4

Bureaucracy-Interest Group Interactions in Comparative Perspective

4.1 Introduction

We compare daily. We pick the best apples in the supermarket by comparing them to the rest; we compare restaurants by the quality of their food. Or, we compare different options to commute by evaluating their comfort, speed and anticipated delays. Either consciously or unconsciously, human judgment seems to be fundamentally grounded on comparisons. The same is true for the scientific endeavours we undertake. One of the most powerful empirical strategies, the experiment, is based on a cautious and rigorous controlled comparison of the presence or absence of independent variables and their impact on the dependent variable, i.e. the phenomenon to be explained. The importance of comparison is not only visible in the natural sciences. Social sciences also rely on comparison to a very large extent. Even in philosophy or theoretical politics, comparison is, albeit indirectly, important for conceptual development. Concepts and ideas in theoretical politics are often judged by comparing and contrasting them to the existing body of literature and evaluating how they contribute to contemporary thinking. Comparing thus seems to be a fundamental research strategy.

This study is explicitly designed to be comparative in nature. It is an attempt to develop a theoretical model to systematically compare bureaucracy-interest group interactions and thereby to contribute both conceptually and empirically to the literature on interest group politics and bureaucratic politics.\(^1\) In what follows, I discuss the comparative design I employed in this research, the methods of data collection, and the dataset on bureaucracy-interest group interactions constructed to test the theoretical model developed in chapter 3.

---

\(^1\) I interpret comparative research more as a general strategy of scientific inquiry. In public administration as well as in political science, however, comparative research is also often used to denote a particular type of research, rather than a general research strategy (Lijphart 1971). Interpreted as such, comparative politics or comparative public administration constitutes a sub-discipline of the political science or public administration disciplines, attracting specialised journals and a body of literature tailored to that end.
4.2 The challenges of the comparative method

In scientific research, we compare theoretical propositions with empirical reality or existing theories to rule out rival explanations, thereby lending confirmatory support to the formulated hypotheses. The field of comparative politics relies heavily on Mill’s methods of experimental inquiry, in particular his method of agreement and method of difference (Mill 1970[1843]). When two situations share a similar feature among many different ones, and when they share a similar outcome, according to Mill’s method of agreement, this similar feature is said to result in the outcome that both situations share. Mills method of difference explains the contrary situation. According to method of difference, the absence or presence of the outcome in a given situation is the result of a deviant feature in two otherwise similar cases.

Mill’s methods of agreement and difference were originally designed for experimental research. And they encourage careful and controlled variation of independent and dependent variables in (quasi-)experimental research (see also Cook and Campbell 1979; King, Keohane, and Verba 1994). As said, the discipline of comparative politics heavily draws upon Mill’s methods (see Lijphart 1971; George and Bennet 2005). Given the non-experimental or quasi-experimental nature of comparative politics, however, studies inspired by Mill’s methods of agreement and difference need to address some of the difficulties that arise when one applies an experimental method in a non-experimental environment. Two of the major difficulties, the complexity in establishing causality and the pitfalls of poor equivalence, are addressed below.

4.2.1 Establishing causality

Mill explicitly noted that the methods of agreement and difference were designed for the experimental settings: “It is very seldom that nature affords two instances, of which we can be assured that they stand in this precise relation to one another (...) so a spontaneous experiment of the kind required by the method of difference, is commonly not to be found” (Mill 1970[1843], 256-157). In particular, these two methods are difficult to apply to a setting with what he calls ‘permanent causes,’ that is, causes that cannot be fully excluded or isolated from the environment. The influence the moon exerts on the surface of the earth is a good example. For such situations, he proposed a method of concomitant variation. This method states that when a particular variation of a phenomenon is related to a particular variation of another phenomenon, “[this particular phenomenon] is either a cause or effect of that phenomenon, or is connected through some fact of causation” (Mill 1970[1843], 263). To return to the example of the moon, we can observe that variations in the position of the moon are followed by variations in the time and place of high water. Mill explicitly warned us, however, that there is no way of knowing which of these phenomena are the effect or cause, until we are able

---

* For reasons of convenience, I will use ‘comparative politics’ to denote the distinct comparative research tradition in political science and public administration mentioned in footnote 1.
to ascertain that the variation of a (particular) set of phenomena produces the
particular variation in the dependent variable (Mill 1970[1843], 263).³

It is a not a big step to conclude that, although highly preferred, the methods of
agreement and difference will be difficult to apply to the quasi-experimental
settings we often encounter in political and public administrative research. Indeed,
how can we remove from the scene the Dutch Prime Minister Balkenende and
Minister Rouvoet, both Christian-Democrats, when we want to establish a causal
relation between the regime of such Christian-Democrats and an increasing
scepticism about current practices of abortion and euthanasia in the Netherlands?
We can only establish, if at all, a relation between a particular variation in political
persuasion and a particular variation in the tone of current public and political
debates on such ethical issues. This very point, the fact that isolating or controlling
independent and dependent variables is hard in real-life politics and public
administration, is used by many authors to conclude that, although comparative
research is necessary for establishing empirical relations and hypotheses, it should
be applied in an extremely careful manner (cf. Brady and Collier 2004; King,
Keohane, and Verba 1994; Lijphart 1971; Prezeworski and Teune 1970; Ragin 1987;
2000). The competitive advantage of experimental research is thus clearly the
ability to isolate or control the (in)dependent variables in order to establish
causality. This, alas, is not fully allowed by the quasi-experimental environment
that political science and public administration usually have to offer. One of the
most complex issues scholars thus have to deal with is how to infer causality.

An important impediment to establishing causality in comparative politics is the
multi-level and multi-context nature of political and administrative phenomena.
According to Prezeworski and Teune (1970) in their classic work The Logic of
Comparative Social Inquiry, for instance, the aim of comparative research, and at the
same time its biggest challenge, should be to express characteristics of different
social or political systems in general variables and in measurements designed to
account for such systemic influences. Individual or meso-level relations may differ
between political or social systems at a more aggregate level, such as nation states.
What is it that makes interest groups behave differently in the Netherlands
compared to the United States, for instance? A fundamental challenge facing
comparativists is thus to incorporate systemic effects into their theories on meso-
or micro-level social behaviour. One way to do this is to identify the different types
of characteristics of such social and political systems, according to Prezeworski and
Teune (1970). Yet, this is precisely why comparative public administration is

³ Although these are the methods most often invoked to legitimate the method of comparison that is applied,
they are not the only methods Mill developed. What he calls the ‘indirect method of difference,’ or the ‘joint
method of agreement,’ applies when it is not possible to generate situations in which two instances agree in
every antecedent but one, and differ in the outcome. What we can do in such cases is first examine instances
that agree in the outcome a and factor A, and then search for situations that agree in not having outcome a
and factor A. This joint method of agreement thus constitutes a double application of the method of
agreement, where each application or proof is independent and confirms the other. Another method Mill
suggested is the ‘method of residues.’ According to this method, when we subtract the antecedents from a
phenomenon, which are previously proved to be the cause, the residue of the phenomenon is the result of the
non-examined antecedents. Suppose we have factors A, B, and C followed by a, b, c, and suppose we
established a causal relation between A and a and B and b, then by subtracting a and b from the total outcome
we know that c is the effect of C (Mill, 1970[1843], 258-260).
difficult. As Peters (1990, 10-11) observes: “What we are less certain about, however, is how differences in structures translate directly into differences in the behaviour of members of the organisation, or in the outputs of the organisation.” In order to engage in comparative studies on administrative phenomena, theorising their multi-level nature is, however, prerequisite.4

So, multi-level and multi-context characteristics of administrative phenomena threaten causal inferences. Additionally, the cumulative nature of political processes poses additional problems to causal inferences and comparative research. Concerning interest representation, for instance, community characteristics may influence to some extent an interest group’s success in gaining access to public decision makers. In such processes, the dependent variable in one part of the process, limited access for interest groups, may be an independent variable in another process; for instance, mobilisation (Lowery and Gray 2004). In such cases, isolating variables is even more difficult. The solution may lie in segmentation of the process, such as the ‘influence production process’, and then comparing these segments across different political systems (Lowery, Poppelaars and Berkhout 2008; Pelassy and Dogan 1990). Such a segmentation strategy does not entirely circumvent the problem, as one still has to theorise the fine lines between two segmented parts of the entire process. Yet, by segmenting, we can control for the independent variables that are associated more directly with individual segments.

The very nature of political and administrative processes, i.e. their multi-level, multi-context and cumulative nature, renders it difficult to isolate or control independent variables in such a way that they allow firm inferences about causal relations. A more modest but still challenging aim is thus to establish relations between variations in (sets of) independent variables and the dependent variables, while at the same time incorporating their multi-level or cumulative nature into the theories characterising the relationships that one is to examine. In other words, the challenge is to embed Mill’s method of concomitant variation in a multi-level, multi-context, and cumulative environment of political-administrative reality.

4.2.2 Establishing equivalence

Inferring causality is not the only threat to solid comparative politics. Unconsciously comparing things that are not similar may also be a problem. This phenomenon is known as the travelling problem, which indicates that concepts or measures based on experiences in one society are not necessarily meaningful when applied to other societies. Raising a hand may in one society indicate a way of

---

4 Another, related problem in comparative research is the total number of cases that can be handled and data collection strategies that are pursued in relation to causal statements. Comparative politics as a distinct sub-discipline usually involves a small or intermediate number of cases and has often been applied at the nation-state level (see Lijphart, 1971; for concrete examples and a theoretical discussion, see also Rueschemeyer and Mahoney 2003). Again, many authors have considered the advantages and disadvantages of using case studies in this respect, paying specific attention to causal relations and what kind of strategy to use (see Lijphart 1971; George and Bonnet 2005; Mahoney and Rueschemeyer 2003; Gerring 2006). An important trait that those studies have in common is a focus on paired comparison or structured, focused comparisons, reminiscent of and inspired by Mill’s methods of difference and agreement. A particular difficulty concerning case study designs in establishing causal relations, is the small N these studies usually entail and thus a lack of necessary variation in (in)dependent variables (King, Keohane and Verba 1994; Ragin 1978; 2000).
saying goodbye, while in other countries it may be an offensive gesture. Cooperation may in a democratic society mean working together on an equal basis, whereas in authoritarian societies it could mean a more asymmetrical relationship. In that sense, the travelling problem seems to operate at two levels (Peters 1998, 86-92). Concepts may not only have different connotations or meanings, but measurement and operationalisation may also apply only to specific cases.

Researchers often engage in so-called concept-stretching to remedy the travelling problem. That is, they tend to broaden the concept so that it becomes more inclusive, without considering the range of characteristics that objects must possess to belong to that broadened concept (Sartori 1970). Concepts have often fallen prey to unconscious stretching attempts, resulting in vague concepts that can no longer be falsified. Ideology is a good example of a concept that has been stretched so much that it “never ceases to apply (it has no opposite)” (Sartori 1991, 249). Thus, hypotheses using the concept of ideology are very hard to falsify. When does a particular idea or set of ideas stop being an ideology or begin to constitute an ideology? Seen from this perspective, defining individual or organisational preferences as ideologies (see Tsebelis 2002) does not contribute to sound comparative research and solid falsification. When does someone’s preference stop being a preference when we cannot decide upon the demarcation of an ideology?

A careful and systematic way of defining and organising concepts is needed to address the travelling problem properly. Sartori (1970) developed a specific instrument to do so, a so-called ladder of abstraction. Its basic rule is that the connotation of a word (the collection of properties determining the things to which the word applies) and its denotation (the class of things the word applies to) are inversely related. So, to increase comparability, one should reduce its characteristics, and thus at the same time enlarge the class of things the concept applies to. Conversely, to make a concept more concrete and thus less comparable, one adds properties or characteristics (Sartori 1970; 1991). When one moves consciously up and down this ladder of abstraction, concepts and terms remain related to one another. These different levels of abstraction should be clearly related to each other, resulting in two pyramidal shapes that are inversely related (see figure 4.1). The further one climbs the connotation pyramid, the further one descends on the denotation pyramid.

Essentially, Sartori’s strategy is about determining the absolutely vital elements of concepts and distinguishing them from those that are secondary or peripheral.
As he (Sartori 1984, 32) puts it, we must focus on “defining as opposed to accompanying properties.” Put differently, the challenge in comparative politics is thus to find the appropriate intersection of the connotation and denotation pyramid. In an early, previously unpublished, manuscript, part of which is now published in *West European Politics* (Mair 2005), Sartori offered an example of how this ladder of abstraction works. In the process of building a classification of political parties, he first identified a core analytical concept of political parties, i.e. organisational networks, and defined three functions central to such organisational networks: participation, electioneering, and expression. While political parties exhibit many other functions, these three are irreplaceable and apply to political parties throughout the Western world (Sartori 2005[1967]). But only the combination of the three denotes political parties. When we remove electioneering, the two remaining characteristics could denote any type of civil society organisation, NGO or ad-hoc protest organisation. And, when leaving out both participation and expression, it could denote such people as Barack Obama, John McCain, Jan-Peter Balkenende, or Wouter Bos running for president or prime minister. This example shows that political parties are somewhere in between the connotation and denotation pyramid shown above, and that it is often the exact combination of characteristics that defines a concept that may travel well.

An important principle underlying Sartori’s strategy to systematically unravel concepts is the idea that categories are hierarchically related. This means that some concepts are subordinated to others. This taxonomy, or hierarchical classification, can mathematically be represented as a tree diagram. Ultimately, all concepts, except for the ones down at the bottom of the classification, have at least one element or one category in common. An often-used example of such a taxonomic hierarchy is the scientific classification of organisms. On top of the tree diagram stands one particular category of organisms, for instance, mammals. A subcategory of mammals is big cats, a subcategory of big cats is tigers, and finally a Siberian tiger is a specific sub-species differing in its colour and striped pattern.

Moving up and down the connotation ladder (the collection of characteristics of organism) and denotation ladder (the class an organism belongs to) is relatively easy for such hierarchically organised concepts. Social categories are, however, not so neatly organised. Corporatism, for instance, is often defined by listing a set of commonalities, but not every commonality is apparent in countries classified as being corporatist (Collier and Mahon 1993, 846-849). Corporatism, in Collier and Mahon’s (1993) words, is a type of family resemblance. Family resemblance occurs when cases each exhibit different combinations of commonalities without sharing one common feature. Another option discussed by these two authors is the idea of radial categories. In this case, secondary categories split up the original category rather than together constituting the overall concept. The authors give the example of democracy; many types have been constructed: liberal democracy, participatory democracy and popular democracy, for instance. They distinguish different types of democracy instead of a sub-type of an overall category of democracy.

---

5 Several scholars have built on Sartori’s strategy to define appropriate strategies to engage in comparative research. Pelassy and Dogan’s (1990) ‘functional equivalence’, for instance, is but one example. Functional equivalence refers to the idea that political systems necessarily fulfil certain fundamental tasks; yet these functions may appear in different guises in different nations. The task then is to define and compare those functions instead of their appearance. Functional equivalence thus refers to those elements of concepts that are, in Satori’s words ‘irreplaceable.’

6 Corporatism, in Collier and Mahon’s (1993) words, is a type of family resemblance. Family resemblance occurs when cases each exhibit different combinations of commonalities without sharing one common feature. Another option discussed by these two authors is the idea of radial categories. In this case, secondary categories split up the original category rather than together constituting the overall concept. The authors give the example of democracy; many types have been constructed: liberal democracy, participatory democracy and popular democracy, for instance. They distinguish different types of democracy instead of a sub-type of an overall category of democracy.
(2003) min-max strategy presents an alternative way to systematically organise the concepts of the phenomenon in question without necessarily relying on the hierarchical assumption that is so apparent in Sartori’s organising structure. Their min-max strategy is a conjoined use of minimal and ideal-type definitions of the concept. Minimal definitions, they argue, contain the ‘bare essentials’ of a concept, are minimal in their attributes and maximal in their phenomenal range. Ideal types include all attributes that together define the concept, and are thus minimal in their phenomenal range (Gerring and Barresi 2003, 201). Both theorising about and systematically organising concepts are thus a matter of distinguishing what is really important about an entity (Goertz 2006). From a comparative perspective, it is not only a matter of what are the important and defining characteristics of a concept, but also of finding the appropriate level of abstraction that both allows falsification and travelling to other political-administrative systems.

Comparative research is important. Poor conceptualisation and overestimating the possibility of establishing causal relations, so, not paying full attention to the complex multi-level and multi-contextual nature of political phenomena seriously threaten sound comparative research in political science and public administration.

4.2.3 Why the resource dependence model does the comparative job

A reconceptualisation of bureaucracy-interest group interactions in terms of resource exchange has a strong potential to meet the demands of comparative research. First, it addresses the conceptualisation issue by defining the characterising elements of bureaucracy-interest group interactions: an exchange of resources under the assumption that bureaucrats are not able to achieve full independence because of the responsive and democratic nature of public administration. This conceptualisation enables us to move beyond different connotations of cooperation or asymmetrical contacts such as capture. Moreover, such resource exchanges can still be falsified. If no resource exchange takes place and there are still interactions, the model will be falsified.

Second, a reconceptualisation in terms of resource exchange addresses the multi-level and multi-contextual nature of bureaucracy-interest group interactions, given a careful consideration of each contextual factor that may affect the individual elements constituting the degree of dependence that characterises those interactions. The dependence model thus tries to incorporate systemic characteristics of the administrative-political environment in which these interactions take place. In doing so, this model attempts to contribute to comparative research on bureaucracy-interest group interactions by explicitly addressing the conceptual difficulties the travelling problem poses by a conscious reconceptualisation. In addition, it incorporates the multi-level and multi-contextual nature of the political-administrative context in which bureaucracy-interest group interactions occur. And, by systematically varying contextual factors, it aims to establish relationships between variations in a set of independent variables and variation in the dependent variable. It is thus an attempt to embed the method of concomitant variation in a multi-level and multi-contextual system.

7 These defining characteristics only hold for interactions in democratic societies, and would need refinement if applied to non-democratic countries.
4.3 A dataset of bureaucracy-interest group interactions

The proof of the pudding is in the eating, and this also applies to the resource dependence model of bureaucracy-interest group interactions. One cannot properly judge a theoretical model without testing it. And most models will need at least some refining based on the findings such analyses generate. Interestingly, both testing and a dialogue back and forth between theory and the data (or, more generally, empirics) are often considered part of two distinct research traditions. The chasm between quantitative and qualitative researchers in public administration and political science reflects a longstanding fundamental debate between positivists and hermeneutics. Put differently, scholarly debates about the question whether we should deduce relationships from theories or construct relationships more inductively often end in stalemate. The argument is essentially about weighing the potential of explanation and understanding in social sciences. In other words, is it better to deduce hypotheses and try to explain (and perhaps even predict) reality based on general patterns of behaviour? Or, is it better to understand the idiosyncrasies of reality by interpreting them based on theories, and then derive patterns of behaviour from the very same idiosyncrasies? And, as in any debate, there are believers and non-believers. Deductive versus inductive approaches in social sciences are even today often juxtaposed as two equally valid yet unbridgeable approaches to knowing.8

This research explicitly proceeds from the proposition that these two traditions should not be juxtaposed, but should supplement each other to genuinely contribute to scientific research. That is, posing and answering research questions rather than methodology as such should lie at the heart of social sciences. This study relies on testing theories and thus proceeds from a positivist perspective. In this way, we can evaluate the theoretical model that has been developed. Yet, it also relies on more interpretative ways of conducting research, both in developing the theoretical model and in suggesting avenues for future research. In general, this study is firmly based on the assumption that answering research questions requires a conscious consideration of, and preferably a combination of, several methods, but at a minimum an objective attitude must be brought to the entire array of existing methods. All in all, this study is grounded in a pluralist philosophy of science that will be reflected in a strategy of mixed methods.

4.3.1 Data collection strategy: online survey

A theoretical model that is specifically designed to engage in comparative research requires a data-collection strategy that meets such demands at an operational level. To collect data on the different elements constituting the degree of dependence between bureaucrats and interest groups, I opted for a cross-sectional online survey instrument. Other instruments or strategies to collect data were less suitable for this particular topic. Information on interactions between bureaucrats and interest groups are not readily observable, so relying on existing databases is not useful here. And if there are such data available, they only contain part of what this model

---

8 The relatively recent trend to mixed methods (Bennet and Braumoeller 2005) is a fruitful attempt to bridge the gap.
tries to explain (Carpenter 2004; Yackee and Yackee 2006). Obtaining the data directly from bureaucrats and interest groups, the subjects under study, seemed therefore the most appropriate choice. In addition, as the aim was to test the theory, I needed a large N to perform the analyses in order to provide statistical leverage, as most of the variables cited in the model are included in the analysis. Although best fitted for the purpose of this study, cross-sectional (online) surveys have some drawbacks that need to be taken into account before they can result in a solid database. Apart from the general cautionary tales regarding the reliability and validity of employing a survey instrument, which is often associated with a low response rate (Fowler 2002; Groves et al. 2004; Litwin 1995), adding online and cross-sectional elements requires perhaps even more caution. Survey research in comparative politics has a long tradition and has contributed to the development of standardised data sets and replicable procedures to test hypotheses on political behaviour (Rokkan 1969). Currently, almost a quarter of all articles and half of the quantitative articles use sample surveys (King et al. 2001).9

True development in theory building in comparative political and administrative behaviour, however, lies in a contextual approach to comparative surveys. Traditional survey research has focused on the individual, assuming that a person’s behaviour stems from his or her preferences and experiences. But a contextual approach seems necessary to develop theories on political behaviour, as such behaviour will undeniably be influenced by the institutional setting in which it occurs (Verba 1969). Such an assumption was also reflected in the multi-level approach of Prezeworski and Teune (1971) to comparative research in general. Comparative surveys are, however, notorious for biased results (King et al. 2004) resulting from two main thresholds: conceptualisation and measurement techniques (King et al. 2004; Verba 1969). Bias, as a result of conceptualisation, stems from unconscious concept formation, which I discussed in section 4.2. Bias as a result of measurement techniques refers to the equivalence of individual questionnaire items and whether respondents interpret questions similarly. Such measurement problems thus relate to equivalence issues, which conceptualisation problems refer to on a theoretical level.

Although the interactive character and options to control question ordering are obvious advantages of online surveys as opposed to mail surveys, two drawbacks are worth mentioning here. Coverage error and the assumption of internet familiarity could result in significant bias (Orr 2005; Winne de, Plevoets, and Sels 2003). If only a certain proportion of the population has access to the internet, together with the fact that only people who are more familiar with the internet will be likely to fill out such a questionnaire, an online survey significantly bias the results. In addition, the possibility of multiple responses, when an individual respondent fills out the questionnaire twice or more, may cause the results to be unreliable. Although the latter issue can be relatively easily addressed, the former two are serious drawbacks that should be taken into account when online surveys are selected as a data collection method.

---

9 These results come from a content analysis conducted by the authors of five years (1991-1997) of the American Political Science Review, the American Journal of Political Science, and the British Journal of Political Science (see footnote 1 in King et al. 2001).
In sum, there are several major challenges in using cross-sectional online surveys. Similar interpretation of questionnaire items and equivalence of concepts are challenges to cross-sectional surveys in general. In addition, online surveys add to this a potential sampling bias as a result of overrepresentation of those respondents who are more familiar with or have access to the internet. And, in general, surveys often result in high non-response or high non-item response. In what follows, I outline how I take these issues into account in developing comparative surveys for bureaucrats and interest groups so as to use them as solid data collection instruments.

4.3.2 Surveying bureaucrats

Who is a bureaucrat is often crystal clear to those outside the public bureaucracy. ‘They have a 9-5 job and drink a lot of coffee,’ seems an appropriate summary of a popular interpretation of ‘the bureaucrat.’ Those working inside public agencies or studying them, however, often point to considerable differences within bureaucracies, not to mention the differences between bureaucracies in different countries. Yet, for comparative research to produce solid findings, we have to establish at least some commonalities or equivalences between what we popularly call the bureaucracy. This section explains the rationale behind the definition of bureaucrats used in this study, and discusses the strategy employed to establish a population and to apply an appropriate sampling method.

Defining bureaucrats

There seems to be a scholarly consensus that public agencies and civil service systems are more alike across different environments than any other part of the political-administrative system. If we proceed from the Weberian ideal type of bureaucracy, there would indeed be very few differences. As Peters (1990, 9) suggests: “When one knocks on the door of bureaucratic offices, one is likely to find more similarities, than when knocking on the doors of legislators or even chief executives.” Recent reform trends also point to similarities or even convergence of national civil service systems (OECD 2003). However, those adhering to the parable of six blind men touching a giant African elephant are perhaps more likely to emphasise the vast differences among state agencies (Lowery and Brasher 2004, 197). Every blind person would indicate that they are touching a different animal, as they cannot imagine its overall shape when touching only its trunk or one of its legs. Comparative studies of national civil service systems indeed suggest that there are significant differences among them (Bekke, Perry, and Toonen 1996). Bringing in meso-level considerations further threatens any comparable endeavour, as the number of such differences may exponentially grow.

To steer a middle course between full variation in all public agencies and imposing overly exact similarities by too narrow sampling, I selected only a particular division within national bureaucracies to include in this research: senior civil services. Many countries have formally organised civil services, including an upper echelon, which I define as the senior civil service. The senior civil service includes the civil servants with highest rankings both in terms of salary and function. As such, these senior civil servants perform equivalent functions in terms of management and often have similar strategic positions within their
organisations (OECD 2003). The comparison in this research thus concerns national senior civil services rather than entire national civil service systems. Selecting senior civil services addresses the equivalence issue, as they are comparable across different national civil service systems.

A drawback of such a selection is, however, that it could result in a bias towards strategic interactions while ignoring other types of exchanges taking place. Strategic interactions are more likely to occur between senior civil servants and interest groups, as the first usually have important strategic positions in terms of policy advice. Moreover, they often have influential positions within the civil service system. Middle or lower-ranking civil servants are assumed to have interactions with interest organisations as well, but are more likely to exchange other types of resources. Their interactions are more likely to concern the substance of particular bills rather than the political strategy behind a set of similar issues or bills. As a lobbyist of a Dutch interest group puts it: “Civil servants that are truly important for us are involved with multiple issues at the same time.” Or, as there may be a lower threshold for accessing middle and lower-level civil servants, this selection could also result in an under-representation of bureaucracy-interest group interactions. I addressed this issue by including the answer option of “junior employees (in Dutch, beleidsambtenaren) usually interact with interest organisations” in the questionnaire (item 4, see appendix I). This option was seldom chosen. Only 8.4 per cent of all respondents indicated that they did not interact with interest groups. Five of them replied that junior employees usually interacted with interest groups. All in all, comparing senior civil services meets the demands of equivalence. Yet, it may result in bias in the type and number of interactions that take place with civil servants.

Establishing a population of senior civil servants

The appropriate population of civil servants consists of those officials belonging to a country’s senior civil service. Comparing senior civil services thus automatically results in a demarcation of the population. I used existing institutional configurations of senior civil services in each country to define the population. For the Netherlands, the population consisted of civil servants belonging to the Dutch senior civil service (Algemene Bestuursdienst (ABD)). The UK population consists of civil servants belonging to the Senior Civil Service (SCS).12

The selection of senior civil servants also controls for different levels of government. Namely, by selecting the senior civil services, I opt for a comparison at a national government level. Some indeed argue that differences exist in types of interactions between different levels of government. Local governments, for instance, would entail much smaller bureaucracies and, therefore, be more responsive to their environment or have a greater flexibility in their response (Scholz and Wei 1986). However, analysing such interactions in the case of immigrant integration policy in the Netherlands also reveals similarities across different levels of government, because the need for certain resources appears to be the same (Poppelbaars 2007). If different levels of government were treated as a full explanatory variable, this variable would become intertwined with one of the explanatory variables in the model, namely the character of national interest group systems. In some cases, interest group representation and involvement could be different at local and regional levels. When including other levels of government, it will become hard to keep other variables constant while examining the level of government. Therefore, the analysis will be limited to the national level of government.

Interview by author.

To ensure a proper mix between advisory agencies and executive agencies (see chapter 5), I also included the Dutch civil-service candidates who have been selected as persons who could (potentially) fulfil a senior
Drawing a random sample from these two populations was hindered by one of the drawbacks of sample surveys: a notoriously low response rate. The initial strategy was to randomly sample both populations. In order to ensure a sufficiently high response rate, given a high non-response, I opted for the following strategy. I sent questionnaires to the total population of Dutch senior civil servants (N = 996). In the case of the UK, I first randomly sampled 1,147 from a total of 3,306 senior civil servants at the time of conducting the initial surveys (based on Dod’s Civil Service Companion 2007). As the initial response rate was very low, I decided to submit questionnaires to the remainder of the SCS population listed in Dod’s Civil Service Companion as well. Given the low response rates, both anticipated and actually experienced, this was at that time the best strategy available. On the other hand, studying the entire population is a strength as well. Nevertheless, I will take this issue into account in the interpretation and conclusions of this study.

4.3.3 Surveying interest groups

Compared to the complexity of equivalence in comparative interest group research, the issues encountered in comparative research on civil service systems seem relatively easy. Quite ironically perhaps, for a study emphasising equivalence, a major and crucial difference among countries traditionally denoted as pluralist and corporatist, is the availability of data on interest group populations. In countries usually identified as neo-corporatist, there are no databases available containing information on individual interest groups, whereas in neo-pluralist systems such databases commonly exist. This difference posed a major challenge in sampling respondents for the interest group survey. Furthermore, there is a general agreement within the interest group scholarly literature that a common definition of interest groups does not exist. In this section, I will address both issues to justify the sampling techniques used for conducting the interest group surveys.

civil service position relatively soon (www.algemenebestuursdienst.nl). For the UK, level 6/7 civil servants were also included in the survey. For the US, civil servants from the Senior Executive Service (SES) were included. More specifically, the population of senior civil servants in the US was defined as those belonging to the SES, who are career employees and whose position is a competitive one, other than the non-career SES members who receive non-competitive positions. This selection was supposed to provide a better equivalence between the nature of the Dutch and US senior civil services (Federal Yellow Book 2007, fall and personal communication with dr. F. M. van der Meer, Leiden University, Department of Public Administration). For the Swedish case, I used the civil servants denoted as senior civil servants in the Swedish State Calendar. Although there is not a formal institutionalisation of a senior civil service, based on existing classifications (Page and Wright 2001) and the listing in the Swedish State Calendar, I could define a population of senior civil servants as equivalent as possible to those of the UK, the US and the Netherlands.

When the senior civil servant’s survey in the UK and the Netherlands is discussed, ‘I’ should be interpreted as ‘we’ as I conducted these surveys together with Caspar van den Berg, a then colleague at the Department of Public Administration, Leiden University. As two populations in our individual research projects overlapped, we jointly conducted the survey.


For the US and the Swedish cases, the sampling methods were as follows. From the Swedish State Calendar, I selected from each national agency the two most senior civil servants, from whom I tried to obtain contact details. I submitted a survey to those from whom I had been able to obtain the email addresses. For the US, I took a random sample of SES members to establish a sample for the survey.
A strategy to define an indefinable population

Despite the need for comparative research to truly assess the nature of individual interest representation systems, there is considerable disagreement and doubt about the feasibility of comparative research in this field. Quite fundamentally, there is no real consensus about what interest groups, the subject in question, exactly are. Part of the fuzziness seems related to the very concepts employed to designate the subjects under scrutiny. “Counting is not the issue”, according to Jordan (2006, 2), “it is categorisation that is the problem.” A listing of existing definitions by Baumgartner and Leech (1998, 29) aptly illustrates Jordan’s conclusion. According to the literature, interest groups are:

- Social or demographic categories of the population; membership organisations;
- any set of individuals with similar beliefs, identifications, or interests; social movements; lobbyists registered in legislatures; political action committees, participants in rule-making or legislative hearings; institutions, including corporations and government agencies; coalitions of organisations and institutions; prominent individuals acting as political entrepreneurs or lobbyists.

This is a colourful collection of definitions containing a wide variety of potential interest groups. What are we to make of such a collection, if one is to engage in comparative research? Most of these concepts refer to specific organisational configurations. What all these loosely (or very formally) organised entities have in common is that they intersect with government actions when they pursue their private interests. When governments adopt certain measures or rules affecting private interests, interest groups will try to influence these measures or regulations through elected and/or appointed government officials (Baumgartner and Leech 1998; Lowery and Brasher 2004; Salisbury 1984). So, what these entities have in common are two fundamental aspects, irrespective of whether they are an ad hoc protest group, a full-fledged peak labour organisation or something else in between. First, they pursue their private interests. And second, in doing so, they intersect with public policy or regulations. Theoretical issues determine which of these two aspects will dominate in a given research project. The aspect of pursuing private interests will merit more attention in studies of mobilisation, maintenance and population studies of interest groups, while the intersection with public policy may be more pronounced in studies of interest groups’ influence, and strategies.

A definition of interest groups in functional terms meets the conceptual demands necessary for comparative research. In Sartori’s (1970) words, we “move up the ladder of abstraction.” A definition in functional terms also eliminates the idiosyncrasies associated with various individual types of interest groups that prohibit any meaningful comparison. When this strategy is consistently applied, groups of any type of organisational configuration that intersect with public policy should be referred to as ‘interest-pursuing entities.’ For reasons of convenience, I will continue to use the term ‘interest groups,’ while keeping in mind that it is not the specific organisational configuration that is of importance, but it is rather the activity of pursuing one’s interests.16

---

16 A complicating factor in this definition may be that the extent to which such ‘interest-pursuing entities’ are successful in pursuing private interests relates to the organisational characteristics of these entities. In other
Chapter 4

How to sample from an unknown population?
Not only do differences in definitions seem to exist, but available data on interest groups differs markedly between the two broad sets of interest representation regimes. In the UK and US, databases of interest group populations are available. In contrast, in the Netherlands and Sweden, such comprehensive databases do not exist. Two fundamental questions arise in this case. What should we sample from and what should we compare? For the Dutch case, I developed a strategy to deal with the non-existing database problem.\(^7\)

There is no census of the existing interest groups in the Netherlands from which to draw a random sample. The important or ‘big players’ among the interest groups are often well-known to policy makers in the Netherlands. Beyond this familiar collection of interest groups, however, it is hard to get an overview of other relevant interest groups. Recent studies of membership organisations in the Netherlands also restrict themselves to the largest ones (de Hart 2005), concern sector specific studies (Akkerman 2005; Huitema 2005), or study a particular type of organisations, such as professional associations (Visser and Wilts 2006). Such studies do not accumulate and thus fail to provide a full overview of the interest group population. As a civil servant noted: “When there are problems with certain health and safety rules in a particular profession, we often meet interest groups that were completely unfamiliar to us before.”\(^8\) In general, there is no representative indication of the existing interest groups, outside those that regularly knock on policy makers’ doors. The scholarly literature on interest representation in the Netherlands very much reflects this political-administrative practice. Put differently, the literature does not provide a full overview of the existing interest groups in the Netherlands either.

By studying only those organisations that already have access, we cannot draw conclusions about the nature of interest representation. We do not know how many others there are, so we do not know (for instance) how severely access is limited. Conceptual models that explicitly draw attention to this issue thus cannot be fully applied. The question, then, is how to measure resource concentration when the boundaries of the interest population are unknown. When applying a resource dependence model, we need to include population dynamics, which require us to be aware of the total interest population (see chapter 6). Moreover, comparative research is further complicated by the fact that we cannot compare different systems when there are no comparably representative estimates of the interest populations available from which we can sample in an equivalent manner.

words, their organisational configuration may be a vehicle to a successful intersection with public policy and thus access. So, by excluding this, we might omit an important intervening variable. Still, such organisational characteristics are not an essential or fundamental characteristic of interest groups. Pursuing private interests and seeking access when interests intersect with public policy are, similar to Sartori’s definition of political parties, the bare essentials of this concept. In explaining differences in success, we should move down the ladder of abstraction and find parameters covering, for instance, organisational characteristics.

\(^7\) For Sweden, a similar full-blown strategy would have been appropriate, but given a lack of time and resources I opted for a satisfying rather than optimising strategy: selecting the organisations reported in the Swedish State Calendar and via an additional, snowball sampling web search adding other major interest organisations to establish a dataset of at least the major Swedish interest groups.

\(^8\) Interview by author.
Indeed, there are no legal requirements for interest groups to register in the Netherlands when they want to consult or contact political parties in parliament, ministers or civil servants. Such requirements to register not only contribute to more transparency in public decision making, but also provide scholars, for instance in the US, with a fairly satisfactory start for collecting data on interest groups (Baumgartner and Leech 1998). And whereas the multiple databases available on EU interest groups generate their own methodological and substantive problems, at least there is some census of the EU interest system (Berkhout and Lowery 2008). These databases allow researchers to define in a relatively satisfactory way a population from which to sample and eventually infer some general conclusions about interest populations at large. A lack of such a database thus results in a fundamental sampling issue when studying interest representation in the Netherlands. Indeed, how are we to generalise findings and results on the national interest population, if we are not familiar with the characteristics and size of this population?

Two databases have been constructed to provide a representative dataset for the Dutch national interest group population. One includes interest groups that are not necessarily known to policy makers and the other one includes those that are presumably familiar to policy makers. The two together provide an estimate of the entire national interest population.

Let me discuss the database with interest groups not necessarily known to the government first. To establish such a database, I used the association and foundation database generated by the National Chamber of Commerce in the Netherlands (NCC). The NCC is an autonomous public agency (in Dutch: zelfstandig bestuursorgaan (zbo)) under the auspices of the Ministry of Economic Affairs. The NCC not only administers the trade register, but also administers an association and foundation register (Register Act 1997). Essentially, every citizen in the Netherlands either planning to establish his or her own company or start any type of foundation or association needs to register with the NCC. These trade, association and foundation registers are available online and accessible via a NCC coding system, which provides codes for every type of firm or association. As all interest groups are formally registered as either a foundation or association, this database provides a good starting point to establish an estimate of the national interest population. To select candidate organisations, I searched the NCC database via their coding system. Based on classifications used by other scholars (Beyers and Kerremans 2004; Mahoney 2004), I categorised different types of interest groups to guide the selection of appropriate NCC codes (see appendix II).

Searching the NCC register on the basis of these codes resulted in an initial database of 58,220 organisations. However, two caveats needed to be addressed before finalising the dataset. First, as the NCC coding system is not mutually exclusive, duplicate organisations needed to be eliminated. Second, the NCC coding system only registers the location, name and type of organisations rather than its scope of activity. It was therefore necessary to distinguish between national and local organisations, as I was only interested in interest groups operating at the national level. In order to make this distinction, I manually searched the initial database of 58,220 organisations by organisations’ names to distinguish the
national from local ones.\textsuperscript{19} Finally, as even this coding system turned out to be incomplete, I checked the database against an overview of the largest interest groups (de Hart 2005), and, whenever necessary, I added missing organisations to the database. This strategy eventually resulted in a database of 7,565 organisations.\textsuperscript{20} From these 7,565 organisations, I took a random sample of 1000 organisations and gathered contact details for each via an internet search and by phone. This information was also used to ensure that organisations had a national scope of activities and to check whether they still existed.

Second, to account for interest groups that were already familiar to policy makers, I examined websites from each central government department and from the organisations that are members of the Social and Economic Council. This did not, of course, provide a full overview of the collection of interest groups with which the government is familiar. Policy makers are not truly familiar with all of the groups, and some with which policy makers are familiar are still excluded. Assuming that familiar and non-familiar organisations are missing at random, this website database provides a fairly good estimate of the characteristics of the interest population known to the government rather than an estimate of the total number of organisations with which the government is familiar.

The two diagrams in figure 4.2 depict the results of the database searches. They show the different characteristics of each population based on random samples from both datasets. That is, they show the proportion of each type of interest group in the random samples drawn from the two datasets: the one containing organisations known by the government (A), and one not necessarily known by the government (B). I checked both samples to assess whether the organisations indeed had a national scope, still existed, or were international organisations. The NCC sample contained 507 national organisations and the website-based sample contained 336 national organisations. The results in figures 4.2 are based upon the number of national organisations, rather than the full sample.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.2.png}
\caption{Interest groups in the Netherlands A (website database)}
\end{figure}

\textsuperscript{19} In Dutch, I coded organisations with the following words in its title as possibly national interest organisations: Nederlandse, nationale, landelijke, vereniging van/verenigde, koepel, platform.
\textsuperscript{20} This database excludes individual corporations as well as advisory councils.
Figures 4.2 The population of interest groups in the Netherlands

Two important aspects of these datasets need attention. First, the sample based on the Chamber of Commerce database contained 317 organisations that were untraceable. They either ceased to exist or could simply not be found via the internet, a phone book, or the yellow pages. The NCC database provided indeed a comprehensive start for establishing a database of national interest organisations, as it registers every type of organisational activity. Yet, it does not provide a complete and up-to-date registration, indicating that the database is not entirely reliable. Second, the website database most probably overestimates the number of NGOs, in particular development and environmental NGOs. For this dataset, I used snowball sampling. Via the websites of the central government departments and peak organisations, I collected organisations and member organisations respectively. The association of development organisations provided a separate, comprehensive website of its member organisations. The listing of the member organisations of the two professional organisations was not complete, however. This strategy may have resulted in an overrepresentation of development NGOs. Yet, these two datasets together provide a comprehensive view on the Dutch interest group population. As such, they offer a fairly equivalent database to those in pluralist systems. In addition, these datasets can provide more information on interest representation in the Netherlands than the case studies and the studies of particular types of interest groups have together offered until now.

I used these databases to take random samples, which are illustrated in figure 4.3 above. For the UK, I used the *Directory of British Associations, Edition 6* (CBD Research 2006). This directory contains a database of British and Irish associations, including 7,368 British associations and their contact details. Some, however, question its completeness and representativeness. Yet, for the purpose of this research, it was the best available option to conduct a comparative survey.

---

*CBDS 2006, Directory of British Associations and associations in Ireland, CD-ROM version, CBD research, Beckenham, Kent. I thank Prof. dr. W. Maloney for his suggestion and reference to this database.

*Research project of Prof. dr. G. Jordan and Prof. W. Maloney: presented at the ESRC Conference on the Politics of Interest Representation: Counting interest groups, Aberdeen University, 02-07-2007.*
From those 7,368 organisations, I randomly sampled 977 organisations. For organisations without current email addresses, I called to ask the email address of the director, chief executive, or general secretary. After the database search and the telephone calls, I was able to collect contact details for 879 organisations, which included either personal or personal assistant’s email addresses.

### 4.4 Assessing the quality of the survey dataset

Based on the surveys submitted to the random samples of bureaucrats and interest groups, I established a dataset of bureaucracy-interest group interactions. To measure each of the theoretical concepts of resource importance and resource concentration, questionnaire items were designed. In the following chapters, I will discuss the operationalisation and measurement of each of the theoretical concepts in more detail, including the individual questionnaire items. I sent the questionnaires to the respondents via an online web tool, which allowed me to fully administer the surveys myself. The questionnaires were sent via an email containing a unique link to the questionnaire. The surveys were conducted between January 2007 and March 2008.

These surveys resulted in a database of bureaucracy–interest group interactions mainly from the UK and the Netherlands. Surveys were also sent to Sweden and the US. This particular set of countries was selected as they provided for variation in interest representation regimes (for a detailed discussion, see chapter 5). Table 4.1 shows the responses in total N.

<table>
<thead>
<tr>
<th>Nations</th>
<th>Bureaucrats</th>
<th>Interest groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Netherlands</td>
<td>418</td>
<td>222</td>
</tr>
<tr>
<td>UK</td>
<td>406</td>
<td>133</td>
</tr>
<tr>
<td>US</td>
<td>82</td>
<td>38</td>
</tr>
<tr>
<td>Sweden</td>
<td>119</td>
<td>49</td>
</tr>
</tbody>
</table>

This was done in two ‘rounds.’ I first randomly sampled 609 organisations, but as the response rate turned out to be very low, I additionally sampled 368 organisations to increase the response rate, as several reminders had not helped to increase the response rate of the first round.

For the US, I used the database of interest groups based on the Lobbying Disclosure Act (see Baumgartner and Leech 1998) and selected a list of interest groups and lobbyist via the fields: issue codes; government entity contacted and year filed. For each issue, I selected the first 100 organisations that were listed after the query based on three fields: policy area (issue), Congress (government entity contacted) and 2007 (year filed). For the Sweden case, I listed the organisations included in the Swedish State Calendar and conducted snowball sampling via a website search.

See chapters 5-7 for further discussion and appendix I for a full list of questions.

See: www.zipsurvey.com; I thank Dr. T. Capelos for this very useful suggestion.

I used a general email address (for the ABD survey: topambtenaren@fsw.leidenuniv.nl; for the UK, US and Sweden SCS surveys and the NLD, UK, US and Sweden interest group surveys I used: governance@fsw.leidenuniv.nl) to which respondents could respond and provide us comments on the questionnaire. Such a reply option allowed me to detect possible technical malfunctions or other questions regarding the survey, and prevented multiple responses from a single respondent. As a result, these options effectively addressed some of the drawbacks of online survey
Whereas in absolute numbers the bureaucrats’ datasets of the UK and the Netherlands are comparable, the interest group dataset of the Netherlands is larger than that of the UK. And the total N generated by the Sweden and US surveys, both for the bureaucrats and the interest groups, is much smaller than the total N generated by the UK and Dutch surveys. Therefore, the Sweden and the US datasets are not used in this study. The UK and the Dutch datasets are, however, sufficient to assess the model’s explanatory value. I can only use these datasets to measure the degree of dependence characterising bureaucracy-interest group interactions, however, when they meet the demands of reliability and validity. Below, I discuss these issues in turn.

### 4.4.1 Assessing the reliability of the questionnaires

One of the measures of the robustness of a dataset is reliability. Reliability in survey research often hinges on how consistently respondents have answered the questions, how well they have interpreted the questions, whether they have interpreted the questions in a similar way, and, finally, whether or not there is a bias resulting from a high non-item response. Testing the reliability of a questionnaire can be done in several ways. An obvious method is the test-retest method, since providing the same item twice to the same respondent with some period of time in-between is an obvious measure of consistency. In the case of multi-item measurements, checking the internal consistency of the respective items is an additional way to measure the reliability of a questionnaire (Fowler 2002; Groves et al. 2004). The questionnaires contained both single item and multiple item measurements. So, the obvious way to test reliability was the test-retest method, but that was not possible given practical limitations. Although the questionnaires did not contain true parallel form or multi-item measures, several items related to similar topics were included. Therefore, to gauge the reliability of the questionnaires, inter-item correlations (de Vaus 2002) were examined.

For the bureaucrats’ dataset, I chose two items that were designed to measure interactions with interest groups and two items on the influence of the EU on interactions with national interest groups. The internal consistency of the interest group dataset was also measured for items related to the importance of the EU and importance of civil servants for getting access to the policy making process. The inter-item correlations vary from 0.38 to 0.82, indicating a satisfactory internal consistency. Table 4.2 shows the inter-item correlations compared between the UK and the Netherlands for both the interest group and bureaucrats’ datasets.

---

28 The number of questions in a questionnaire is limited, as a questionnaire should not take too long to complete. This helps to increase both the response rate and reliability (the longer it takes, the less people will be concentrating at the end of the questionnaire). For this particular research project, the questionnaire contained items measuring various aspects of interactions with interest organisations as well items measuring the Europeanisation of various competencies and activities of senior civil servants. These items were part of the empirical research of C. van den Berg (see footnote 13). As a consequence, the number of questions for our individual research projects was restricted.
Chapter 4

Table 4.2 Inter-item correlations questionnaire-items

<table>
<thead>
<tr>
<th>Datasets</th>
<th>Inter-item correlations</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UK</td>
<td>NLD</td>
</tr>
<tr>
<td>Bureaucracy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interactions interest groups</td>
<td>0.83</td>
<td>0.82</td>
</tr>
<tr>
<td>Influence EU on interactions</td>
<td>0.69</td>
<td>0.53</td>
</tr>
<tr>
<td>Interest Groups</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Importance civil servants</td>
<td>0.44</td>
<td>0.50</td>
</tr>
<tr>
<td>Importance EU</td>
<td>0.38</td>
<td>0.51</td>
</tr>
</tbody>
</table>

The inter-item correlations for the bureaucrats’ datasets are generally higher than for the interest group datasets. Whereas the values for the bureaucrats’ dataset vary from 0.53-0.83, the values of the interest group dataset vary from 0.38-0.50. Possibly, this difference is due to variance in uniformity. A dataset of senior civil servants may be less heterogeneous than a dataset of representatives of interest groups. Accordingly, this heterogeneity may be reflected in a lower internal consistency of the interest group dataset. For both questionnaires, the inter-item correlations for the EU are less high than those related to the importance of interest groups or civil servants respectively. This is most likely a result of sequencing. The EU questions followed those on the importance of civil servants and interactions with interest groups, respectively, and were asked almost at the end of the questionnaires. This could reflect lack of concentration, which often occurs at the end of questionnaires (Fowler 2002; Groves et al. 2004). Also, interviews revealed that the EU may be a complex topic, which could also entail lower consistency. Still, the inter-item correlations suggest a satisfactory internal consistency of the questionnaires. And as the interview data are generally in line with the survey results, we can infer that this survey dataset is fairly reliable.

4.4.2 Assessing the validity of the questionnaires

Validity, one could argue, is always in the eye of the beholder. The theoretical framework that is adopted very much determines whether questionnaire items are valid measurements or not. So, in the end, validity remains much more a matter of theoretical judgment rather than a rigorously tested characteristic of the dataset. Given the nature of the subject under study, I was able to address its face and content validity by conducting pilot tests. Pilot testing included a group of peers in the field of public administration and political science who critically assessed whether the questionnaire items measured what they were supposed to measure. This phase included both native Dutch and native English speakers to ensure that

29 According to several authors (de Vaus 2002, Field 2005), inter-item consistency, or internal consistency, can be measured by inter-item correlations and Cronbach’s alpha. Cronbach’s alpha is a measure of the reliability of a scale and it measures how well the different items or components reflect a similar concept. According to a similar logic, that different items are designed to measure the same thing, we can use Cronbach alpha to check the reliability of multi-item measurement. The values of Cronbach’s alpha vary from 0.68 to 0.91 for the civil service dataset and from 0.54 to 0.66 for the interest groups dataset. These values indicate a good to satisfying internal consistency.

30 More sophisticated analyses of validity, such as criterion and construct validity, could not be used given the nature of the topic at hand (see Litwin 1995).
interpretation of the questions was as similar as possible. I refined the questionnaires based on their comments. I also pre-tested the senior civil servants’ questionnaire with a small group of civil servants to check whether the items were meaningful and whether the questionnaire was not too long. Again, their comments resulted in improvement of the questionnaires. In addition, several experts in survey methods provided feedback on the wording and order of the questionnaire items.\[31\] I did not assess the internal validity of the survey dataset as such, yet further analyses (see chapters 5-7) did not point to major internal contradictions that would indicate poor internal validity. I addressed the internal validity of the entire database, however, by including both surveys and elite interviewing as corroborative methods of data collection.

4.4.3 Assessing the response of the SCS survey
Assessing external validity in the case of survey research is usually based on the quality of the sampling method and an assessment of the relative response rate. The Dutch senior civil service, including the members of the candidate-programme, consisted of 996 members at the time the survey was conducted. I sent questionnaires to the total population and received 416 questionnaires (a response rate of 41.4 percent). Such a response rate allows generalisation of the results to the entire population of senior civil servants in the Netherlands.

To further assess the external validity of the Dutch dataset, I examined the differences between the dataset and the population in terms of gender, age, minority and grades. Figures 4.3 show the results of this analysis for the Dutch senior civil service. The proportion male/female is roughly similar, as are the figures representing age. Until the age of 55, the percentages are more or less similar. Above the age of 55, the survey dataset shows smaller percentages of these age categories. In terms of grades, the results show some minor differences. The survey dataset contains less people with a grade under 15, somewhat more people with grades 15 and 16, and less people with grades 17, 18 and 19 than the entire population. The results are quite satisfactory for the Dutch senior civil service dataset and they confirm the possibility to generalise the results to the entire population of senior civil servants in the Netherlands.

\[31\] I thank both Prof. dr. P.L.H. Scheepers (Nijmeg University) and Dr. T. Capelos (at that time Leiden University, currently University of Leeds) for very useful feedback on the structuring and wording of the items.
Figures 4.3 Representativeness of the Senior Civil Service (SCS) survey in the Netherlands
With regard to the UK SCS, I received 406 completed questionnaires from a total of 3,306 senior civil servants listed in Dod’s Civil Service Companion (2007). This would have resulted in a response rate of 9.1 per cent, which is too low to justify generalisations for the entire senior civil service.

A further assessment of the external validity of the UK dataset, similar to the Dutch dataset, included figures of gender, age, minority and grades. Generalisation appeared to be somewhat difficult for this dataset given an almost completely inversed picture in terms of grade. The population of senior civil servants (SCS and level 6/7) contains 1.7 percent SCS civil servants, whereas the survey dataset includes 69.3 percent SCS civil servants. Clearly, something unexpected had happened during the process of collecting data for the UK civil servants. I had used Dod’s Civil Service Companion (2007) to establish the population of senior civil servants in the UK, assuming that those listed in the database actually were senior civil servants. A further examination of the differences in grades between the survey database and figures of the entire UK SCS (National Statistics 2006; ORC International 2006) suggested something else. Dod’s database encompassed some of the level 6/7 civil servants as well, rendering the response rate even lower.

To deal with this inconvenient situation, I decided to examine how representative the SCS respondents in the survey dataset were compared to the true population of SCS members in the UK. The respondents had indicated grades corresponding to their jobs in the questionnaire. Dod’s Civil Service Companion, appeared to explicitly denote civil servants who were members of the SCS but not those belonging to level 6/7 and included about 600 SCS members. From other sources (National Statistics 2006; ORC International 2006), I was able to infer that the UK has a total population of about 4,455 senior civil servants. Comparing the number of SCS civil servants in the survey dataset (124) to the full population of senior civil servants, the response rate is only 2.7 per cent. This is, of course, far too low for any legitimate and meaningful generalisation. Based on demographic figures, however, the set of SCS respondents appeared to be quite representative. Figures 4.4 show the results of the analysis.

32Apparently, the organisation did not offer a systematic database of SCS-members based on specific criteria. The organisation works with the database of the previous year, updates it and sends it to the department directors for approval. Some civil servants will add information or request to be omitted from the database. Very likely, civil servants who have an external profile or have a very responsive disposition will be overrepresented in the database. Instead of a systematic and representative database of the UK SCS, it turned out to be a rather unsystematic database that contained both SCS members and level 6/7 civil servants (telephone call to Dod’s, 12-01-2008 after observing differences in grades).

33 Inferred from ORC International, SCS survey 2006. The response rate of the latter was 67% of the entire SCS, including a total N of 2,985, which results in a full population of about 4,455. Since Dod’s database includes about 600 SCS, this represents 13.2 percent of the full SCS population. Then, I figured out the percentage of true SCS members in the survey. The demographic and ranking questions were posed at the end of the questionnaire. As a result of non-item response throughout the questionnaire, I had to apply the ratio of level 6/7 and SCS-level answers to the initial set of respondents. About 30.5 percent of the respondents had answered that they were SCS members. Applying this percentage to the initial number of respondents, I estimated the number of senior civil servants in our dataset at about 124. The ratio of level 6/7 and SCS-level in the UK dataset rendered it less comparable to the Dutch SCS dataset, in which the percentage of true SCS-level civil servants is higher. This may affect the analyses. Yet, answers to the questions of whether civil servants interacted with interest groups and if not for what reasons, did not reveal significant differences between the two samples concerning interactions with interest groups.
Figures 4.4 Representativeness of the Senior Civil Service (SCS) survey in the UK
As can be observed from the figures, the numbers in terms of gender and minority are roughly similar between the population and the survey dataset. In terms of age, the figures are also rather similar, yet the survey produced a somewhat different proportion of younger and older people than the proportion apparent in the population. This may be a result of the online survey, as younger people may perhaps be more inclined to fill out an online based survey. Based on the demographic statistics, we may, very cautiously, generalise the results to the UK SCS, while keeping the low response rate in mind. For the entire dataset, the results only apply to the civil servants that participated in the survey, as I was unable to determine the representativeness of the level 6/7 civil servants. However inconvenient Dod’s Civil Service Companion database may be, it was the only database of UK (senior) civil servants available at the time of collecting data. So, acknowledging the low response rate and its consequences for generalisations is generally the best option for this study. But, given the fact that in demographic terms our SCS dataset is rather similar to the full population of SCS members, we may, if cautiously, generalise the results to the entire SCS population.

4.4.4 Assessing the response of the interest groups survey

Response rates for the UK and Dutch interest group datasets were also quite low, as is common in interest group research (Gray and Lowery 1996a). I received 133 questionnaires from a sample of 879 organisations from a total population of 7,368 interest groups in the UK. The eventual response rate was 1.8 percent. For the Netherlands, however, it is difficult to formulate an exact response rate. Both samples consisted of national and non-national organisations. After establishing the number of national organisations and extrapolating the ratio to the population databases, I found a response rate of about 3 percent of the (estimate of) the total population of interest groups in the Netherlands. Both datasets are thus not really satisfactory in terms of response rates and representativeness.

34 It can also be the result of a difference in grades, as the SCS is more likely to consist of older people. Yet, the relatively large number of SCS in our survey suggests otherwise.

35 The NCC sample contained 507 (51.0 percent) national organisations, and the website sample contained 316 (84.0 percent) national organisations. Extrapolating this to the entire database, this means that the full NCC database contains about 5,718 national organisations and the website database contains about 1,737 national organisations. Adding these two numbers and comparing it with the total number of respondents would result in an initial estimate of the response rate. Yet, we still have to take into account that both datasets may overlap. Comparing the two samples (only national organisations), 2 percent of the organisations in the website database overlapped with the NCC database. This means that we have to subtract about 35 organisations from the total website database before adding the total number of organisations to the NCC database. Based on these results, we can infer that there are roughly 7,420 national interest groups in the Netherlands.
Another estimate of the representativeness of the dataset was to compare the survey dataset and the population datasets in terms of characteristics of the interest group population. In other words, we could compare the datasets in terms of the relative contribution of each type of interest group to the entire dataset. When we compare the Dutch survey dataset with the two databases, we see that the distribution of different types of interest groups in the survey dataset falls between the NCC database and the website database. Some proportions of the types of interest groups are more similar to the NCC database and others are more similar to the website database. Estimating the representativeness of the UK interest group dataset was also difficult, as the CBD research database does not contain a mutually exclusive classification system. Some organisations are classified more than once, but it is impossible to discover which ones. If we compare the relative contributions of the types of interest groups between the survey dataset and the full CBD research database, the different categories are relatively similar. However, a major difference is that there is no category for NGOs in the CBD research database. So for those organisations, it is hard to determine the representativeness of the survey dataset. Figures 4.5 show the relative contributions of each type of interest group to the survey datasets and the other databases.

We observe a larger number of employers’ organisations and NGOs in the UK interest group database than in the CBD database. The CBD dataset contains more organisations in the fields of science and culture/sports/recreation. In the case of the Netherlands, the survey dataset contains a similar percentage of consumer, employers, and minorities’ organisations as the NCC database. The percentage of development NGOs and labour organisations is similar to those of the website database. In the case of idealistic, environmental and culture/sports/recreation organisations, the survey dataset is comparable to both the website and the NCC databases. The survey dataset contains fewer ‘other’ organisations and more voluntary organisations than the two population datasets. Easy generalisation of the results is not possible, because the survey dataset of Dutch interest groups, in terms of relative proportion of types of interest groups, falls between the two databases. With regard to the UK dataset, the relative contribution of each type of interest group renders the survey dataset to be roughly similar to the CBD research database. But as the response rate is also low, we need to be cautious with generalisations. When we compare the UK and NLD samples, there are some differences. In general, however, the ratio between professional organisations and NGO’s is roughly similar. In the NLD dataset the percentage of NGOs is somewhat higher, whereas in the UK dataset the percentage of professional associations is somewhat higher.
Figures 4.5 Representativeness NLD and UK interest group surveys
In sum, generalisation of the results seems possible for the Dutch senior civil service, given the relatively high response rate and similar demographic figures of both the survey dataset and the entire population. In case of the UK SCS, generalisation is somewhat more complex, given the low response rate, yet similar demographic figures in the case of the senior civil servants seem to justify cautious generalisations. In the case of the interest groups surveys, the response is low, but this is common in interest group research. What is different in this study, at least for the dataset of Dutch interest groups, is that an attempt has been made to get a detailed picture of the broader population. This dataset allows for a comparison in the distribution of types of interest groups and permits a better judgement of the possibility to generalise findings. More broadly, despite a relatively low response rate, the resulting datasets allow a more systematic analysis of bureaucracy-interest group interactions than case studies would have allowed.

4.4.5 The quality of the dataset

So, what can we conclude about the quality of the datasets? In terms of validity, I relied on pilot testing, including peers from within the scholarly field as well as several civil servants and experts on survey methodology. Their comments and feedbacks substantively improved the questionnaires. Satisfactory inter-item correlations suggest a reliable database. The low response rate seems worrisome for generalisations. Generalising the results for the NLD senior civil servants dataset is possible given the relatively high response rate and representativeness in terms of demographic characteristics. Yet, we can only make cautious generalisations about the representativeness of the UK SCS survey in terms of the demographic characteristics, and in comparing the proportion of different types of interest groups between the populations and the datasets. All in all, this dataset can be used to test hypotheses on bureaucracy-interest group interactions and to make cautious inferences to the entire populations of civil servants and interest groups. The discussion of the findings of each of the analyses will refer to ‘civil servants’ and ‘interest groups’ more generally for convenience. Despite the usage of these general terms, the limits of easy generalisation are acknowledged.

Missing data

There is one important issue concerning survey research in comparative politics and administration that I have not dealt with so far. Surveys are notorious for their relatively high non-response or non-item response rate, resulting in incomplete datasets, which in turn complicate the robustness of statistical analyses. A recent survey of the literature shows that almost 94 percent of scholars use listwise deletion to remove entire observations from the analysis to address non-item response, resulting in missing data (King et al 2001). Generally speaking, we can observe three mechanisms of missing data. First, data can be missing completely at random (MCAR), which means that a missing response occurs only by chance. A second mechanism of missing data entails the option of missing at random (MAR). Responses are still missing at random, but may depend on other observed characteristics. Finally, data can be missing not at random (MNAR), meaning that missing responses are related to a value that could have been observed. For instance, people with a high income are more likely not to report their income.
(Sinharray, Stern, and Russel 2001). Listwise deletion is only appropriate if the proportion of missing data is relatively small and if the assumption of MCAR holds; a situation that rarely occurs. Otherwise, using listwise deletion could result in biased results and inefficient data analysis (King et al 2001). In addition, listwise deletion means that potentially valuable data cannot be used for analyses.

The databases in this study are also characterised by a relatively large non-item response (see also chapter 5 and 6). To avoid biased results, I will use multiple imputation to generate complete datasets to test the hypotheses. Various methods exist to handle missing data, and multiple imputation is but one. Although contested, there seems to be a consensus that multiple imputation is generally the best solution (Sinharray, Stern and Russell 2001; Tabachnik and Fidell 2007). Due to the contested nature of multiple imputation (Gelman, King, and Chuanhai 1998), I will compare results based on both datasets: the original incomplete dataset and the complete datasets after multiple imputation. Further details and the diagnostics of the multiple imputation process will be discussed in the context of the civil service (chapter 5) and interest groups analyses (chapter 6) respectively.

4.5 Approximating a longitudinal perspective

The dataset generated by the cross-national survey enables us to analyse current interactions between a bureaucracy and interest groups. The impact of time, however, can only be traced to a limited extent. Recall that the model developed in chapter 3 revealed the existence of alternative types of rationality that can only be unravelled by a long-term perspective. That is, habitual and anticipatory rationality may also determine resource exchanges between interest groups and bureaucrats, in addition to the strategic rationality implicit in resource dependence theory. A full appreciation of a time dimension in bureaucracy-interest group interactions, and thus the ability to examine the complete set of different types of rationality underlying those interactions, requires longitudinal data. Longitudinal data, preferably panel data, would enable us to comprehensively measure the nature of the degree of dependence characterising bureaucracy-interest group interactions over time. That is, measurement over time would make it possible to examine whether the resource exchanges are based on strategic, anticipatory, or habitual rationality or a combination of these types of choices. Unfortunately, establishing a longitudinal database was not possible in this study. Nevertheless, to address such a time perspective, I supplemented the dataset described above by semi-structured interviews with both civil servants and representatives of interest groups. I added this method of elite interviewing to the survey instrument to collect data, thereby employing a mixed-method strategy in the data analysis.

The underlying assumption of the mixed-method strategy in this research is twofold. First, I used a combination of survey research and semi-structured elite interview methods to collect data. This approach allows for a comprehensive understanding of the complex interactions between bureaucracies and interest groups. Additionally, the use of mixed-method strategies in this research enables a more holistic analysis, combining quantitative and qualitative data to provide a more nuanced understanding of the phenomena under study. This approach is particularly valuable in the context of this research, as it allows for a more detailed examination of the dynamic and multifaceted nature of bureaucracy-interest group interactions.
interviewing because this was the best mix of data collection methods available to generate the data necessary for testing the hypotheses within the boundaries of a relatively small-scale research project. I was able to use the resulting dataset with a number of analysis techniques to test the hypotheses discussed in the previous chapter. Combining statistical analyses with a version of fuzzy-set analysis, an in-depth content analysis and counterfactual analysis allowed me to measure the different types of rationality underlying bureaucracy-interest group interactions.

Second, and equally important, combining various data collection and analysis methods helps to strengthen the internal validity of the research by establishing true triangulation. The semi-structured interviews with bureaucrats and interest group representatives provided an additional check on the survey data. More importantly, these interviews enabled me to allow respondents to reflect on their interactions with bureaucrats or interest groups, respectively, when explicitly considering them over time. Relying on several methods of data collection also allowed me to construct a dataset that would serve the application of multiple analysis methods necessary for testing the different types of rationality underlying the resource exchanges between bureaucrats and interest groups.

I restricted the interviews to the Netherlands. Yet, to bring in the comparative element so important to this study, I chose two policy areas, macro-economics and public health, from which to select respondents. The choice of these two policy areas was based on an initial analysis of the survey results, which showed that these two policy areas involved civil servants interacting with a relatively large number of interest groups. I further opted for these specific policy areas for the following reasons. First, macro-economics, and social-economics, is the subject of most studies of corporatism and international political economy (Molina and Rhodes 2002). Explicitly addressing the two rival explanations, habitual and anticipatory rationality, for bureaucracy-interest group interactions regarding this policy area could result in an interesting contribution to this strand of literature. Second, public health in the Netherlands is a policy field that is not only characterised by a high number of interest groups, but also by a diverse array of interests. Many patient or consumer organisations are active in this field, as well as the very influential private firms, such as insurance companies and pharmaceutical firms, and associations representing them. The existence of many diverse organisations provides an interesting case to study how organisational characteristics may influence differences in access and the degree of dependence. Both policy areas thus allowed for an interesting additional perspective on bureaucracy-interest group interactions to that already offered by the survey dataset. In total, I

(Peters 1998; term coined by Webb et al. 1967). What is often missing in existing mixed-method work is attention to the testing of rival explanations by using several research strategies. And this is exactly the point where the implicit logic behind mixed methods - using a supplementary research method to strengthen causal inferences regarding one theoretical model or set of hypotheses - no longer seems to apply. Put differently, various methods are often combined to test the same hypotheses, but various methods are seldom combined to test different hypotheses related to a single phenomenon.

38 Interviews were held with SCS-level civil servants in the UK via telephone (N = 8). The quality of the data resulting from these interviews was not similar to the quality of the Dutch interview data. Therefore, I decided to limit the interviews to civil servants and interest groups in the Netherlands.

39 The aim of these semi-structured interviews was to collect data on the impact of time on bureaucracy-interest group interactions. I therefore did not select on the basis of the dependent variable, for instance
conducted 39 interviews with civil servants and 18 interviews with representatives of interest groups. This resulted in a dataset of bureaucracy-interest group interactions that included information on these interactions over time. As I opted for an individual-level analysis (see chapter 7), the dataset includes information on 57 different instances of bureaucracy-interest group interactions varying across policy areas, agency types, and types of interest groups, with a set of indicators for different types of choices (see appendix II and chapter 7 for further details).

4.5.1 **Reliability and validity in elite interviewing**

Interviewing, in particular elite interviewing, poses challenges to the reliability and validity of the data collected by these interviews. To address these issues, I opted for the following two strategies: semi-structured interviews and a combination of interviewing civil servants and interest groups active in a similar policy area to collect multiple interpretations concerning a single topic. Using multiple sources is an obvious yet fundamental strategy to estimate whether the things respondents telling you are close to the truth (Berry 2002). I also used the results of the survey analysis to check whether answers to interview questions corresponded to the survey results, and vice versa. The semi-structured nature of the interviews enabled me to construct a dataset of comparable data so as to allow sound analyses. The additional room for manoeuvre such interviews offer allowed me to probe further when it was necessary and to dig deeper into issues that from the outset seemed interesting, yet were not at the core of the research project. Additionally, semi-structured questions and open-ended questions allow respondents to organise their answers within their own frame, which seems essential in the case of elite interviewing. Elite interviewees apparently do not like “to be put in a straitjacket of closed-ended questions” (Aberbach and Rockman 2002). Semi-structured interviews are thus very suitable for elite interviewing when used for exploring alternative explanations of a particular phenomenon, as they both allow a researcher to establish a sound database and at the same time provide enough flexibility to engage in what should be interpreted as a good conversation by the respondents (Aberbach and Rockman 2002; Odendahl and Shaw 2002).

To address the political sensitivity of the topic, I promised confidentiality and anonymity and took into account the ordering of the questions (Leech 2002). That is, I started with general questions and gradually asked those that could reveal the choosing one area with no or relatively few interactions and a policy area with many interactions. Such a selection would not serve the aim of the data analysis eventually to be conducted: examining the existence of different types of rationality revealed by long-term interactions. In addition, the comparison of two policy areas does not involve a detailed, structured, focused comparison (George and Bennet 2005). Rather, the selection serves to add an initial comparison of policy areas, types of interest organisations, and agency types to the in-depth analysis of the interview data so as to probe whether a mixture of different types of rationality would vary along these dimensions.

* The authors argue as follows: in the case of elite interviewing, or interviewing other highly educated people, respondents often comment on the rationale behind the answer options and try to suggest other ways of measuring what you want to measure (Aberbach and Rockman 2002, 674). Tellingly, the option of replying to our general email address in the survey research, as well as the option provided to the respondents to comment on the survey in a final open question, provoked exactly such commentary.
true nature of their interactions with bureaucrats or interest groups, respectively. I taped each interview and transcribed them so as to generate an accurate database for coding. A small-scale research project does not allow for multiple coders, so to deal with the intercoder-reliability issue, I coded half of the interviews twice to check the consistency of my coding (see also Chapter 7).

### 4.6 A summary of the research design

This study is truly comparative in nature, although I consider comparative research a general research strategy rather than constituting a distinct sub-discipline within public administration and political science. Establishing causal relations and sound equivalence both at a conceptual and measurement level require special attention in comparative research. I designed the research in such a way as to carefully consider both demands and to offer the best available options possible within the scope of this research project to address these issues properly.

First, in order to draw conclusions on bureaucracy-interest group interactions, I will rely on Mill's method of concomitant variation to examine whether a particular pattern of variation among a set of independent variables coincides with a particular pattern of variation in the dependent variable. As such, we can draw conclusions about how contextual factors relate to resource importance and resource concentration and how a degree of dependence may vary under different circumstances. Given the nature of this research project, it will be difficult to clearly establish causal relations, i.e. to infer which contextual factor results in which particular degree of dependence. Conclusions will thus be put in probabilistic rather than causal terms.

I address equivalence, a second important threat to comparative research, both at a conceptual and measurement level. In terms of conceptual equivalence, I use a reconceptualisation of resource exchange to measure bureaucracy-interest group interactions. To compare, we need to define a fundamental set of characteristics of a concept. The fundamental characteristic of bureaucracy-interest group interactions is the exchange of resources, be they tangible or intangible. Such a resource exchange renders it possible to distinguish between the different natures of such interactions as hypothesised by the literature on bureaucratic politics, on interest group politics, and on policy networks. Moreover, as the two elements determining such a resource exchange, resource concentration and resource importance, can be measured under different circumstances, we can also compare the nature of such interactions under such different circumstances. That is, we can compare the impact of contextual factors on the degree of dependence characterising bureaucracy-interest group interactions. In addition, the definitions of senior civil servants (civil servants belonging to a nation's senior civil service),

---

When respondents indicate that they want to tell you something ‘off the record’, you know that they are quite frank in general and in such cases are willing to give away some sensitive details. This happened in several interviews, which can be interpreted as a sign of relatively sound reliability. Yet, it is always important to keep in mind, as is well known, that each individual’s answer offers a particular perspective to a given situation.
and interest groups (private interesting-pursuing entities that intersect with public policy) also take into account equivalence issues.

In terms of data collection and measurement, I opted for comparative online surveys, an instrument that perfectly fits the subject under study, yet has some drawbacks in terms of equivalence and reliability. Careful operationalisation, pre-testing, the usage of an online survey instrument and multiple imputation have been applied to address these issues. To generate equivalent populations of bureaucrats, this study has been restricted to the senior civil services. A lack of comparable databases of interest groups populations posed, however, a serious challenge to this comparative endeavour. Databases of interest groups were available for the UK (and the US), but the Netherlands (and Sweden) lack such databases. Therefore, I developed a strategy to establish a dataset of interest groups in the Netherlands to overcome this particular comparative problem. By examining the foundation and association database of the Netherlands Chamber of Commerce (NCC) and an additional website search, I constructed a comprehensive dataset of a national Dutch interest group population. This dataset was rather equivalent to the existing datasets in pluralist systems. It thus allowed for both comparative research and for random sampling, a difficulty in comparative interest group research in general and, in particular, in corporatist regimes.

The surveys resulted in a reliable dataset of bureaucracy-interest group interactions, reflected in satisfying inter-item correlations. In terms of representativeness, the Dutch senior civil service dataset offers a sound dataset as well. The UK SCS dataset is a bit more complicated, as are both interest group datasets. For these three datasets, the response rates are very low. By examining demographic figures in the case of the UK SCS survey and comparing the types of organisations included in the survey and the population datasets for interest groups, I provided an additional check on external validity. Such constraints in terms of validity will be taken into account in the final interpretations and conclusions. Yet, as this study set out to generate a theoretical model to better compare bureaucracy-interest group interactions, an internally valid and reliable dataset is most important, and the survey dataset meets these demands.

To examine bureaucracy-interest group interactions over time and to provide a measure for the internal validity of this study’s dataset, the survey dataset has been supplemented with 57 interviews of Dutch senior civil servants and representatives of interest groups in the areas of macro-economics and public health. A format of semi-structured interviews and partial recoding of the interview transcripts were used to address the issues of validity and reliability in elite interviewing.

All in all, by systematically addressing the methodological issues generated by comparative research, I have been able to construct a useful dataset of bureaucracy-interest group interactions containing both survey and interview data. As such, the data collection strategy has been a combination of different methods, just as the analysis strategy will also combine different analysis techniques. Chapters 5-7 measure the degree of dependence from a bureaucratic perspective, examine resource concentration within the interest group environment via population dynamics, and examine the set of multiple rationalities underlying bureaucracy-interest group interactions.