addressees of the IOs' activities, a tensions is emerging. On the one hand, the external environment of the system pushes it towards higher representation of the powerful states; on the other hand, the internal functioning concerns push the IO towards an equal distribution, or even to a distribution that will disproportionately favour the staff from its main client countries.

The reason is that the work of the IOs depends to a considerable degree on thick understanding of the situations in the client countries and in general on a lot of contextual information. This information, however, is very difficult to be transferred in some standardized form, its collection can be rarely based on supposedly more objective technical processes,

rather it is based on development of thick intuitions and understandings (Petersen, 2004).

Recently, Hooghe and Marks built on the distinction between hard and soft information a theory of geographical composition of states (Hooghe & Marks, 2012). They argue that effective governance requires not only hard information but also a considerable amount of soft contextual information, and since this is notoriously difficult to transmit and aggregate, functional demands for de-centralization of governance emerge. It is the soft information problem, not geography, what determines the principal constitutional composition of states. To show this, they demonstrate empirically that whether states are unitary or federal (or to what degree) does not depend on their geographical characteristics, but rather on the distribution of their populations.

In general, if the political systems do not possess mechanisms with the capacity to collect and process also the soft information on the level on which the decisions are made, they are doomed to serious problems. In particular, the Systems One will be crippled in their management function, as they will not be able to collect all the information they need, and neither will the System Three. Hence, no level will be able to consider in the decision-making all the information it should consider (cf. Beer, 1979).

Consequently, I hypothesize that the IOs will have to counter-act the principal-agent logic. More precisely, my argument is that while we can expect the distribution of the staff in the IOs to follow roughly the distribution of power in the given area, there will be important deviations caused by the IOs' efforts to secure their internal functioning, i.e. to secure efficient flows of information within the System One.

I will test the hypothesis in two steps: First, I assess the big picture, i.e. the unevenness of the states' representation in the IOs. Second, I move to a more nuanced analysis, where I focus on the important deviations from this broad pattern.

The analysis is conducted on data capturing four of the most important IOs in the contemporary global governance of socio-economic matters: these are the International Monetary Fund (IMF), the United Nations Development Programme (UNDP), UNESCO, and the World Health Organization (WHO). These four IOs are selected to achieve variation on two potentially significant dimensions: First, I distinguish IOs that belong to the family of United Nations Programmes, as opposed to those more distant from the UN centre, the specialized UN agencies (Koenig-Archibugi, 2003, p. 64). Since there is a widespread perception of the United Nations as of a body seeking broad and relatively proportional representation (but see further in section 9.2), it is plausible to expect that the UN programmes would follow the representation principle more than the IOs with more independence from the central UN bodies. While UNDP and UNICEF are the UN programmes, both WHO and IMF are classified as specialized agencies by the UN.

Second, I distinguish between the IOs' whose work is likely to raise more political contestation from the rather technical ones. In particular, I ask whether the IOs produce significant outcomes that the member states may like to oppose. From my sample these are the IMF (with its high level of politicization) and UNDP (due to the rather political nature of a number of its activities, ranging from democratization and human to institutional reforms and transparency). WHO and UNICEF are then considered relatively less political, with clear functional mandates.

When combined, these two criteria produce a 2x2 matrix, where each of the IOs falls, roughly, into one category (table 9.1).

The analysis is dominantly quantitative, using a number of variables to explore the dynamics underlying the staffing distribution in these IOs. Most of the data used in the analysis are centred at year 2007 for which the data on staff composition by nationality from

	Partially political mandate	Dominantly technical mandate
UN Programmes	UNDP	UNESCO
Specialized Agencies	IMF	WHO

Table 9.1: Classification of the cases

all the four IOs were available.

9.2 Empirical analysis: the built-in inequality of representation

I will start the empirical analysis with observation of the commonly expressed fact that the member states are in the IOs' secretariats distributed highly unevenly. This unevenness is given both by the formal rules that stipulate how the staff should be distributed regionally and across states, and in the empirical application of these rules, i.e. in what does the staff composition actually look like. After presenting this big picture in this section, I turn to a more detailed analysis in section 9.3 where I show that in fact the picture is much more nuanced.

9.2.1 The design of the personnel rules

For many IOs, it holds that the composition of their staff in terms of its geographical distribution matters, and hence explicit rules regulate what should, on average, the distribution look like. The most prominent example of this phenomenon is the United Nations secretariat, as well as its other bodies. In Article 101, the Charter stipulates that

> [t]he paramount consideration in the employment of the staff and in the determination of the conditions of service shall be the necessity of securing the highest standards of efficiency, competence, and integrity[,]

but also that

> [d]ue regard shall be paid to the importance of recruiting the staff on as wide a geographical basis as possible.

So although the members of the IOs' staff are expected to conduct their tasks impartially and without view on the national interests or the countries of their origin, it is acknowledged that even geographical distribution is *per se* desirable. This principle clearly corresponds to the information-based logic presented in the previous section, where the diversity of the staff is supposed to bring to the IOs' administrative bodies deeper understanding of the social, cultural, and institutional realities in the countries. This is the knowledge that is not easily transferable.

In practice, this principle takes on different forms. Let us see the arrangements in the four IOs under review in this text. In IMF the appropriate level of representation is broadly given by the quota subscription (International Monetary Fund, 2003, p. 7), i.e. it reflects the economic power of the individual members. This would suggest that from all the four IOs the distribution of staff should be in the IMF most closely given simply by power of the individual members. However, this rather unusual (in its explicitness) approach is moderated by two conditions. First, the developing countries should together account for 40% of all the staff. This is not negligible, but given that out of more than 180 member countries only about 30 are classified as advanced, the 40% share for the rest is still rather low. Second,

African countries should together account for 8% of the staff; this is approximately double the share of its quote subscription.

In WHO, the standards for what is appropriate representation are given by the standards of the UN Secretariat (Bouayad-Agha & Hernandez, 1996). This means that 40% of all staff positions is distributed evenly across all members, so each state is automatically awarded $\frac{0.4*staffsize}{N}$ staff positions, where $staffsize$ gives the total size of the staff of the given IO, and N gives the number of member states. 55% of the positions is distributed according to the size of the budget contributions, which are in turn determined by the size of the economies. These positions are then distributed highly unevenly, favouring large (and rich) economies. 5% of the staff is awarder with respect to the size of the states' populations.

Finally, in UNDP as well as in UNESCO no formal quota is established (Bouayad-Agha & Hernandez, 1996). This means that strictly speaking these bodies do not have a firm staff distribution policy. However, given that they are both among the most prominent UN bodies I will stick in their assessment to the standard usual across the UN, i.e. 40% for membership, 55% for contributions, and 5% for population. As we will see, the results of the analyses for UNDP and UNESCO show patterns consistent with the other two IOs, suggesting that this simplification is not supported only theoretically, but also empirically.

9.2.2 Under-representation of the developing countries in the secretariats

Given the distribution of wealth and of populations across countries of the World, we can expect a highly uneven pattern of representation. Figure 9.1 shows descriptively the unevenness of the states' representation in the IOs' secretariats. Most states have extremely

low shares, while very few states enjoy high representation. In fact, as few as 12 out of 145 countries in the dataset account for more than 50% of all the staff, as indicated by the vertical red line.[1]
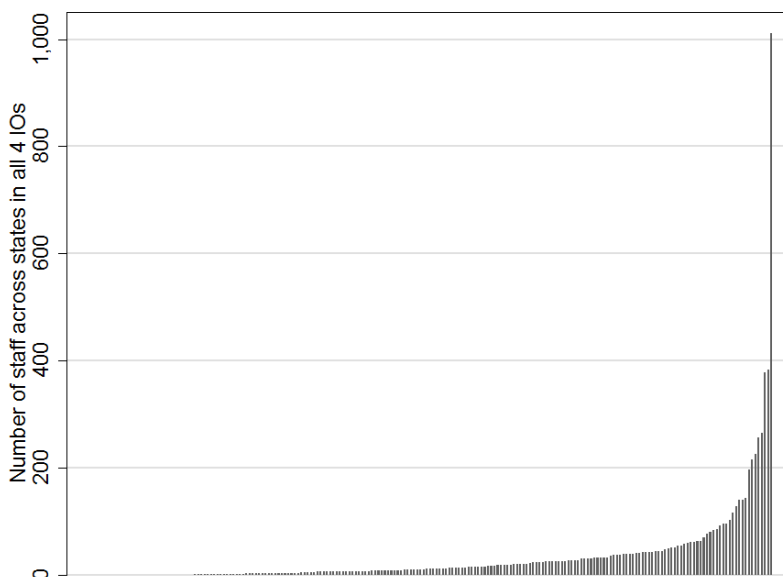
Figure 9.1: Distribution of states' share of staff across IMF, UNDP, UNESCO, and WHO

This pattern can be expected to be obtained due to two major factors: the size of the countries measured in the size of the populations, and their economic performance, measured in terms of the GNP per capita. Figure 9.2 plots the relationship between these two factors and the number of staff in a 3-dimensional plot, with the size of representation measured as the absolute number of staff in all five IOs for the given state (on the vertical axis).

Note that all three variables are measured in decadic logarithms because they are all extremely non-normally distributed, in particular they all have a heavy positive skew (e.g. most countries have very small populations, while there are only few large countries, but these are extremely large). This means that even though figure 9.2 correctly shows a strong

---

[1]These are, in the decreasing order: USA, FRA, GBR, IND, DEU, CAN, ITA, JPN, BEL, PHL, NLD, ESP, AUS, BRA. Note that exactly the same finding holds also when we weight the representation by the size of the IOs, i.e. shares in the IOs, rather than absolute numbers of staff, are considered.
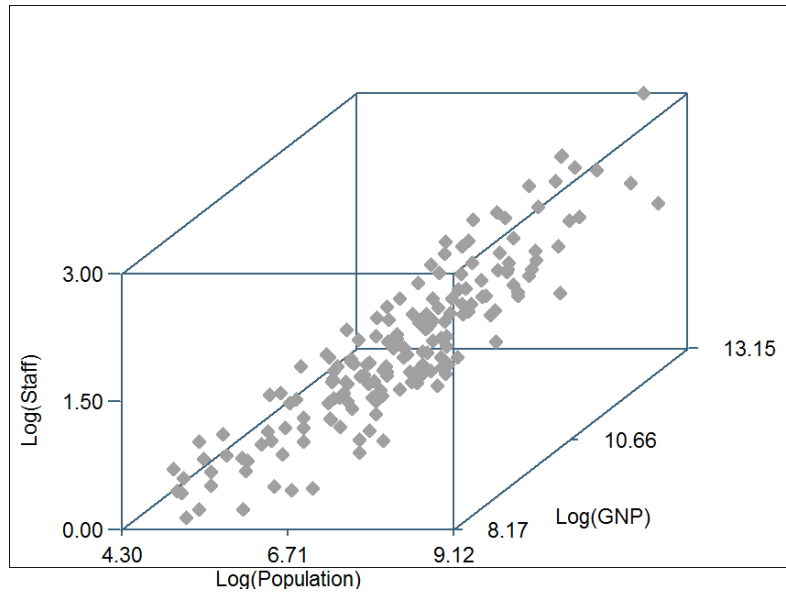
Figure 9.2: Staff as given by the size of economy and population (all in logs)

linear relationship between the variables, this is only due to the fact that in the graph all of them are in the logarithmic form. In practice this means that every unit increase in one of the variables in the plot increases its true value 10 times. So a state with the value of Log(Population) 7 has 10^7 inhabitants, or ten million; a states with value 8 on the same dimension has 10^8 inhabitants, one hundred million.

If we test the effect of population and economy size on the size of representation in the IOs numerically, we see that indeed both have a significant positive effect and that they jointly account, on average across all the five IOs, for 57% in the staff size variation. Table 9.2.2 summarizes this results as well as the results for the individual IOs.

The interpretation of the logarithmic coefficients is not straightforward so I report the standardized Beta coefficients instead. This allows us to see already now a phenomenon that becomes important in the next section, namely that the relative importance of these two predictors in the regression analysis changes across the IOs. While for the IMF the log of

	(1)	(2)	(3)	(4)	(5)
	logtotalstaff	logimfstaff	logundpstaff	logunescostaff	logwhostaff
logpopulation	0.450***	0.120	0.227*	0.417***	0.460***
	(6.49)	(1.44)	(2.47)	(5.14)	(5.54)
loggnp	0.397***	0.639***	0.408***	0.339***	0.304***
	(5.73)	(7.66)	(4.44)	(4.17)	(3.67)
R2	0.622	0.531	0.343	0.493	0.499
Adjusted R2	0.617	0.524	0.334	0.487	0.492
N	173	139	147	163	144

Standardized beta coefficients; t statistics in parentheses

$* \ p < 0.05$, $** \ p < 0.01$, $*** \ p < 0.001$

GNP per capita is more important than for the other IOs (i.e. the Beta coefficients show higher substantive effect of the predictors on the staff size), the log of population is more important for UNDP, UNESCO and WHO, as compared to IMF. So there appears to be some variation in terms of what matters for how many staff members a particular country has.

Overall, the findings in this section give us the big picture, one corresponding to the almost common knowledge according to which the developed countries are massively over-represented in the IOs' staff, considering the sizes of their populations. Indeed, from the 12 countries with largest representations that account for more than 50% of all staff only India and Philippines are not OECD-members.

Yet, as indicated in the previous section, this implies a potentially serious problem. Given that the main direct beneficiaries of many of these IOs' activities are first and foremost the developing countries, the fact that they are significantly under-represented in these bodies' secretariats may lead to information problems as the secretariats will lack the kind of soft information discussed above precisely about the countries they should know best. What we see is a tension between, on the one hand, an effort of the system to be aligned with the international environment and the demands of the powerful states for control and, on the other hand, its need to possess functioning mechanisms that will secure efficient access to the information that the system needs for performance of its tasks.

9.3 Under- and over-representation in practice: IOs counter-acting the external pressure

The direct implication of the findings in the previous section, in the light of the theoretical debate presented in section 9.1, is that the system can be expected to build mechanisms that would counter-act the one-sided pressure on over-representation of staff from the powerful member states. If this external pressure leads to internal inconsistencies, one can expect that the leaders of the IOs will try to make use of the autonomy they have in personnel policy to lower its negative effects.

In this section I try to demonstrate that this is actually what we observe empirically, if we explore the evidence in more detail. I elaborate on three pieces of evidence: First, at a closer look we see the IOs have a tendency to over-represent their client countries relative to their economic performance and population size, i.e. they suppress their under-representation. Second, this holds even if we take as the criterion of the appropriate level of representation the IOs' own criteria. Third, the variation across the IOs in the sample appears to be explained by the degree of actual importance of the IOs, so in the IOs whose rulings are more consequential, and hence arguably that have lower autonomy, the under-representation of the developing countries is more persistent than in the more technical or a-political IOs.

9.3.1 Poorer countries relatively over-represented

The big picture identified in the previous section – one corresponding to the general view of the developing countries being under-represented – is not fine-grained enough to capture an interesting phenomenon, namely the fact that while population and economy sizes do indeed matter a lot for the level of representation, there is an important amount of variation

that they do not explain. In particular, when we control for the states' power or importance by controlling for its population and economy size, we find that, on average, the weakest and poorest countries are in the IOs' secretariats, relative to their power positions, over-represented.

To see why, let us extend on the analysis presented in the previous section. There we established the role of the two key predictors of states' representation in the IOs: their population size, and their economy size. Figure 9.3 plots the residuals of the Model 1 in table 9.3, i.e. of the model that summarizes the results across all the five studies IOs. The residuals, on the vertical axis, are plotted against the variable GNP per capita, indicating the wealth of the country. Such a residual plot allows us to see whether there is any variation in the residuals along the values of some other variable of interest, in this case the measure for countries' wealth. This then shows whether this additional variable of interest may also have an effect on the dependent variable[2]. The locally weighted regression function plotted through the data (the u-shaper curve) gives the best-fit non-linear description of the data, for easier interpretation. If the residuals are not distributed evenly along the horizontal zero-line, the new independent variable potentially has an effect on the dependent variable.

From the plot we see that in this case this is certainly so. As expected, the most wealthy countries are in the secretariats over-represented, even when we control for their size of economy and population. This is the cluster of mostly the OECD countries in the upper-right corner of the graph. What is more important, though, is that a majority of the developing

---

[2]The results of this plot are numerically identical to the results captured in the so-called added variable plot, provided that the additional variable of interest − in this case GNP per capita − was also present in the original regression analysis. That means, technically, the residuals plot shows the effect of the additional independent variable has on the dependent variable, while controlling for all the other independent analysis included in the regression. The reason why I use here the residuals plot instead of the added variable plot is that a regression model with log of population, log of GNP, and log of GNP per capita, suffers from extreme multicollinearity, that potentially threatens unbiasedness of the beta coefficient and certainly prevents reliable inference (due to the inflation of standard errors).

Figure 9.3: Residuals plot: effect of GNP per capita on staff representation

countries is actually relatively over-represented as well. In the graph, all countries above the zero-line (indicating no deviation from explanation based on population and economy size) are relatively over-represented, while those below it are under-represented. Clearly, most under-represented countries are those with GNP per capita approximately between 1000 and 10000 USD (in 2007), i.e. formally the middle-income countries. The low income countries, with GNP per capita below 1000USD, are relatively over-represented.

Again, this obviously should not be understood as an argument that these countries are over-represented in absolute terms, as for any of these the actual number of staff members will be in single digits. The over-representation is only relative, given the sizes of their populations and economies. It should be noted, however, that the same patterns is observed when we switch in the analysis GNP per capita and GNP, i.e. when we control in the regression for population size and wealth (measured as GNP per capita), and then plot the residuals against the size of GNP (measuring size of economy, irrespective of per capita wealth). As

indicated in figure 9.4, the same u-shaped pattern ensues also in this case. That means, it is not only the poorest countries in per capita terms that are relatively over-represented, but it is also in general smaller economies, including some lower-middle income countries. The conceptual distinction between poorer and smaller economies notwithstanding, in practice these two groups overlap to a high extent.
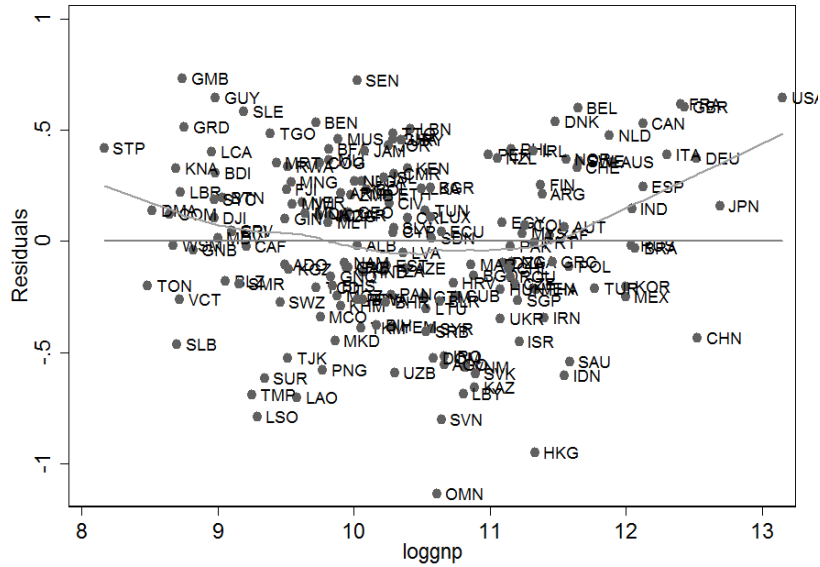

Figure 9.4: Residuals plot: effect of GNP on staff representation

This is obviously a somewhat cynical view, assuming that the appropriate (reference) levels of representation are given by power positions of the countries, i.e. by the size of their populations and especially of their economies. As a matter of fact, though, the results holds also when we consider the staff distribution from the perspective of the rules of the IOs themselves.

As I discussed above, the four IOs I review in this study vary in terms of what they consider as an appropriate regional distribution of the staff, we need to consider them separately or in groups. While IMF follows quite closely the principle of correspondence between the

size of the economy and share of the staff, WHO determines the appropriate distribution on the basis of the size of economy (55%), the size of population (5%), and of the simple fact that each member has a right for some number of staff, by the virtue of its very membership (40%). UNESCO and UNDP do not have strict formal rules defining appropriate distribution, but the criteria common in the UN system can be plausibly applied to them as well.

If we reproduce the analyses from the previous section with the standards set by the criteria just described, and hence measuring the deviations in representation from these standards, we obtain a fairly clear picture in which, again, the poorest countries are over-represented relative to their power positions and to their 'assigned' numbers of positions. Figure 9.5 presents the overall pattern, across all four IOs. As in the previous case, the plot show the negative effect of rising GNP per capita on the level of over- or under-representation. This time the criterion of appropriateness is given by the formulas reported by the IOs themselves. What we see is, again, a systematic relative over-representation of the least wealthy countries (with GNP per capita below 1000 USD) – as indicated by the linear trend fitted through the data. Hence, clearly the IOs do in practice somewhat counter-act the official guidance for staffing, reducing the bias built into the formulas which give enormous weight to 1) membership as such and to 2) the size of budget contributions, with little respect for the country's population sizes.

9.3.2   Functional demand for information

So far we have established that, either given the criteria of the IOs themselves or the criteria of size of economy and population, the poorest countries are in the IOs' secretariats relatively over-represented. To be sure, this over-representation is only relative, in absolute terms these
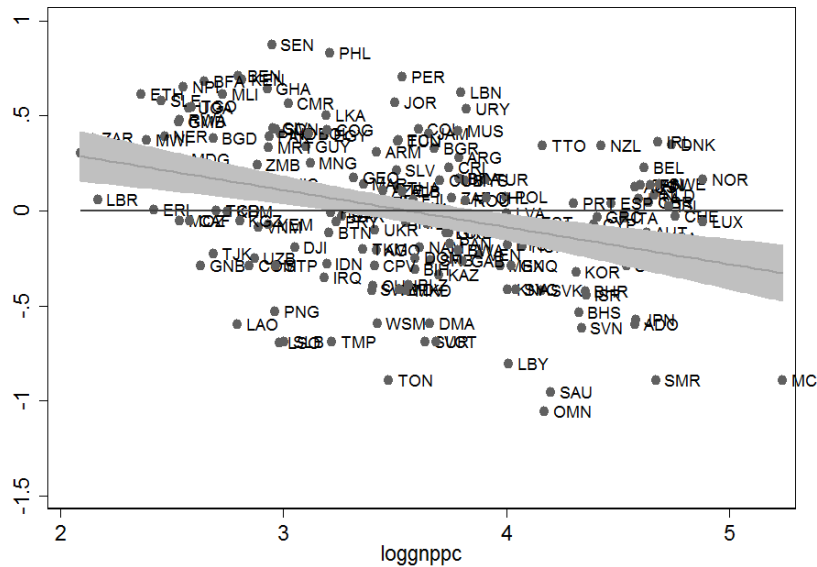
Figure 9.5: Residuals from desired representation regression

countries have very few staff members. But if there is this relative, over-representation, how can we explain it? What drives the fact that the weakest countries are more represented, than would correspond to their weakness?

In section 9.1 I made an argument that highly uneven power standing of the countries, if translated directly into their representation in the IOs, would likely prove dysfunctional. In particular, I argued that the IOs need for their functioning staff from the pertinent countries, to provide the thick understanding of the local conditions and in general to secure efficient collections and processing of the soft information, or information that is very difficult to be systematically collected and transmitted through some standardized channels, yet may prove crucial.

If we refer back to the IMF Action Plan for diversity, we see that for the African countries their desirable representation on the staff is set at about double the level of their subscriptions. The explanation given for this 'special treatment' speaks directly to the information

problem (International Monetary Fund, 2003, p. 7, emphasis added):

> In general, the indicators are based on Fund financial quotas. For Africa, however, three factors would suggest a higher representation for sub-Saharan Africa. First, under alternative indicators (e.g., *number of member countries in each region, the number of program countries, and the number of staff days spent on a region*), Africa accounts for a much larger share than its quota. Second, *it is important to have staff familiar with the institutional and other factors at play in Africa* to help the member countries in designing and implementing growth-oriented policies to reduce poverty. Finally, a critical mass of staff of sub-Saharan origin is considered desirable.

In other words, the IMF does a lot of work in Africa, and it needs staff familiar with the specifics in the region. So here I will present the empirical evidence showing that this argument is indeed substantiated; I will show that part of the remaining variance in the countries' representation in the secretariats can be explained by the IOs' demand for staff from their client countries, i.e. those countries with which they are likely to interact.

We should see that already the debate above indicates this result. I showed that there is some deviation in the IOs from the standards set by various possible criteria, and that this deviation tends to relative favour the less developed countries, i.e. those who are often among the primary addressees of the IOs' assistance. It is difficult, though, to what extent this deviation may be driven by the need for staff with knowledge of the frequent partner countries, as opposed to some broader organizational concerns, e.g. concerns with some elementary levels of credibility and legitimacy in the eyes of the developing countries.

To at least partially disentangle this, I introduce into the analysis a further set of control

variables, that for each of the four IOs serve as proxies capturing how likely a given country is to become a cooperation partner of the IO, or how intensely it is. In other words, these variables indicate some broad structural conditions that make the given country more or less relevant for the given IOs.

For the IMF, this purpose is served by the variable that captures the size of the current account balance of the country relative to its GDP. This variable measures to what extent a particular country has a problem, or a potential problem, with its imports/exports balance, and hence with its ability to pay for its imports. From the perspective of the IMF, the size of current account deficit, relative to the country's size of economy , serves as a key indicator of the country being in danger[3].

As far as UNDP is concerned, I proxy the potential for cooperation between the IO and the states with the share of population that is not immunized against measles. Although fairly specific, this is a broadly used indicator capturing well the general level of development-related problems in the given country. An obvious alternative to consider would be the Human Development Index published by UNDP itself, but this indicator correlates very highly with the countries' GNP per capita and hence could not be used in the analysis.

For UNESCO I use the number of UNESCO heritage sites in the country as the relevant proxy. This variable is used because it captures the unique aspect of the UNESCO's activity, while other potentially relevant indicators – e.g. illiteracy levels – fall within the realm of several different IOs (e.g. UNICEF, UNDP).

Finally, for WHO, I use average life expectancy, where the countries with low life expectancy are considered as the primary clients of the WHO. Again different indicators can

---

[3]The possible exception being the United States, that have run significant current account deficits, especially with China, for many years.

be suggested, but life expectancy captures the level of health problems in general very well[4].

Table 9.3.2 summarizes the results of the four regression models, each column capturing one of the IOs. The table shows that the variables indicating the interest of the IO in a country (the probability of becoming or the intensity of being a client) are consistently significant across all the IOs.

In Model 1 the higher the current account balance, i.e. the more positive it is and the less the country faces a balance of payments problem, the less it is represented on the IMF staff, *ceteris paribus*. The countries that do run significant current account deficits, and are in potential danger, are represented more. In Model 2 we see that the countries with a higher share of population not immunized against measles have higher representation in the UNDP staff. Taking levels of immunization as a proxy for serious of the development challenges, the countries facing more problems are represented in the UNDP staff relatively more. The same is true with Model 3 and UNESCO. Again, the more cultural heritage sites the country hosts, the more represented it is likely to be in the UNESCO staff. Finally, in Model 4 we see that the countries with higher life expectancy tend to be relatively less represented in the WHO secretariat, *ceteris paribus*. Those countries with serious health problems are represented more.

All four IOs show some level of ability (and willingness) to counter-act in their practical personnel policy the formal rules that define the appropriate levels of each country's representation. Yet, we should bear in mind that for each of the IOs the appropriate level is defined differently (only UNDP and WHO share the same criteria), and hence that how the desired representation in the models is established varies. In the last part of the empirical analysis I try to explore whether some of the IOs over-represent the developing countries

---

[4]I ran the analysis also with the variable capturing share of population not immunized against measles – the same variable I use for UNDP – and the results are essentially identical, even somewhat stronger.

	(1)	(2)	(3)	(4)
	logimfstaff	logundpstaff	logunescostaff	logwhostaff
logimfdesired	0.417*			
	(2.02)			
loggnppc	0.236	0.054	-0.034	0.102
	(1.41)	(0.55)	(-0.38)	(1.00)
logpopulation	0.348	0.220*	0.362***	0.382***
	(1.89)	(2.49)	(3.80)	(5.37)
cabalancesharegdp0608mean	-0.197*			
	(-2.51)			
logundpdesired		0.508***		
		(4.70)		
logmeaslesnonimmun		0.153*		
		(2.21)		
logunescodesired			0.257**	
			(2.94)	
logunescoheritage			0.275**	
			(2.99)	
logwhodesired				0.555***
				(6.36)
lifeexpectancy07				-0.272**
				(-3.17)
R2	0.515	0.458	0.542	0.638
Adjusteed R2	0.494	0.443	0.528	0.628
N	97	147	139	144

Standardized beta coefficients; t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

more than the others do.

9.3.3 Variation across the IOs

I indicated above that the four IOs were selected for variation on two criteria: 1) are they UN programmes, or are they independent (UN) agencies? 2) Do they deal in their work with politically salient matter, or with broad socio-economic problems, or is their work rather seen as impartial, a-political. For this part of the analysis, we can presume that the more political IOs will be more directly controlled by the powerful states, i.e. that also their personnel policy will be more accountable to only the key stakeholders. So we expect that especially the IMF, and to a lesser extent perhaps also UNDP, will maintain relatively more uneven distribution of staff, favouring the more powerful countries. The less political IOs, UNESCO and especially WHO, can then be expected to enjoy more autonomy in their personnel policies, and are likely to counter-act the uneven distribution more significantly.

Figure 9.6 indicates that this logic of higher or lower autonomy, due to higher or lower political saliency of the IOs, is justified. The figure plots the distribution of staff of each of the IO along the different levels of GNP per capita of the member states. Clearly in IMF the rich countries hold their over-representation most strongly. We should note, however, that the result is to some extent drive by the United States, who are in the IMF massively over-represented. In the WHO, on the contrary, the rich countries are not as significantly over-represented, relative to the poorer ones. UNDP and UNESCO show approximately the same results, in between IMF and WHO.

This graphical impression can be supported also numerically. Table 9.3.3 presents results of three regression models where the staff representations from the individual countries (in logarithms) serve as the independent variables 'predicting' the values of three different de-
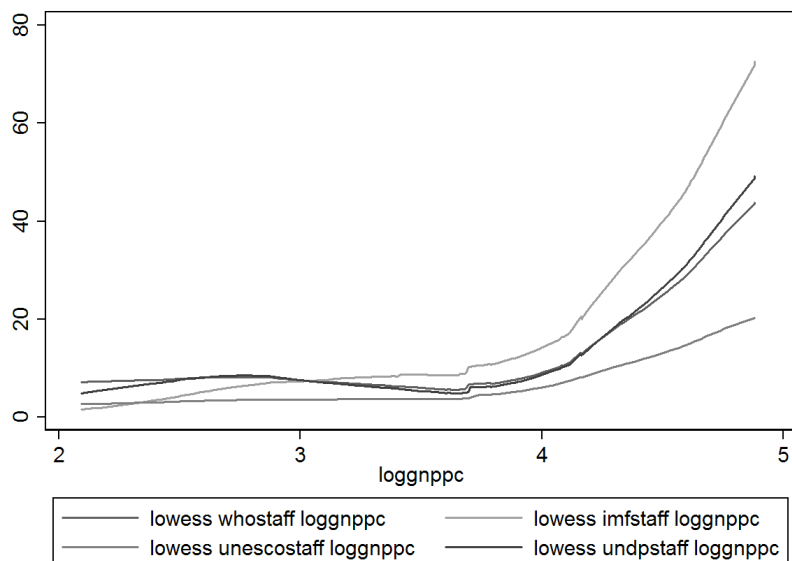
226

Figure 9.6: Comparison of staff composition dependence on GNP per capita

pendent variables: GNP per capita, GNP, and population size (again, all in logs). Obviously I do not presume a causal effect of the states' representation sizes on their economic performance or even population sizes, but this reverse approach can help us compare which of the four IOs corresponds in their structure of representation most closely to the distribution of each of the three 'dependent' variables.

What we see is that the distribution of votes in the IMF closely corresponds to the distribution of both wealth measured in GNP per capita (Model 1) and economy size (Model 2). On the other hand, WHO most closely follows in its composition the distribution of population across the member states. Although informal and not theoretically grounded, this simple analysis supports the earlier intuitions about that the more political IOs (in particular the IMF) are likely to follow more closely in their composition the distribution of power, while the more technical ones (WHO) are not so strictly bound.

	(1)	(2)	(3)
	loggnppc	loggnp	logpopulation
logimfstaff	0.609***	0.590***	0.183
	(4.94)	(5.99)	(1.66)
logunescostaff	0.060	0.177	0.170
	(0.48)	(1.77)	(1.53)
logwhostaff	-0.211	0.170	0.525***
	(-1.31)	(1.32)	(3.86)
logundpstaff	-0.014	-0.167	-0.237
	(-0.10)	(-1.44)	(-1.84)
R2	0.261	0.526	0.384
Adjusted R2	0.235	0.509	0.362
N	115	115	119

Standardized beta coefficients; t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

9.4 Conclusions

In general, this chapter showed how the organizational-cybernetic framework based on the problem of information transmission can help us understand the important problem of how the major international organizations are staffed.

I showed that the empirical pattern of staffing is influenced by the concern of the IOs with their ability to access not only hard – quantified, standardized – information about their client countries, but also the information that is inherently difficult to transmit, the soft information on local cultural, institutional, and political contexts. To be able to work with this information, the IOs need staff from the countries they deal with. The IOs, as political systems, need to have the capacity to collect and process all important information about the realities in their client countries. If they do not, they will not be able to deliver the goals they are expected to deliver.

Yet, it is in the direct interest of the powerful states to maintain as high a control over the internal functioning of these IOs as possible, and securing high representation of the staff from them is an important means to that goal. These two counter-acting tendencies result in an interesting empirical pattern, in which the interest of the powerful determine the overall pattern, but where the functional needs of the institutions lead them to act against the pattern at the margin, or in the cases where the unevenness of the countries' representation is most pronounced.

Chapter 10

Conclusions

I have presented in the course of the dissertation a set of several relatively complex theoretical propositions. The major purpose of this concluding chapter needs to be to take a step back and to provide some space for a broader reflection of these arguments. I will try to achieve this in three steps. The first and ultimately the most important step is to review the arguments in a way that highlights its essential components that form together a coherent picture. In the second step, I discuss what my arguments, as they stand, imply for the existing discussions on the design of international institutions and what are the possible avenues of future research. Third, I would like to discuss the broader implications of my findings.

10.1 The organizational-cybernetic framework for the design of international institutions

In the dissertation, I present an organizational-cybernetic framework for the design of international institutions. It is a framework that stresses two principal factors that – I argue – are essential for functioning of the international institutions: the key role of information transmission, and the systemic nature of international institutions. The core of the framework lies in its ability to explicitly deal, in a somewhat sophisticated way, with the problem of flows of information among the actors that are trying to achieve cooperative outcomes in international politics. I argue that if we consider power and interest distributions as given, the problem of cooperation lies in whether the existing institutional frameworks ensure all the information transmission that the actors need for successful cooperation. For analysis of this problem two components are necessary: a well defined dependent variable that captures this ability to ensure information transmission, and a measure, or a criterion, of how much of this ability to ensure information transmission is enough.

I propose the concept of capacity to serve the first purpose. If we want to understand how well different designs of institutions contribute to their functioning, we need a concept of the institutions' ability to ensure information transmission as the dependent variable. It is not at all trivial to define such a dependent variable. We need that it effectively tells how well the institutions we study work, but we want to make sure that it also clearly distinguishes the positive or negative effects of the design as such from the effects that are brought to cooperation by the given distributions of powers and interests. The concept of capacity does that; it reflects the powers and interests of the key actors, but it also directly captures the role played in the actors' interaction by the institutional framework within which it takes

place.

As for the second purpose, once we have this dependent variable of capacity, we need to know how much capacity the institutions we study need to provide the space for successful cooperation. I find the solution to this enormous problem in the viable system model (VSM) of cybernetician Stafford Beer (Beer, 1972). Derived from the cybernetic law of requisite variety, VSM identifies all the key information channels that any system needs to contain, and stipulates in a general manner the capacity these channels need to possess, for the entire cooperative system to be viable. In other words, by importing from the field of organizational cybernetics the VSM, we obtain a framework with the help of which we can answer the essential question of what exactly the targets of our design efforts need to be.

Having both the notion of capacity and the VSM, we really have the two keys to the analysis of design. On a conceptual level, we know exactly what institutional design needs to ensure. In any specific case, we are able to determine whether or not the targets have been met, and how much is lacking.

On the basis of the theoretical framework, we are also able to identify the types of design arrangements that are likely to prove effective in making the institutions work better, that means to increase their capacity. Some of these will have to do with securing that sufficiently powerful actors support the institution's efforts to transmit the information it is supposed to transmit. Delegation of powers to the supranational bodies whose task it is to collect, process, and disseminate information, is a prime example of such tools. Other changes in institutional arrangements will simply deal with the technical capacity of the institutions to transmit information. It is essential that what design arrangements are adopted should not be directly determined by whatever immediate specific targets we may have, but rather by consideration of their long-term impact on information flows within the system. Therefore,

even where creation of new rules that prohibit certain behaviour and induce other seems desirable from some specific perspective, the primary concern of the design always needs to lie in information transmission.

10.2   Research current and future

Throughout the text I have referred to a great variety of other academic works dealing with the problem of design but it is important that I stress again at this point how my work relates to the current debates on the design of international institutions. In particular, there are two key propositions I put forward in the dissertation, and these are – I would argue – critical for how the analysis of institutional design is conducted.

First, I argue that we should focus in the analysis of how the design of international institutions contributes to their functioning on the information flows and only on the key system functions of the institutions. That in effect means that the analyses should deal primarily with the information flows that are vital for cooperation. My argument is that if we want to take power and interests as the ultimate drivers of outcomes in international politics, and if we consider these as given exogenously, the only function the international institutions can perform lies in better transmission of information. This is not a contingent proposition; I argue that on the basis of the rationalist theories of international relations this is a *necessary* conclusion.

Similarly, I maintain that if we want to take seriously the enormous complexity of the problems that are dealt with in international politics and global governance, we need to limit our attention only to how the institutions perform their vital functions. We cannot hope to be able to model outcomes in international politics through sophisticated and detailed sets of

rules, the best we can hope for is that the institutions ensure reasonable levels of transmission of information that the actors always need for cooperation. Again, I try to show that this is a *necessary* step we need to make, if we believe the problems of international governance to be complex in their nature, i.e. involving lots of actors with very diverse interests, and deciding under the conditions of high uncertainty.

Both these steps, I have argued in part I, are necessary, we cannot avoid them if we assume both power and interests on the one hand, and the inherent complexity of the international problems on the other, to be important.

In part II, I started by taking this as given, and proposed a framework that can deal with these challenges in a fruitful way. This framework is derived from the field of organizational cybernetics and in particular from the viable system model. The key advantage of using the viable system model is that it has been developed precisely to deal, in a general manner, with the problem of information flows within systems of regulation. I should stress, thought, that it does not mean that the framework I propose here is the only logically possible solution to the theoretical challenges I identify in the earlier chapters. There may be other fruitful approaches. So while the theoretical argument I make in part I follows necessarily from the assumptions I make about the nature of international politics, the organizational cybernetic framework is not necessarily the only solution. Yet, given the exact fit of the theoretical challenges I identify and of the cybernetic approach and specifically of the organizational cybernetic theory, I believe that application of the organizational cybernetic approach to the analysis of international institutions' design promises to be a fruitful research endeavour. The ability of the viable system model to identify the information channels, and their capacity, that are logically necessary for viability of any governance system, and in general its deductive power, provide the kind of a firm theoretical basis for empirical research that is not often

seen in political institutional analysis.

Whether the general insights of the viable system model and of the framework I build on it can be fruitfully translated into empirical research remains to be seen. The two applications of the framework presented in part III indicate, though, that with its help we can address empirically interesting problems of international institutions' design. And importantly, we can do that within a framework that is theoretically sound, i.e. that focuses on the key problem of information transmission and that approaches the problem of design form the perspective of the vital system functions.

Any future research using the framework I develop here should, first and foremost, derive from the framework testable hypotheses and test them empirically. Given the very general nature of the framework, an uncountable number of empirical applications can be devised. It is possible, though, to identify at least two main avenues the future empirical research should take. First, we clearly need systematic empirical studies of how some of the standard design features discussed in the literature fair in terms of information transmission, that means to what extent they contribute to, or hinder, the key information transmission function. In chapter 7 I identified a number of specific design features that may plausibly be expected to improve transmission of information, and their true impact needs to be tested. Many other potentially useful design features can be listed, and empirical research that would systematically explore effects of existing design on information transmission is needed.

Second, future research should also be oriented more politically-substantively, and explore the reasons for which some of the existing international institutions do not achieve the results many hope for. The deadlocked negotiations of trade liberalization, or the uneasy negotiations on the climate change package are just two examples of key issue areas where success has been lacking, and where careful scholarly institutional analysis is needed. The

235

framework I propose here can serve as the basis for such substantively important research endeavours.

10.3   Some broader implications of my findings

The entire argumentation I have presented was very explicitly positive, in the sense that I did not consider at all the normative dimension of the design of international institutions. Clearly the conclusion of the dissertation is not the place to start considering the normative problems of international governance from scratch. But there is one important normative problem that directly follows from my argumentation and that it might be worth stressing now. If we believe that international politics are driven by power and interest, we may face in discussing the problem of the design of international institutions an unhappy trade-off between how we would like the institutions to look like, from some normative stand point, and what they need to look like to be viable.

I have argued at some length that the power- and interest-driven nature of international politics means that ultimately the only contribution international institutions can make to cooperation lies in improved information transmission. This is what they can deliver. Yet, for various normative reasons we may want to design the institutions in specific ways that will actually go against performance of this function. We may believe that certain design features are, by their nature, more normatively justifiable than others. What choice should one make, then, when we realize that these features not only do not contribute to better information transmission, but perhaps that they even prevent it?

I do not propose any specific answer to this problem, but what my arguments certainly do imply is that whenever possible, such design solutions should be chosen that maximize

performance of the information transmission function; if several normatively similarly desirable design arrangements can be devised, those that ensure better information transmission need to be chosen. Moreover, even if we do have some normative considerations about how the institutions are designed, we should understand very well that in the anarchical realm of international politics, viability of cooperation is always the paramount concern. Securing more even weights for countries' votes in some international decision-making body may be normatively desirable. Yet, if it leads to increased by-passing of the formal procedures and to move of the core political negotiations into exclusive arenas, the interests of the weaker countries may as a result not even be heard. Power politics matter for what outputs the international political processes produce, and the design of international institutions needs to reflect this functionally.

Finally, my argumentation in the dissertation implies that perhaps the way we approach the problem of the design and functioning of international institutions should shift more from studying of how the institutions are designed to how specifically they function, or more precisely where their vital functions are lacking. We have more and more sophisticated studies of why states design institutions the way they do and we also have studies of how institutions work and fail to work. The framework I present in this text should serve the direct and explicit integration of these research endeavours. We need systematic theoretically grounded studies of what design makes for better international institutions, why it does so, and how confident we can be it will continue to do so even when power and interest relations among the actors change.

I have argued that to be able to achieve this, we need an approach that is very directly oriented at the causal connection between design and capacity of the institutions but at the same time that is able to truthfully reflect the enormous complexities inherent in the

problems the international institutions need to deal with. We cannot hope to design good international institutions by assuming these complexities away, as much as we cannot hope to design good international institutions while neglecting the underlying power and interest realities. An approach that will stress one of these aspects of international politics at the expense of the other is doomed to fail. We need a synthetic framework that will fruitfully integrate power and interests with the more technical problems involved in solutions of global problems. I hope this dissertation has offered such a framework.

Bibliography

Abbott, Kenneth W., Keohane, Robert O., Moravcsik, Andrew, Slaughter, Anne-Marie, & Snidal, Duncan. 2000. The Concept of Legalization. *International Organization*, **54**(3), 401–419.

Achen, Christopher H. 2006. Institutional Realism and Bargaining Models. *Pages 86–123 of:* Thomson, Robert, Stokman, Frans N., Achen, Christopher H., & König, Thomas (eds), *The European Union Decides.* Cambridge: Cambridge University Press.

Akerlof, George A. 1970. The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, **84**(3), 488–500.

Albert, Mathias, & Cederman, Lars-Erik. 2010. Introduction : Systems Theorizing in IR. *Pages 3–22 of:* Albert, Mathias, Cederman, Lars-Erik, & Wendt, Alexander (eds), *New systems theories of world politics.* Basingstoke: Palgrave Macmillan.

Albert, Mathias, Cederman, Lars-Erik, & Wendt, Alexander. 2010. *New systems theories of world politics.* Basingstoke: Palgrave Macmillan.

Alberts, David Stephen, & Czerwinski, Thomas J. 1997. *Complexity, global politics, and national security.* National Defense University.

Allison, Graham T. 1969. Conceptual Models and the Cuban Missile Crisis. *American Political Science Review*, **63**(03), 689–718.

Almond, Gabriel A., Powell, Bingham G., Strom, Kaare, & Dalton, Russell J. 2003. *Comparative Politics: A Theoretical Framework*. 4th edn. London: Longman.

Andrews, David M. 1994. Capital Mobility and State Autonomy: Toward a Structural Theory of International Monetary Relations. *International Studies Quarterly*, **38**(2), 193.

Ashby, William Ross. 1956. *Introduction to Cybernetics*. London: Routledge Kegan & Paul.

Axelrod, Robert. 1981. The Emergence of Cooperation among Egoists. *The American Political Science Review*, **75**(2), 306–318.

Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic Books.

Baldwin, David A. (ed). 1993. *Neorealism and Neoliberalism*. New York: Columbia University Press.

Banzhaf, John F. III. 1964. Weighted Voting Doesn't Work: A Mathematical Analysis. *Rutgers Law Review*, **19**, 317.

Barnett, Michael, & Duvall, Raymond (eds). 2005. *Power in Global Governance*. Cambridge University Press.

Barnett, Michael, & Finnemore, Martha. 2004. *Rules for the World: International Organizations in Global Politics*. Cornell University Press.

BBC. 2012 (July). *Russian NGO bill signed into law*. http://www.bbc.co.uk/news/world-europe-18938165.

Beer, Stafford. 1972. *Brain of the Firm.* London: Wiley.

Beer, Stafford. 1979. *The Heart of Enterprise.* London: Wiley.

Beer, Stafford. 1984. The Viable System Model: Its Provenance, Development, Methodology and Pathology. *The Journal of the Operational Research Society*, **35**(1), 7–25.

Bertalanffy, Ludwig von. 1968. *General System Theory: Foundations, Development Applications.* George Braziller.

Best, Edward, & Settembri, Pierpaolo. 2008. Surviving enlargement: how has the Council managed? *Pages 34–53 of:* Best, Edward, Settembri, Pierpaolo, & Christiansen, Thomas (eds), *The institutions of the enlarged European Union: continuity and change.* London: Edward Elgar Publishing.

Bouayad-Agha, Fatih, & Hernandez, Homero L. 1996. *Comparison of methods of calculating equitable geographical distribution within the United Nations system.* Tech. rept. JIU/REP/96/7. United Nations Joint Inspection Unit.

Brambor, Thomas, Clark, William Roberts, & Golder, Matt. 2006. Understanding Interaction Models: Improving Empirical Analyses. *Political Analysis*, **14**(1), 63–82.

Breitmeier, Helmut, Young, Oran R., & Zürn, Michael. 2006. *Analyzing international environmental regimes: from case study to database.* Cambridge: MIT Press.

Bueno de Mesquita, Bruce. 2009. *Principles of International Politics.* 4th edn. Washington D.C.: CQ Press.

Bueno de Mesquita, Bruce, Smith, Alastair, Siverson, Randolph M., & Morrow, James D. 2003. *The Logic of Political Survival.* Cambridge: The MIT Press.

Busch, Marc L., & Reinhardt, Eric. 2003. Developing Countries and General Agreement on Tariffs and Trade/World Trade Organization Dispute Settlement. *Journal of World Trade*, **37**(4), 719–735.

Carr, Edward Hallett. 1939. *The twenty years' crisis, 1919-1939: an introduction to the study of international relations*. Basingstoke: Palgrave Macmillan.

Cederman, Lars-Erik. 1997. *Emergent Actors in World Politics*. Princeton University Press.

Cederman, Lars-Erik. 2010. Complexity and Change in World Politics: Resurrecting Systems Theory. *In:* Albert, Mathias, Cederman, Lars-Erik, & Wendt, Alexander (eds), *New systems theories of world politics*. Basingstoke: Palgrave Macmillan.

Chayes, Abram, & Chayes, Antonia Handler. 1993. On Compliance. *International Organization*, **47**(2), 175–205.

Clarke, Bruce. 2011. Heinz von Foerster and Niklas Luhmann: The Cybernetics of Social Systems Theory. *Cybernetics & Human Knowing*, **18**(3-4), 95–99.

Coase, R. H. 1960. The Problem of Social Cost. *Journal of Law and Economics*, **3**(Oct.), 1–44.

Conant, Roger C., & Ashby, William Ross. 1970. Every good regulator of a system must be a model of that system. *International Journal of Systems Science*, **1**(2), 89–97.

Conceição Heldt, Eugénia da. 2010. Who Controls Whom? Dynamics of Power Delegation and Agency Losses in EU Trade Politics. *JCMS: Journal of Common Market Studies*, **48**(4), 1107–1126.

Council of the European Union. 2009. *Council Decision of 1 December 2009 adopting the Council's Rules of Procedure. 2009/937/EU.*

Cox, Robert W. 1983. Gramsci, Hegemony and International Relations : An Essay in Method. *Millennium - Journal of International Studies*, **12**(2), 162 –175.

Desmedt, Axel. 2001. Proportionality in WTO Law. *Journal of International Economic Law*, **4**(3), 441–480.

Deutsch, Karl W., Burrell, Sidney A., & Kann, Robert A. (eds). 1957. *Political community and the North Atlantic area. International organization in the light of historical experience.* New York: Greenwood Pr.

Deutsch, Karl Wolfgang. 1963. *The Nerves of Government: Models of Political Communication and Control.* London: Free Press.

Downs, George W., Rocke, David M., & Barsoom, Peter N. 1996. Is the Good News About Compliance Good News About Cooperation? *International Organization*, **50**(3), 379–406.

Drezner, Daniel W. 2009. The Power and Peril of International Regime Complexity. *Perspectives on Politics*, **7**(01), 65–70.

Easton, David. 1975. A Re-assessment of the Concept of Political Support. *British Journal of Political Science*, **5**(04), 435–457.

Egeberg, Morten. 1999. Transcending intergovernmentalism? Identity and role perceptions of national officials in EU decision-making. *Journal of European Public Policy*, **6**(Jan.), 456–474.

Egeberg, Morten, Schaefer, Günther, & Trondal, Jarle. 2003. The many faces of EU committee governance. *West European Politics*, **26**(3), 19–40.

European Court of Auditors. 2012. *Structural Funds: Did the Commission Successfully Deal with the Deficiencies Identified in the Member States' Management and Control Systems? Special Report 3/2013.* Accessed 20. 3. 2013.

Evera, Stephen van. 1985. Why Cooperation Failed in 1914. *World Politics*, **38**(1), 80–117.

Fearon, James D. 1994. Domestic Political Audiences and the Escalation of International Disputes. *The American Political Science Review*, **88**(3), 577–592.

Fearon, James D. 1995. Rationalist Explanations for War. *International Organization*, **49**(3), 379–414.

Fearon, James D. 1998. Bargaining, Enforcement, and International Cooperation. *International Organization*, **52**(2), 269–305.

Fearon, James D., & Wendt, Alexander. 2002. Rationalism v. Constructivism: A Skeptical View. *In:* Carlsnaes, Walter E. (Emmanuel), Risse, Thomas, & Simmons, Beth A (eds), *Handbook of International Relations.* London: Sage Publications Ltd.

Finer, S. E. 1970. Almond's Concept of The Political System: A Textual Critique. *Government and Opposition*, **5**(1), 3–21.

Fukuyama, Francis. 2004. *State-Building : Governance and World Order in the 21st Century.* Ithaca: Cornell University Press.

Gell-Mann, Murray. 2002. *The Quark and the Jaguar: Adventures in the Simple and the Complex.* Owl Books.

George, Alexander L., & Bennett, Andrew. 2005. *Case Studies and Theory Development in the Social Sciences.* Mit Press.

George, David Lloyd. 1938. *War Memoirs.* 2nd edn. London: Odhams.

Geyer, Felix, & Zouwen, Johannes van der. 2001. *Sociocybernetics : complexity, autopoiesis, and observation of social systems.* London: Greenwood Publishing Group.

Gilligan, Thomas W., & Krehbiel, Keith. 1989. Asymmetric Information and Legislative Rules with a Heterogeneous. *American Journal of Political Science*, **33**(2), 459.

Glaser, Charles L. 1994. Realists as Optimists: Cooperation as Self-Help. *International Security*, **19**(3), 50–90.

Gleick, James. 1988. *Chaos : making a new science.* Penguin.

Golub, Jonathan. 2007. Survival Analysis and European Union Decision-making. *European Union Politics*, **8**(2), 155–179.

Goodin, Robert E. 1998. *The Theory of Institutional Design.* Cambridge University Press.

Greif, Avner. 1993. Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders' Coalition. *The American Economic Review*, **83**(3), 525–548.

Grieco, Joseph, Powell, Robert, & Snidal, Duncan. 1993. The Relative-Gains Problem for International Cooperation. *The American Political Science Review*, **87**(3), 727–743.

Grieco, Joseph M. 1988. Anarchy and the Limits of Cooperation: A Realist Critique of the Newest Liberal Institutionalism. *International Organization*, **42**(3), 485–507.

Grieco, Joseph M. 1993. Understanding the Problem of International Cooperation: the Limits of Neoliberal Institutionalism and the Future of Realist Theory. *In:* Baldwin, David A. (ed), *Neorealism and Neoliberalism.* New York: Columbia University Press.

Gutner, Tamar, & Thompson, Alexander. 2010. The politics of IO performance: A framework. *The Review of International Organizations,* **5**(July), 227–248.

Häge, Frank M. 2007. Committee Decision-making in the Council of the European Union. *European Union Politics,* **8**(3), 299 –328.

Häge, Frank M. 2008. Who Decides in the Council of the European Union? *JCMS: Journal of Common Market Studies,* **46**(3), 533–558.

Hagemann, Sara. 2007. Applying Ideal Point Estimation Methods to the Council of Ministers. *European Union Politics,* **8**(2), 279 –296.

Hagemann, Sara, & Hoyland, Bjorn. 2008. Parties in the Council? *Journal of European Public Policy,* **15**(8), 1205–1221.

Halás, Matúš. 2009. Post Scriptum on Relative and Absolute Gains. *Perspectives. Review of International Affairs,* **17**(1), 27–56.

Hardin, Garrett. 1968. The Tragedy of the Commons. *Science,* **162**(3859), 1243 –1248.

Hasenclever, Andreas, Mayer, Peter, & Rittberger, Volker. 1997. *Theories of International Regimes.* Cambridge: Cambridge University Press.

Hawkins, Darren G., Lake, David A., Nielson, Daniel L., & Tierney, Michael J. (eds). 2006. *Delegation and Agency in International Organizations.* Cambridge: Cambridge University Press.

Hayek, F. A. 1945. The Use of Knowledge in Society. *The American Economic Review*, **35**(4), 519–530.

Hayes-Renshaw, Fiona, & Wallace, Helen. 2006. *The Council of Ministers*. 2nd edn. Basingstoke: Palgrave Macmillan.

Hayes-Renshaw, Fiona, Van Aken, Wim, & Wallace, Helen. 2006. When and Why the EU Council of Ministers Votes Explicitly. *JCMS: Journal of Common Market Studies*, **44**(1), 161–194.

Held, David, & McGrew, Anthony G. (eds). 2002. *Governing globalization: power, authority and global governance.* Cambridge: Polity.

Held, David, & McGrew, Anthony G. 2003. The Great Globlaization Debate: an Introduction. *Pages 1–50 of:* Held, David, & McGrew, Anthony (eds), *The Global Transformations Reader: An Introduction to the Globalization Debate*, 2 edn. Cambridge: Polity.

Helm, Carsten, & Sprinz, Detlef. 2000. Measuring the Effectiveness of International Environmental Regimes. *Journal of Conflict Resolution*, **44**(5), 630 –652.

Héritier, Adrienne. 1999. *Policy-Making and Diversity in Europe: Escape from Deadlock.* Cambridge: Cambridge University Press.

Hertz, Robin, & Leuffen, Dirk. 2011. Too Big to Run? Analysing the Impact of Enlargement on the Speed of EU Decision-Making. *European Union Politics*, **12**(2), 193–215.

Herz, John H. 1950. Idealist Internationalism and the Security Dilemma. *World Politics*, **2**(02), 157–180.

Hix, Simon. 2005. *The Political System of the European Union*. 2nd edn. Basingstoke: Palgrave Macmillan.

Hoffmann, Matthew J., & Riley, John. 2002. The Science of Political Science: Linearity or Complexity in Designing Social Inquiry. *New Political Science*, **24**(2), 303–320.

Holland, John H. 1999. *Emergence : From Chaos To Order*. Basic Books.

Holmstrom, Bengt, & Myerson, Roger B. 1983. Efficient and Durable Decision Rules with Incomplete Information. *Econometrica*, **51**(6), 1799.

Hooghe, Liesbet, & Marks, Gary. 2003. Unraveling the Central State, but How? Types of Multi-Level Governance. *The American Political Science Review*, **97**(2), 233–243.

Hooghe, Liesbet, & Marks, Gary. 2009. A Postfunctionalist Theory of European Integration: From Permissive Consensus to Constraining Dissensus. *British Journal of Political Science*, **39**(01), 1–23.

Hooghe, Liesbet, & Marks, Gary. 2012. Beyond Federalism: Estimating and Explaining the Territorial Structure of Government. *Publius: The Journal of Federalism*, July.

Hosli, Madeleine O., & Machover, Moshé. 2004. The Nice Treaty and Voting Rules in the Council: A Reply to Moberg (2002). *JCMS: Journal of Common Market Studies*, **42**(3), 497–521.

Hovi, Jon, Sprinz, Detlef F., & Underdal, Arild. 2003. The Oslo-Potsdam Solution to Measuring Regime Effectiveness: Critique, Response, and the Road Ahead. *Global Environmental Politics*.

Huck, Steffen, & Weizsäcker, Georg. 2002. Do players correctly estimate what others do?: Evidence of conservatism in beliefs. *Journal of Economic Behavior & Organization*, **47**(1), 71–85.

Hurwicz, Leonid. 1972. On informationally decentralized systems. *In:* Arrow, Kenneth J., & Hurwicz, Leonid (eds), *Studies in Resource Allocation Processes*. Cam: Cambridge University Press.

Hurwicz, Leonid. 1973. The Design of Mechanisms for Resource Allocation. *The American Economic Review*, **63**(2), 1–30.

International Monetary Fund. 2003. *Enhanced Diversity Action Plan*. Tech. rept. UN-DOCI03/149.

Janssen, Marco A., & Ostrom, Elinor. 2006. Resilience, vulnerability, and adaptation: A cross-cutting theme of the International Human Dimensions Programme on Global Environmental Change. *Global Environmental Change*, **16**(3), 237–239.

Jervis, Robert. 1976. *Perception and misperception in international politics*. Princeton: Princeton University Press.

Jervis, Robert. 1978. Cooperation Under the Security Dilemma. *World Politics*, **30**(2), 167–214.

Jervis, Robert. 1982. Deterrence and Perception. *International Security*, **7**(3), 3–30.

Jervis, Robert. 1988. Realism, Game Theory, and Cooperation. *World Politics*, **40**(3), 317–349.

Jervis, Robert. 1998. *System Effects*. Princeton: Princeton University Press.

Johnston, Alastair. 2001. Treating International Institutions as Social Environments. *International Studies Quarterly*, **45**(4), 487–515.

Karlas, Jan. 2008. *Forma mezinarodnich instituci: teoreticka analyza*. Praha: Karolinum.

Keohane, Robert O. 1982. The Demand for International Regimes. *International Organization*, **36**(2), 325–355.

Keohane, Robert O. 1984. *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton: Princeton University Press.

Keohane, Robert O. 1989. Neoliberal institutionalism. *In:* Keohane, Robert O. (ed), *International Institutions And State Power: Essays In International Relations Theory*. Westview Press.

Keohane, Robert O. 1993. Institutional Theory and the Realist Challenge after the Cold War. *In:* Baldwin, David A. (ed), *Neorealism and Neoliberalism*. New York: Columbia University Press.

Keohane, Robert O. 2003. Global Governance and Democratic Accountability. *In:* Held, David, & Koenig-Archibugi, Mathias (eds), *Taming Globalization: Frontiers of Governance*. Cambridge: Polity Press.

Keohane, Robert O. 2006. Accountability in World Politics. *Scandinavian Political Studies*, **29**(2), 75–87.

Keohane, Robert O., & Martin, Lisa L. 1995. The Promise of Institutionalist Theory. *International Security*, **20**(1), 39–51.

Keohane, Robert O., & Nye, Joseph S. 1977. *Power and Interdependence.* New York: Longman.

Kindleberger, Charles P. 1986. *The world in depression, 1929-1939.* University of California Press.

Kleine, Mareike O. 2010. Making Cooperation Work: Informal Governance in the EU and Beyond. *SSRN eLibrary.* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1641953.

Koenig-Archibugi, Mathias. 2003. Mapping Global Governance. *In:* Held, David, & Koenig-Archibugi, Mathias (eds), *Taming Globalization: Frontiers of Governance.* Cambridge: Polity Press.

König, Thomas, & Bräuninger, Thomas. 2004. Accession and Reform of the European Union. *European Union Politics*, **5**(4), 419 –439.

König, Thomas, & Junge, Dirk. 2009. Why Dont Veto Players Use Their Power? *European Union Politics*, **10**(4), 507 –534.

König, Thomas, Luetgert, Brooke, & Dannwolf, Tanja. 2006. Quantifying European Legislative Research. *European Union Politics*, **7**(4), 553 –574.

Koremenos, Barbara, Lipson, Charles, & Snidal, Duncan. 2001. The Rational Design of International Institutions. *International Organization*, **55**(4), 761–799.

Krasner, Stephen D. 1982. Structural Causes and Regime Consequences: Regimes as Intervening Variables. *International Organization*, **36**(2), 185–205.

Krasner, Stephen D. 1983. *International regimes.* Ithaca: Cornell University Press.

Krasner, Stephen D. 1991. Global Communications and National Power: Life on the Pareto Frontier. *World Politics*, **43**(3), 336–366.

Krehbiel, Keith. 2001. Plausibility of Signals by a Heterogeneous Committee. *American Political Science Review*, **95**(02), 453–457.

Krishna, Vijay, & Morgan, John. 2001. Asymmetric Information and Legislative Rules: Some Amendments. *American Political Science Review*, **95**(2), 435–452.

Lake, David A., & Powell, Robert (eds). 1999. *Strategic Choice and International Relations*. Princeton: Princeton University Press.

Lewis, Jeffrey. 2000. The methods of community in EU decision-making and administrative rivalry in the Council's infrastructure. *Journal of European Public Policy*, **7**(2), 261–289.

Lewis, Jeffrey. 2003. Informal integration and the supranational construction of the Council. *Journal of European Public Policy*, **10**(Dec.), 996–1019.

Lewis, Jeffrey. 2005. The Janus Face of Brussels: Socialization and Everyday Decision Making in the European Union. *International Organization*, **59**(04), 937–971.

Lewis, Jeffrey. 2010. How institutional environments facilitate co-operative negotiation styles in EU decision-making. *Journal of European Public Policy*, **17**(Aug.), 648–664.

Lipson, Charles. 1984. International Cooperation in Economic and Security Affairs. *World Politics*, **37**(1), 1–23.

Luhmann, Niklas. 1995. *Social Systems*. Stanford: Stanford University Press.

Lukes, Steven. 2005. *Power : A Radical View*. 2nd edn. Palgrave Macmillan.

Malik, Constantin. 2011. *Ahead of Change: How Crowd Psychology and Cybernetics Transform the Way We Govern.* Campus Verlag.

Martin, Lisa L. 1992. Institutions and Cooperation: Sanctions during the Falkland Islands Conflict. *International Security,* **16**(4), 143–178.

Mattila, Mikko. 2004. Contested decisions: Empirical analysis of voting in the European Union Council of Ministers. *European Journal of Political Research,* **43**(1), 29–50.

Mattila, Mikko. 2009. Roll call analysis of voting in the European Union Council of Ministers after the 2004 enlargement. *European Journal of Political Research,* **48**(6), 840–857.

Mattila, Mikko, & Lane, Jan-Erik. 2001. Why Unanimity in the Council? A Roll Call Analysis of Council Voting. *European Union Politics,* **2**(1), 31–52.

McCarty, Nolan, & Meirowitz, Adam. 2007. *Political Game Theory: An Introduction.* Cambridge: Cambridge University Press.

Mearsheimer, John J. 1994. The False Promise of International Institutions. *International Security,* **19**(3), 5–49.

Miller, James Grier. 1978. *Living Systems.* 2nd edn. Colorado: Univ Press of Colorado.

Miller, John H., & Page, Scott E. 2010. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life: An Introduction to Computational Models of Social Life.* Princeton University Press.

Mitchell, Ronald B. 1994. Regime Design Matters: Intentional Oil Pollution and Treaty Compliance. *International Organization,* **48**(3), 425–458.

Mitchell, Ronald B. 2006. Problem Structure, Institutional Design, and the Relative Effectiveness of International Environmental Agreements. *Global Environmental Politics*, **6**(3), 72–89.

Momani, Bessma. 2004. American politicization of the International Monetary Fund. *Review of International Political Economy*, **11**(5), 880–904.

Moravcsik, Andrew. 1997. Taking Preferences Seriously: A Liberal Theory of International Politics. *International Organization*, **51**(4), 513–553.

Morgenthau, Hans Joachim, Thompson, Kenneth W., & Clinton, W. David. 1948. *Politics among nations: the struggle for power and peace.* London: McGraw-Hill Higher Education.

Morrow, James D. 2001. The Institutional Features of the Prisoners of War Treaties. *International Organization*, **55**(04), 971–991.

Mosher, James S. 2003. Relative Gains Concerns when the Number of States in the International System Increases. *Journal of Conflict Resolution*, **47**(5), 642–668.

Mulford, Matthew, & Berejikian, Jeffery. 2002. Behavioural Decision Theory and the Gains Debate in International Politics. *Political Studies*, **50**(2), 209–229.

Muthoo, Abhinay. 1999. *Bargaining Theory with Applications.* Cambridge: Cambridge University Press.

Muthoo, Abhinay. 2000. A Non-technical Introduction to Bargaining Theory. *World Economics*, **1**(2), 145–166.

Myerson, Roger B. 1991. *Game Theory: Analysis of Conflict.* Harvard: Harvard University Press.

Myerson, Roger B. 2008. The Autocrat's Credibility Problem and Foundations of the Constitutional State. *American Political Science Review*, **102**(01), 125–139.

Myerson, Roger B. 2009. Fundamental theory of institutions: a lecture in honor of Leo Hurwicz. *Review of Economic Design*, **13**(1-2), 59–75.

Myerson, Roger B, & Satterthwaite, Mark A. 1983. Efficient mechanisms for bilateral trading. *Journal of Economic Theory*, **29**(2), 265–281.

Nash, John F. 1950. The Bargaining Problem. *Econometrica*, **18**(2), 155–162.

Nielson, Daniel L., & Tierney, Michael J. 2003. Delegation to International Organizations: Agency Theory and World Bank Environmental Reform. *International Organization*, **57**(2), 241–276.

Niskanen, William A. 1971. *Bureaucracy and representative government*. Aldine, Atherton.

North, Douglass C. 1990. *Institutions, Institutional Change and Economic Performance*. Cambridge University Press.

Nye, Joseph S. 2008. Public Diplomacy and Soft Power. *The ANNALS of the American Academy of Political and Social Science*, **616**(1), 94–109.

Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press.

Ostrom, Elinor. 2005. *Understanding Institutional Diversity*. Princeton University Press.

Ostrom, Elinor. 2009. A General Framework for Analyzing Sustainability of Social-Ecological Systems. *Science*, **325**(5939), 419 –422.

Ostrom, Elinor, Gardner, Roy, & Walker, Jimmy. 1994. *Rules, Games, and Common-Pool Resources.* Ann Arbor: University of Michigan Press.

Oye, Kenneth A. (ed). 1986. *Cooperation under Anarchy.* Princeton: Princeton University Press.

Panke, Diana. 2011. Small States in EU Negotiations: Political Dwarfs or Power-Brokers? *Cooperation and Conflict,* **46**(2), 123–143.

Parízek, Michal. 2011. Kapacita mezinárodních institucí. Konceptuální rámec pro empirickou analýzu. *Mezinárodní vztahy,* **46**(3), 55–76.

Parízek, Michal. 2012a. Design mezinárodních institucí: organizačně-kybernetický přístup. *Mezinárodní vztahy,* **47**(3), 53–73.

Parízek, Michal. 2012b. A 'need to know': how institutional design impairs articulation of interests in trade negotiations. Political Economy of International Organizations Conference, Heidelberg and Mannheim.

Parsons, Talcott. 1951. *The Social System.* New York: Free Press.

Petersen, Mitchell A. 2004. Information : Hard and Soft. *Northwestern University Woring paper.* www.disas.unisi.it/mat_did/gabbi/729/10.1.1.126.8246%5B1%5D.pdf.

Plechanovová, Běla. 2004. *Institucionální vývoj Evropské unie od maastrichtské smlouvy k východnímu rozšíření.* Praha: Karolinum.

Plechanovová, Běla. 2011a. Coalitions in the EU Council: Pitfalls of Multidimensional Analysis. *AUCO Czech Economic Review,* **5**(3), 249–266.

Plechanovová, Běla. 2011b. The EU Council enlarged: North-South-East or core-periphery? *European Union Politics*, **12**(1), 87 –106.

Plott, Charles R. 2000. Markets as information gathering tools. *Southern Economic Journal*, **67**(1), 2.

Pollack, Mark A. 1997. Delegation, agency, and agenda setting in the European Community. *International Organization*, **51**(01), 99–134.

Pollack, Mark A. 2003. *The Engines of European Integration: Delegation, Agency, and Agenda Setting in the Eu*. Oxford: Oxford University Press.

Popovski, Vesselin. 2000. International Criminal Court: A Necessary Step towards Global Justice. *Security Dialogue*, **31**(4), 405–419.

Powell, Robert. 1991. Absolute and Relative Gains in International Relations Theory. *The American Political Science Review*, **85**(4), 1303.

Putnam, Robert D. 1988. Diplomacy and Domestic Politics: The Logic of Two-Level Games. *International Organization*, **42**(3), 427–460.

Ragin, Charles C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. University of California Press.

Ragin, Charles C. 2000. *Fuzzy-Set Social Science*. University of Chicago Press.

Rasmusen, Eric. 2006. *Games and Information: An Introduction to Game Theory*. 4 edn. Wiley-Blackwell.

Raustiala, Kal, & Slaughter, Anne-Marie. 2002. International Law, International Relations and Compliance. *Pages 538–558 of:* Carlsnaes, Walter E., Risse, Thomas, & Simmons, Beth A. (eds), *Handbook of International Relations*. Sage Publications Ltd.

Raustiala, Kal, & Victor, David G. 2004. The Regime Complex for Plant Genetic Resources. *International Organization*, **58**(02), 277–309.

Reh, Christine, Héritier, Adrienne, Bressanelli, Edoardo, & Koop, Christel. 2013. The Informal Politics of Legislation: Explaining Secluded Decision Making in the European Union. *Comparative Political Studies*, **46**.

Reinalda, Bob, & Verbeek, Bertjan (eds). 1998. *Autonomous policy making by international organizations*. Routledge.

Reinicke, Wolfgang H. 1998. *Global Public Policy: Governing Without Government?* Washington D.C.: Brookings Institution Press.

Rind, D. 1999. Complexity and Climate. *Science*, **284**(5411), 105–107.

Rittberger, Volker, & Zürn, Michael. 1990. Towards regulated anarchy in East-West relations: causes and consequences of East-West regimes. *In: International regimes in East-West politics*. Pinter Publishers.

Rosenau, James N. 1995. Governance in the Twenty-First Century. *Global Governance*, **1**, 13–43.

Rosenau, James N. 1997. Many Damn Things Simultaneously: Complexity Theory and World Affairs. *Pages 32–41 of:* Alberts, David S., & Czerwinski (eds), *Complexity, Global Politics, and National Security*. Washington D.C.: National Defense University.

Rosenau, James N. 2000. Change, Complexity, and Governance in a Globalizing Space. *In:* Pierre, Jon (ed), *Debating governance.* Oxford: Oxford University Press.

Rosenau, James N., & Durfee, Mary. 2000. *Thinking Theory Thoroughly: Coherent Approaches to an Incoherent World, Second Edition.* Westview Press.

Rosendorff, Peter B. 2005. Stability and Rigidity: Politics and Design of the WTO's Dispute Settlement Procedure. *American Political Science Review*, **99**(3), 389–400.

Rosser, J. Barkley. 1999. On the Complexities of Complex Economic Dynamics. *Journal of Economic Perspectives*, **13**(4), 169–192.

Ruggie, John Gerard. 1982. International Regimes, Transactions, and Change: Embedded Liberalism in the Postwar Economic Order. *International Organization*, **36**(2), 379–415.

Sawyer, R. Keith. 2005. *Social Emergence: Societies As Complex Systems.* Cambridge University Press.

Scharpf, Fritz W. 1997. *Games Real Actors Play: Actor-centered Institutionalism In Policy Research.* Westview Press.

Schelling, Thomas C. 1960. *The strategy of conflict.* Harvard: Harvard University Press.

Schelling, Thomas C. 1969. Models of Segregation. *American Economic Review*, **59**(2), 488–493.

Schneider, Gerald, Gleditsch, Nils Petter, & Carey, Sabine. 2011. Forecasting in International Relations One Quest, Three Approaches. *Conflict Management and Peace Science*, **28**(1), 5–14.

Schulz, Heiner, & König, Thomas. 2000. Institutional Reform and Decision-Making Efficiency in the European Union. *American Journal of Political Science*, **44**(4), pp. 653–666.

Schwaninger, Markus. 2006. Theories of viability: a comparison. *Systems Research and Behavioral Science*, **23**(3), 337–347.

Seidl, David, & Schoeneborn, Dennis. 2010. Niklas Luhmann's Autopoietic Theory of Organisations: Contributions, Limitations, and Future Prospects. *Institute of Organization and Administrative Science Working Paper No. 105.*

Shannon, Claude E. 1948. A Mathematical Theory of Communication. *Bell System Technical Journal*, **27**(3), 379–423.

Shannon, Claude E., & Weaver, Warren. 1949. *The Mathematical Theory of Communication.* University of Illinois Press.

Shepsle, Kenneth A. 1989. Studying Institutions Some Lessons from the Rational Choice Approach. *Journal of Theoretical Politics*, **1**(2), 131–147.

Shepsle, Kenneth A. 2006. Rational Choice Institutionalism. *In:* Rhodes, R. A. W., Binder, Sarah A., & Rockman, Bert A. (eds), *The Oxford Handbook of Political Institutions.* Oxford: Oxford University Press.

Shepsle, Kenneth A. 2010. *Analyzing Politics: Rationality, Behavior and Instititutions.* W. W. Norton & Company.

Simmons, Beth A. 2009. *Mobilizing for Human Rights: International Law in Domestic Politics.* 1 edn. Cambridge University Press.

Singer, David J. 1961. The Level-of-Analysis Problem in International Relations. *World Politics*, **14**(01), 77–92.

Smith, James. 2004. Inequality in international trade? Developing countries and institutional change in WTO dispute settlement. *Review of International Political Economy*, **11**(3), 542–573.

Smith, Michael E. 2003. *Europe's Foreign and Security Policy: The Institutionalization of Cooperation.* Cambridge University Press.

Snidal, Duncan. 1991. Relative Gains and the Pattern of International Cooperation. *The American Political Science Review*, **85**(3), 701–726.

Sofer, Sasson. 1988. Old and new diplomacy: a debate revisited. *Review of International Studies*, **14**(3), 195–211.

Stein, Arthur A. 1982. Coordination and collaboration: regimes in an anarchic world. *International Organization*, **36**(2), 299–324.

Steinbruner, John D. 2002. *The Cybernetic Theory of Decision: New Dimensions of Political Analysis.* Princeton: Princeton University Press.

Stiglitz, Joseph E. 1975. The Theory of "Screening," Education, and the Distribution of Income. *American Economic Review*, **65**(3), 283–300.

Stiglitz, Joseph E. 2000. The Contributions of the Economics of Information to Twentieth Century Economics. *The Quarterly Journal of Economics*, **115**(4), 1441–1478.

Stone-Sweet, Alec, & Sandholtz, Wayne. 1997. European integration and supranational governance. *Journal of European Public Policy*, **4**, 297–317.

Taagepera, Rein. 2008. *Making Social Sciences More Scientific: The Need for Predictive Models.* Oxford University Press, USA.

Tallberg, Jonas. 2002. Paths to Compliance: Enforcement, Management, and the European Union. *International Organization,* **56**(3), 609–643.

Thompson, Alexander. 2010. Rational design in motion: Uncertainty and flexibility in the global climate regime. *European Journal of International Relations,* **16**(2), 269–296.

Thompson, Leigh L. 1991. Information exchange in negotiation. *Journal of Experimental Social Psychology,* **27**(2), 161–179.

Thomson, Robert, Boerefijn, Jovanka, & Stokman, Frans. 2004. Actor alignments in European Union decision making. *European Journal of Political Research,* **43**(2), 237–261.

Tsebelis, George. 2002. *Veto players: how political institutions work.* Princeton: Princeton University Press.

Tsebelis, George. 2008. Thinking about the Recent Past and the Future of the EU. *JCMS: Journal of Common Market Studies,* **46**(2), 265–292.

Tsebelis, George, & Yataganas, Xenophon. 2002. Veto Players and Decision-making in the EU After Nice. *JCMS: Journal of Common Market Studies,* **40**(2), 283–307.

Underdal, Arild. 1992. The Concept of Regime 'Effectiveness'. *Cooperation and Conflict,* **27**(3), 227–240.

UNFCCC. 2009. *Report of the Conference of the Parties on its fifteenth session, held in Copenhagen from 7 to 19 December 2009.* Tech. rept. FCCC/CP/2009/11/Add.1.

Urpelainen, Johannes. 2012. Unilateral Influence on International Bureaucrats An International Delegation Problem. *Journal of Conflict Resolution*, **56**(4), 704–735.

Vaubel, Roland. 1986. A public choice approach to international organization. *Public Choice*, **51**(1), 39–57.

Wallerstein, Immanuel. 1979. *The Capitalist World-Economy*. Cambridge: Cambridge University Press.

Waltz, Kenneth N. 1954. *Man, the State, and War: A Theoretical Analysis*. New York: Columbia University Press.

Waltz, Kenneth N. 1979. *Theory of International Politics*. Columbus: McGraw-Hill.

Weinberg, Gerald M. 1975. *An Introduction to General Systems Thinking*. London: John Wiley & Sons.

Weinberg, Gerald M., & Weinberg, Daniela. 1980. *On the Design of Stable Systems*. John Wiley & Sons Inc.

Wendt, Alexander. 1999. *Social Theory of International Politics*. Cambridge: Cambridge University Press.

Wendt, Alexander. 2006. Social Theory as Cartesian Science: an Auto-Critique from a Quantum Perspective. *Pages 178–216 of:* Guzzini, Stefano (ed), *Constructivism And International Relations: Alexander Wendt And His Critics*. London: Routledge.

Wiener, Norbert. 1948. *Cybernetics : or the Control and Communication in the Animal and the Machine*. Cambridge: The MIT Press.

Wilhelm II. 1922. *The Kaiser's Memoirs*. Harper and Brothers.

Williamson, Oliver E. 1979. Transaction-Cost Economics: The Governance of Contractual Relations. *Journal of Law & Economics*, **22**, 233.

Wilson, James. 2002. Scientific Uncertainty, Complex Systems, and the Design of Common-Pool Institutions. *Pages 327–360 of:* Ostrom, Elinor, Dietz, Thomas, Dolsak, Nives, Stern, Paul C., Stonich, Susan, & Weber, Elke U. (eds), *The Drama of the Commons*. Washington D.C.: National Academies Press.

World Bank. 1997. *World Development Report 1997: The State in a Changing World*. Tech. rept. World Bank.

Yolles, Maurice. 2003. The political cybernetics of organisations. *Kybernetes*, **32**(9/10), 1253–1282.

Yolles, Maurice. 2005. Revisiting the political cybernetics of organisations. *Kybernetes*, **34**(5), 617–636.

Young, Oran R. 1989. The Politics of International Regime Formation: Managing Natural Resources and the Environment. *International Organization*, **43**(3), 349–375.

Young, Oran R. 1991. Political Leadership and Regime Formation: On the Development of Institutions in International Society. *International Organization*, **45**(3), 281–308.

Young, Oran R. 2001. Inferences and Indices: Evaluating the Effectiveness of International Environmental Regimes. *Global Environmental Politics*, **1**(1), 99–121.

Zielonka, Jan. 2007. Plurilateral Governance in the Enlarged European Union. *JCMS: Journal of Common Market Studies*, **45**(1), 187–209.

Zimmer, Christina, Schneider, Gerald, & Dobbins, Michael. 2005. The Contested Council: Conflict Dimensions of an Intergovernmental EU Institution. *Political Studies*, **53**(2), 403–422.

Zürn, Michael, & Joerges, Christian. 2005. *Law and governance in postnational Europe: compliance beyond the nation-state.* Cambridge: Cambridge University Press.

Zürn, Michael, Binder, Martin, & Ecker-Ehrhardt, Matthias. 2012. International Authority and Its Politicization. *International Theory*, **4**(1), 69–106.