Explaining Participation Bias in the European Commission's Online Consultations

Maiken Røed

Master thesis,

Department of Political Science

UNIVERSITY OF OSLO

May 2015

Explaining Participation Bias in the European Commission's Online Consultations

2015

Explaining Participation Bias in the European Commission's Online Consultations

Maiken Røed

http://www.duo.uio.no/

Print: Reprosentralen

Abstract

This thesis is an expansion of Rasmussen & Carroll (2014). The theme is bias, a prominent one in the interest group literature, and the focus is on the European Commission's online consultations. The research question is: Which features pertaining to online consultations can explain participation bias? Bias is defined as the distribution of different types of interests, and interest groups are thus the participants studied. Rasmussen & Carroll's (2014, pp. 447-448) theoretical framework based on pluralism lays the foundation to which specifications are added. These include a cost-benefit and an exchange theoretical perspective. Pluralism expects every affected interest to be represented. The specifications add that groups representing affected interests might not participate if the costs of doing so are high compared to the potential benefits, or if they do not possess relevant information. The new hypotheses concern the organizer, format and technical complexity of the online consultations. The effect of government activity is moreover considered. One new hypothesis, concerning salience, related to the initial pluralist framework is added as well. Rasmussen & Carroll's (2014, pp. 447-448) expectations are included, but they are secondary since they have been explored in depth previously.

The hypotheses are examined by running various regressions. In addition to Rasmussen & Carroll's (2014, pp. 447-448) consultation features, two new aspects are found to explain participation bias. First, technically complex consultations are generally more biased than less complex ones. Second, online consultations organized by regulatory Directorate-Generals (DGs) are generally biased towards business interests. One of the European Commission's goals with the consultations is inclusiveness (European Commission, 2002, pp. 16-17). The findings indicate that further measures need to be undertaken to ensure this in technically complex consultations or those organized by regulatory DGs.

Compared to Rasmussen & Carroll (2014), the study is improved methodologically as well. Misspecified statistical models are respecified and improved, a quasi-separated estimate is pointed out and considered, and the statistical power to detect effects of different sizes is examined.

Acknowledgments

I am very grateful to my main supervisor Elin Haugsgjerd Allern for taking the time to oversee this thesis from start to finish, including answering my questions and providing helpful suggestions and guidance. Her thorough and useful comments, especially regarding the structure, theoretical framework, variable measures and findings, helped me move forward with the thesis. I would also like to thank Bjørn Høyland, who co-supervised this thesis, for his feedback in general and in particular for answering my questions concerning the statistical analyses, as well as directing my attention to King & Roberts (2014).

Oslo, May 2015
Maiken Røed

Number of words (excluding the front page and references): 33,949

Contents

List of Figures

List of Tables

1 Introduction

The study of bias has been a prominent part of the interest group literature since its inception. A bias towards business interests has been the overall refrain, including in the European Union (EU) (Coen, 2007, p. 335; Wonka et al., 2010, pp. 467-469). Little attention has been given to how the characteristics of participatory activities affect bias, however. One exception is Rasmussen & Carroll (2014) who examine participation bias in the European Commission's online consultations. This thesis is an expansion of their study, with the research question: Which features pertaining to online consultations can explain participation bias?

Rasmussen & Carroll's (2014, pp. 447-448) theoretical framework is based on pluralist ideas, where all affected interests are assumed to be represented. However, being affected does not necessarily lead to participation. Some may for instance be less affected than others. If participating then is costly in terms of time, information or other resources – or if the potential gains are small – groups representing less affected interests might not contribute. This cost-benefit aspect of participating is included in this thesis, thus taking the initial pluralist framework one step further by considering when affected groups are less likely to participate. Furthermore, an exchange theoretical view is added, where interest groups' supply and DGs' demand for information are the focal point. Groups might be affected but still not participate because they do not possess relevant information to do so.

Several new hypotheses are derived from the expanded theoretical framework. Two of these are supported by the results of the regression analyses. First, technically complex consultations are generally more biased than less complex ones, as expected based on cost-benefit reasoning. In addition, the type of Directorate-General (DG) organizing a consultation is found to affect the direction of the bias. Consultations organized by regulatory DGs are generally biased towards business interests. This hypothesis is based on the exchange theoretical framework.

These findings are relevant in a broader sense. One of the Commission's goals is for the con-

sultations to be inclusive (European Commission, 2002, pp. 16-17). The findings here indicate that to be able to achieve these, it is not sufficient to have the consultations be formally open for everyone. If the matters they concern are technically complex or if the consultations are organized by regulatory DGs, further action need to be taken to incite wider participation. This is furthermore related to how interest groups function as a channel between civil society and decision-makers, in addition to the electoral channel (Persson, 2007, p. 225). When only some types of interests are represented, how well this channel functions can be questioned. Decision-makers are unlikely to only take the interests of those participating into account (Lowery, 2007, pp. 34-36), but biased participation might still lead to less representative proposals.

Having introduced the thesis' theme, research question and main findings, the rest of this introductory chapter is structured as follows. The underlying interest group perspectives in this thesis are presented, including how they relate to the subject matter of this thesis. A summary of Rasmussen & Carroll (2014) is then provided, in addition to a motivation for why examining additional explanatory factors is relevant. The limits of choosing to study online consultations in the EU, in terms of being able to generalize the findings to national settings, are also discussed. Lastly, a brief summary of this introductory chapter is given before providing an outline of the rest of the thesis.

1.1 Interest group perspectives and online consultations

The study of bias constitutes a considerable part of the interest group literature. The pluralist and transactions perspectives have defined the field since the 1950s. In later years, an additional perspective – neo-pluralism – has emerged, which reads as a fusion between the two initial ones (Lowery, 2007, pp. 31-34). These perspectives, and how they relate to this thesis, are discussed here. It is furthermore argued that the online consultations lay the groundwork for pluralistic participation patterns. It is worth noting that bias is understood mainly as an empirical, not normative, concept in this thesis.

During the 1950s and '60s, pluralism was the dominant force in interest group studies (Baumgartner & Leech, 1998, pp. 48-50). Pluralist theory expects that when someone's interests are being threatened, the affected individuals will mobilize and form interest groups. When one group is formed, new interests are potentially threatened, resulting in the formation

of additional groups (Truman, 1951, pp. 40-44). The interest group population in a given society will accordingly reflect all salient interests and be free of bias (Baumgartner & Leech, 1998, p. 54; Lowery, 2007, p. 31). Because of this, interest groups are viewed as an asset to democracy. The entire range of relevant interests is available to supply their views to the government (Baumgartner & Leech, 1998, p. 48). This relates to the first step of mobilization, the formation of interest groups. The second step concerns political participation (Schlozman et al., 2012, p. 272). Here, pluralism would expect every interest affected by the subject matter of a consultation to be represented.

In the transactions perspective, the idea that all actors whose interests are threatened will mobilize and form interest groups is opposed. To mobilize, selective incentives are needed. Individuals will form or join interest groups if they can gain something from doing so which they would otherwise miss out on. This is not the case for collective and diffuse interests. The benefits arising from mobilizing here will befall everyone whether one joins or not. Each individual is therefore better off by free riding on the efforts of others, resulting in a collective action problem. When everyone follows this logic, no groups will be formed to represent collective and diffuse interests. The result is a biased interest group population. Certain types of interests are represented while others remain latent, making the population unrepresentative of the distribution of interests in society (Lowery, 2007, p. 32; Olson, 1971, pp. 1-2, 11-16, 21).

Translated to online consultations, an interest group has selective incentives to participate if it expects that no other group representing the same interests is likely to contribute. If this is not the case, the group in question could free ride on the contributions of other groups with similar interests. If each of these groups follows this logic, however, their interests would not be represented at all. The result would therefore be a more biased participation pattern. If no or only a few other groups represent the same interests, the group would have selective incentives to participate since their interests would otherwise not be represented.

Neo-pluralism merges insights from the pluralist and transactions perspectives. Empirical findings have shown that the collective action problem is not as severe as expected (Lowery & Gray, 2004, p. 166; Lowery, 2007, p. 34). Citizen groups, representing collective and diffuse interests, do mobilize and constitute more than a negligible part of interest group populations (Hojnacki et al., 2012, p. 382). Incentives other than selective ones can lead to mobilization,

such as purposive or solidary ones. In the latter case, groups are formed because of the social aspects they offer. Purposive incentives motivate mobilization based on the end goals of the group that are of no direct benefit to the members themselves (Clark & Wilson, 1961, pp. 134-136). The interest group population is not as unbiased as pluralism expects, but not as biased as assumed in the transactions perspective either (Lowery & Gray, 2004, p. 166). In the online consultations, different incentives might motivate groups to participate, resulting in a participation pattern that lies somewhere between unbiased and biased.

Of these three interest group perspectives, the online consultations lay the groundwork for pluralistic participation patterns. In contrast to the Commission's other consultation activities, such as expert groups, forums and conferences, there are few hurdles to overcome for interest groups that want to participate. No invitation is needed and the groups can contribute from wherever they are, as long as they have Internet access or are able to send their submission by postal mail (Greenwood, 2011, p. 209; Quittkat & Kotzian, 2011, pp. 401-404; Rasmussen & Carroll, 2014, p. 449). That no invitation is required leaves the decision of whether to contribute or not entirely in the hands of each interest group. Contributing to online consultations is also possibly less costly, since it can take place from wherever a group finds itself. Being in Brussels, for instance, is not necessary. This makes participating possible for less resourceful interest groups as well. Online consultations thus constitute a most likely case for pluralistic participation patterns.

The openness of online consultations could mean that they are less attractive lobbying tools. This is not likely to be the case, however. Online consultations are intended for the Commission's policy proposals that are expected to have a substantial impact on affected parties, or that concern larger policy reforms. Lobbying at this early stage is important. It gives an actor the possibility of contributing to shape a proposal before it is debated by the Council and the European Parliament. For these institutions, accepting a proposal is easier than modifying it. This makes potentially exerting influence at the proposal stage pertinent (European Commission, 2002, p. 15; Klüver, 2011, p. 485; Quittkat & Kotzian, 2011, p. 405). Interest groups are therefore assumed to be interested in participating in online consultations.

Studying participation bias in online consultations can thus shed light on whether any characteristics facilitate pluralistic participation patterns. The features that lead to more biased outcomes – in line with the picture painted by the transactions or neo-pluralist perspective –

are also uncovered. In addition, two of the new hypotheses relate to the direction of the bias – whether the consultations are biased towards business or civil society interests. The related results here can show whether the expected interest group types dominate, or if the participation is not as biased as predicted and leaning more towards a pluralistic or neo-pluralistic pattern.

1.2 Explaining participation bias

As mentioned, Rasmussen & Carroll (2014) lays the foundation for this thesis. In this section, a summary of this study is provided. Furthermore, a motivation for why it is relevant to continue studying this is given. Lastly, the limits of choosing to study online consultations in the EU, in terms of being able to generalize the results to national settings, are discussed.

Rasmussen & Carroll (2014) study participation bias in the Commission's online consultations. The degree of bias in each consultation is determined based on the distribution of represented substantive interests. To measure this, it is necessary to know which interests a participant represents. Interest groups are thus the relevant actors. To explain participation bias, pluralist theory is used. Four expectations are derived from this, where the main focus is how broad the range of affected interests is likely to be. These concern the policy type and cost structure of the proposal a consultation[1] deals with, as well as the number of interest group participants and how biased a consultation is when it comes to the types of actors participating. The expectations concern the general degree of bias, but whether the variables affect the bias towards business interests is examined as well. This is because bias is often posited as a bias towards business. All of their hypotheses are supported by the results of their ordinary least squares (OLS) regression. Every variable, except the number of interest group participants, is also found to affect the bias towards business (Rasmussen & Carroll, 2014, pp. 446-448, 450, 453-456).

In this thesis, a number of new hypotheses for why online consultations are biased are examined. Rasmussen & Carroll's (2014, pp. 447-448) expectations are furthermore kept, but are secondary since they have been examined in depth before. Space is instead given to the new hypotheses. The research question is: Which features pertaining to online consultations can explain participation bias? One of the Commission's goals is for the consultations to be

[1] "Online consultations" and "consultations" are used interchangeably throughout the thesis.

inclusive (European Commission, 2002, pp. 16-17). The aim is to gain a better understanding of the consultation characteristics that decrease and increase the likelihood of achieving these goals. In a broader sense, this concerns how well the interest groups function as a channel between civil society and decision-makers. Less biased participation patterns can make it easier for the Commission to take the interests in society into account when continuing work on the proposal a consultation concerns. This relates to the concept of deliberative democracy (Finke, 2007, pp. 4, 15). If the participation instead is biased, less representative proposals might be the outcome. In these cases, whether the interest group channel functions as a source to democratic, input legitimacy can be questioned.

The motivation for continuing to examine potential explanatory factors behind participation bias is to see whether non-pluralistic aspects affect bias as well. Interests are not necessarily affected to the same degree by a consultation. If the costs associated with participating are relatively high, or the potential benefits small, groups representing less affected interests might not participate. The costs may exceed the benefits of contributing for these groups, leading to a more biased participation pattern. A cost-benefit framework can thus shed further light on what explains participation bias. The supply and demand for information is moreover considered. The focus here is on what the organizers of the consultations – the DGs – typically demand and what the different types of interest groups supply. Some groups might be affected by a consultation but not possess the relevant information to participate, making it less likely that they do. In sum, there might be other relevant factors that can explain why consultations are biased. That the R^2 from Rasmussen & Carroll's (Rasmussen & Carroll, 2014, p. 455) OLS regression is around 0.33 underlines that there is more variance left to explain.

In addition, a demarcation is made between what is likely to affect the direction of the bias and the general degree of it. Rasmussen & Carroll (2014, pp. 447-448, 453-456) hypothesize about what is likely to affect the general degree of bias, but they also examine whether any of their variables affect the bias towards business interests. Here, two new hypotheses are added that specifically concern the direction the bias is expected to take.

The explanatory factors that are found to affect participation bias in the EU might apply in national settings as well. Online consultations are not unique for the EU. They are also organized in some national states (Rasmussen, 2015, pp. 3-4, 11-12). They furthermore resemble written consultations but with the additional possibility of participating online. There are

nevertheless differences between national settings and the EU that might limit the possibility of generalizing the results. The prevalence of EU wide associations and groups from different national settings at the EU level is one such difference. The EU also mainly deals with regulatory matters, with distributive ones for largely left to the discretion of each member state (Hix & Høyland, 2011, pp. 189, 218). This may make it more relevant for business interests to try to influence the Commission (Coen, 1997, p. 24). The broader responsibilities of national authorities might attract a more diverse range of interests. The generalization possibilities are therefore limited.

1.3 Summary

In this introductory chapter, the theme and research question were presented, the interest group perspectives underlying the thesis were elaborated upon, and a motivation for why continuing to explore this is relevant was given. Building on Rasmussen & Carroll (2014), the aim is to explore which features pertaining to online consultations that can explain participation bias. The online consultations investigated are the ones organized by the European Commission. There are few barriers to entry in these consultations, making them a most likely case for a low degree of bias and pluralistic participation patterns. Rasmussen & Carroll's (2014, pp. 447-448) pluralist framework is broadened. Groups might be affected to different degrees, and how resource demanding participating is can therefore also affect bias. To take this into account, a cost-benefit as well as an exchange theoretical framework are added.

1.4 Outline of the thesis

The rest of the thesis is outlined as follows. In the second chapter, following this introductory one, the theoretical framework is detailed. A brief literature review of relevant empirical research follows each theoretical section. Rasmussen & Carroll's (2014, pp. 447-448) initial framework is first presented. The main idea here is that consultations that affect a larger part of the public are less biased. Seeing that their expectations have been examined in depth before, they are secondary in this thesis to make room for discussing the new hypotheses. The new add-ons to the initial framework are subsequently discussed, where the overarching theme is

the resources needed to participate. The new hypotheses concern the salience, technical complexity and format of the consultations, an interaction between complexity and salience, as well as the type of organizer and EU activity.

In Chapter 3, the data, coding procedure and variable measures are presented and discussed. 98 new consultations are added to the initial sample. The new observations were coded by the author twice to be able to test the intra-coder reliability, which is found to generally be high. The population is every online consultation organized by the European Commission between December 2001 and August 2011, whereas the sample includes the consultations where information about the participants has been published. This information is necessary to be able to code the dependent variable measures. How the sample might differ from the population is furthermore discussed. Concerning variable measures, Rasmussen & Carroll's (2014, pp. 453-454) operationalizations are kept in this thesis. The new independent variables are measured using operationalizations found in the interest group literature.

In Chapter 4, a first look at the data is provided. Descriptive statistics for the variables are presented. Whether the new hypotheses are supported by the results of bivariate analyses is furthermore considered. Technical complexity, format and the type of DG organizing a consultation are found to affect bias as hypothesized. The hypotheses concerning salience and activity are not supported. These results give an indication of what the outcome of the regression analyses is likely to be. Lastly, which statistical models that are appropriate given the dependent variable measures are discussed. OLS, multilevel linear and fractional regression models are found to be suitable.

In Chapter 5, the results of the regression analyses are presented and discussed. The hypotheses concerning the general degree of bias are first considered. Technical complexity and Rasmussen & Carroll's (2014, pp. 447-448) variables have the expected effect on bias here. Rasmussen & Carroll (2014, pp. 453-456) also examine whether any of their variables affect the bias towards business, since this type of bias is a common refrain in the literature. This is therefore done in this thesis as well. Some of Rasmussen & Carroll's (2014, pp. 447-448) variables, as well as technical complexity, are found to affect the bias towards business. The results indicate that there might be some differences between what explains bias in general and what explains the bias towards business interests. Lastly, the two new hypotheses concerning the direction of the bias specifically are examined. Consultations organized by regulatory DGs

are found to be biased towards business interests, as expected. The results concerning the interaction between technical complexity and salience are uncertain, but some are in the expected direction.

Chapter 6 concludes the thesis. The larger implications of the findings are discussed. The participation in technically complex consultations, as well as those organized by regulatory DGs, is biased. In these cases, the input legitimacy that can be derived from the interest group channel is reduced. If the Commission wants these consultations to be more inclusive, extra measures to ensure this might be needed. Furthermore, how the participation patterns align with the pluralist, transactions and neo-pluralist perspectives is considered. When fewer interests are assumed to be affected, or the costs associated with participating are assumed to be high, the patterns are mostly biased. When more interests are assumed to be affected, or the costs are lower, neo-pluralistic participation patterns are generally the outcome. Few consultations are characterized by fairly unbiased participation, in line with the pluralist perspective. Lastly, the main findings of the thesis are summarized. Technical complexity and the type of DG organizing a consultation are the two new characteristics found to explain participation bias. To be able to conclude with more certainty that the other new hypotheses are not supported, however, a larger sample is needed.

2 Theoretical framework and literature review

In this chapter, the theoretical framework of the thesis is presented. The unit of analysis is online consultations. Bias is defined based on the distribution of different types of interest groups in a given consultation (Rasmussen & Carroll, 2014, p. 446). If the distribution is skewed in any one direction, the participation pattern is described as biased. This is a descriptive, rather than normative, definition. It does furthermore not take the distribution of interests in the interest group population into account.

First, Rasmussen & Carroll's (2014, pp. 447-448) framework is presented and discussed. The main underlying aspect of this framework is pluralistic. The degree of conflict between different types of interests is in focus. Their expectations are secondary in this thesis since they have been examined in depth previously. Priority is instead given to the new hypotheses. The expectations are nevertheless kept, since they have been found to explain bias. Second, the add-ons to the initial framework are discussed. This is more specifically a cost-benefit and an exchange theoretical perspective. Whether affected interests participate is assumed to depend on how resource demanding participating is. The hypotheses are presented after the theoretical accounts, and are generally based on theory rather than empirical findings. Relevant previous empirical findings are, however, also presented in this chapter.

2.1 Rasmussen and Carroll's theoretical framework

The main theme in Rasmussen & Carroll's (2014, pp. 447-448) theoretical framework is conflict. Different degrees of conflict are expected to result in differences in how biased the participation is. Underlying this is pluralism. Actors whose interests are affected by a consultation

are assumed to want to contribute to protect their interests. Consultations that impinge on a broader set of interests are therefore expected to be less biased than those that affect fewer interests. The expectations derived from this framework are expanded upon in this section. They are secondary here since they have been examined in depth previously. They are nevertheless kept, since Rasmussen & Carroll (2014, pp. 454-456) find support for all of them. Examining whether the expectations are still supported when both the sample and model are expanded in this thesis is therefore relevant.

2.1.1 Distribution of costs across actors and sectors

The type of policy proposal a consultation concerns can potentially affect participation bias. Rasmussen & Carroll (2014, p. 447) distinguish between regulatory and distributive policies. Distributive policies require public spending. A larger pool of actors is assumed to carry the costs of such proposals. The number of affected interests will therefore be higher. Regulations, on the other hand, impinge on the behavior of a specific and smaller set of the public. Fewer interests are assumed to have to carry the costs of such proposals and be affected by them. Consultations concerning regulatory proposals are therefore expected to be more biased than those concerning distributive ones (Rasmussen & Carroll, 2014, p. 447).

Rasmussen & Carroll (2014, p. 447) also hypothesize that the proposals' cost structure affects bias. This structure is either characterized by concentrated or diffuse costs. Concentrated ones are limited to a smaller number of sectors, while diffuse costs are spread among several sectors and/or taxpayers in general. Based on this, consultations concerning proposals with diffuse costs are assumed to affect a broader range of interests. They are therefore expected to be less biased. Proposals with concentrated costs are assumed to affect fewer interests, leading to more biased consultations (Rasmussen & Carroll, 2014, p. 447).

Discussion

Based on the transactions perspective described in Chapter 1, Rasmussen & Carroll's (2014, p. 447) abovementioned expectations can be challenged. Consultations concerning distributive proposals, or ones with a diffuse cost structure, are assumed to impinge on a broader range of interests. However, affected interest groups might not participate if they do not have selective

incentives to do so, as the ideas in the transactions perspective imply. Groups representing diffuse interests may assume that they can free ride on the contributions of others with similar interest positions[1]. If every group of this type follows this logic, the result is that these interests go unrepresented. The participation bias in these consultations may therefore not be any lower than in those concerning regulatory proposals, or ones with a concentrated cost structure.

A hypothetical example that illustrates this is a consultation concerning farm subsidies – a distributive policy. The interests of consumers are affected here, since subsidies affect food prices. A consumer group is therefore a relevant potential contributor. However, the possible gains from contributing – lower food prices – are diffused among all consumers, including the members and constituencies of other consumer groups. Each consumer group's best response is therefore to free ride on other consumer groups' contributions. If every group follows this logic, this type of diffuse interest will not be represented at all. In contrast to consumer groups, farmer groups might have selective incentives to contribute. Contributing can have consequences for their specific members and constituencies, which no other group represents. Each farmer group can thus not count on other farmer groups to represent their members' particular interests. Their selective incentives are therefore assumed to drive them to participate.

Based on the transactions perspective, this hypothetical consultation concerning a distributive proposal is likely to be biased. The groups with selective incentives to participate will do so. Groups that face diffuse benefits, on the other hand, will not be incited to participate because they can free ride on the contributions of similar groups. If all these groups follow this logic, their interests will not be represented. It can therefore be theorized that consultations concerning distributive proposals are not necessarily less biased than those concerning regulatory matters. The same applies to consultations with a diffuse compared to concentrated cost structure. When controlling for new factors – as is done in this thesis – the correlations between the initial independent and dependent variables might be affected. The effects of policy type and cost structure might be overestimated because relevant explanatory factors are left out. Whether this is the case for the policy type and cost structure estimates will be seen here.

[1]This argument hinges on that there are numerous interest groups representing the same kind of interest. Otherwise, coordination between similar kinds of groups may take place.

2.1.2 Actor bias and interest group density

In addition to the theoretical expectations concerning the distribution of costs, Rasmussen & Carroll (2014, pp. 447-448) expect that the degree of bias will vary with the distribution of the actors participating in a consultation. Truman (1951, pp. 52-55) argues how changes in one institution likely results in changes in adjacent institutions. Complexity in one area usually goes hand in hand with complexity in other areas. When society is evolving and becoming more complex, more interest groups will be established and the diversity in the interest group population will increase. Societal development thus also results in interest group development. Given that the EU is the only system examined, Truman's (1951, pp. 52-55) hypothesis cannot be used (Rasmussen & Carroll, 2014, p. 448).

Instead, the same logic is applied to online consultations. Consultations are regarded as complex if they attract participation from different kinds of actors, such as companies, individuals, public organizations and interest groups. If consultations are complex in this regard, they are also assumed to be complex when it comes to the types of interest groups participating. That is, a lower degree of actor bias is likely to correlate with a lower degree of participation bias (Rasmussen & Carroll, 2014, p. 448).

Lastly, the concept of countervailing power may also be relevant to explain participation bias (Rasmussen & Carroll, 2014, p. 448). Originally, this concept concerns the formation of interest groups. When one kind of group is formed, individuals with conflicting interests are affected. These will mobilize and form interest groups to countervail the views of the initial groups. All salient and relevant interests will therefore be represented in the interest group population. These ideas, presented by Truman (1951, pp. 40-44), are at the core of the pluralist view of interest groups. In the context of online consultations, interest groups are assumed to be motivated by wanting to countervail the perspectives of other groups. When one type of group participates, other groups might be incited to participate. Thus, when the interest group density in a consultation is higher, bias is expected to be lower (Rasmussen & Carroll, 2014, p. 448).

Discussion

While the theoretical account behind the actor bias expectation starts out as interesting, the result borders on being trivial. Gerring (2005, p. 170) lists different criteria for what makes a causal[2] argument interesting. One of them is "differentiation" (Gerring, 2005, pp. 174-175). It should be possible to differentiate between the independent and dependent variables. In this case, both the independent and dependent variables concern bias. Actor bias and participation bias can therefore be said to be variations on the same theme. Another criteria is "priority" (Gerring, 2005, pp. 170, 175-176). The independent variable should not be closely related in time to the dependent variable. This is not the case for actor bias and participation bias. Both are the outcome of actors choosing to participate or not, and they may therefore said to be concurrent. This is also the case for the expectation regarding interest group density, which concern the number of interest group participants in a consultation. The actor bias expectation in particular, but also the interest group density one, is thus less interesting than the other expectations.

2.2 Theoretical specifications

In this section, the add-ons to Rasmussen & Carroll's (2014, pp. 447-448) theoretical framework are presented. An exchange theoretical and a cost-benefit perspective are added. The first relates to interest groups' supply and DGs' demand for information, from which a hypothesis concerning the type of DG organizing a consultation is derived. The cost-benefit perspective mostly concerns how costly participating is likely to be. More specifically, hypotheses related to EU activity, the format and technical complexity of a consultation and an interaction effect between salience and technical complexity are derived from this perspective. The main assumption is that less affected groups might refrain from participating if they do not possess the relevant information to do so, or if the costs associated with contributing is too high compared to the potential benefits. This is expected to lead to more biased participation patterns. A new feature – salience – derived from the initial, pluralist framework is also added and discussed.

[2]Rasmussen & Carroll's (2014, p. 454) analysis is quantitative, as is this thesis'. *Causal* relationships between the variables are therefore not explored – correlations are. The aim is still explanation, however, which makes the criteria relevant.

Relevant previous research is considered after each hypothesis.

Whereas all of the expectations of Rasmussen & Carroll (2014, pp. 447-448) concern the general degree of bias, some of the new ones concern the direction of it. That is, whether consultations are expected to be biased towards for instance business or civil society interests. There are for this reason several dependent variable measures, which is discussed in Chapter 3. However, the presentation of the theoretical specifications is structured according to the independent, not the dependent, variables. The hypotheses related to the direction of the bias are, in other words, interspersed among the ones related to the general degree of it.

2.2.1 Salience

The first new aspect that is added to the expanded model in this thesis is the salience of the consultations. Salient consultations are likely to affect the interests of a larger part of the public and be more prominent on the policy agenda than less salient consultations (Gormley, 1986, p. 598; Lowry & Joslyn, 2014, p. 154; Rasmussen et al., 2014, pp. 253-254). This relates to Rasmussen & Carroll's (2014, pp. 447-448) theoretical framework. Salient consultations are assumed to affect a larger range of interests than less salient ones, and are thus expected to be less biased as well.

Hypothesis 1 *Consultations concerning less salient matters are more biased than consultations concerning salient matters.*

Of relevant previous findings, Rasmussen et al. (2014, pp. 253-254, 257, 260-262) carry out a quantitative study where they investigate whether salience affects the number of interest group participants in the Commission's online consultations. They find a correlation between a high degree of salience and more participants. Rasmussen & Carroll (2014, pp. 454-456), as mentioned, find a correlation between the number of interest group participants and bias. These findings indicate that salience might affect the degree of bias as well.

2.2.2 Supply and demand for information

Having introduced salience, the additional hypothesis derived from the same pluralist framework as that of Rasmussen & Carroll (2014, pp. 447-448), the exchange theoretical perspective is considered here.

The majority of the Commission's staff is concerned with translation and interpretation, whereas only about 20% of the 28,000 employees work in policy-related areas. For comparison, this is about the same amount as a medium-sized European national government department (Corbett et al., 2012, p. 53). The policy-related resources of the Commission are therefore limited. When the Commission is working out legislative proposals, collecting relevant information is resource demanding. Interest groups can lessen this burden by providing information – for instance via online consultations. In return they might expect to influence outcomes. If policies can be broadly categorized as either distributive or regulatory, DGs can be distinguished by which of these policy types they mainly deal with. As described above when discussing Rasmussen & Carroll's (2014, pp. 447-448) theoretical framework, distributive policies require public spending whereas regulations address the behavior of certain actors.

Since the Commission is a democratic institution, it depends on having democratic legitimacy. This can take on different forms. Output legitimacy involves the quality of policies. Input legitimacy relates to making decisions that are supported by the public. DGs will have a demand for mainly input or output legitimacy based on whether they primarily deal with regulatory or distributive policies. Based on this, they have different demands for information. Output legitimacy goes hand in hand with technical information, which may enhance the quality of the policies. Input legitimacy is associated with political information concerning the public opinion (Coen & Katsaitis, 2013, pp. 1105-1106, 1108-1109).

This theoretical perspective thus adds the DGs to the equation. DGs that primarily prepare regulatory proposals are assumed to need output legitimacy and demand technical information. This work requires specialized knowledge in the relevant fields that are going to be regulated. For instance, if the Commission is working on a proposal concerning the regulation of some specific aspect of light bulbs, input from experts on lighting may enhance the quality of the proposal. Distributive DGs are assumed to mostly need input legitimacy and therefore demand political information. Distributive policies are generally more politicized and less technically complex compared to regulations (Coen & Katsaitis, 2013, pp. 1105-1106, 1108-1109). An example here is a consultation regarding the EU's state aid, where what the public thinks may be pertinent.

Interest groups are on the supply side. It is assumed that civil society groups generally are able to supply more political than technical information. Civil society groups represent

more diffuse and broader interests than for instance business groups. They will therefore have information about what one part of the public – their constituency – will accept. For example, interest groups which favor a lesser role for the EU are able to supply information about what this constituency will accept concerning state aid levels. This relates to input legitimacy – what distributive DGs are assumed to mainly be demanding.

Business groups, on the other hand, are assumed to chiefly offer technical know-how. They usually represent a narrower set of interests pertaining to economic or occupational areas and likely possess technical expertise about their specific field (Coen & Katsaitis, 2013, pp. 1105-1106, 1108-1109). Given a regulatory proposal concerning light bulbs, business groups representing light bulb manufacturers are likely to have an inkling of how the proposition will affect their industry in practice. The different expectations are summarized in Figure 2.1.

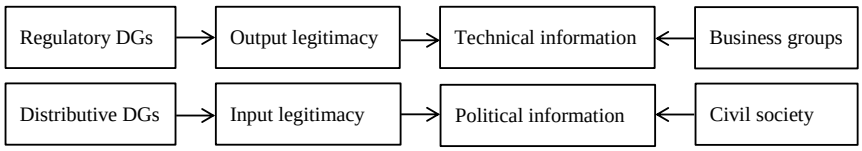


Figure 2.1: DGs' demand for legitimacy and information and interest groups' supply

From this, it follows that the distribution of interests in consultations organized by both types of DGs are expected to be biased. Consultations organized by distributive DGs are expected to be biased towards civil society interests. They demand political information, which mainly civil society groups are assumed to supply. Consultations organized by regulatory DGs, on the other hand, are expected to be biased towards business interests. They demand technical information, which mainly business groups are assumed to supply.

Hypothesis 2a *Consultations organized by DGs mainly concerned with regulatory matters are biased towards business interests.*

Hypothesis 2b *Consultations organized by DGs mainly concerned with distributive matters are biased towards civil society interests.*

This is one of the hypotheses for which a dependent variable measure that considers the direction of the bias is needed. The point is not to examine whether the consultations organized by regulatory DGs are more biased than ones organized by distributive DGs, but whether the direction of the bias is different.

As for relevant empirical results, Coen & Katsaitis (2013, pp. 1110-1111) examine hypotheses related to the supply and demand for information. Their unit of analysis is business and civil society groups which have self-registered in the Commission's Register of Interest Representatives. They examine whether business groups have registered the policy areas that regulatory DGs deal with as fields of interest in this register. The same is done for civil society groups and distributive DGs. Coen & Katsaitis (2013, pp. 1111-1114) find a correlation between the two. Business groups have mainly registered the same policy areas as those regulatory DGs work with. Civil society groups and distributive DGs also have concurrent interests.

However, the hypothesis has not been examined with regards to a specific activity, such as online consultations. There are fewer costs associated with self-registering as an interest representative than there is to actually participate. The picture emerging from Coen & Katsaitis' (2013) study may be more optimistic than what is the case for participatory activities. Moreover, there are some potential problems with the measurement validity in the study. There are no limits on how many policy domains the actors can select as interesting (Coen & Katsaitis, 2013, p. 1110). This can result in a skewed correlation. The registered actors can have ticked off many areas as relevant, when in reality they are more specialized (Greer et al., 2008, pp. 427-428). Examining the abovementioned hypothesis in relation to online consultations can therefore illuminate whether the results hold when these aspects are different.

2.2.3 EU activity

Moving on to the new, broader cost-benefit perspective added in this thesis, several studies have suggested and found that there is a connection between the levels of government activity and interest group activity (Coen, 2007, p. 336; Greenwood, 1997, p. 9; Mahoney, 2004, pp. 457, 460; Rasmussen et al., 2014, p. 255). Applying this insight here, EU activity may have an effect on the degree of bias. Interest groups have limited resources. Assuming that one of their goals is to be heard and have influence, they will attempt to sway the decision-makers in the EU. One way to do this is by contributing to online consultations (Quittkat & Kotzian, 2011, pp. 404-405). To make the most of this strategy, it may be more relevant to contribute to consultations pertaining to areas where the EU is more active.

In the areas where the EU is less active, it potentially also shares competences with the

member states or only have a supporting function. Less might be gained by participating in consultations that fall in these policy areas, because the EU does not make binding decisions here. Participating in these consultations is therefore assumed to be more relevant for those that are affected to a large degree, as for instance an indirect strategy to gain influence at the national level. For groups that are less affected, a more cost-effective strategy could be to seek influence at the national level instead. Consultations in areas where the EU is less active are therefore expected to be more biased. At the EU level, resources can be more wisely used by contributing to consultations in areas where the EU is more active and makes binding decisions. All affected groups may therefore be more inclined to contribute in consultations that fall in areas where the EU is more active, resulting in less biased participation here.

Hypothesis 3 *Consultations in areas where the EU is less active are more biased than consultations in areas where the EU is more active.*

Of relevant previous research pertaining to this hypothesis, Mahoney (2004, p. 460) finds a strong correlation between interest group activity and EU activity. EU activity is measured as the number of staff in each of the Commission's DGs. In areas with a higher number of employees, there is generally more registered interest groups. The lobby register of the Commission – Consultation, the European Commission and Civil Society (CONECCS) – is used to measure interest group activity (Mahoney, 2004, p. 445). In contrast to the register used by Coen & Katsaitis (2013, p. 1110) described above, the groups that have chosen to register in CONECCS are regularly participating in committees or hearings organized by the Commission (Wonka et al., 2010, p. 464). However, committees and hearings may not be as open as online consultations since participation usually requires an invitation. Online consultations are, on the other hand, generally open for everyone (Quittkat & Kotzian, 2011, p. 402). Exploring this hypothesis therefore sheds light on activity's effect on bias in a specific participatory activity, and one that is a most likely case for unbiased participation in addition.

Furthermore, Rasmussen et al. (2014, pp. 259-260) find that more legislative activity generally leads to a higher number of interest group participants. Again, the focus is on the number of groups rather than the distribution of interests. As argued in relation to the salience variable, however, a correlation between activity and the number of interest group participants could also imply that activity affects the degree of bias.

2.2.4 Resource demanding consultations

Some characteristics may make participating in consultations more resource demanding, affecting the groups' cost-benefit analyses in a similar vein as with the degree of EU activity. Contributing to online consultations requires time and information, and potentially staff or money to obtain relevant information (Quittkat, 2011, pp. 660, 668). Here, how the format of a consultation and how technically complex it is make participating more or less resource demanding is elaborated upon.

Format

The format of a consultation is one of the features that might affect how resource demanding contributing is. Contributing to a standardized consultation, where one is to fill out a questionnaire, may require fewer resources than participating in non-standardized consultations, where the questions are open-ended. First of all, filling out standardized questionnaires likely requires less time. Checking a box is less time-consuming than formulating an answer to a question. Second, standardized questionnaires require less information since the starting point is not a blank sheet. The possible answers can act as a frame and provide guidance. This implies that one can participate even without much prior knowledge about the particular subject of a consultation (Quittkat, 2011, pp. 661-662, 665). Standardized consultations therefore open up for more positive cost-benefit analyses compared to non-standardized ones. Less affected interests might be more likely to contribute, resulting in less biased participation.

Hypothesis 4 *Non-standardized consultations are more biased than standardized consultations.*

This has been examined empirically in relation to participation rates, but not the distribution of substantive interests. Quittkat (2011, pp. 656-657) finds that all types of actors are more prone to participate in standardized online consultations (Quittkat, 2011, pp. 661-662, 665). That all types of actors are more prone to participate indicates that the actor bias is lower, which Rasmussen & Carroll (2014, p. 454) find is associated with a low degree of participation bias. This implies that the format of a consultation could affect the distribution of interests as well. Moreover, Quittkat (2011, pp. 656-657) only include online consultations organized by

two particular DGs. These were chosen for their eagerness to involve civil society, and may therefore display the most optimistic tendencies. This thesis includes a wider range of DGs, making the results more general.

Technical complexity

The technical complexity of a consultation is another aspect that may affect how resource demanding participating is. Technical complexity is defined by the participants' need to possess specialized knowledge to be able to participate in a meaningful way (Gormley, 1986, p. 598; Ringquist et al., 2003, p. 145). It thus refers to the content of a consultation, and not the consultation process. Rasmussen & Carroll (2014, p. 447) highlight the aspect of technical complexity when distinguishing between policy types, mentioning that regulations often are more technically complex than distributive policies. They do not go beyond this mention, however, as is done here. How technical complexity can affect actors' cost-benefit analyses and their incentives to participate in lobbying have been explored previously (Broscheid & Coen, 2007, pp. 351-352, 358; Ferretti & Lener, 2008, p. 513; Quittkat, 2011, pp. 660-661; Steffek & Ferretti, 2009, p. 47; Yackee & Yackee, 2006, p. 131). The theoretical insight from these studies is here applied to participation bias in online consultations.

To understand the content of a technically complex consultation and be able to provide relevant comments, interest groups with expertise in the given field have an advantage. They are likely to already possess the relevant information needed to participate. For instance, interest groups representing paper manufacturers can most likely easily participate in a consultation concerning the chemicals contained in paper. To be able to participate without already possessing relevant information, resources like money or staff may be needed to gather it. This will make contributing more costly. Less affected groups may in such a case find that the costs exceed the potential benefits of contributing. This makes wide-ranging interest representation in technically complex consultations less likely. Participating in less complex consultations, on the other hand, is not assumed to require specialized knowledge. This makes them less resource demanding, which opens up for more positive cost-benefit analyses and less biased participation outcomes.

Hypothesis 5 *Consultations concerning technically complex matters are more biased than consultations concerning less complex matters.*

Technical complexity has not been explored systematically in relation to participation bias previously. There are nevertheless some relevant case studies, examining e-participation rates in consultations concerning the authorization of GMO products. This is a technically complex area, where being able to bring relevant contributions to the table requires expertise. These studies have shown that equal participation opportunities – openness – do not necessarily result in participation from other actors than experts (Ferretti & Lener, 2008, pp. 507-508, 510-511, 521-522; Quittkat, 2011, pp. 660-661; Steffek & Ferretti, 2009, pp. 44, 47-49).

2.2.5 Technical complexity and salience

Technical complexity's effect on bias may depend on how salient the content of a consultation is. Gormley (1986, pp. 597-598, 603) develops a theory in the field of regulatory politics, where the interaction between technical complexity and salience is in focus. As mentioned above, salient issues are understood as issues that are prominent on the policy agenda and affect a large number of people – the general public (Gormley, 1986, p. 598; Lowry & Joslyn, 2014, p. 154). The theory is used to deduce which actors are likely to participate in policy processes concerning regulations. Eshbaugh-Soha (2006, p. 239), however, finds that hypotheses derived from the theory also hold water for other types of policies. It is therefore not necessarily limited to regulatory politics. In other words, the theory can be used to derive expectations regarding participation patterns in policy processes. Of the participants, citizen, business and occupational groups are relevant here[3] (Gormley, 1986, pp. 603-605).

Expectations about who will participate when are based on the different groups' incentives and abilities to engage in the policy process. This relates to the cost-benefit perspective – participation depends on the groups' resources and concerns. Citizen groups are expected to be more likely to participate if the issues are salient but not complex. With technically complex issues they face more barriers – they are not assumed to possess a lot of technical know-how. They are, however, assumed to participate when issues are salient because they represent the concerns of citizens (Gormley, 1986, p. 604). An example Gormley (1986, p. 600) uses here is regulations concerning gun control. This issue is not especially technically complex, but it is

[3]Gormley (1986, pp. 603-605) includes "professionals". Occupational groups organize professionals (Rasmussen & Carroll, 2014, p. 450), however, and this type of group is therefore relevant to operationalize these participants here.

highly relevant for the general public.

Concerning business groups, the expectation is that they will participate in the policy process irrespective of the issue's degree of technical complexity and salience. They are, however, assumed to be especially prone to contribute if the issue is less salient. Business groups are likelier to have the necessary resources – such as information, money or staff – to participate in policy processes concerning technically complex issues, but their participation is expected to be stigmatized. By participating in less salient consultations, they are not as likely to draw attention to themselves (Gormley, 1986, pp. 604-605). Antitrust regulation is used by Gormley (1986, p. 600) as an example of an issue that is complex and less salient, and thus where business groups would be expected to participate.

Lastly, occupational groups are expected to participate if the issue is technically complex as well as salient. An example of such an issue is where power plants should be situated. Occupational groups are likely to possess expertise in the field of the profession they represent. They are therefore assumed to have the resources to participate in complex consultations. The participation of occupational groups is furthermore not likely to be as frowned upon by the public, since they are participating based on their expertise. This expertise may furthermore help justify unpopular outcomes. They therefore do not need to shy away from participating in consultations concerning salient matters (Gormley, 1986, pp. 600, 603-606).

Applied to the context of online consultations, participation bias is expected to vary with a consultation's degree of salience and technical complexity – an interaction effect. As summarized in Table 2.1, each combination is expected to result in biased consultations, but in different directions. The ensuing hypothesis is therefore also related to the direction of the bias, rather than the general degree of it. Business groups are expected to participate in consultations concerning less salient and both less complex and complex issues. Citizen groups are expected to participate when issues are salient and less complex. Occupational groups are expected to participate in consultations concerning technically complex and both less salient and salient issues.

Hypothesis 6a *Technically complex and salient, as well as complex and less salient, consultations are biased towards occupational interests.*

Hypothesis 6b *Technically complex and less salient, as well as less complex and less salient,*

Table 2.1: Technical complexity, salience and the direction of the bias

	High degree of salience	**Low degree of salience**
High degree of technical complexity	Occupational groups	Business and occupational groups
Low degree of technical complexity	Citizen groups	Business groups

consultations are biased towards business interests.

Hypothesis 6c *Salient and less complex consultations are biased towards citizen interests.*

Gormley's (1986) theory has previously been applied in different contexts, such as the political control of regulatory agencies (Ringquist et al., 2003, pp. 141-142, 144-146). The ideas have to a lesser extent been applied in interest group studies (Eshbaugh-Soha, 2006, p. 239), which will be the case here.

2.3 Summary of the expanded model

Table 2.2 shows all the variables included in the initial and new model. These two are merged together, and this expanded model is used when running regression analyses. There are several dependent variable measures, which are detailed in Chapter 3, since both the general degree and the direction of the bias are explored. The same expanded model is used no matter the dependent variable measure, however. Even though a variable is hypothesized to affect the general degree of bias, it can potentially affect the direction of the bias as well – and vice versa. Furthermore, Rasmussen & Carroll (2014, p. 453) examine both the general degree of bias and the bias towards business interests using their initial model – even though their expectations do not concern the direction of the bias. Whether the new variables, that are expected to affect the general degree of bias, also affect the bias towards business is therefore explored in this thesis.

As Table 2.2 shows, most of the variables are seen from the perspective of interest groups. It is assumed that interest groups decide whether or not to participate based on mainly these factors. Two of the variables – actor bias and DG type – are seen from other perspectives. The actor bias variable is not connected directly to interest groups' decision to participate or not. Instead, it connects complexity in one area (actor bias) with complexity in another (participation bias) and is thus seen from more of a macro perspective. The DG type variable concerns

Table 2.2: Variables, from which/whose perspective they are seen and their connection to the consultations

Initial independent variables	**Policy type** Distributive or regulatory Perspective: interest groups Consultations: content	**Cost structure** Diffuse or concentrated Perspective: interest groups Consultations: content	**Actor bias** Distribution of the different kinds of participating actors Perspective: macro Consultations: participants	**Interest group density** Number of interest group participants Perspective: interest groups Consultations: participants
New independent variables	**DG** Type: distributive or regulatory Perspective: DGs, interest groups Consultations: organizer	**DG** Activity Perspective: interest groups Consultations: organizer	**Format** Degree of standardization Perspective: interest groups Consultations: directly	**Technical complexity and salience** (Separate and interaction terms) Perspective: interest groups Consultations: content
Dependent variable (participation bias)	**General degree of bias** Distribution of the different types of interest groups participating in the consultations	**Direction of the bias** Proportion of business group participants	**Direction of the bias** Proportion of occupational group participants	**Direction of the bias** Proportion of civil society group participants

the interest groups' supply and DGs' demand for information, and is arguably seen from both the interest groups' and DGs' perspectives.

Table 2.2 also clarifies how the variables are connected to the consultations. Policy type, cost structure, technical complexity and salience are connected to the consultations via their content. Actor bias and interest group density are connected via their participants. DG type and EU activity are connected to the consultations by way of their organizers. Lastly, the format of the consultations concerns them more directly by way of the questions asked.

2.3.1 The chronological order of the variables

There is arguably some chronology in the ordering of the independent variables in the expanded model. The hypothesized order is shown in Figure 2.2. Each consultation is announced on the website of the DG that organizes it (Quittkat, 2011, p. 657), making the organizer its first known characteristic. Whether the DGs mainly deal with regulatory or distributive matters, as well as the EU activity in the policy area, is therefore determined before the other variables. After these, the independent variables related more directly to the consultations can be said to be determined. These are furthermore assumed to be more or less concurrent. Lastly, actor bias and the number of interest group participants are assumed to be determined closer in time to the dependent variable – participation bias – as discussed above. They are all related to who

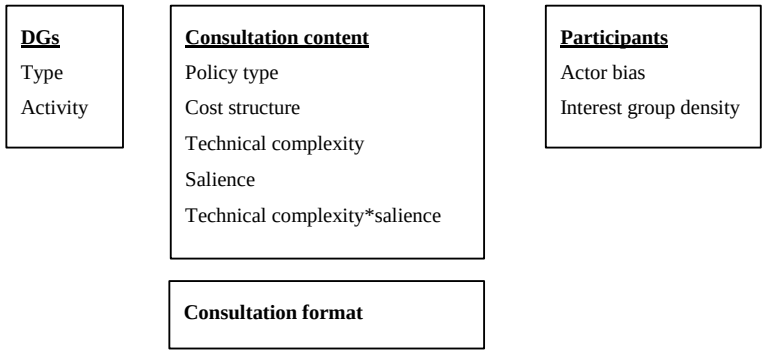


Figure 2.2: Hypothesized chronological order of the independent variables

participates in the consultations.

2.4 Summary

In this chapter, the theoretical framework has been presented and discussed. Rasmussen & Carroll's (2014, pp. 447-448) theoretical framework, mainly based on pluralist ideas, was first presented. The expectations derived from this are secondary in this thesis – space is instead given to the new hypotheses. They are nevertheless kept since they have been found to affect bias. Furthermore, alternative expectations were presented with regards to the initial policy type and cost structure variables, based on the ideas in the transactions perspective. The actor bias and interest group density expectations were moreover argued to be less interesting because of their close proximity to the dependent variable. Second, an additional aspect – salience – based on the same pluralistic reasoning was included in the expanded model.

Third, the new perspectives – an exchange theoretical and a cost-benefit one – were presented. These take the pluralistic framework one step further by considering when less affected interests are likely to be represented. The resources needed to participate are the main theme. Hypotheses concerning the type of DG organizing a consultation, EU activity, the format and technical complexity of a consultation, as well as an interaction between salience and complexity are derived from these perspectives. Most of them concern the general degree of bias, but two concern the direction of it. A demarcation is therefore made between what is likely to affect the degree and direction of bias in this thesis. All the variables are, however, included in one expanded model.

3 Data and variable measures

Having introduced the theoretical framework and hypotheses in the previous chapter, the practicalities related to examining these hypotheses are presented in this chapter. First, the datasets and adequacy of the sample size are discussed. Information about the population and sample is then provided, and potential systematic differences between them are considered. The coding of the data is subsequently described. Second, how the dependent and independent variables are measured is discussed. Some of the new independent variables get two measures, with the second one serving as a robustness check. Lastly, the control variables are described.

3.1 Data and coding

In this section, the data and coding process are described. First, the datasets and how the new consultations were identified are presented. Whether the sample size is adequate is addressed by testing the statistical power to detect effects of different sizes. Second, the population and sample are discussed. Lastly, the coding process is described.

3.1.1 Data

Two datasets are used. The first one is Rasmussen & Carroll's (2014, pp. 445, 449-450) dataset, which includes consultations from between December 2001 and April 2010 (N = 210). The second dataset is an expanded version of this, where consultations up to and including August 2011 are added. The online consultations organized after Rasmussen & Carroll's (2014, p. 449) cutoff point were identified in June 2014 by using the "Your Voice in Europe"-website, as well as the specific DG websites dedicated to online consultations. A total of 101 consultations included identifiable information about the participants between April 2010 and August 2011.

Rasmussen & Carroll (2014, p. 450) forgo coding one consultation because it is an outlier in terms of the number of received contributions. This is also the case for three[1] of the 101 new consultations. Overall, 98 new consultations are therefore added to the sample[2]. The total number of consultations is thus expanded from 210 to 308.

More consultations were coded to make up for the new variables that are added to the model. To evaluate whether the sample size is adequate given the number of variables (degrees of freedom), power analysis can be used. Such an analysis indicates one's ability to reject a false null hypothesis – that there is no correlation when, in reality, there is. An analysis with greater power is likelier to detect actual correlations compared to one with low power (Osborne, 2013, pp. 20, 26). Examining this, the power to detect small effects is shown to be poor. If there is a small correlation between the independent and dependent variables in reality, the effect will only be detected around 20% of the time in the statistical analyses. Using the same test for Rasmussen & Carroll's (2014, pp. 455-456) sample reveals a similarly low power (18%) to detect small effects. A power level of at least 80% is desirable (Osborne, 2013, p. 26). To achieve this power level regarding small effects, a sample consisting of around 1,118 consultations would be needed. Coding this amount of consultations has not been possible here, however, due to time constraints.

The power to detect medium-sized and large effects is much higher. The power levels here are 0.99 and 1, respectively. This means that small effects are likely to go undetected because of the limited sample size, whereas medium-sized or large effects almost certainly will be detected. This is discussed further in the ensuing chapters.

3.1.2 Population, sample and representativeness

By expanding the cutoff point, the population in this thesis is every online consultation organized by the Commission between December 2001 and August 2011. Every possible consultation organized in this time period has been included in the sample, following Rasmussen & Carroll (2014, pp. 449-450). The ones that are not included are ones where information about

[1] One concerning a possible revision of the Tobacco Products Directive received over 85,000 contributions, another concerning a review of the Markets in Financial Instruments Directive received 4,200 contributions, and the last one dealing with collective redress received 18,000 contributions (Directorate-General for Health and Consumers, 2011, p. 6; Directorate-General for the Internal Market and Services, 2011; Hess et al., 2011, p. 4).

[2] The participation bias in these consultations could still be affected by the factors discussed in this thesis. They have not been coded and this can therefore not be investigated.

the participants has not been published, and ones that are no longer available online. The participant information is necessary to be able to measure the dependent variable, bias. Without knowing who participated, it is not possible to identify the distribution of the different types of interest group participants. This information can be missing because the Commission has not published the results of a consultation, or because the contributions are anonymous. This is, for instance, usually the case when standardized questionnaires are used. Furthermore, the consultations have to be available online to be able to code any of the variable measures. The ones that have been deleted or lost when websites are changed or moved are therefore not included in the sample.

The sample has thus not been randomly drawn. There might therefore be systematic differences between those included in the sample compared to the population. There are few indications that this is the case when it comes to the consultations that are still available online compared to those that are not. One potential difference here could be that some DGs are more prone to remove consultations from the Internet compared to other DGs. Based on the Commission's list of online consultations, this does not seem to be the case, however. There furthermore does not seem to be a pattern when it comes to unpublished results.

The sample also differs from the population in that information about the participants has to be available. There may be systematic differences between consultations where information is disclosed and those where this is not the case. First of all, consultations without participant information often make use of standardized questionnaires which renders participants anonymous. The consultations in the sample might therefore differ systematically from the ones in the population when it comes to the format of them. The consultations not included are more likely to be standardized. This can also mean that the consultations in the sample have a systematically higher degree of actor bias compared to the ones that are not included. As mentioned, Quittkat (2011, pp. 661-662) finds that standardized consultations are more likely to lead to participation from a wider range of actors.

Given that the consultations that are not included in the sample are more likely to be standardized and have a lower degree of actor bias, a wider range of interests might be represented as well. Rasmussen & Carroll (2014, p. 454) find that a low degree of actor bias is associated with less biased consultations. If the non-included consultations are likely to be characterized by lower degrees of actor bias in general, this effect could potentially be even stronger in the

population as a whole. Furthermore, if standardized consultations are found to lead to less biased participation patterns as well, the effect could likewise have been even stronger if the standardized consultations in the population could have been included.

Overall, these factors imply that the sample may not be entirely representative of the population of online consultations. They might differ from each other when it comes to format and actor bias, which potentially means that they are different when it comes to the degree of participation bias as well. However, there are no indications that there are other systematic differences between the consultations included in the sample and those that are not. The sample might therefore still be representative of the population with regards to the other variables.

3.1.3 Coding

Since both new observations and variables are added, coding these has been necessary. The data in the initial dataset were coded by "two independent human coders" (Rasmussen & Carroll, 2014, p. 450), with an additional coder who dealt with disagreements between the main ones. This should ensure inter-coder reliability with consistently coded measures. All the new variable measures, and all the measures for the new observations, were coded by the author. This was done intermittently between September 2014 and the beginning of March 2015. Given that the coding of some of the measures requires the use of judgment, the data was coded for a second time between the end of March and in April 2015. The first coding is the one used in the analyses. The second coding is used to evaluate whether the measures have been consistently coded, in other words the reliability (Bryman, 2004, pp. 70-71). Having an independent second coder would be preferable, but has not been a possibility here. At least one month passed between the two codings to ensure some independence between them. The results generally indicate that the intra-coder reliability is good. The cost structure measure scores slightly below the threshold for what is considered satisfactory, however. The reliability of each measure will be elaborated upon below.

To code the initial variable measures, the instructions presented in Rasmussen & Carroll's (2014, pp. 450, 453) paper and their accompanying codebook were used. The new measures were coded according to a new codebook developed by the author, which can be found in the appendix. Coding the independent variable measures was for the most part straightforward.

The ones involving an element of judgment were more challenging, however, which is reflected in the intra-coder reliability of these measures. In addition, finding information about the consultations in the initial sample proved difficult in cases where the relevant consultation websites were no longer accessible. They were often retrievable via the Internet Archive, but at other times no information could be found, resulting in missing values.

3.2 Variable measures

How the variables are measured and coded is discussed in this section. The dependent variable measures are first presented and discussed, including the preparatory steps involved in the coding of them. The operationalizations of the initial and new independent variables are then detailed. Rasmussen & Carroll's (2014, pp. 453-454) measures are maintained in this thesis. They are fairly straightforward and common. The measurement validity – whether the operationalizations measure the concepts they are intended to measure (Adcock & Collier, 2001, pp. 529-531) – are therefore assumed to be good. Using the same measures also has the advantage of making the statistical results comparable to the initial ones.

3.2.1 The dependent variable

The dependent variable, bias, refers to the distribution of different types of interests in a given consultation (Rasmussen & Carroll, 2014, p. 446). This is operationalized based on the interest group participants. If every type of group is represented equally, it is by definition unbiased. If this is not the case, the consultation is defined as biased. To be able to measure this, some preparatory coding is needed. This is discussed below, before detailing the specific dependent variable measures.

Preparatory steps

To measure participation bias, it is necessary to define and separate interest groups from the other types of participants. The participants in the online consultations are therefore divided into different actor categories. Each contributor is classified as a company, public organization, European body or institution, private person, interest group, international organization or "other" (Rasmussen & Carroll, 2014, p. 450).

An interest group is defined as "(...) an association of individuals or organizations, usually formally organized, that on the basis of one or more shared concerns, attempts to influence public policy" (Rasmussen & Carroll, 2014, p. 450), which is *Encyclopedia Britannica's* definition. There is no main agreed upon definition of interest groups in the literature. However, two others, which are put forward in the context of trying to arrive at one unifying definition, are similar to the one used here. Interest groups are defined as organizations that try to influence public policy, but do not compete in elections (Beyers et al., 2008, p. 1106; Yoho, 1998, pp. 231-234). Yoho (1998, pp. 233-234) also adds that interest groups are non-governmental. This aspect is also implicitly used by Rasmussen & Carroll (2014, p. 450). When classifying the participants, public organizations and interest groups are two distinct types of actors.

These two alternative definitions are similar to the one used by Rasmussen & Carroll (2014, p. 450). Their definition can therefore be described as a fairly standard, behavioral one. It focuses on what interest groups do rather than organizational aspects such as membership (Baroni et al., 2014, pp. 143-144). Using an organizational definition would be unnecessarily restrictive in this context. The point is to be able to discern which interests are being represented, which both non-member and member groups grant.

After determining which type of actor each participant in a given consultation is, the interest groups among them are further divided into seven categories. These are trade and other business groups, occupational groups, trade unions, public interests groups, identity groups, groups including public authorities as members and "other" interest groups (Rasmussen & Carroll, 2014, p. 450). A few of these types are not as obvious as the others, and merit some additional comments. Occupational groups gather people who have the same profession that requires higher education, such as investment managers or police officers. Trade unions differ from these by collectively negotiating on behalf of their members. Public interests include groups that are concerned with the environment, consumer rights, civil liberties, government reform or tax reform, among others. Identity groups represent interests based on race, ethnicity, nationality, religion, age, sexual orientation or gender (Schlozman et al., 2012, pp. 326, 328, 333, 336).

Using this classification scheme allows for a more fine-grained categorization of the interest group participants based on substantive categories. This is fitting because the dependent variable is meant to measure the distribution of different types of interest group participants

in consultations. More categories therefore make for a more specific measure of participation bias.

For the 98 new consultations coded by the author, each actor was classified based on information found on the participant's website. Alternative sources, such as the actor's contribution, were used in cases where no website was found. Second, to code the substantive interests an interest group represents, Rasmussen & Carroll's (2014, pp. 450-451) population dataset was used. In this dataset, the interest groups registered in the EU's Transparency Register are coded according to the interest group type scheme described above. This makes coding the task of finding its assigned group type code in this dataset. The groups that are not registered here were classified based on information found on their websites.

Dependent variable measures

These two preparatory steps – classifying the actors and interest groups – are done to be able to measure the dependent variable, participation bias. Bias is measured in two different ways by Rasmussen & Carroll (2014, p. 453). The first measure is the Herfindahl-Hirschman Index (HHI), and the second is the proportion of business group participants in the consultations. HHI is calculated by taking the sum of the squared interest group type proportions for each consultation. The index ranges from a minimum of 1/7, where each of the seven interest group types are participating to an equal degree, to a maximum of 1, where only one type of interest group is participating. Higher values therefore equal a more biased participation pattern (Rasmussen & Carroll, 2014, p. 453).

Rasmussen & Carroll (2014, pp. 451-453) use the proportion measure to gauge whether any of their variables affect the bias towards business. This is done since lobbying is often posited as being dominated by business interests. The proportion of business participants is based on the sum of the trade and business as well as occupational group participants. Occupational groups are included because it makes for a more effective measure (Rasmussen & Carroll, 2014, p. 453), and likely because this type of group also represents interests related to the economic sphere. This proportion measure is also useful to examine one of the specific hypotheses related to the direction of the bias. This is H2a, that consultations organized by regulatory DGs are biased towards business interests.

H2b and H6 also concern the direction of the bias. The two initial dependent variable

measures are less useful here, however. New proportion measures are therefore used in these cases. H2b and H6c concern the bias towards civil society and citizen interests, respectively. These are here defined by concerning the collective, more diffuse interests of a large number of individuals, and by not being related to occupational or other professional interests (Berry, 1999, p. 369; Pollack, 1997, pp. 572-573). Of the interest group types in this thesis, identity and public interests are relevant to capture this. Both types decidedly represent non-economic, more diffuse interests. Using the proportion of identity and public interests group participants therefore makes it possible to examine whether the bias is tending towards civil society and citizen interests. Lastly, H6a and H6b concern bias towards business and occupational interests, respectively. The initial business proportion measure merges these two types of groups, whereas they are considered separately in the mentioned sub-hypotheses. Two separate measures are therefore coded. One for the proportion of trade and business group participants, and one for the proportion of occupational group participants.

The intra-coder reliability is high for all of the dependent variable measures. The HHI measure has an intraclass correlation coefficient (ICC) of 0.94, whereas the ICC for the proportion of business and occupational group participants is 0.99. For the three new proportion measures, the ICC is 1 for business groups, 0.99 for occupational groups and 0.98 for civil society and citizen groups. Some of the participants have in other words been coded differently the second time around, but the majority are consistently coded.

3.2.2 Independent variables

How the independent variables are operationalized and the reliability of the measures are discussed here. The measures of the initial independent variables are first presented, before discussing the operationalizations of the new ones added in this thesis.

Rasmussen & Carroll's variable measures

The independent variables in the initial model are interest group density, policy type, cost structure and actor bias. How these are measured is briefly described here, as well as how reliably they are coded for the new consultations included in the sample. The measures are summarized in Table 3.1.

Table 3.1: Rasmussen & Carroll's independent variable measures

Variable	**Measure**
Interest group density	Number of interest group participants.
Policy type	Expenditure: costs are covered by the EU or the member states. Regulatory: costs are covered by private actors. Administrative: no significant costs.
Cost structure	Diffuse: costs are spread among several industries, sectors or taxpayers. Concentrated: costs are limited to a small number of industries or sectors.
Actor bias	Sum of the squared actor type proportions.

The first variable mentioned – density – is measured as the number of interest group participants, in hundreds, in each consultation (Rasmussen & Carroll, 2014, pp. 448, 453). The intra-coder reliability is high, with an ICC of 0.998. The measure of the second variable – policy type – is categorical. The policy proposal a given consultation concerns is categorized as an expenditure, regulatory or administrative one. Expenditure policies involve costs meant to be covered mainly by the EU or the member states, whereas the costs of regulations are covered by private actors. Administrative policies do not involve any significant costs (Kardasheva, 2009, p. 115; Rasmussen & Carroll, 2014, pp. 447, 453). The Scott's pi and Cohen's kappa reliability coefficients for this measure are 0.77, which is low but acceptable. The cost aspect of a proposal is often not mentioned explicitly in the consultation document or elsewhere, and the coder therefore has to make a judgment. As evident by the relatively low reliability coefficient, this is not always straightforward. A clearer operationalization could therefore have been useful here, to establish the trustworthiness of the data to a larger degree.

The third variable is the cost structure. To measure this, the costs of the policy proposal a given consultation concerns are classified as either diffuse or concentrated. The structure is diffuse if the costs are spread among several industries, sectors or taxpayers in general. If the costs are limited to a small number of industries or sectors, the cost structure of the proposal is concentrated (Rasmussen & Carroll, 2014, p. 447). The coding of this measure is done in two steps. First, the actors the proposal is directed at are identified. Second, the coder evaluates whether the costs are diffuse or concentrated (Rasmussen & Carroll, 2014, p. 453).

This procedure makes the classification more deliberate, since step one requires the coder to consider how many actors are affected by the consultation's subject matter. The intra-coder reliability is still slightly below what is considered acceptable, with a Scott's pi and Cohen's kappa coefficient of 0.67.

The reason for this is that the costs of the proposals are not always explicitly mentioned, and coding the measure thus involves some judgment. This is also mentioned by Rasmussen & Carroll (2014, p. 453), and is the reason why the coding procedure is done in two steps. The exact number that constitutes several or few sectors or industries are not explicitly clear, and up to the coder to judge. This operationalization therefore makes it harder to reliably code the data. A clearer cost structure measure could therefore be advantageous to ensure more consistent coding, and thus more trustworthy data.

The final independent variable in the initial model is actor bias. This measure is based on the seven different actor types[3]. The Herfindahl-Hirschman Index is used, here calculated as the sum of the squared actor type proportions (Rasmussen & Carroll, 2014, pp. 447-448, 453). The intra-coder reliability for this measure is high, with an ICC of 0.99.

New variable measures

The new independent variables that are added in this thesis are the format of a consultation, the type of DG organizing it, EU activity, technical complexity and salience. How each of these is operationalized is summarized in Table 3.2. The first variable – the format of a consultation – relates to the questions the Commission wants answered. These are usually presented in the accompanying consultation document. If the questions are closed-ended, and the participants are meant to tick boxes, the consultation is classified as standardized. Non-standardized consultations have open questions, where the participants have to formulate their own answers. Consultations containing a mixture of both closed-ended and open questions are coded as semi-standardized (Quittkat, 2011, p. 661). The intra-coder reliability for this measure is good; the Cohen's kappa and Scott's pi coefficients are both 0.88. The inconsistency arises from judging whether a consultation is semi-standardized or open in cases where the consultation document is no longer available online. The format then has to be derived from other sources, such as the

[3]These are companies, public organizations, European bodies or institutions, individuals, international organizations, interest groups and "others".

contributions received or the published results.

Table 3.2: New independent variable measures

Variable	**Measure(s)**
Format	Whether the questions the Commission wants answered are open, closed-ended or a mixture of both.
DG type	1. Broscheid and Coen's (2007) categorization of regulatory and distributive DGs. 2. The proportion of distributive consultations each DG has organized.
Activity	1. Each DG's number of staff. 2. Legislative activity in the policy area a consultation falls in.
Technical complexity	How technically complex the policy area a consultation falls in is.
Salience	The EU level percentage of respondents who find a certain issue one of the two most important ones facing their country at the moment. Each consultation is linked to one issue.

The second variable concerns the organizers of the consultations, the DGs, and whether they mainly deal with regulatory or distributive matters. The DG that announces the consultation on their website is considered the organizer, following Quittkat (2011, p. 656). Each DG is classified as regulatory or distributive based on Broscheid & Coen's (2007, p. 354, 362) scheme. This is originally used in the context of exploring why the number of interests groups differs between different policy domains (Broscheid & Coen, 2007, p. 347). Three DGs – Budget, Informatics and Secretariat-General – that have organized online consultations are left uncategorized here, however.

To classify these, I use Kardasheva's (2009, pp. 102-103, 116) grouping of legislative proposals[4] in the EU between 1999 and 2007. Of the proposals that DG Budget made, 94.3% are classified as distributive. DG Secretariat-General, on the other hand, made 30% distributive proposals. The rest are categorized as either regulatory or administrative ones. DG Budget is therefore grouped among the distributive DGs, whereas DG Secretariat-General is classified as regulatory. Kardasheva (2009, p. 116) does not include DG Informatics in her sample. Based on the information this DG presents on its website, it is coded as regulatory since it mainly

[4]The 973 legislative proposals Kardasheva (2009, pp. 102-103) classifies are ones where the European Parliament have proposed amendments, because she examines the EP's influence on policy outcomes. She does not categorize all of the 1,465 legislative proposals in her sample. This might bias the results of the classification, if for instance either regulatory or distributive proposals generally are more contentious and likely to incite amendment proposals from the EP. However, the resulting classification of the DGs corresponds to what one would expect based on the DGs' responsibilities.

deals with administrative matters related to the EU institutions (Directorate-General for Informatics (DIGIT), 2014). The intra-coder reliability for this measure is 1, indicating that the first and second coding is fully consistent.

One issue with this measure, however, is that the DGs that are classified as regulatory here might mainly organize distributive consultations, or vice versa. If this is the case, the measure might not adequately capture the theoretical concept – the type of information demanded – in this context. An alternative measure, specifically related to the online consultations each DG has organized, is therefore also used. This is done to examine whether the results are robust. Inspired by Kardasheva (2009, p. 123), the number of consultations concerning distributive matters each DG has organized is divided by its total number of consultations. The intra-coder reliability here is high, with an ICC of 0.97[5].

The third new variable is EU activity – consultations in areas where the EU is less active are expected to be more biased. Activity is measured as the organizing DG's number of staff, with more employees indicating more activity, following Mahoney (2004, pp. 457, 460). The intra-coder reliability for this measure is good. The ICC is 0.98[6].

Ideally, each consultation would be ascribed the organizing DG's number of staff for the year the consultation was organized[7]. I have only been able to find staff numbers from 2006 and 2009, however (Broscheid & Coen, 2007, pp. 362-363; European Commission, 2009). This means that changes in activity levels over time will be poorly reflected. An alternative measure is therefore used to check the robustness of the results. Rasmussen et al. (2014, pp. 255-256, 259) measure government activity as legislative activity in their study of the Commission's online consultations. Each consultation is given the number of passed regulations, directives and decisions in the policy area it falls in.

I use Toshkov's (2013) policy area specific dataset of EU's legislative production for this. The policy areas are based on the 20 different directory codes used in EUR-Lex, EU's legislation database. Each consultation is connected to one of these based on its content. If it concerns

[5]The policy type measure described above is the basis for the number of distributive consultations a DG has organized. The slightly inconsistent coding of this measure is the reason why the DG type measure is not completely consistently coded.

[6]This is the result of an error. The consultations from 2006 are consistently coded with staff numbers from 2009 – not 2006 – in the first coding. This error is not repeated the second time around. Fixing this error, and running the regressions anew, does not change the conclusions related to the hypothesis, however.

[7]The staff numbers for a given year are published in January (European Commission, 2009), making lagging unnecessary.

a regulation, directive or decision, it is ascribed to the policy area the act is classified under in EUR-Lex. Otherwise, the category I deem most relevant is chosen. The consultation is then given the number of legislative acts in this policy area from the year prior to when it was organized. Lagging ensures that the activity levels precede the consultation in time (Rasmussen et al., 2014, p. 259). This measure has also been coded twice, and the resulting ICC is 0.93. This indicates that the consistency between the first and second coding is high, even though an element of judgment is involved in the process.

The fourth new variable is the technical complexity of a consultation, which is defined by the participants' need to possess specialized knowledge to be able to participate (Gormley, 1986, p. 598; Ringquist et al., 2003, p. 145). This is measured based on the complexity of the policy area a consultation falls in. Jakobsen & Mortensen (2014, pp. 1-2, 10-12) do this in a Danish context when examining how political factors affect rule production. They "(...) categorize policy domains based on a pragmatic evaluation of the degree to which it is required to have technical knowledge of the domain to be able to design effective rules." (Jakobsen & Mortensen, 2014, p. 12). I use their classification as guidance when categorizing the DGs based on whether they mainly deal with technically complex or less complex matters. This means that consultations organized by DG Taxation and Customs Union, for instance, are coded as technically complex since they concern tax or customs matters. Jakobsen & Mortensen's (2014, p. 12) classification builds on previous similar ones. The measure should therefore be able to capture the concept of technical complexity to a reasonable degree. This measure is consistently coded with Cohen's kappa and Scott's pi coefficients of 1.

An alternative measure of technical complexity could be the number of pages of the consultation document, inspired by Yackee & Yackee (2006, p. 128, 132-133) and Jakobsen & Mortensen (2014, p. 12) who use the length of rules to measure technical complexity. The consultation document contains information about the proposal, questions and other relevant instructions. However, most consultation documents are more or less the same length whether they in reality are likely to be technically complex or less so[8]. This might be because the DGs have instructions to keep the document within a certain limit, or that there is a more general norm for how long it should. Furthermore, not every consultation document follows the same

[8]For example, the document for a consultation concerning short selling is 18 pages long compared to 16 for a consultation concerning less bureaucracy for citizens.

formula. Some only include the consultation questions, for instance. What at first sight might look like a forthright measure, is therefore not found to be suited for the purposes.

The final variable in the expanded model is salience. Issues are defined as salient if they are prominent on the policy agenda and affect a large number of people (Gormley, 1986, p. 598; Lowry & Joslyn, 2014, p. 154). One way to measure this is as the media attention to a given policy area at the EU level, as for instance Neshkova (2014, pp. 22-23) does. This is not done here, however, since it is unclear whether a European public opinion exists and not just separate national ones (Rasmussen et al., 2014, pp. 256-257).

Instead, I follow the example of Rasmussen et al. (2014, p. 258) who device a salience measure to use in their study of the Commission's online consultations. Starting in 2003, Eurobarometer surveys include a question where respondents – citizens in the EU member states – are asked to choose what they consider to be the two most important issues their country faces at the moment. To do this, they are presented with a list of 15 issues. The salience measure is based on the answers to this question. Each consultation is linked, based on its content, to one of the policy areas. It is then given the EU level percentage[9] of respondents who consider this one of the two most important issues. The measure is lagged by one year to ensure that the salience levels precede the consultation in time. The intra-coder reliability for the measure is fairly high with an ICC of 0.88. This reflects that the coder's judgment is involved in the coding through linking the consultations to a policy area.

One limitation with this salience measure is that the subject matter of some consultations does not, or barely, fit into one of the 15 categories that the respondents can choose from. The additional "other" category is used when none of the others are deemed relevant. This is the same strategy Rasmussen et al. (2014, online appendix A) use. In addition, the question was first included in the Eurobarometer surveys in 2003. Consultations organized before this point are given the numbers from 2003 as a substitute. This is not optimal, but the percentages for most of the policy areas are fairly stable, indicating that it is not likely to be too much of a stretch. Despite these disadvantages, the measure is an adequate solution. It should be able to capture the first part of the salience concept fairly well – the importance the general public attaches to an issue. The second part – prominence on the policy agenda – might not be captured

[9]These are reported in the Eurobarometer reports and are made up of the average values for each member state, weighted according to population size.

as well with this operationalization. However, the list of possible issues the respondents can choose from are all fairly prominent. Examples are the economic situation, crime, terrorism and the environment. Both salience aspects are therefore included to a reasonable degree.

3.3 Control variables

Three control variables are included in the statistical analyses as well. Two of these are from Rasmussen & Carroll (2014, p. 454) and one is new. The first initial one is which year a consultation was organized. The use of online consultations has increased over time, and a learning effect might take place among potential participants. This can result in the mobilization of a broader range of interests and thus less biased consultations over time (Rasmussen & Carroll, 2014, p. 454). The second initial control variable is the length of the consultation period. This is measured as the number of weeks that passes from the date the consultation opens for contributions to the submission deadline. A shorter consultation period can make contributing challenging for groups that have to coordinate among their members before formulating an answer. Shorter consultations periods can therefore lead to more biased consultations (Rasmussen & Carroll, 2014, p. 454). Both of these control variables are consistently coded with reliability coefficients of 1.

The new control variable concerns the openness of a consultation. The target groups of a consultation are specified on each consultation's website. Most are open for everyone, some single out actors who are particularly welcome to contribute, whereas others only accept submissions from specific actors (Quittkat, 2011, pp. 659-660). The initial sample includes consultations with specific target groups, but this is not controlled for or otherwise taken into account. The openness of a consultation might affect the degree of participation bias, however. Selective consultations where some are singled out might dissuade non-targeted groups from participating, whereas closed consultations can disable groups from participating. In the latter case, the Commission organizes bias by specifying who they seek submissions from (Beyers et al., 2008, p. 1118). In addition to controlling for openness, three separate regressions are run for open, selective and closed consultations. What affects bias in open consultations might be different from selective or closed ones, because not everyone is equally welcome to contribute.

To measure openness, consultations are classified based on whether they are open for ev-

eryone, selective or closed (Quittkat, 2011, pp. 653-654, 659-660). The intra-coder reliability for this measure is acceptable, with Scott's pi and Cohen's kappa coefficients of around 0.80. In some cases, a long list of actors is included in the target group-section, making it necessary to judge whether the consultation is open for everyone or not. This is the reason why the measure is not more consistently coded.

3.4 Summary

In this chapter, the data, coding and operationalizations of the variables were discussed. The sample has been expanded by 98 consultations to a total of 308. Every accessible online consultation organized between December 2001 and August 2011, with available information about the participants, is included in the sample. There might be some systematic differences between the consultations included and those that are not, which limits the possibility of generalizing the results to the population of online consultations as a whole.

The new observations have been coded twice by the author, and the intra-coder reliability is overall good. However, the policy type and cost structure measures of Rasmussen & Carroll (2014, p. 453) are less consistently coded because of the judgment involved in the process. More explicit operationalizations could therefore have been useful here.

Two different kinds of dependent variable measures are used. The first is the Herfindahl-Hirschman Index, which is used to measure the general degree of bias. The second is four different proportion measures, used to gauge the direction of the bias. For the initial independent variables, Rasmussen & Carroll's (2014, p. 453) measures are kept. These are fairly straightforward and the measurement validity is therefore assumed to be good. The new independent variables are measured using operationalizations found in the interest group literature. The two initial control variables are also used in this thesis, as well as a new one measuring the openness of the consultations.

4 Descriptive statistics, bivariate correlations and statistical models

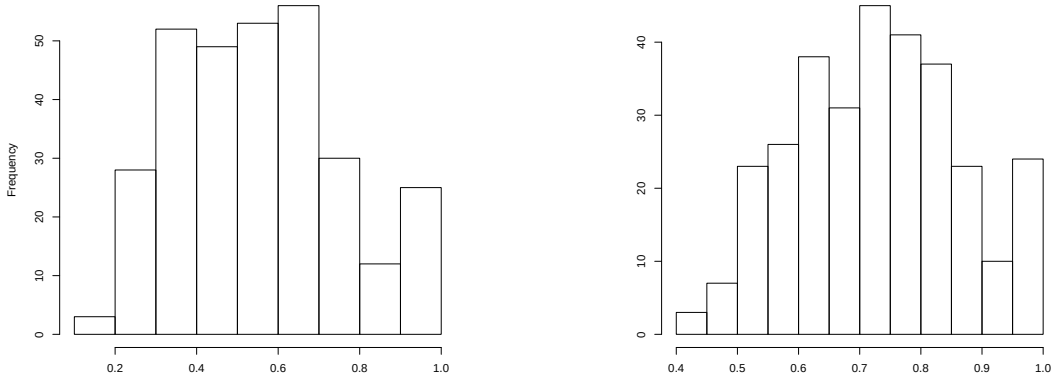
In this chapter, a first look at the data is provided. First, the descriptive statistics for the variable measures described in the previous chapter are considered. Second, the bivariate correlations between the independent and dependent variables are explored, to get an indication of whether the hypotheses are supported. Lastly, the statistical models that are used are discussed. An OLS and multilevel linear model are used when bias is measured as the HHI. Fractional regression models are used when the direction of the bias is examined using one of the proportion measures.

4.1 Descriptive statistics

Descriptive statistics for the variable measures, indicating how the data look, are presented in this section. The distribution of the dependent variables is first considered, and is furthermore compared to the distribution in the population. How the independent variable measures look is then discussed.

4.1.1 Dependent variable measures

The first dependent variable measure, the HHI, ranges from around 0.18 to 1. In theory, the lowest possible value is around 0.14. This would be a case where every type of group participated to an equal degree. When bias is measured as the HHI, none of the consultations can thus be described as completely unbiased. Figure 4.1a shows the slightly skewed distribution of this measure. Most of the consultations are characterized by a somewhat biased participa-



(a) Distribution before transformation

(b) Distribution after square root transformation

Figure 4.1: Distribution of the HHI measure before and after transformation

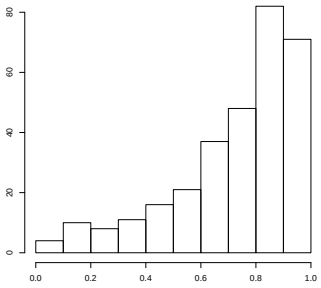
tion pattern. There is also a fair concentration of consultations that are highly biased with a HHI value of between 0.9 and 1. To get a more normally distributed measure[1], a square root transformation is found to be the most effective. The resulting distribution is shown in Figure 4.1b.

The distributions of the four different proportion measures, used to gauge the direction of the bias, are shown in Figure 4.2. As can be seen, most of the consultations are characterized by a high proportion of business and occupational group participants. Business is the main driver here, as Figure 4.2b shows. Neither occupational nor civil society interests are as prominently represented in consultations as business ones.

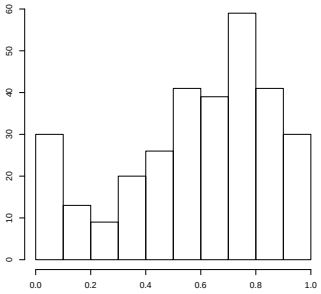
None of the dependent variable measures take the distribution of interests in the population into account. To see whether any interests are under- or overrepresented in the consultations, Rasmussen & Carroll (2014, pp. 450-453) compare the distribution of interests in the consultations to that of the population. The population is operationalized as the EU's Transparency Register, where interest groups active at the EU can self-register. All operationalizations of interest group populations are bound to have flaws, but this is one of the better existing sources (Rasmussen & Carroll, 2014, pp. 450-451).

A table comparing the distributions can be found in the appendix. The comparison shows that business interests are overrepresented in the consultations compared to the population, as Rasmussen & Carroll (2014, pp. 451-453) also find. Trade unions and groups including public

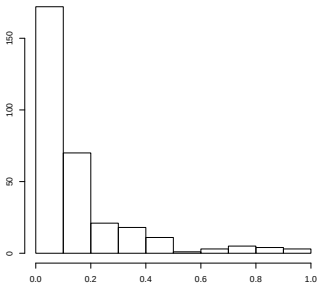
[1]This is done to fulfill an OLS requirement and is discussed below.



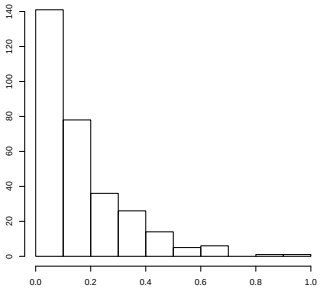
(a) Proportion of business and occupational group participants



(b) Proportion of business group participants



(c) Proportion of occupational group participants



(d) Proportion of public interests and identity group participants

Figure 4.2: Distribution of the proportion measures

authorities as member are better represented in the consultations than in the population as well. Public interests groups are, on the other hand, underrepresented in the consultations. These results indicate that the representation of interests in online consultations are biased compared to the population as well, and not just in the general sense that the dependent variable measures capture.

4.1.2 Independent variable measures

Table 4.1 shows the minimum, maximum and mean values, as well as the standard deviation, of each independent variable measure. First, the mean value of the salience measure is approximately 18, which is a fairly high degree of salience. The standard deviation of around 14 indicates variation as well, however. Second, both the main and robustness DG type measure, with mean values of 0.08 and 0.21, respectively, show that most consultations are organized by regulatory DGs. Third, the activity measures indicate that most consultations are orga-

Table 4.1: Descriptive statistics for the independent variable measures

Variable	**Minimum**	**Maximum**	**Mean**	**Std. dev.**
Salience	1	41	17.99	14.17
DG type	0	1	0.08	0.27
DG type: Proportion of distributive consultations	0	1	0.21	0.20
Activity: DG staff	355	1,893	745.50	276.53
Activity: Legislative activity	8	493	82.65	66.07
Format	0	2	1.84	0.37
Technical complexity	0	1	0.80	0.40
Policy type	1	3	1.61	0.80
Cost structure	0	1	0.55	0.50
Actor bias	0.23	0.92	0.40	0.11
Interest group density	0.01	4.29	0.45	0.53

nized by DGs with a high number of employees, or fall in policy areas where a fair amount of legislation is passed. Fourth, the format statistics show that most of the consultations are standardized. Only one of the consultations in the sample is completely standardized, and is therefore merged with the semi-standardized ones to get a more representative comparison in the subsequent analyses. Lastly, technical complexity ranges from 0 (less complex) to 1 (complex). The mean value here is 0.8, which shows that most of the consultations fall in technically complex policy areas.

The remaining four independent variable measures in Table 4.1 concern Rasmussen & Carroll's (2014, pp. 447-448, 453) initial independent variables. First, the mean and standard deviation of the policy type measure show that a fair amount of consultations concern regulatory proposals. Fewer are either distributive or administrative. Second, the cost structure measure is more evenly distributed with a mean of 0.55. Nearly half are thus coded as diffuse (0) and the other as concentrated (1). Third, the consultations are generally characterized by some degree of actor bias and, as the low standard deviation shows, this does not vary much. Lastly, the mean number of interest group participants in the consultations is 45, with a fair amount of variation.

4.2 Bivariate correlations

Exploring the bivariate correlations between the independent and dependent variables gives a basic idea of whether the hypotheses are supported, and an indication of what the results of the regression analyses will be. Here, the correlations related to the new hypotheses concerning the general degree and direction of the bias are discussed. The bivariate correlations between the initial independent and dependent variables are all in the expected direction. They are not discussed, however, since the related expectations have been examined by Rasmussen & Carroll (2014) previously. Priority is instead given to the new hypotheses. Furthermore, H6 concerns an interaction effect and is therefore not considered here.

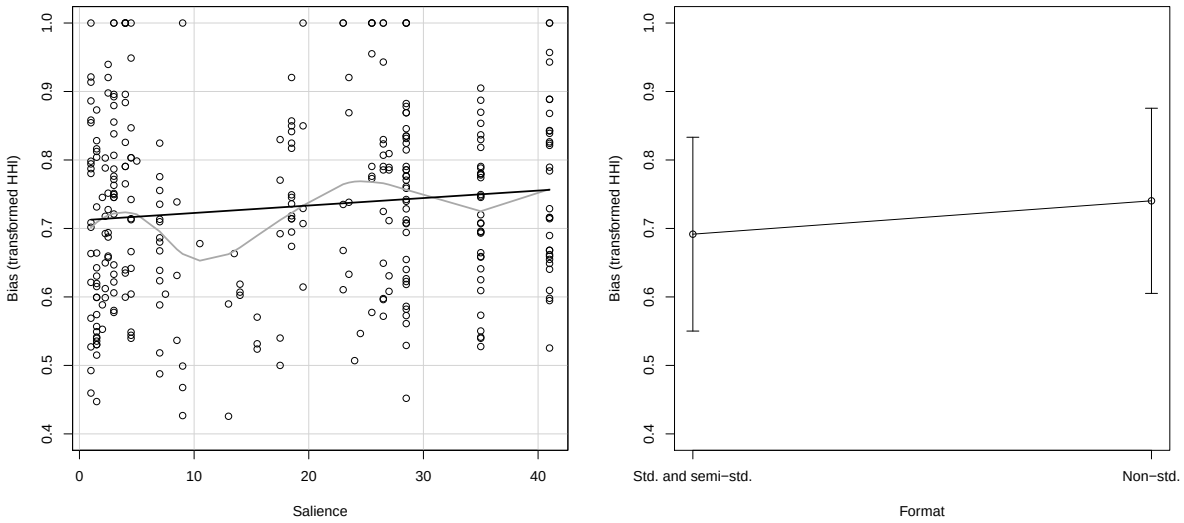
General degree of bias

The general degree of bias is measured using the square root transformed HHI measure. The first of the new hypotheses, H1, concerns salience. The scatterplot, Figure 4.3a, shows that it is not supported. Salient consultations are generally slightly more biased than less salient ones, which is the opposite of the expected effect. Figure 4.3b shows the results related to H4. Standardized consultations are expected to be less biased than non-standardized ones. The mean of standardized and semi-standardized consultations is slightly lower than that of non-standardized ones, lending some support to H4.

H3 states that consultations in areas where the EU is less active are more biased than consultations in areas where the EU is more active. As can be seen in Figure 4.4a, the very slight correlation is negative, as expected, when activity is measured as the number of DG employees. However, the lowess line shows a non-linear correlation. Transforming the measure to take this into account, results in a slightly positive correlation. As seen in Figure 4.4b, there is also a very slight, positive correlation between activity and bias when using the robustness measure of activity[2]. Overall, this does not lend support to H3.

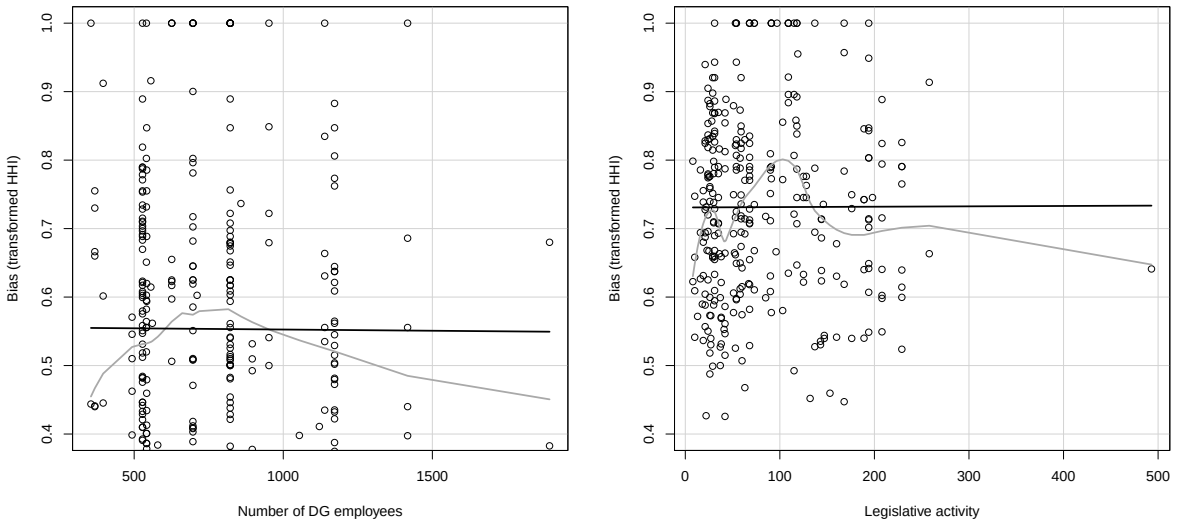
Lastly, H5 is the expectation that technically complex consultations are more biased than less complex ones. This is supported by the bivariate analysis. As Figure 4.5 shows, going from less complex to technically complex consultations results in a relatively strong increase

[2]Transforming the measure to take the form of the lowess line into account, results in a slightly stronger positive correlation.



(a) Salience (b) Format

Figure 4.3: Bivariate correlation between a) salience and bias and b) format and bias



(a) Number of employees (b) Number of legislative acts

Figure 4.4: Bivariate correlation between activity and bias

in bias.

The direction of the bias

H2 concerns the direction of the bias and is divided in two parts. H2a expects that consultations organized by regulatory DGs are biased towards business interests. This is supported, as can

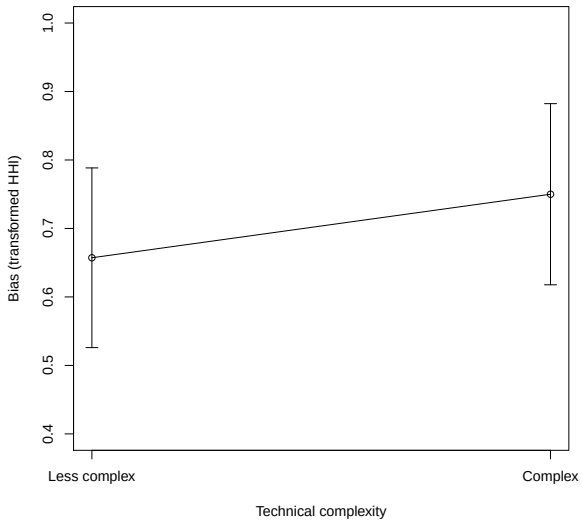


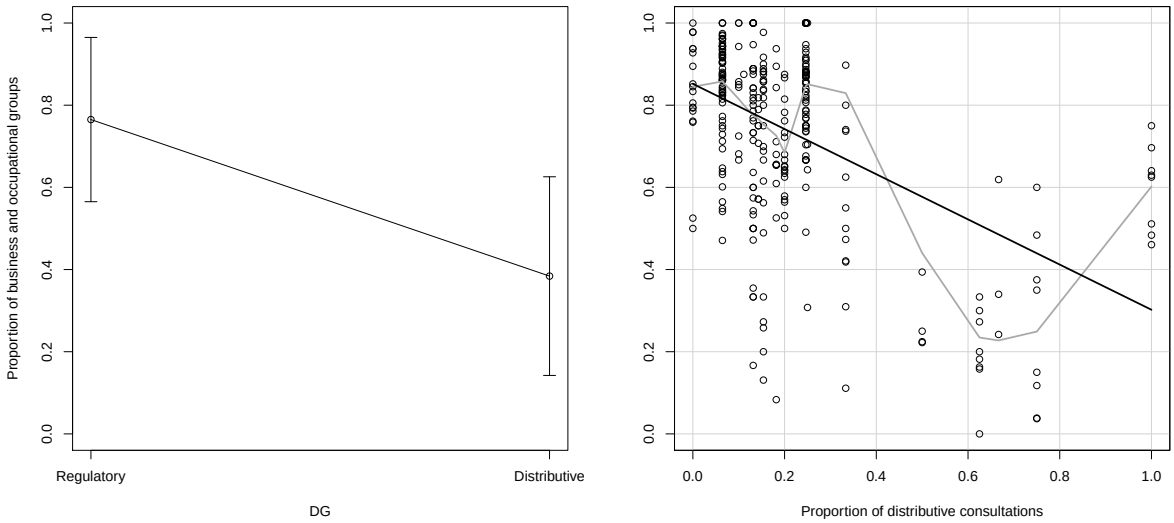
Figure 4.5: Bivariate correlation between technical complexity and bias

be seen in Figure 4.6a. The mean proportion of business interests in consultations organized by regulatory DGs is slightly below 0.8. Figure 4.6b shows that the same is the case when using the alternative, robustness measure of DG type. However, the lowess line turns upward when the proportion of distributive consultations is 1. This indicates that consultations organized by distributive DGs are fairly dominated by business interests as well.

H2b expects that consultations organized by distributive DGs are biased towards civil society interests. This gets some support, as can be seen in Figure 4.7a. The mean proportion of civil society group participants is higher for consultations organized by distributive DGs compared to regulatory ones. It is not especially high overall, however. As Figure 4.7b shows, this also holds for the robustness DG type measure. Instead, as Figure 4.6b shows, they are fairly biased towards business interests.

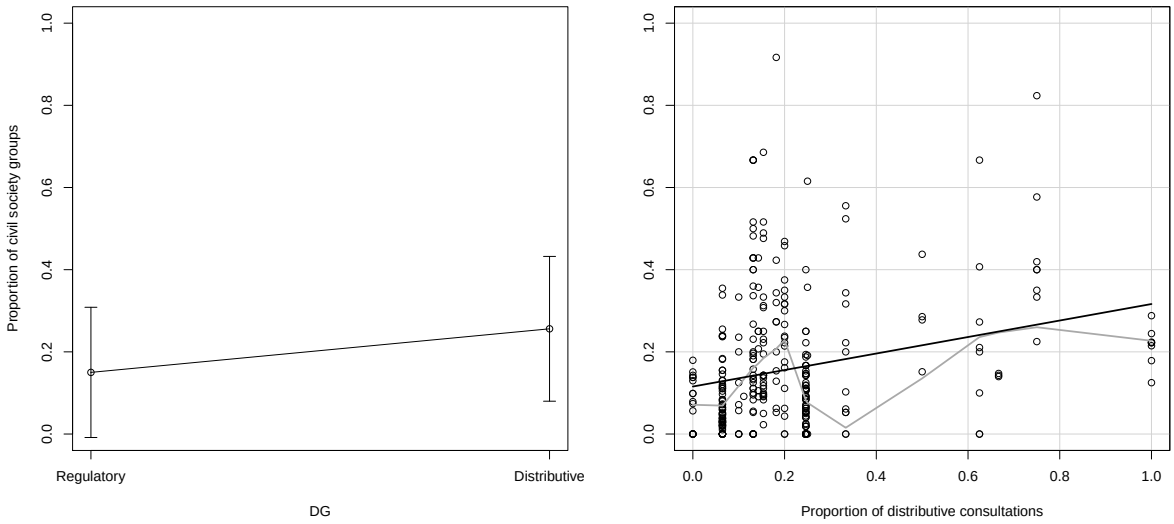
Power

The results of the bivariate analyses indicate that most of the new variables only slightly affect bias. Small effects are, however, unlikely to be detected in the expanded model because of a low power level. This increases the chances of keeping a false null hypothesis, and concluding that a variable does not affect bias when it does. To take this possibility into account, sparser models, including fewer variables, are also used when running regressions in the ensuing chapter. This



(a) Regulatory or distributive DG (b) Proportion of distributive consultations

Figure 4.6: Bivariate correlation between DG type and the proportion of business and occupational groups



(a) Regulatory or distributive DG (b) Proportion of distributive consultations

Figure 4.7: Bivariate correlation between DG type and the proportion of civil society groups

increases the degrees of freedom and thus the power to detect small effects.

4.3 Statistical models

The statistical models are considered in this section. Rasmussen & Carroll (2014, p. 454) run an OLS and a fractional logit regression. The properties that make these models suitable here as well are discussed. Furthermore, a multilevel linear model is used to take the hierarchical nature of the data into account.

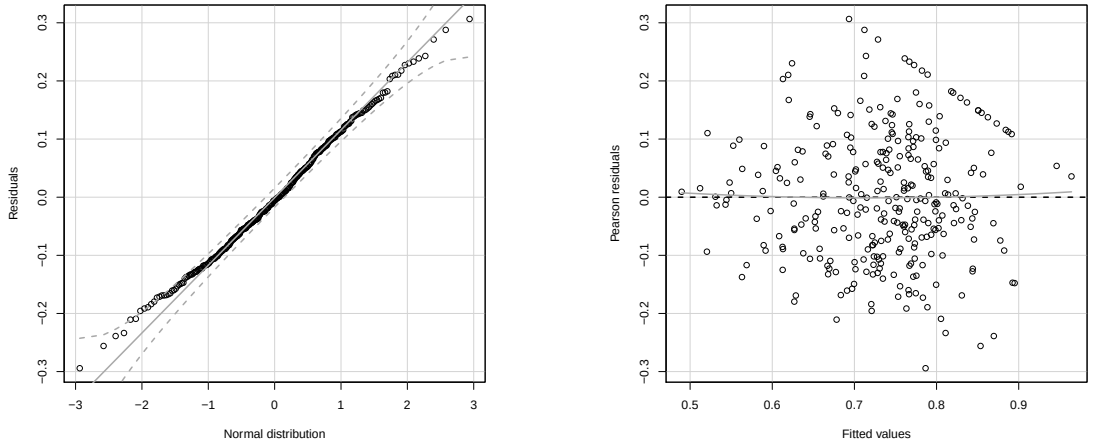
4.3.1 OLS regression

An OLS regression is run when bias is measured as the Herfindahl-Hirschman Index, following Rasmussen & Carroll (2014, p. 454). The values of this measure range from around 0.18 to 1. Even though the OLS predictions may lie outside of this bounded range, Rasmussen & Carroll (2014, p. 454) argue that it does not bias the results because of the measure's normal distribution[3]. For the expanded sample, the mentioned square root transformation of the HHI corrects the slightly skewed distribution to the largest extent possible. This approximately normally distributed measure is therefore the one used.

Other OLS requirements are also fulfilled. First, the residuals are normally distributed, as Figure 4.8a shows. Second, Figure 4.8b shows that the errors have a constant variance, fulfilling the homoskedasticity requirement (Long, 1997, p. 13). The residuals are spread out evenly. Third, the residuals are not correlated, as evident by the Durbin-Watson statistic value of 1.97. A value close to 2 indicates non-correlation. Lastly, most of the independent variables do not correlate to a large degree. All the generalized variance inflation factor (GVIF)-values are below 5, indicating that they are linearly independent (Long, 1997, p. 12). This does not hold for salience, and the interaction term between salience and technical complexity, however. Their GVIF-values are 12.87 and 16.02, respectively. That salience is correlated with itself via the interaction term is not surprising – it is entered into the model twice. The correlation can result in standard errors that are biased upward, which might yield non-significant results. This does not justify omitting the interaction term, however (Brambor et al., 2006, p. 70).

One requirement is not fulfilled. The consultations are clustered according to the DGs. The

[3]A tobit model takes the censoring and slight concentration of observations at the upper limit of the dependent variable into account (Tobin, 1958, pp. 24-25), and is run to see if the results are similar to that of the OLS regression. This is found to be the case.



(a) Residuals (b) Homoskedasticity

Figure 4.8: OLS requirements

intraclass correlation coefficient for the square root transformed HHI measure is 0.26. 26% of the variance in bias can thus be attributed to the DGs (Bliese, 2013, pp. 53, 71). This indicates that the consultations organized by the same DG correlate and are dependent[4]. This violates the assumption that errors are independently and identically distributed (Primo et al., 2007, pp. 447-448). Ignoring the clustering can lead to standard errors that are biased downward (Moulton, 1990, p. 334). Rasmussen et al. (2014, pp. 259-260) face the same issue in their analysis of the European Commission's online consultations. Their solution is to cluster the standard errors according to policy areas. However, King & Roberts (2014, pp. 1-3, 6-7) point out that this is not a sound strategy. Clustered standard errors should be used as a diagnostic tool rather than a fix. A significant difference between clustered and classical standard errors indicates that the model is misspecified – which usually also affects and biases the estimates.

4.3.2 Multilevel linear model

Instead, a multilevel linear regression is run to take the hierarchical structure of the data into account. One issue is that the sample is fairly small, with 308 consultations and 25 DGs. The average group mean reliability, which indicates whether the DG means can be reliably differentiated from each other, is 0.6. It should ideally be above 0.7 (Bliese, 2013, pp. 52-54). To improve the model performance, DGs that have organized a high number of consultations are split

[4]This dependence is also present in the initial sample – the ICC is 0.28.

into several policy area categories. Those that have organized few consultations are grouped together where possible. Details can be found in the codebook. This results in 30 DG/policy area units[5], and increases the average group mean reliability to 0.77. Some individual level two units still fall below 0.7, however, indicating that the conditions to run a multilevel model are still not optimal. The results are nonetheless presented and compared to the OLS outcomes in Chapter 5.

4.3.3 Fractional regression

The remaining dependent variable measures are the proportion of business, occupational and civil society group participants. These are bounded by 0 and 1. Rasmussen & Carroll (2014, p. 454) use a fractional regression model for their proportion measure, which is also done for the new ones added in this thesis. Fractional regression models are meant to be used for proportional dependent variable measures, and thus give predictions that lie inside of the 0 to 1 interval. Furthermore, as Figure 4.2 shows, the measures are not normally distributed. A large number of observations are gathered at the boundaries. This further underlines the appropriateness of using fractional regression models (Ramalho et al., 2011, pp. 19-20).

The consultations are also clustered according to the DGs here. The ICC is 0.59 for the proportion of business and occupational groups combined – 0.53 and 0.18, respectively, when separated – and 0.18 for the proportion of civil society groups. A fair amount of the variation in the direction of the bias can therefore be ascribed to the DGs and policy areas the consultations are nested within. However, the sample is too limited to run a multilevel model – they do not compute. Using sparser models yield random intercepts estimated to be 0, and is therefore not a relevant alternative. To get an indication of whether the results of the fractional regressions hold when taking the hierarchical structure into consideration, multilevel linear regressions are run. The proportion measures are arcsine square root transformed – a common approach when transforming proportions (Fox & Weisberg, 2011, p. 133) – to make them more normally distributed.

[5]The ICC between the dependent variable measure and the DG/policy area variable is 0.33.

4.4 Summary

This chapter gave a first look at the data, first by presenting descriptive statistics. Comparing the distribution of interests in the consultations to that of the interest group population furthermore showed that the consultations are biased compared to this benchmark as well. Second, the bivariate analyses showed that salience (H1) and activity (H3) are not supported. The format of the consultations (H4) behaves as expected, but the correlation is weak. Technical complexity has a relatively strong effect on bias, and is in the expected direction, lending support to H5. The same is the case for H2a – consultations organized by regulatory DGs are generally biased towards business interests. The bivariate analysis also lends slight support to H2b. The bias towards civil society interests in consultations organized by distributive DGs is generally not especially high, however. These results give an indication of the hypotheses that are likely to be supported by the regression results. The non-supported ones are unlikely to garner support when the other variables are included in the model as well. Furthermore, most of the correlations were weak. The power to detect small effects in a regression analysis is low when using the expanded model. To take this into account, sparser models are also used in the ensuing chapter. An OLS and multilevel linear regression are run when bias is measured as the HHI. When the direction of the bias is examined, fractional regression models are run.

5 Regression analyses

The results of the OLS, multilevel linear and fractional regression analyses are presented and discussed in this chapter. First, whether the hypotheses concerning the general degree of bias are supported by the OLS results is examined. The multilevel results are then considered. Both H5, concerning technical complexity, and Rasmussen & Carroll's (2014, pp. 447-448) expectations are supported. H4, concerning the format, is supported when a sparser model, including the new variables only, is used. H1 and H3, concerning salience and activity respectively, are not supported.

Second, the direction of the bias is considered. Rasmussen & Carroll (2014, pp. 453-456) examine whether their variables affect the bias towards business interests. Whether the new variables, expected to affect the general degree of bias, also affect the bias towards business is examined. Improving the somewhat misspecified fractional logit model, results in support to three of the initial variables. None of the new variables, with the slight exception of technical complexity, are found to affect the bias towards business interests. This indicates that there might be some differences between what explains the general degree of bias and bias towards business interests. Third, the hypotheses concerning the direction of the bias specifically are examined. Of these, H2a is supported. Consultations organized by regulatory DGs are generally biased towards business interests. H2b and H6 are not supported. Lastly, the main results are summarized.

5.1 The general degree of bias

In this section, the general degree of bias is examined. The results of the OLS regression[1], using the Herfindahl-Hirschman Index as the dependent variable, are discussed. The new hypotheses

[1] Running a tobit regression leads to the same conclusions as those based on the OLS results.

are first considered, before briefly examining the initial, secondary expectations and the control variables. One of the control variables is the openness of the consultations. As explained in Chapter 3, what explains bias in open, selective and closed consultations might differ. The results of running separate regressions for each category are also discussed. The missing data is then addressed briefly, before considering the model fit. Lastly, the results of the multilevel linear regression are discussed. The same hypotheses are found to be supported here as with the OLS results.

5.1.1 OLS regression

Table 5.1 shows the OLS results. The first model is a replication of Rasmussen & Carroll (2014, p. 455), the second is the results for the initial model using the expanded sample and the third is the results for the expanded model[2]. To illustrate the effects, Figure 5.1 shows the simulated expected value of the HHI for the expanded model as each measure changes from its minimum to maximum value. The other measures are held at their mean or median values. 95% confidence intervals are displayed.

Main hypotheses

Of the new hypotheses concerning the general degree of bias, H5 is supported. Technically complex consultations are generally more biased than less complex ones. As Table 5.1 and Figure 5.1 show, when going from less complex consultations to complex ones, bias generally increases. The effect is relatively strong. The HHI value for less complex consultations is expected to be around 0.48 on the original scale, which increases to 0.55 for complex ones. The estimate is furthermore statistically significant. Comparing the distribution of interests in technically complex consultations to that of the population furthermore indicates that they are biased in this respect as well. The distribution of interests in less complex consultations is more similar to that of the population. These results can be found in the appendix.

None of the remaining new hypotheses concerning the general degree of bias are supported. As Table 5.1 and Figure 5.1 shows, the salience (H1) estimate is nearly 0. It is furthermore not statistically significant, and thus not supported. The same is the case for H4, concerning the

[2]Year has been omitted to save space and can be found in the appendix. This is also the case for the subsequent statistical results.

Table 5.1: OLS results for the initial model (1), initial model and expanded sample (2) and expanded model (3)

	Bias (HHI)		Transformed HHI
	(1)	(2)	(3)
Policy type			
Expenditure policy	-0.108^{***}	-0.089^{***}	-0.047^{**}
	(0.037)	(0.028)	(0.019)
Administrative policy	-0.017	-0.012	-0.001
	(0.040)	(0.031)	(0.021)
Cost structure			
Concentrated costs	0.089^{***}	0.082^{***}	0.038^{**}
	(0.031)	(0.024)	(0.017)
Actor bias	0.576^{***}	0.381^{***}	0.249^{***}
	(0.121)	(0.094)	(0.063)
Density	-0.132^{***}	-0.094^{***}	-0.071^{***}
	(0.048)	(0.021)	(0.015)
Duration	-0.002	-0.004^{*}	-0.001
	(0.002)	(0.002)	(0.001)
Salience			-0.0002
			(0.002)
DG type			-0.088^{***}
			(0.028)
Number of staff			0.00005
			(0.00003)
Format			
Non-standardized			0.025
			(0.019)
Technical complexity			0.057^{**}
			(0.026)
Tech. compl. $\times$ sal.			0.0004
			(0.002)
Openness			
Selective			0.029^{*}
			(0.017)
Closed			0.017
			(0.027)
N	210	308	302
R^2	0.328	0.296	0.369
Adjusted R^2	0.276	0.258	0.315
Residual Std. Error	0.181 (df = 194)	0.176 (df = 291)	0.113 (df = 277)
F Statistic	6.323^{***} (df = 15; 194)	7.653^{***} (df = 16; 291)	6.757^{***} (df = 24; 277)

Note: The baseline categories are regulatory policy, diffuse costs, regulatory DGs, standardized and semi-standardized consultations, less complex consultations, and open consultations.
$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$

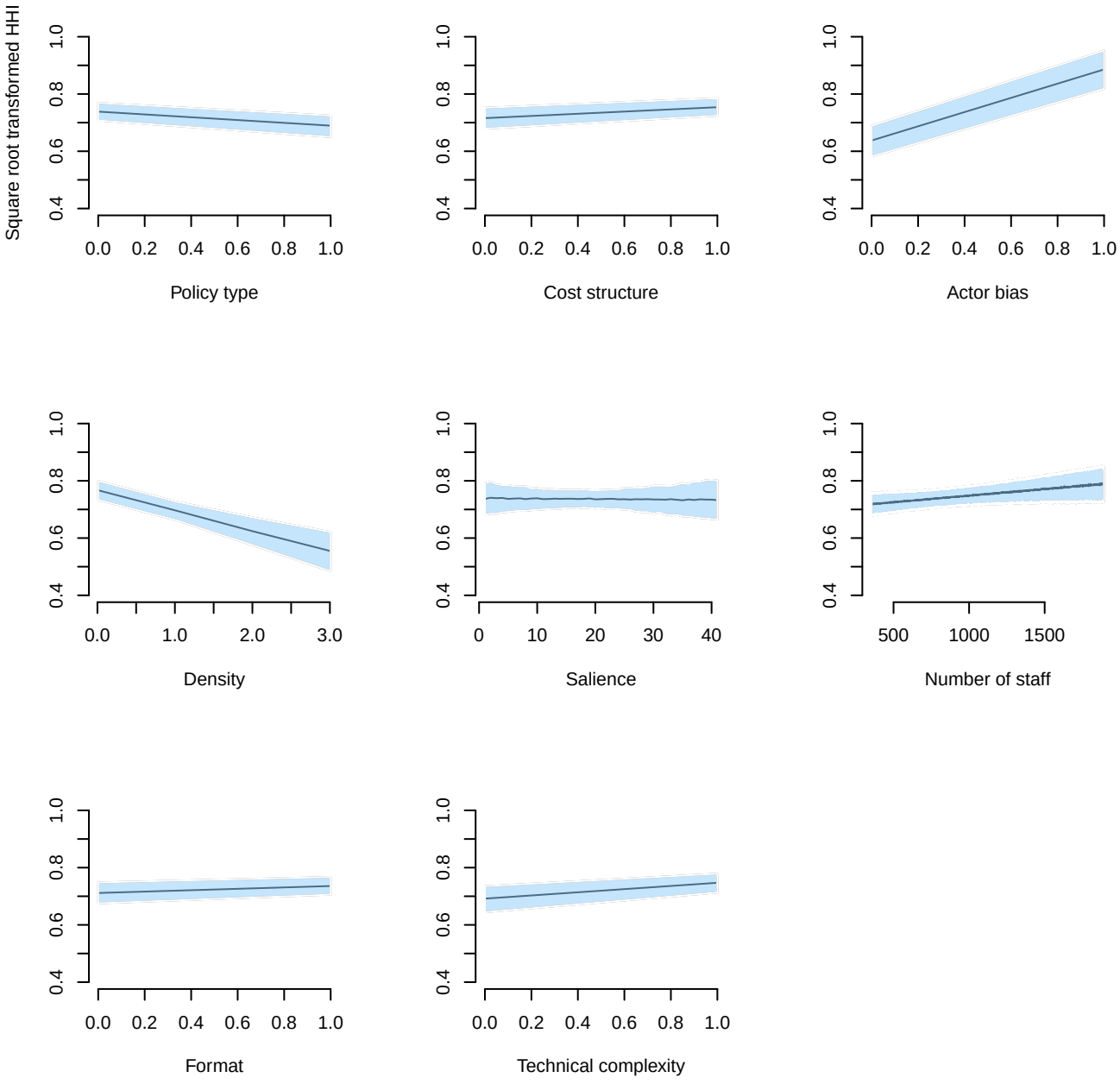


Figure 5.1: Simulated expected values of the square root transformed Herfindahl-Hirschman Index, 95% confidence intervals

format of the consultations. The estimate is in the expected direction – bias increases by around 0.03 units on the original HHI scale when going from semi-standardized and standardized consultations to non-standardized ones. The effect is neither strong nor significant, however. H4 is thus not supported.

Unlike the salience and format estimates, the activity estimate is not in the expected direction. H3 is the expectation that consultations in areas where the EU is less active are more biased than ones in areas with a higher level of activity. As Table 5.1 shows, the opposite is estimated to be the case. Consultations organized by more active DGs are associated with a higher degree of bias. The estimate is not significant, however. These results are mimicked when using the robustness measure – legislative activity. The correlation is again positive and non-significant[3]. H3 is overall not supported, as the bivariate analysis also indicated.

One thing these three non-supported hypotheses have in common is that the estimated effects are fairly small, as the bivariate analyses showed as well. The power to detect small effects is low. They are only likely to be detected around 20% of the time given the sample size and degrees of freedom. The chances of keeping a false null hypothesis – concluding that there is no correlation when there is – in the cases of H1, H3 and H4 are therefore high. To get an indication of whether this is the case, an OLS regression is run for a model including only the new variables. Including fewer variables increases the degrees of freedom, which improves the power to detect small effects to 0.35. Both the DG staff and format estimates are small, but significant in this model. The format estimate is furthermore in the expected direction, lending support to H4. H3 is again not supported – the activity estimate is not in the expected direction. The salience estimate is not significant. Running an OLS regression including only salience, results in a significant, positive estimate. This is the opposite of the expected effect. Overall, the sparser model lends some support to H4. The conclusion that H1 and H3 are not supported can furthermore be drawn with slightly more certainty.

Rasmussen & Carroll's expectations

All of Rasmussen & Carroll's (2014, pp. 447-448) expectations, related to the policy type, cost structure, actor bias and interest group density, are supported. As Table 5.1 shows, the effects

[3]Transforming the activity measures to take the potential non-linearity of the effects into account, results in the same conclusions. This also holds for the subsequent statistical results.

are slightly weaker when using the expanded sample. The same is the case when including the new independent variables. Figure 5.1 shows that the policy type and especially cost structure of a consultation do not affect the degree of bias to a large extent. A low degree of actor bias or a high number of interest group density, on the other hand, lead to a fairly steep decline in bias. That the estimates are statistically significant and fairly certain is indicated by the slim confidence intervals.

As discussed in Chapter 2, however, actor bias and interest group density are closely related to the dependent variable. They all concern the participants of the consultations, and actor bias furthermore somewhat measures the same concept as the dependent variable. These factors make the expectations less interesting than the others. They are therefore removed from the expanded model before running the regression anew. The results can be found in the appendix. The conclusions are overall the same – H5 and Rasmussen & Carroll's (2014, p. 447) expectations concerning the policy type and cost structure of the proposals are supported. Two additional estimates are moreover also significant – DG staff (H3) and format (H4). This lends further support to the conclusion that activity does not have the expected effect on bias, whereas format does.

Control variables

The effects of the control variables from the initial model – year and the duration of the consultation period – are not found to systematically affect bias. Furthermore, consultations are estimated to become more biased over time. This is the opposite of the expected effect. The duration estimate is in the expected direction, as Table 5.1 shows. As for the new control variable, openness, Table 5.1 shows that consultations with target groups are more biased than open ones. This result is statistically significant. Closed consultations are also estimated to be more biased, but not significantly. This control variable therefore behaves somewhat as predicted.

However, what affects bias could differ between open, selective and closed consultations. The consultations are therefore divided in three according to their openness, and separate OLS regressions are run for each type. The results can be found in the appendix. For the open consultations, the results of the OLS regression are very similar to the results reported above. The conclusions still hold.

Only 93 and 23 consultations are coded as selective and closed, respectively. These sample

sizes limit the power to detect both small and medium-sized effects. The results for selective consultations differ from the results for open ones. Technical complexity is no longer significant, but still in the expected direction. Interestingly, activity has a significant, positive effect on bias in selective consultations. This is the opposite of the hypothesized relationship. Another notable difference is the small, non-significant actor bias estimate. Rasmussen & Carroll's (2014, p. 448) third expectation is therefore not supported for selective consultations. None of the results for the closed consultations are statistically significant, which is not surprising with only 23 observations. Some are also in the opposite of the expected direction. Overall, these results indicate that there may be differences between what affects bias in open, selective and closed consultations. To be able to conclude anything with more certainty, however, a larger number of observations is needed.

Missing data

Six consultations are deleted when running the OLS regression for the expanded model, due to missing format and openness values. The missingness is a result of defunct websites, making the information necessary to classify these consultations unavailable. Using the multiple imputation method of King et al. (2001), with five imputed datasets, leads to the same conclusions as those described above. The same hypotheses are supported with and without imputing the missing data. The same is the case when using the alternative, robustness measures.

Model fit

Concerning the model fit, the statistics reported in Table 5.1 indicate that all three models do reasonably well. The F statistic is above 1 and statistically significant for all three models, implying that there is a relationship between the measures and bias (James et al., 2013, p. 76). Moreover, the classical and robust standard errors (not reported) are more or less the same, which indicates that the models are well-specified (King & Roberts, 2014, pp. 1-3). Some of the statistics furthermore imply that the expanded model does better than the initial one. The residual standard error is lower, indicating a better fit. Bias is expected to deviate less from the true regression line (James et al., 2013, pp. 68-69). The expanded model also explains more of the variance in bias than the initial model, as shown by the higher R^2 and adjusted R^2 values.

Since all relevant consultations have been included in the sample, the estimated standard

errors are less informative (Ward et al., 2010, p. 372). The out-of-sample prediction performance of the models is therefore also considered. The results are described in the appendix. Comparing the out-of-sample performance of the initial and expanded model indicates that the latter is slightly better at predicting new samples. Activity and salience are found to worsen the performance, however. A sparser model, excluding these, gives the best prediction rates. Overall, this indicates that more of the variance is explained by the expanded model – and even more without activity and salience – compared to the initial one on its own.

Even though the robust and classical standard errors are found to be more or less the same, this is not the case when the standard errors are clustered according to the DGs. These deviate more from the classical ones, which can bias the estimates. As shown in Chapter 4, most of the OLS requirements are fulfilled[4]. However, the assumption that errors are independently and identically distributed is violated by the multilevel structure of the data. That the consultations are clustered according to the DGs is not taken into account by an OLS model.

5.1.2 Multilevel linear model

A multilevel model takes the hierarchical structure of the data into account. Table 5.2 shows the results of the multilevel linear regression for Rasmussen & Carroll's (2014, pp. 447-448) initial model using the initial and expanded sample, as well as the results for the expanded model. The model statistics show that the initial, sparser model fits better than the expanded one. The main conclusions discussed above still hold, however[5]. H5, concerning technical complexity, is again supported but only significant at the 10%-level. The standard error is larger, reflecting that the technical complexity measure concerns the policy areas, rather than the proposals directly. The effect is furthermore estimated to be slightly stronger.

The estimates related to H1, H3 and H4 are again not significant. The format (H4) estimate is in the expected direction, however. Running a regression for a model including just the new independent variables furthermore indicates that it is significant. H4 therefore receives some support. This is not the case for H1 (salience) and H3 (activity). The salience estimate is nearly nonexistent and non-significant, as Table 5.2 shows. Furthermore, an increase in

[4]One exception is the high correlation between salience and the interaction term. Removing the interaction term still results in diverging classical and clustered standard errors.

[5]This is also the case when imputing the missing values.

Table 5.2: Multilevel linear results for the initial model (1), initial model and expanded sample (2) and expanded model (3)

	HHI		Transformed HHI
	(1)	(2)	(3)
Fixed effects			
Expenditure policy	-0.077^{**}	-0.077^{***}	-0.044^{**}
	(0.039)	(0.029)	(0.020)
Administrative policy	0.012	0.0004	0.003
	(0.040)	(0.031)	(0.021)
Concentrated costs	0.067^{**}	0.060^{**}	0.033^{**}
	(0.031)	(0.024)	(0.017)
Actor bias	0.515^{***}	0.318^{***}	0.233^{***}
	(0.121)	(0.094)	(0.062)
Density	-0.128^{***}	-0.095^{***}	-0.069^{***}
	(0.048)	(0.021)	(0.015)
Duration	0.0001	-0.002	-0.0003
	(0.002)	(0.002)	(0.001)
Salience			0.0002
			(0.002)
DG type			-0.084^{**}
			(0.034)
Number of staff			0.00004
			(0.00004)
Non-standardized			0.013
			(0.020)
Technical complexity			0.070^{*}
			(0.036)
Tech. compl. $\times$ sal.			-0.00001
			(0.002)
Selective			0.029^{*}
			(0.017)
Closed			0.012
			(0.027)
Random effect			
Intercept variance	0.006	0.006	0.002
	(0.077)	(0.075)	(0.044)
Consultation level N	210	308	302
Policy area level N	29	30	30
Log Likelihood	34.080	67.902	157.482
AIC	-32.159	-97.805	-260.965
BIC	28.089	-26.933	-160.783

Note: Standard deviation in the parentheses for the random intercepts.
$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$

salience is here expected to lead to a slight increase in bias. This is the opposite of the expected effect. The activity estimate related to H3 is also small, not in the expected direction and non-significant. This is also the case when using the alternative measure, legislative activity. These results mimic those of the OLS regression.

Rasmussen & Carroll's (2014, pp. 447-448) expectations, concerning policy type, cost structure, actor bias and interest group density, are all supported by the multilevel results as well. The estimates and standard errors are more or less the same as the ones in the OLS model. Overall, the conclusions from the OLS regression are upheld when running a multilevel linear regression.

5.2 The direction of the bias

In this section, what affects the direction of the bias is examined. Four different proportion measures are used to gauge this, and fractional regressions are run. Rasmussen & Carroll (2014, pp. 447-448, 453) do not have any specific expectations regarding the direction of the bias. They nevertheless examine whether any of their variables affect the proportion of business interests represented in the consultations, since bias is often posited as a bias towards business. Whether any of the new variables, expected to affect the general degree of bias, also affect the bias towards business is therefore examined here as well. The fractional logit model Rasmussen & Carroll (2014, p. 454) use is found to be slightly misspecified, and an improved model is therefore also used. Subsequently, the two new hypotheses that specifically concern the direction of the bias are examined.

5.2.1 Bias towards business interests

Table 5.3 shows the results of the replication of Rasmussen & Carroll's (2014, p. 456) initial model, the results for the initial model using the expanded sample, as well as the results for the expanded model. The dependent variable measure here is the proportion of business and occupational group participants.

To better be able to interpret the results, Table 5.4 shows the odds ratios. The ones labeled RC are the odds ratios in the initial model, whereas the rest are for the expanded model. Furthermore, a visualization of the results can be seen in Figure 5.2. This shows the simulated

Table 5.3: Fractional logit results for the initial model (1), initial model and expanded sample (2) and expanded model (3)

	Proportion of bus. and occ. group participants		
	(1)	(2)	(3)
Expenditure policy	-0.774^{*}	-0.966^{***}	-0.549
	(0.458)	(0.343)	(0.390)
Administrative policy	-0.500	-0.342	-0.038
	(0.501)	(0.395)	(0.430)
Concentrated costs	0.443	0.429	0.389
	(0.412)	(0.317)	(0.355)
Actor bias	3.294^{*}	1.091	1.100
	(1.951)	(1.355)	(1.410)
Density	-0.236	-0.026	-0.114
	(0.624)	(0.261)	(0.312)
Duration	-0.002	-0.011	0.017
	(0.029)	(0.025)	(0.029)
Salience			-0.008
			(0.031)
DG type			-1.197^{**}
			(0.526)
Number of staff			-0.0001
			(0.001)
Non-standardized			0.244
			(0.393)
Technical complexity			0.147
			(0.506)
Tech. compl. $\times$ sal.			0.033
			(0.033)
Selective			0.134
			(0.376)
Closed			0.065
			(0.586)
N	210	308	302
Log Likelihood	-77.513	-126.887	-106.058
AIC	187.027	287.774	262.116
BIC	240.581	351.185	354.877

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$

expected proportion of business and occupational group participants as each independent variable measure in turn changes from its minimum to maximum value, holding the other variables at their mean or median values. 95% confidence intervals are displayed.

Table 5.4: Odds ratios in the initial and expanded model

	Odds ratio	Std. Err.	z	P>|z|
Expenditure policy (RC)	0.461	0.211	-1.691	0.091
Expenditure policy	0.577	0.225	-1.407	0.159
Administrative policy (RC)	0.607	0.304	-0.997	0.319
Administrative policy	0.963	0.414	-0.089	0.929
Concentrated costs (RC)	1.557	0.641	1.076	0.282
Concentrated costs	1.476	0.523	1.098	0.272
Actor bias (RC)	26.958	52.609	1.688	0.091
Actor bias	3.005	4.237	0.780	0.435
Density (RC)	0.790	0.493	-0.379	0.705
Density	0.892	0.279	-0.365	0.715
Duration (RC)	0.998	0.029	-0.055	0.956
Duration	1.017	0.029	0.575	0.565
Salience	0.992	0.031	-0.247	0.805
DG type	0.302	0.159	-2.273	0.023
Number of staff	1.000	0.001	-0.090	0.928
Non-standardized	1.276	0.501	0.621	0.535
Technical complexity	1.158	0.586	0.290	0.772
Tech. compl. $\times$ sal.	1.033	0.034	0.998	0.318
Selective	1.143	0.429	0.356	0.722
Closed	1.067	0.625	0.111	0.912

Note: RC indicates the odds ratios in the initial model.

New variables

As Table 5.3 shows, none of the new variables that are found to affect the general degree of bias systematically affect the bias towards business interests. Technical complexity is, for instance, consistently found to affect the general degree of bias. Here, the estimate is in the expected direction – the proportion of business participants increases by around 54% when going from less complex to complex consultations – but it is not significant. The increase is slight and uncertain, as Figure 5.2 illustrates. Technical complexity is thus not found to affect the bias towards business interests. The same is the case for salience, activity and format. Running a fractional logit regression for the new variables only, confirms these results.

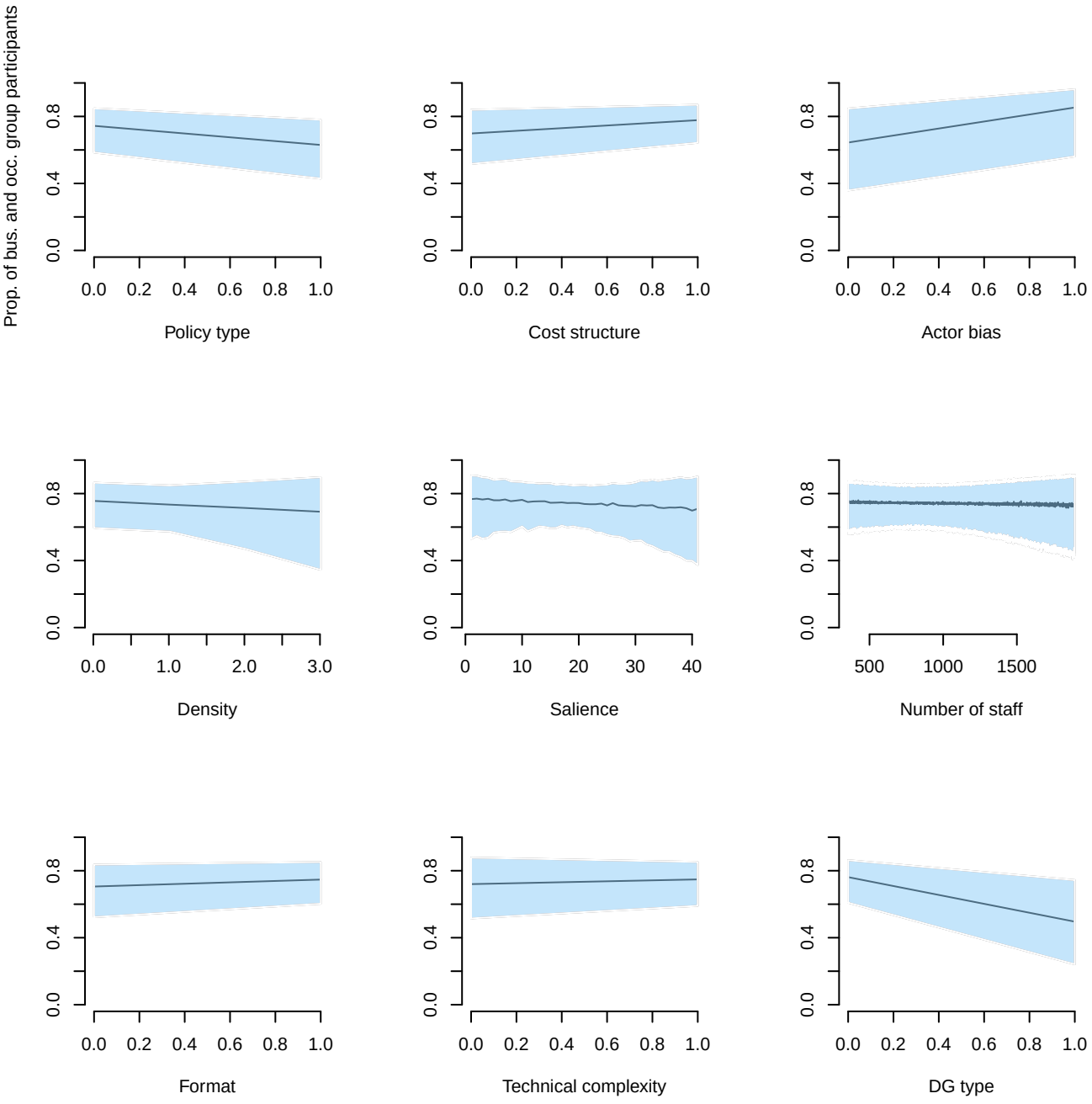


Figure 5.2: Simulated expected values of the proportion of business and occupational group participants, 95% confidence intervals

Rasmussen & Carroll's variables

The estimates related to Rasmussen & Carroll's (2014, pp. 447-448) expectations are all in the predicted directions, but are not statistically significant in the expanded model. The actor bias estimate is furthermore quasi-separated. This can be seen by the large odds ratios in Table 5.4. Quasi-separation occurs when, for instance, a high degree of actor bias is consistently associated with a high proportion of business and occupational group participants. This results in an independent variable, actor bias, that separates the dependent variable nearly perfectly (UCLA: Statistical Consulting Group, n.d.). Rasmussen & Carroll (2014, pp. 455-456) do not comment on the quasi-separation issue, but treat it as a standard estimate. That the estimate is trustworthy when it is quasi-separated is doubtful.

The quasi-separation issue may stem from the fact that the measure concerns the same concept as the dependent variable. Both measure bias, albeit different aspects. That actor bias is able to predict the proportion of business and occupational group participants to a large degree is therefore not very surprising. Removing the variable from the analysis changes the strength of some of the estimates, but not the overall conclusions. The same is the case when interest group density, which is closely related to the dependent variable as well, is removed. The results can be found in the appendix.

Improving the model

To summarize, none of the variables that affect the general degree of bias is found to affect the bias towards business interests in the fractional logit model[6]. This stands in contrast to the conclusions of Rasmussen & Carroll (2014, pp. 454, 456), who find that all but their interest group density variable affect it. One reason for this discrepancy is that they use robust, rather than classical, standard errors. There is a fairly large difference between the two kinds of standard errors in these models (not reported). In most cases the robust standard errors are at least half the size of the classical ones. Even though various tests[7] indicate that the functional form is appropriate, differences between classical and robust standard errors indicate that models are misspecified. This can bias the estimates. Misspecified models are furthermore a problem robust standard errors do not fix (King & Roberts, 2014, pp. 1-3, 6-7).

[6]This also holds when imputing the missing data.

[7]The RESET- and goodness-of-functional-form tests proposed by Ramalho et al. (2011, pp. 28, 30-33).

To improve the expanded model, the proportion measure is arcsine square root transformed – a common approach when transforming proportions (Fox & Weisberg, 2011, p. 133). This results in a fairly normally distributed measure. Subsequently, an OLS and a multilevel linear regression are run. The discrepancy between the classical and robust standard errors is lower, indicating a better fit. Several of the estimates are statistically significant. This is the case for three of Rasmussen & Carroll's (2014, p. 453) variables – actor bias, cost structure and policy type. Furthermore, technical complexity is statistically significant at the 10%-level in the multilevel model – indicating that generally, a higher proportion of business interests are represented in technically complex consultations. In sum, these more trustworthy results indicate that some of the variables expected to affect the general degree of bias also affect the bias towards business interests. There also seems to be some differences, however. This could be a result of the low power to detect small effects[8]. A larger sample is therefore needed to be able to conclude with more certainty that there are differences.

5.2.2 The direction of the bias

Whether the two new hypotheses concerning the direction of the bias are supported is considered in this section. First, H2 related to the type of DG organizing a consultation is examined. Second, the interaction effect between technical complexity and salience – H6 – is explored.

DG type

H2 is split in two sub-hypotheses. The first part, H2a, states that consultations organized by DGs that mainly deal with regulatory matters are expected to be biased towards business interests. When business interests are measured as the proportion of both business and occupational group participants[9], H2a is supported. As Figure 5.2 shows, consultations organized by regulatory DGs are generally expected to be highly biased towards business interests. The expected proportion lies slightly below 0.8. Furthermore, the slim confidence intervals indicate that this result is fairly certain. Using the alternative, robustness DG type measure – the proportion of

[8]The results of running a sparser model including the new variables only – thus increasing the power – do not differ much from the results for the expanded model. Format is, however, found to significantly affect the bias towards business (estimate 0.126, standard error 0.040) in the OLS results, but not the multilevel ones.

[9]The conclusions also hold when business interests are measured as the proportion of trade and other business groups only, excluding occupational groups.

distributive consultations each DG has organized – yields similar results, with the expected proportion of business interests lying around 0.8. An increase in the proportion of distributive consultations leads to a 16%[10] decrease in the proportion of business participants, which is significant with a p-value of 0.035. Consultations organized by regulatory DGs are thus found to generally be biased towards business interests, as expected.

These results also hold when taking the missing data into account by using the imputed datasets. Furthermore, comparing the distribution of interests in consultations organized by regulatory DGs to that of the interest group population in the EU, reveals that business interests are overrepresented in this respect as well. This comparison can be found in the appendix.

The second part of this hypothesis, H2b, is the expectation that consultations organized by distributive DGs are biased towards civil society interests. As can be seen in Table 5.4, the DG type odds ratio is 0.302. When going from consultations organized by regulatory DGs to those organized by distributive ones, the proportion of business interests generally decreases by 23%. Compared to regulatory DGs, consultations organized by distributive DGs are not expected to be as biased towards business interests. However, as Figure 5.2 shows, the expected proportion of business participants is still above 0.5 – although this is somewhat uncertain, as the wide confidence intervals display. Based on this result, which mimics that of the bivariate analysis, it is already clear that consultations organized by distributive DGs are not biased towards civil society interests to a large degree.

Using the proportion of civil society group participants as the dependent variable measure, Figure 5.3 shows the simulated expected value of this proportion when going from consultations organized by regulatory to distributive DGs. The increase in the proportion of civil society interests is small – around 59% with an odds ratio of 1.423. This is furthermore not statistically significant, as the wide confidence intervals show. The results are similar when using the alternative DG type measure[11]. Comparing the distribution in these consultations to that of the interest group population shows that identity groups are overrepresented, however. This lends some support to the hypothesis. Public interests, on the other hand, are slightly underrepresented. Overall, H2b is therefore not supported. The consultations are instead found to be characterized by a fairly high proportion of business interests, albeit less so than for those

[10]The odds ratio is 0.188.

[11]An increase in the proportion of distributive proposals leads to a 63% increase in the proportion of civil society group participants (odds ratio 1.706). The increase is not significant.

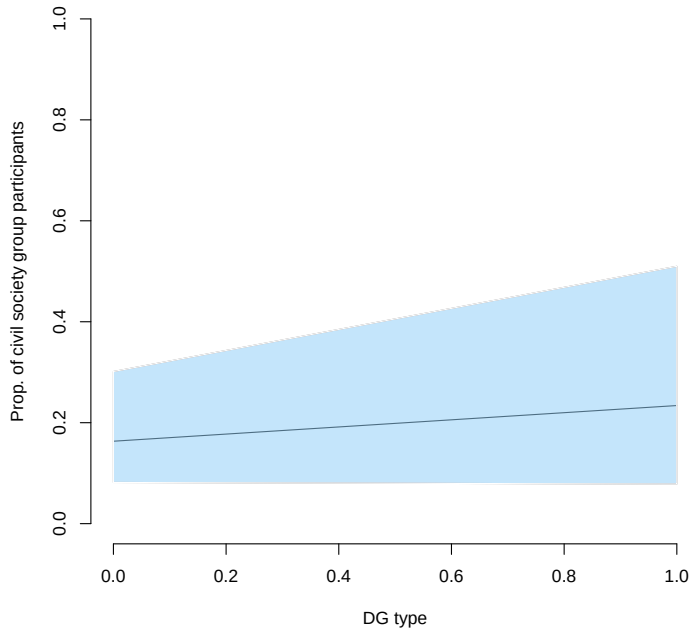


Figure 5.3: Simulated expected values of the proportion of civil society group participants, 95% confidence intervals

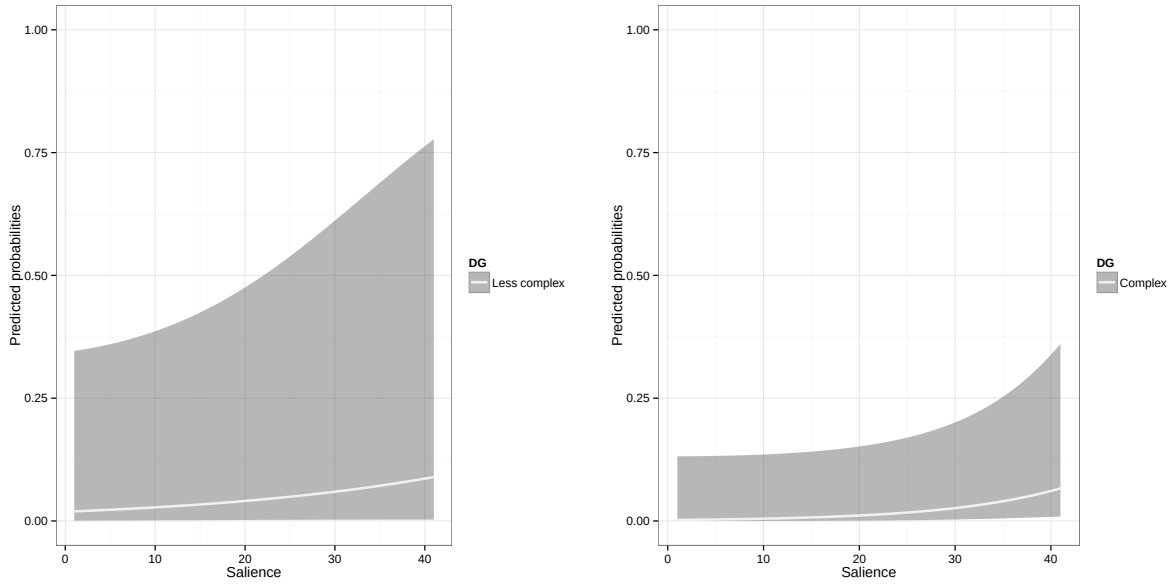
organized by regulatory DGs.

These results also hold when transforming the dependent variable measures to run multilevel linear regressions. Although this is not ideal, it is the only alternative that computes and thus gives an indication of whether the results would hold when the hierarchical structure of the data is taken into account.

Interaction between technical complexity and salience

H6 also concerns the direction of the bias. Different combinations of technical complexity and salience are expected to lead to a bias towards different types of interests. H6 consists of three sub-hypotheses, which are examined in turn here. A fractional logit regression is run for each model, with the exception of the proportion of occupational group participants where a fractional cauchit regression is run[12]. The conclusions discussed below also hold when running multilevel linear regressions, with arcsine square root transformed proportion measures.

[12]The distribution of the proportion of occupational group participants is highly skewed. The Cauchy distribution is more robust to outliers compared to, for instance, the logistic distribution (Ramalho et al., 2011, p. 23). A fractional cauchit regression is therefore found to be a better fit in this case.

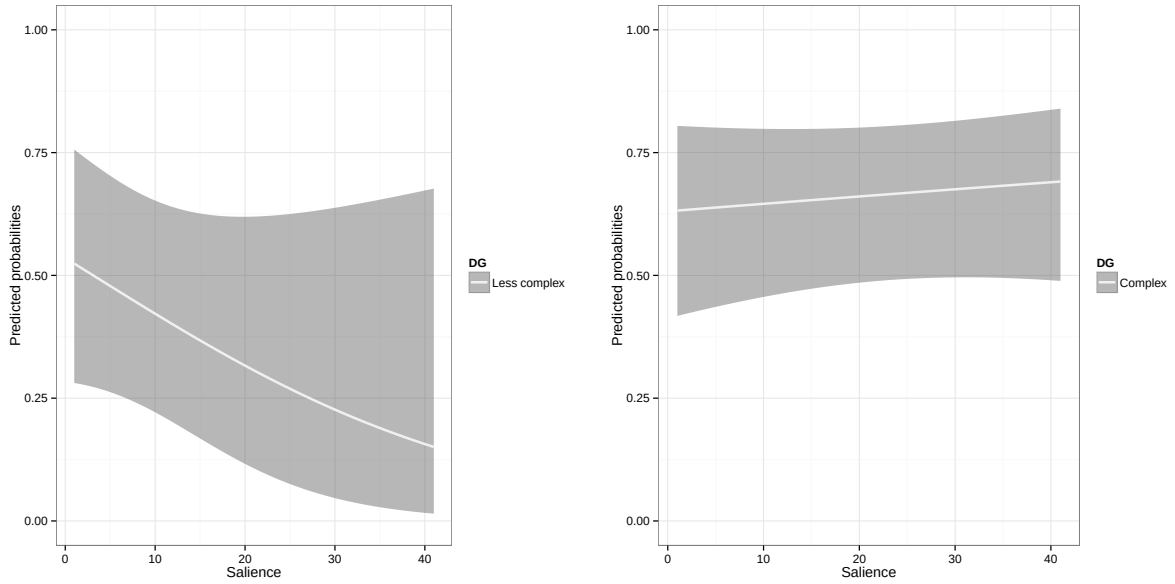


(a) Less complex (b) Complex

Figure 5.4: The effect of salience on the proportion of occupational group participants depending on technical complexity, predicted probabilities, 95% confidence intervals

H6a concerns bias towards occupational interests. Figure 5.4 shows the simulated predicted probabilities of the proportion of occupational group participants. Salience varies from its minimum to maximum value, whereas technical complexity and non-complexity are held constant. 95% confidence intervals are displayed. First, technically complex and less salient consultations are expected to be biased towards occupational interests. As can be seen, the predicted proportion is nearly nonexistent for consultations that fulfill these criteria. Second, technically complex and salient consultations are also expected to be biased towards occupational interests. The predicted proportion is slightly higher compared to less salient consultations, but is still low. Furthermore, Figure 5.4a shows that occupational groups are likelier to participate in *less* complex consultations that are both salient and less salient, than they are to participate in complex ones. The confidence intervals are furthermore wide and overlapping, indicating that the results are not significantly different from each other. Overall, H6a is therefore not supported.

H6b concerns bias towards business interests, here operationalized as trade and other business group participants only. Figure 5.5 shows the predicted probabilities. First, technically complex and less salient consultations are expected to be biased towards business interests. A high proportion of business groups is predicted to participate in such consultations, as Figure 5.5b shows. However, the proportion increases as salience increases, which is unexpected. Sec-



(a) Less complex (b) Complex

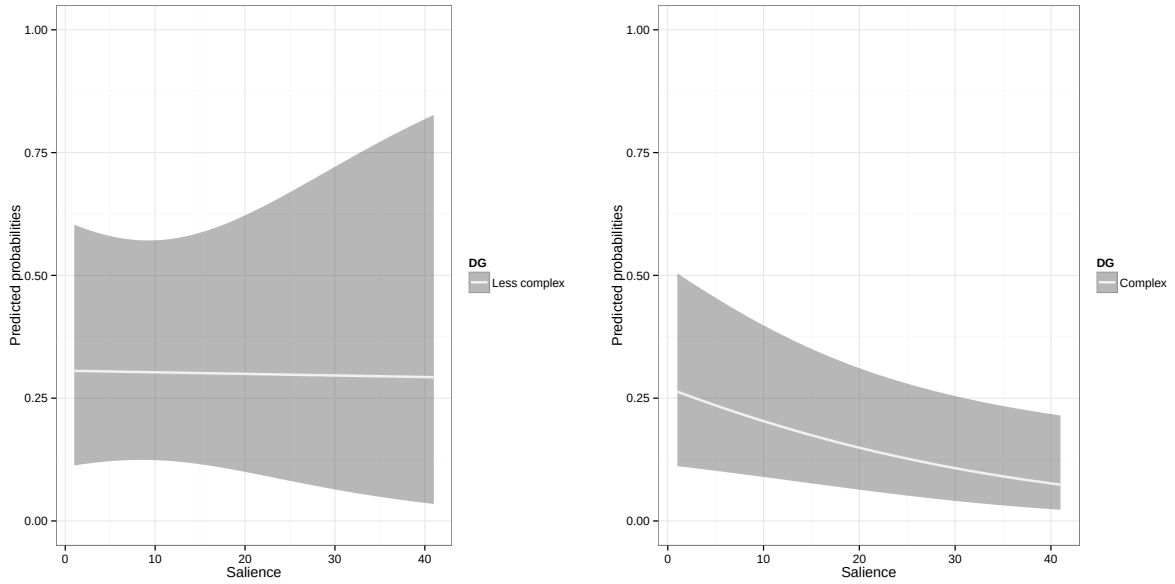
Figure 5.5: The effect of salience on the proportion of business group participants depending on technical complexity, predicted probabilities, 95% confidence intervals

ond, less complex and less salient consultations are also expected to be biased towards business interests. The predicted proportion of business groups is fairly high for less complex and less salient consultations – and it decreases as salience increases. However, the confidence intervals display that the results are uncertain and not significantly different from each other. H6b is thus not supported.

Lastly, H6c is the expectation that salient and less complex consultations are biased towards citizen interests. Figure 5.6, which displays the predicted probabilities, shows that the proportion of citizen group participants is predicted to be higher for less complex consultations compared to complex ones. The proportion does not increase with salience in less complex ones, however, but remains fairly stable with a very slight decrease instead[13]. The predicted proportion is moreover not especially high. The results are uncertain, however, as the wide confidence intervals show. They furthermore overlap, and are thus not significantly different from each other. H6c is overall not supported.

To summarize, H6 and its sub-hypotheses are not supported. The business (H6b) results are somewhat in the expected direction, which is not the case for H6a and H6c. The results are,

[13]Interestingly, the proportion of citizen groups *decreases* as salience increases for technically complex consultations. This is contrary to the expectation that citizen groups will participate when issues are important to a large part of the public.



(a) Less complex

(b) Complex

Figure 5.6: The effect of salience on the proportion of citizen group participants depending on technical complexity, predicted probabilities, 95% confidence intervals

however, not significant. One potential reason for this is that the power to detect small effects is low. To get an indication of whether this might be the case, sparser models including only salience, technical complexity and the interaction term are used. This increases the degrees of freedom and thus the power to detect small effects. Figures displaying the results can be found in the appendix.

The results of these regressions are more certain. The confidence intervals overlap to a lesser degree. First, H6a is still not supported. Occupational groups are predicted to participate more in *less* complex – rather than complex – consultations that are either salient or less so. Second, the results related to H6b are similar to the ones for the complete model. As expected, the proportion of business interests is predicted to decrease as salience increases in less complex consultations. However, the proportion is once again predicted to increase as salience increases in complex consultations. One part of H6b is therefore supported, whereas the other is not. Lastly, the results for H6c are slightly different to the ones reported above. They are now in the expected direction. The predicted proportion of citizen groups increases as salience increases in less complex consultations. The proportion is furthermore again predicted to be higher in less complex consultations compared to complex ones. These results are, however, still fairly uncertain. Overall, this lends some support to H6b. The results related to H6c are in the

expected direction, but uncertain. H6a is not supported.

5.3 Summary

A summary of the results discussed in this chapter follows. First, the hypotheses concerning the general degree of bias were examined based on the results of an OLS and a multilevel linear regression. Technically complex consultations were found to generally be more biased than less complex ones, lending support to H5. It furthermore constituted one of the stronger effects. H1, H3 and H4 were not supported in the expanded model. None of the related estimates were significant. They were, however, significant when using sparser models where the power to detect small effects was higher. Neither salience (H1) nor activity (H3) had the expected effect on bias in these cases. That they are not supported can therefore be concluded with more certainty. The format of a consultation (H4), on the other hand, behaved as expected. Going from more standardized to non-standardized consultations generally leads to a slight increase in bias. It might therefore affect bias in reality, but go undetected in the expanded model because of the low power level. To be able to conclude more decisively that it does not systematically affect bias when the initial variables are included in the model as well, would require more observations. Lastly, all of Rasmussen & Carroll's (2014, pp. 447-448) expectations were still supported, although the estimates were weaker, when the expanded sample was used, as well as after having included the new variables in the model. This indicates that their findings are robust.

Second, whether the variables expected to affect the general degree of bias also affect the bias towards business interests were examined, following Rasmussen & Carroll (2014, pp. 453-456). Based on the results of a fractional logit regression, none were found to significantly do so. This model was somewhat misspecified, however. This was taken into account. The findings based on the results of the improved models are summarized in Table 5.5. Three of Rasmussen & Carroll's (2014, pp. 447-448) variables – policy type, cost structure and actor bias – were found to affect the bias towards business. Technical complexity also had a significant effect on it, albeit only in the multilevel model. These findings indicate that there might be some differences between what explains the general degree and the direction of the bias. To be able to conclude this with more certainty, a larger sample is needed so that the power to detect

small effects is higher.

Table 5.5: Findings regarding the bias towards business interests, based on the improved models

Variable	**Affects bias towards business**	**Does not affect bias towards business**
Salience		X
Activity		X
Format		X
Technical complexity	X	
Policy type	X	
Cost structure	X	
Actor bias	X	
Interest group density		X

Third, the hypotheses concerning the direction of the bias were examined based on the results of fractional logit and cauchit regressions. H2a was supported. Consultations organized by regulatory DGs were found to generally be biased towards business interests. H2b was not supported. Consultations organized by distributive DGs were not found to generally be biased towards civil society groups. H6 was generally not supported either. Using the expanded model led to uncertain results, whereas using a sparser model led to more certain ones. H6a expects technically complex and salient consultations, as well as complex and less salient ones, to be biased towards occupational interests. Instead, the findings indicated that a higher proportion was represented in *less* complex and either salient or less salient consultations. H6b got some support. Less complex and less salient consultations were found to generally be more biased towards business interests than less complex and salient ones. However, complex and *salient* consultations were found to generally be more biased towards business interests than complex and *less* salient ones. This part was therefore not supported. Lastly, H6c was not supported. The results of the sparser model indicated that complex and less salient consultations were slightly more biased towards citizen interests, but the results were uncertain.

Other findings in this chapter included that the initial control variables – year and duration – still did not have a significant effect on the general degree of bias. Openness was found to somewhat affect it, with selective consultations generally being slightly more biased than open ones. The regressions that were run for open, selective and closed consultations separately

moreover indicated that there might be differences between them when it comes to relevant explanatory factors. However, more observations would be needed to conclude this with more certainty. Lastly, two of Rasmussen & Carroll's (2014, p. 447) initial variables – actor bias and the number of interest group participants – were removed from the model to see whether the results still held without them. This was done because they are closely related to the dependent variable. The actor bias estimate is furthermore quasi-separated in the fractional logit model. The main conclusions were found to still hold.

6 Discussion and conclusion

In this chapter, the implications of the findings are looked at more in depth. First, theoretical and normative aspects of the findings are discussed. What light the results sheds on previous relevant findings is also considered. Second, how the participation patterns align with the expectations of the pluralist, transactions and neo-pluralist perspectives presented in Chapter 1 is discussed. Lastly, the main conclusions of the thesis are summarized.

6.1 Implications of the findings

The larger theoretical and normative implications of the findings, and how they align with previous research, are discussed in this section. This is first done for the results related to the general degree of bias. Given that the conclusions regarding the initial expectations are upheld in this thesis, the implications are not discussed as this has been done previously by Rasmussen & Carroll (2014, p. 457). Instead, how the findings align with the new, transactions based expectations proposed in this thesis regarding policy type and cost structure is discussed. Second, the implications of the results related to the two new hypotheses concerning the direction of the bias are discussed.

6.1.1 General degree of bias

Four of the new hypotheses concern the general degree of bias in online consultations, in addition to Rasmussen & Carroll's (2014, pp. 447-448) expectations. Table 6.1 gives an overview of these, their theoretical foundations and whether they are supported.

Table 6.1: The hypotheses concerning the general degree of bias, their theoretical foundations and whether they are supported

Hypothesis and theory	General degree of bias
H1: Consultations concerning less salient matters are more biased than consultations concerning salient matters. **Theory:** pluralism. All affected groups are assumed to participate.	Not supported.
H3: Consultations in areas where the EU is less active are more biased than consultations in areas where the EU is more active. **Theory:** cost-benefit. The potential gains from contributing are assumed to be lower in areas where the EU is less active.	Not supported.
H4: Non-standardized consultations are more biased than standardized consultations. **Theory:** cost-benefit. Participating in standardized consultations is assumed to be more resource demanding.	Slightly supported.
H5: Consultations concerning technically complex matters are more biased than consultations concerning less complex matters. **Theory:** cost-benefit. Participating in technically complex consultations is assumed to be more resource demanding.	Supported.
Policy type (Rasmussen & Carroll, 2014): Consultations concerning regulatory policies are more biased than consultations concerning distributive policies. **Theory:** pluralism. All affected groups are assumed to participate.	Supported.
Cost structure (Rasmussen & Carroll, 2014): Consultations with concentrated costs are more biased than consultations with diffuse costs. **Theory:** pluralism. All affected groups are assumed to participate.	Supported.
Actor bias (Rasmussen & Carroll, 2014): A higher degree of actor bias is associated with a higher degree of participation bias **Theory:** pluralism. The idea that societal development leads to a more diverse interest group population is applied to consultations. When consultations are complex with respect to the types of actors participating, they are also assumed to be complex when it comes to the interests represented.	Supported.
Interest group density (Rasmussen & Carroll, 2014): Consultations are more biased if the number of participating interest groups is low. **Theory:** pluralism. All affected interests are assumed to be represented based on their wish to countervail the views of other participating groups.	Supported.

Technical complexity

As Table 6.1 shows, H5 is the only new hypothesis concerning the general degree of bias that is fully supported. Consultations concerning technically complex matters are found to gen-

erally be more biased than those concerning less complex matters. The theoretical reasoning behind this hypothesis concerns the costs of contributing to a consultation. Contributing to complex ones are assumed to be more resource demanding. This may result in a higher number of negative cost-benefit analyses, and subsequently more biased participation patterns in complex consultations. If this is what causes the correlation, the finding indicates that a factor based on a cost-benefit perspective also affects bias. Whether this is the actual causal chain is not confirmed, however, since statistical correlations do not equal causality. An aspect that nevertheless lends some support to it, is that technical complexity is determined before bias in time. It is therefore probable that technical complexity, rather than bias, is the cause of the correlation.

The two case studies described in Chapter 2 find that e-participation opportunities for the public regarding GMO products are dominated by experts (Ferretti & Lener, 2008, pp. 507-508, 510-511, 521-522; Steffek & Ferretti, 2009, pp. 44, 47-49). The broader study conducted here adds to this by finding that technically complex online consultations are biased in general. It is not limited to a specific case. This general finding arguably also has normative implications. If the Commission wants input from a wide range of interests, it is not sufficient to have the consultations be formally open for everyone. If they deal with technically complex matters, participation is still likely to be fairly biased. The interest group channel is thus not as good a source to democratic, input legitimacy in complex consultations. If a DG dealing with technically complex issues wants to attract contributions from a broader range of interest groups, taking specific action to do this can be necessary.

Furthermore, when continuing work on the proposal a consultation concerns, it may be challenging for the Commission to map out and take the interests of those that did not participate into account. If the groups participating in complex consultations are able to influence the proposals, less representative ones might therefore be the outcome. In the US, lobbying has been found to be successful "when only one or a handful of organizations lobby on a narrow technical issue of little concern to the public" (Lowery, 2007, p. 36). However, Klüver (2011, pp. 496-497) does not find that lobbying on technically complex issues in the EU systematically leads to an increase in lobby success. The estimate is in the expected direction, but it is not significant. This could nevertheless imply that the participants in complex consultations are likelier to influence the proposals in some cases, albeit not generally, compared to those

participating in less complex ones.

Overall, the bias in consultations concerning technically complex matters might be somewhat normatively problematic. The interest group channel does not provide a representative picture of the interests in society, and the outcomes may be less representative in the cases were the participants are able to influence the proposals.

Salience, activity and format

Table 6.1 shows that, of the three remaining new hypotheses, H1 and H3 are not supported. H4 is slightly supported – the format estimate is in the expected direction, and also found to systematically affect bias when using a sparser model.

First, H1 concerns the salience of the consultations. The theoretical reasoning behind this variable relates to the initial, pluralistic framework of Rasmussen & Carroll (2014, pp. 447-448). All affected interests are assumed to participate. Salient matters are important to a large part of the public, and are therefore assumed to affect a broader range of interests. Consultations concerning such issues are therefore expected to be less biased than consultations concerning less salient matters. Salience is only found to have a systematic effect on bias when using a model where this is the only variable included. Here, a higher level of salience is associated with a very slight increase in bias – the opposite of the expected effect. One reason for this could be that the less affected – or those without selective incentives, in line with the transactions perspective – do not contribute. A more in-depth analysis would be needed to be able to discern the causal workings behind salience's effect on participation bias, however.

Concerning previous results, Rasmussen et al. (2014, pp. 260-262) find that a high degree of salience is correlated with a higher number of interest group participants in the Commission's online consultations. That salience only slightly affects the degree of bias, and not in the expected direction, implies that there is a difference between what explains the number of interest group participants in a consultation and its degree of bias.

Second, H3 is the expectation that consultations in areas where the EU is less active are more biased than ones in areas with more activity. Interest groups have limited resources, making it a potentially more rewarding strategy to focus on consultations in areas where the EU is more active and assumed to have more competences. However, the findings indicate that more activity is associated with more – not less – biased consultations. One reason for this

could be that well-endowed groups might participate no matter the activity levels, which could mitigate the effect of activity on bias. This does not explain why bias is estimated to increase when activity increases, however. The positive correlation indicates that the theoretical account behind H3 is not the whole story. Other forces is likely to be at work here, which this thesis does not uncover. The finding furthermore highlights that what explains the number of interest group participants is not necessarily the same as what explains participation bias. Both Rasmussen et al. (2014, pp. 259-260) and Mahoney (2004, p. 460) find a positive correlation between EU activity and interest group activity, whereas higher levels of activity potentially leads to more biased consultations.

Lastly, H4 concerns the format of an online consultation and gets slight support here. The theoretical reasoning behind this hypothesis relates to the costs associated with participating. Consultations with a more standardized format are assumed to demand less resources than non-standardized ones, opening up for more positive cost-benefit analyses. Since only one consultation in the sample is standardized, the effect might have been stronger had it been possible to include more of them. That H4 gets slight support nevertheless implies that organizing consultations that are more standardized is one way to increase the inclusiveness.

Concerning previous research, Quittkat (2011, pp. 661-662, 665) finds that more standardized consultations are associated with a higher number of different types of participants. In other words, that the actor bias is likely to be lower in standardized consultations. Format therefore seems to have a similar effect on participation bias as it does on actor bias. This finding also hold more generally for DGs organizing consultations, as opposed to that of Quittkat (2011, pp. 656-657), who only includes consultations organized by two DGs.

Rasmussen & Carroll's expectations

Rasmussen & Carroll's (2014, pp. 447-448) four expectations are all supported, as Table 6.1 shows. They are based on different aspects from pluralist theory. Again, whether these are the actual causal workings behind the correlations cannot be concluded based on the results here.

The focus here is on the alternative expectations, elaborated upon in Chapter 2, related to policy type and cost structure. These are based on the transactions perspective. Rasmussen & Carroll (2014, p. 447) assume that distributive proposals affect a broader range of interests than regulatory ones. The same is the case for proposals with a diffuse cost structure compared to

ones with concentrated costs. Based on this, consultations concerning distributive proposals or those with a diffuse cost structure are expected to be less biased. The alternative to this, based on the transactions perspective, is that some groups have selective incentives to participate. The interests of their specific constituencies are not likely to be represented by other, similar groups otherwise. Groups representing diffuse interests, on the other hand, might not participate even if they are affected by a consultation. Several other groups could represent these interests. If every group of this type follows this logic, their interests are not going to be represented at all. Consultations concerning distributive proposals or with diffuse costs are therefore not necessarily less biased than those concerning regulations or with concentrated costs.

The statistical results lend support to the original expectations, however. Rasmussen & Carroll's (2014, p. 454) findings are upheld when including the new variables. Still, the effects are slightly weaker and the resulting decrease in bias is small in comparison to the initial findings. Moreover, the proportion of civil society group participants – which generally represent diffuse interests – is not estimated to be systematically higher in consultations concerning distributive proposals or with diffuse costs[1]. These results indicate that the alternative expectations, based on the transactions perspective, might have something to them as well. It could potentially explain why the degree of bias in consultations characterized by diffuse costs or distributive proposals is fairly similar to that in consultations characterized by concentrated costs or regulatory proposals.

This may moreover mean that not all affected interests are likely to be represented, as the transactions perspective expects. The democratic, input legitimacy that can be derived from the interest group channel can then be questioned. However, whether it is a lack of selective incentives to participate that leads to the more similar bias patterns cannot be derived from these statistical analyses. This also goes for the degree to which all affected interests are represented. Determining this would require a more in-depth analysis.

[1]Going from regulatory to distributive proposals leads to a non-significant 54% decrease in bias – odds ratio 1.18, standard error 0.56. Going from diffuse to concentrated costs leads to a non-significant 42% increase in bias – odds ratio 0.72, standard error 0.31.

6.1.2 The direction of the bias

The two remaining new hypotheses – H2 and H6 – concern the direction of the bias. Table 6.2 gives an overview of the expectations and their theoretical foundations, as well as whether they are supported.

Table 6.2: The hypotheses concerning the direction of the bias, their theoretical foundations and whether they are supported

Hypothesis and theory	Direction of the bias
H2a: Consultations organized by DGs mainly concerned with regulatory matters are biased towards business interests. **H2b:** Consultations organized by DGs mainly concerned with distributive matters are biased towards civil society interests. **Theory:** exchange. Interest groups' supply and DGs' demand for information.	H2a: supported. H2b: not supported.
H6a: Technically complex and salient, as well as complex and less salient, consultations are biased towards occupational interests. **H6b:** Technically complex and less salient, as well as less complex and less salient, consultations are biased towards business interests. **H6c:** Salient and less complex consultations are biased towards citizen interests. **Theory:** Gormley (1986), cost-benefit. Participation is assumed to depend on the groups' resources and concerns.	H6a: not supported. H6b: slightly supported. H6c: not supported.

DG type

As Table 6.2 shows, H2a is the only hypothesis concerning the direction of the bias that is fully supported. Consultations organized by regulatory DGs are found to generally be highly biased towards business interests. This is thus another expectation not derived from the initial pluralist framework that is supported in this thesis. The theoretical reasoning behind this hypothesis relates to the supply and demand for information. Business groups are assumed to be the most relevant supplier of the technical information demanded by regulatory DGs (Coen & Katsaitis, 2013, pp. 1105-1106, 1108-1109). This can make it less costly for this type of group to contribute. Although a statistical analysis cannot reveal the exact workings behind a correlation, the chronological order of DG type and bias corroborates that it is the type of DG that causes the effect. The finding furthermore indicates that the regulatory DGs are likely to get the infor-

mation they demand, which might enhance the quality of the policies and be a source to output legitimacy.

The EU mainly deals with regulatory matters (Hix & Høyland, 2011, p. 189). As the descriptive statistics presented in Chapter 4 furthermore showed, most of the DGs organizing consultations are categorized as regulatory. That the consultations organized by this type of DG are biased towards business interests to a high degree is therefore arguably normatively problematic as well. The finding highlights that the interest group channel does not reflect the range of interests in society in these consultations. That the majority of online consultations are organized by this type of DG underlines that this participatory exercise might not be as inclusive as it is meant to be. The finding therefore implies that if regulatory DGs want a more diverse set of interest groups to participate in their consultations, they might need to consider specific strategies to achieve this.

H2b is not supported. The proportion of civil society groups it not found to be systematically higher in consultations organized by distributive DGs compared to regulatory ones. Instead, a fairly high proportion of the participants is expected to be business interests. Again, the interest group channel leads to some interests being better represented than others. Theoretically, distributive DGs are assumed to demand political information, which civil society groups are mainly assumed to possess (Coen & Katsaitis, 2013, pp. 1105-1106, 1108-1109). Given the lack of support, this reasoning is not likely to be the whole story. It could for instance be that distributive DGs mainly demand technical information when organizing consultations, and that the political information they demand is supplied through other means. The findings imply, however, that distributive DGs do not mainly get the political information they are assumed to demand. The input legitimacy that can be derived from organizing consultations can therefore be questioned.

The findings related to H2 furthermore shed light on a previous study. Coen & Katsaitis (2013, pp. 1111-1114) find that business groups list the same types of policy areas that regulatory DGs deal with as interests when registering in the Commission's Register of Interest Representatives. The same holds for civil society groups and distributive DGs. However, the costs of registering are likely to be lower than participating in, for instance, online consultations. This could affect the results. The support to H2a corroborates at least part of the idea that the DG type does have an effect on the direction of the bias. The expectation is therefore

somewhat supported when a specific activity is the unit of analysis as well.

Interaction between technical complexity and salience

H6 concerns the interaction between salience and technical complexity and consists of three sub-hypotheses, as can be seen in Table 6.2. These expectations are based on Gormley (1986, pp. 597-598, 603). Participation is assumed to depend on the groups' resources and concerns. These are presumed to be divergent, which leads to the expectation that different combinations of technical complexity and salience result in a bias towards different types of interests.

H6a is not supported. Occupational groups are expected to participate in technically complex consultations that are either salient or less so, because of their assumed expertise in the field of the profession they represent. Since their participation is assumed to be based on this know-how, they do not need to shy away from participating in salient consultations. The results show that occupational groups are likelier to participate in *less* complex consultations that are either less salient or salient. Something other than the abovementioned explanation therefore seems to be at work here. In Gormley's (1986, p. 603) theory, professionals are the actor type discussed, whereas occupational groups are studied here. The results might therefore indicate that what motivates a professional on his or her own is not the same as what makes an occupational group participate.

H6b is slightly supported. The findings give some indication that less complex and less salient consultations are biased towards business interests, although these results are uncertain. However, contrary to the hypothesized effect, complex and *salient* consultations are found to be the most biased towards business interests in general. Theoretically, business groups are assumed to possess the resources necessary to be able to participate in either less complex or complex consultations. They are, however, assumed to want to avoid drawing attention to their trying to influence public decision-making (Gormley, 1986, pp. 604-605). Achieving this is presumed to be likelier in less salient consultations. That the proportion of business interests is predicted to be high in salient and complex consultations contradicts this. One potential reason for this finding is that the complexity makes the issues more isolated from the public eye. It could therefore be that they are not as likely to draw attention to themselves, even if the issues are more salient. The analysis here does not uncover whether this is the case, however.

The high proportion of business interests in less complex and less salient, as well as com-

plex and salient, consultations implies that the interest group channel does here not provide the decision-makers with a representative picture of the interests in society. This could potentially affect the representativeness of the proposals as well, if the participants are able to wield influence. As discussed above, complexity has not been found to systematically affect the chances of lobby success in the EU (Klüver, 2011, pp. 496-497). However, both Mahoney (2007, pp. 49-50) and Klüver (2011, p. 497) find that an increase in salience leads to a lesser chance of lobby success. Based on this, that business interests are well represented in less complex and less salient consultations could potentially lead to more biased proposals as well. On the other hand, the business dominance in complex and salient consultations might not be as normatively problematic since the chances of being successful are likely to be lower.

Lastly, H6c is not supported. Less complex and salient consultations are only found to be slightly more biased towards citizen interests when using a sparser model. The finding is furthermore uncertain. Theoretically, citizen groups are assumed to face more barriers in technically complex consultations, making it likelier that they participate in less complex ones. Furthermore, salient consultations are assumed to be relevant for citizen groups' constituencies, which make it pertinent for them to participate here (Gormley, 1986, pp. 603-604). Although a statistical analysis cannot indicate whether this is actually the case, the results indicate that it might be part of the story. Normatively, the finding – albeit uncertain – is arguably not particularly problematic. The proportion of citizen group participants is generally not predicted to be especially high. There is room left for other types of interests to be represented as well.

6.2 Interest group perspectives

In Chapter 1, three interest group perspectives were presented – the pluralist, transactions and neo-pluralist perspectives. It was furthermore argued that pluralistic, unbiased participation patterns were likely to be the outcome because of the relatively few barriers to entry in online consultations. Here, how the findings relate to these perspectives is discussed.

Based on the pluralist perspective, the participation patterns in the consultations are expected to be characterized by a low degree of bias. The participation patterns in consultations characterized by a low degree of actor bias or by a high number of interest group participants – two of Rasmussen & Carroll's (2014, p. 448) variables – are found to be fairly unbiased. How-

ever, they are closely related to the dependent variable in both time and what they measure, making these findings less interesting.

The second perspective – transactions – would expect the participation patterns to be highly biased. This is found to generally be the case in consultations characterized by being technically complex or organized by regulatory DGs. In the latter case business groups dominate, as expected. Furthermore, consultations characterized by different combinations of technical complexity and salience are also found to be biased towards business interests to a fairly high degree – with the exception of salient and less complex consultations. Whether the consultations are biased because affected groups do not have selective incentives to participate, as the transactions perspective assumes, cannot be determined based on the results in this thesis, however.

In addition, the consultations characterized by what Rasmussen & Carroll (2014, pp. 447-448) assume will affect fewer interests are also biased to a large degree. This is the case for consultations concerning regulatory proposals, with concentrated cost structures, a high degree of actor bias and fewer interest group participants. This could, however, be in line with the pluralist perspective. Fewer interests are assumed to be affected, meaning that they could potentially all be represented.

Lastly, neo-pluralism reads as a fusion between the pluralist and transactions perspectives. The expectation here is a somewhat biased participation pattern. This is arguably the outcome for most of the consultations characterized by the factors that are expected to mitigate bias. This holds for less technically complex consultations, ones concerning distributive proposals, as well as consultations with diffuse cost structures. The latter two characteristics are assumed to affect more interests (Rasmussen & Carroll, 2014, p. 447), but this does thus generally not lead to even more unbiased participation patterns in line with the pluralist perspective. In addition, consultations characterized by being salient and less complex are also not highly biased towards any one of the group types examined. The participation in consultations organized by distributive DGs is furthermore also somewhere between biased and unbiased.

To summarize, online consultations are generally not characterized by pluralistic participation patterns, even though they arguably constitute a most likely case for this. The characteristics that are assumed to lead to less biased outcomes do this, but only to a certain degree. The patterns are therefore more similar to what one would expect based on the neo-pluralist

perspective. Two exceptions here are consultations characterized by a low degree of actor bias and a high number of interest group participants. The participation patterns here are closer to unbiased, as the pluralist perspective expects. The characteristics that are assumed to lead to more biased outcomes also do this. The participation patterns are therefore more in line with the expectation based on the transactions perspective. This is also the case for most of the characteristics that are expected to affect the direction of the bias. However, to determine more decisively whether the findings are in line with one or the other perspective would require further research. Neither the degree to which all affected interests is represented, nor what motivates the participants, can be determined here.

6.3 Conclusion

The research question that has guided this thesis is which features pertaining to the European Commission's online consultations that can explain participation bias. A biased consultation is understood as one where not every type of group participates to an equal degree. Rasmussen & Carroll (2014) also examine this, and their study lays the foundation for this thesis. Furthermore, the pluralist, transactions and neo-pluralist interest group perspectives are more generally underlying this exploration. Online consultations are open for everyone and participating is likely to be less costly than in other participatory activities. They therefore constitute a most likely case for pluralistic participation patterns, characterized by a low degree of bias. Instead, however, patterns resembling what one would expect based on neo-pluralism are the outcome for most of the characteristics that are expected to lead to a lower degree of bias.

Rasmussen & Carroll's (2014, pp. 447-448) expectations are based on pluralist theory. This framework is taken one step further in this thesis, by adding an exchange theoretical and a broader cost-benefit perspective. The main idea here is that less affected interests might not contribute if they do not already possess the relevant information to do so, or if the costs associated with participating is high compared to the potential benefits. New hypotheses concerning the type of DG organizing a consultation, EU activity, format, technical complexity and an interaction between technical complexity and salience are derived from these perspectives. An additional hypothesis that is based on pluralist ideas is also added, concerning salience. Statistical analyses of 308 online consultations organized between 2001 and 2011 were used to

examine these hypotheses.

Overall, two new characteristics – technical complexity and DG type – are found to explain participation bias. First, technically complex consultations are found to generally be more biased than less complex ones to a relatively high degree. Second, consultations organized by regulatory DGs are found to generally be highly biased towards business interests. The remaining new features examined in this thesis were not found to have the expected effect on bias, with the slight exception of the format of a consultation.

Theoretically, these findings indicate that non-pluralist perspectives can be relevant to explain participation bias as well. Less affected interests might not be represented if the costs of participating are high, which is assumed to be the case for technically complex consultations. Similarly, possessing relevant information might make participating easier and less costly. Business groups are assumed to possess the type of information regulatory DGs demand, which may be why business interests are better represented in consultations organized by this type of DG. Furthermore, in a normative sense, the biased participation in these consultations implies that the interest group channel here functions more poorly as a source to input legitimacy. The Commission's goal for the consultations to be inclusive is not likely to be achieved (European Commission, 2002, pp. 16-17). Regulatory DGs' demand for output legitimacy, on the other hand, may be supplied.

How the openness of the consultations affects bias is also considered in this thesis. Most consultations are open for everyone, but some have specific target groups or are closed to all but select actors. The results give some support to the idea that selective and closed consultations are more biased than open ones. Furthermore, whether different factors affect bias in open, selective and closed consultations were considered. The results, although uncertain, indicate that this could be the case.

Rasmussen & Carroll's (2014, pp. 447-448) expectations are all supported here as well. This indicates that their findings are robust. They still hold after adding new observations, taking the hierarchical structure of the data into account and controlling for new variables. However, the two alternative, transactions based expectations proposed in this thesis for the policy type and cost structure characteristics might be relevant as well. The ideas here could explain why the effects of these variables are estimated to be weaker in the expanded model. Moreover, how interesting the actor bias and interest group density findings are have been

questioned. These variables are closely linked to the dependent variable in time, which could make the direction of the potential causal effects more clouded. They could be caused by the dependent variable as well.

Rasmussen & Carroll's (2014, pp. 453-456) analysis of what affects the bias towards business interests was also carried out in this thesis. Their conclusions also hold here. Of the new variables, only technical complexity was found to lead to a somewhat higher proportion of business group participants. Comparing these results to what explains the general degree of bias shows that there are some differences.

As mentioned in Chapter 1, there are differences between the EU and national settings which lessen the possibility of generalizing the findings in this thesis. One is the prevalence of groups from different national settings and European wide associations at the EU level. Another difference is that the EU mainly deals with regulatory matters, which might be more relevant for business interests. These aspects imply that there might be systematic differences between the EU and its member states when it comes to participation bias. Furthermore, the potential to generalize to other types of consultations, such as expert groups or conferences, is also likely to be somewhat limited. As mentioned, there are fewer barriers to entry in online consultations. Participation in the other types of consultation exercises might for instance require an invitation, which does not leave the decision to participate up to each actor on its own.

In addition to the substantial findings, two methodological aspects are also considered in this thesis. First, the cases where there were substantial differences between the classical and robust standard errors were taken into account by improving the models. Differences between these errors indicate that the model is misspecified, which can affect the estimates (King & Roberts, 2014, pp. 1-3, 6-7). Second, the power to detect effects of different sizes was considered. The lack of support to some of the new hypotheses might be explained by the limited power to detect small effects. That some of the estimates are significant when using sparser models lends some support to this idea. A much larger sample would be needed to conclude with more certainty that the non-significant estimates do not systematically affect bias.

Bibliography

Adcock, R., & Collier, D. (2001). Measurement Validity: A Shared Standard for Qualitative and Quantitative Research. *American Political Science Review*, *95*, 529-546.

Baroni, L., Carroll, B. J., Chalmers, A. W., Marquez, L. M. M., & Rasmussen, A. (2014). Defining and classifying interest groups. *Interest Groups & Advocacy*, *3*, 141-159.

Baumgartner, F. R., & Leech, B. L. (1998). *Basic Interests: The Importance of Groups in Politics and in Political Science*. Princeton: Princeton University Press.

Berry, J. M. (1999). The rise of citizen groups. In T. Skocpol & M. P. Fiorina (Eds.), *Civic Engagement in American Democracy* (p. 367-394). Washington, D.C.: Brookings Institution Press.

Beyers, J., Eising, R., & Maloney, W. (2008). Researching Interest Group Politics in Europe and Elsewhere: Much We Study, Little We Know? *West European Politics*, *31*, 1103-1128.

Bliese, P. (2013). *Multilevel Modeling in R (2.5).* Retrieved March 19, 2015, from `http://cran.r-project.org/doc/contrib/Bliese_Multilevel.pdf`

Brambor, T., Clark, W. R., & Golder, M. (2006). Understanding Interaction Models: Improving Empirical Analyses. *Political Analysis*, *14*, 63-82.

Broscheid, A., & Coen, D. (2007). Lobbying activity and fora creation in the EU: empirically exploring the nature of the policy good. *Journal of European Public Policy*, *14*, 346-365.

Bryman, A. (2004). *Social Research Methods*. Oxford: Oxford University Press.

Clark, P. B., & Wilson, J. Q. (1961). Incentive Systems: A Theory of Organizations. *Administrative Science Quarterly*, *6*, 129-166.

Coen, D. (1997). The European Business Lobby. *Business Strategy Review*, *8*, 17-25.

Coen, D. (2007). Empirical and theoretical studies in EU lobbying. *Journal of European Public Policy*, *14*, 333-345.

Coen, D., & Katsaitis, A. (2013). Chameleon pluralism in the EU: an empirical study of the European Commission interest group density and diversity across policy domains. *Journal of European Public Policy*, *20*, 1104-1119.

Corbett, R., Peterson, J., & Bomberg, E. (2012). The EU's Institutions. In E. Bomberg, J. Peterson, & R. Corbett (Eds.), *The European Union: How Does it Work?* (p. 47-72). Oxford: Oxford University Press.

Directorate-General for Competition. (2014, March 24). *Closed consultations and comments received.* Retrieved March 15, 2015, from http://ec.europa.eu/competition/consultations/closed.html

Directorate-General for Economic and Financial Affairs. (2013, October 30). *Consultation on the Europe 2020 Project Bond Initiative.* Retrieved March 15, 2015, from http://ec.europa.eu/economy_finance/articles/consultation/europe_2020_en.htm

Directorate-General for Employment, Social Affairs and Inclusion. (2010, July 7). *Public consultations.* Retrieved March 15, 2015, from http://ec.europa.eu/social/main.jsp?catId=700&langId=en&consultId=3&visib=0&furtherConsult=yes

Directorate-General for Health and Consumers. (2011). *Report on the public consultation on the possible revision of the Tobacco Products Directive (2001/37/EC).* European Commission.

Directorate-General for Informatics (DIGIT). (2014, October 14). *Mission Statement.* Retrieved January 11, 2015, from http://ec.europa.eu/dgs/informatics/identity_en.htm

Directorate-General for the Environment. (2015, March 12). *Closed consultations.* Retrieved March 15, 2015, from http://ec.europa.eu/environment/consultations_en.htm

Directorate-General for the Internal Market and Services. (2011, June 28). *Consultation on the review of the Markets in Financial Instruments Directive (MiFID).* Retrieved February 24, 2015, from http://ec.europa.eu/internal_market/consultations/2010/mifid_en.htm

Directorate-General for the Internal Market and Services. (2014, October 23). *Closed Consultations.* Retrieved March 15, 2015, from http://ec.europa.eu/internal_market/consultations/closed/index_en.htm

Eshbaugh-Soha, M. (2006). The Conditioning Effects of Policy Salience and Complexity on American Political Institutions. *The Policy Studies Journal*, *34*, 223-243.

European Commission. (2002). *Towards a reinforced culture of consultation and dialogue – General principles and minimum standards for consultation of interested parties by the Commission.* (COM(2002) 704 final)

European Commission. (2009). *Key Figures Card: Staff Members.* (Directorate-General Personnel and Administration)

Ferretti, M. P., & Lener, M. (2008). Lay Public or Experts? e-Participation in Authorization for GMO Products in the European Union. *Review of Policy Research*, *25*, 507-525.

Finke, B. (2007). Civil society participation in EU governance. *Living Reviews in European Governance*, *2*, 1-31.

Fox, J., & Weisberg, S. (2011). *An R Companion to Applied Regression*. Thousand Oaks: Sage.

Gerring, J. (2005). Causation: A Unified Framework for the Social Sciences. *Journal of Theoretical Politics*, *17*, 163-198.

Gormley, W. T. (1986). Regulatory Issue Networks in a Federal System. *Polity*, *18*, 595-620.

Greenwood, J. (1997). *Representing Interests in the European Union*. Basingstoke: Palgrave Macmillan.

Greenwood, J. (2011). *Interest Representation in the European Union*. Basingstoke: Palgrave Macmillan.

Greer, S. L., da Fonseca, E. M., & Adolph, C. (2008). Mobilizing Bias in Europe: Lobbies, Democracy and EU Health Policy-Making. *European Union Politics*, *9*, 403-433.

Hess, B., Pfeiffer, T., Mertens, M., & Sabanogullari, L. (2011). *Study JUST/2010/JCIV/CT/0027/A4: Evaluation of contributions to the public consultation and hearing: "Towards a Coherent European Approach to Collective Redress"*. Ruprecht-Karls-Universität Heidelberg.

Hix, S., & Høyland, B. (2011). *The Political System of the European Union*. Basingstoke: Palgrave Macmillan.

Hojnacki, M., Kimball, D. C., Baumgartner, F. R., Berry, J. M., & Leech, B. L. (2012). Studying Organizational Advocacy and Influence: Reexamining Interest Group Research. *Annual Review of Political Science*, *15*, 379-399.

Jakobsen, M. L. F., & Mortensen, P. B. (2014). How Politics Shapes the Growth of Rules. *Governance*, 1-23.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning: with Applications in R*. New York: Springer.

Kardasheva, R. (2009). *Legislative package deals in EU decision-making: 1999-2007* (PhD Thesis). London School of Economics and Political Science.

King, G., Honaker, J., Joseph, A., & Scheve, K. (2001). Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation. *American Political Science Review*, *95*, 49-69.

King, G., & Roberts, M. E. (2014). How Robust Standard Errors Expose Methodological Problems They Do Not Fix, and What to Do About It. *Political Analysis*, 1-21.

Klüver, H. (2011). The contextual nature of lobbying: Explaining lobbying success in the European Union. *European Union Politics*, *12*, 483-506.

Long, S. J. (1997). *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks: Sage.

Lowery, D. (2007). Why Do Organized Interests Lobby? A Multi-Goal, Multi-Context Theory of Lobbying. *Polity*, *39*, 29-54.

Lowery, D., & Gray, V. (2004). A Neopluralist Perspective on Research on Organized Interests. *Political Research Quarterly*, *57*, 163-175.

Lowry, W. R., & Joslyn, M. (2014). The Determinants of Salience of Energy Issues. *Review of Policy Research*, *31*, 153-172.

Mahoney, C. (2004). The Power of Institutions: State and Interest Group Activity in the European Union. *European Union Politics*, *5*, 441-466.

Mahoney, C. (2007). Lobbying Success in the United States and the European Union. *Journal of Public Policy*, *27*, 35-56.

Moulton, B. R. (1990). An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units. *The Review of Economics and Statistics*, *72*, 334-338.

Neshkova, M. I. (2014). Salience and Complexity in Supranational Policymaking: The Case of Subnational Interests. *Governance*, *27*, 9-36.

Olson, M. (1971). *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge: Harvard University Press.

Osborne, J. W. (2013). *Best Practices in Data Cleaning: A Complete Guide to Everything You Need to Do Before and After Collecting Your Data*. Thousand Oaks: Sage.

Persson, T. (2007). Democratizing European Chemicals Policy: Do Consultations Favour Civil Society Participation? *Journal of Civil Society*, *3*, 223-238.

Pollack, M. A. (1997). Representing diffuse interests in EC policy-making. *Journal of European Public Policy*, *4*, 572-590.

Primo, D. M., Jacobsmeier, M. L., & Milyo, J. (2007). Estimating the Impact of State Policies and Institutions with Mixed-Level Data. *State Politics and Policy Quarterly*, *7*, 446-459.

Quittkat, C. (2011). The European Commission's Online Consultations: A Success Story? *Journal of Common Market Studies*, *49*, 653-674.

Quittkat, C., & Kotzian, P. (2011). Lobbying via Consultation – Territorial and Functional Interests in the Commission's Consultation Regime. *Journal of European Integration*, *33*, 401-418.

Ramalho, E. A., Ramalho, J. J. S., & Murteira, J. M. R. (2011). Alternative Estimating and Testing Empirical Strategies for Fractional Regression Models. *Journal of Economic Surveys*, *25*, 19-68.

Rasmussen, A. (2015). Participation in Written Government Consultations in Denmark and the UK: System and Actor-level Effects. *Government and Opposition*, 1-29.

Rasmussen, A., & Carroll, B. J. (2014). Determinants of Upper-Class Dominance in the Heavenly Chorus: Lessons from European Union Online Consultations. *British Journal of Political Science*, *44*, 445-459.

Rasmussen, A., Carroll, B. J., & Lowery, D. (2014). Representatives of the public? Public opinion and interest group activity. *European Journal of Political Research*, *53*, 250-268.

Ringquist, E. J., Worsham, J., & Eisner, M. A. (2003). Salience, Complexity, and the Legislative Direction of Regulatory Bureaucracies. *Journal of Public Administration Research and Theory*, *13*, 141-164.

Schlozman, K. L., Verba, S., & Brady, H. E. (2012). *The Unheavenly Chorus: Unequal Political Voice and the Broken Promise of American Democracy*. Princeton: Princeton University Press.

Steffek, J., & Ferretti, M. P. (2009). Accountability or "Good Decisions"? The Competing Goals of Civil Society Participation in International Governance. *Global Society*, *23*, 37-57.

Tobin, J. (1958). Estimation of Relationships for Limited Dependent Variables. *Econometrica*, *26*, 24-36.

Toshkov, D. (2013). *Legislative production in the EU, 1967-2012.* Retrieved October 7, 2014, from `http://www.dimiter.eu/Data.html`

Truman, D. B. (1951). *The Governmental Process: Political Interests and Public Opinion*. New York: Knopf.

UCLA: Statistical Consulting Group. (n.d.). *What is complete or quasi-complete separation in logistic/probit regression and how do we deal with them?* Retrieved March 17, 2015, from `http://www.ats.ucla.edu/stat/mult_pkg/faq/general/complete_separation_logit_models.htm`

Ward, M. D., Greenhill, B. D., & Bakke, K. M. (2010). The perils of policy by p-value: Predicting civil conflicts. *Journal of Peace Research*, *47*, 363-375.

Wonka, A., Baumgartner, F. R., Mahoney, C., & Berkhout, J. (2010). Measuring the size and scope of the EU interest group population. *European Union Politics*, *11*, 463-476.

Yackee, J. W., & Yackee, S. W. (2006). A Bias Towards Business? Assessing Interest Group Influence on the U.S. Bureaucracy. *The Journal of Politics*, *68*, 128-139.

Yoho, J. (1998). The Evolution of a Better Definition of "Interest Group" and Its Synonyms. *The Social Science Journal*, *35*, 231-243.

Appendices

7 Codebook

DG_bin: regulatory or distributive DGs. Regulatory DGs are coded as 0: Climate Action; Communication; Communications Networks, Content and Technology; Competition; Economic and Financial Affairs; Energy; Enterprise and Industry; Environment; Health and Consumers; Home Affairs; Informatics; Internal Market and Services; Justice; Mobility and Transport; Secretariat-General; Taxation and Customs Union; Trade. Distributive DGs are coded as 1: Agriculture and Rural Development; Budget; Education and Culture; Employment, Social Affairs and Inclusion; EuropeAid Development and Cooperation; Maritime Affairs and Fisheries; Regional Policy; Research and Innovation.

DG_omf: The proportion of distributive consultations each DG has organized. The number of consultations concerning distributive policies a DG has organized is divided by its total number of consultations.

DG_ans: The number of employees the DG organizing a consultation had in 2006 or 2009. Consultations organized in 2006 or earlier are given the numbers from 2006. Consultations organized after this are given the numbers from 2009. Where numbers from 2006 are missing, the number of employees in 2009 is assigned.

legact: The number of passed regulations, directives and decisions in the policy area a consultation falls in, lagged by one year. Each consultation is connected to the most relevant policy area and assigned its number of legislative acts.

consform: The format of a consultation. Consultations are coded as 0 if they only include closed questions; 1 if they are semi-standardized with both open and closed questions; 2 if they

are non-standardized and only include open questions.

DG_tc: Whether the policy area a consultation falls in is technically complex. Less complex policy areas are coded as 0: Education and Culture; Employment, Social Affairs and Inclusion; Home Affairs; Justice; Mobility and Transport; Research and Innovation. Complex policy areas are coded as 1: Agriculture and Rural Development; Budget; Climate Action; Communication; Communications Networks, Content and Technology; Competition; Economic and Financial Affairs; Energy; Enterprise and Industry; Environment; EuropeAid, Development and Cooperation; Health and Consumers; Informatics; Internal Market and Services; Maritime Affairs and Fisheries; Regional Policy; Secretariat-General; Taxation and Customs Union; Trade.

cons_oc: Whether a consultation calls for comments from everyone. Open consultations are coded as 0; consultations where some actors are singled, but that otherwise are open for everyone are coded as 1; consultations where only specific actors are asked to contribute are coded as 2.

cons_sal: The salience of the policy area a consultation falls in, lagged by one year. This is based on a question in the Eurobarometer surveys that asks the respondents to pick two issues that they consider to be the most important ones facing their country at the moment. Each consultation is linked to one of the 15 policy areas and given the EU level percentage of respondents deeming this issue the most important. The "other" category is used when none of the substantive issues are relevant. If there are more than one Eurobarometer survey in a given year, the mean of the percentages is assigned. The collective transport category is discontinued from 2006. If a consultation concerns this issue later, it is given the mean of the percentages from between 2003 and 2006. The question was introduced in 2003 – consultations organized before this are given the relevant percentage from this year.

pro_biz: The proportion of trade and other business group participants.

pro_occ: The proportion of occupational group participants.

pro_civ: The proportion of public interests and identity group participants.

DG_navn: The name of the DG that organized the consultation.

DGpolicy: The DG or policy area a consultation falls in. The consultations organized by DG Competition, DG Internal Market and DG Environment are classified based on the policy activity a given consultation is filed under on their websites (Directorate-General for Competition, 2014; Directorate-General for the Environment, 2015; Directorate-General for the Internal Market and Services, 2014). A requirement is that more than one consultation is filed under the same activity. The consultations organized by DG Justice and DG Mobility and Transport are categorized based on the themes of the consultations. The consultations organized by DG Budget, DG Informatics and DG Communication are coded as "EU matters". The consultations organized by DG Agriculture and Rural Development and DG Maritime Affairs are coded as "Agriculture and fisheries". The consultation organized by DG Economic and Financial Affairs is merged with DG Internal Market, since the consultation relates to the completion of the internal market (Directorate-General for Economic and Financial Affairs, 2013). The consultation organized by DG Employment, Social Affairs and Inclusion is merged with DG Internal Market, since this is one of the policy fields listed on the consultation website (Directorate-General for Employment, Social Affairs and Inclusion, 2010).

The rest of the measures for the new consultations were coded according to the instructions in Rasmussen & Carroll's (2014) codebook. Some specifications were made for year and density. First, if a consultation takes place during the turn of the year, it is coded as the year it ended. Second, if there are multiple contributions from the same participant in a consultation, it is only counted once.

8 Comparing interest distributions

Table 8.1: The interest group type distribution in the online consultations and in the EU interest group population (Transparency Register)

	Trade and other business	Occupational	Trade unions	Public interests	Identity	Public authorities as members	Others	Total
Total consultations, Rasmussen & Carroll	3,880 (59.17%)	989 (15.08%)	170 (2.59%)	792 (12.08%)	161 (2.46%)	266 (4.06%)	299 (4.56%)	6,557 (100%)
Total consultations, expanded sample	7,691 (56.02%)	1,961 (14.28%)	417 (3.04%)	1,799 (13.10%)	358 (2.61%)	516 (3.76%)	988 (7.20%)	13,730 (100.01%)
Complex	6,728 (59.05%)	1,565 (13.74%)	327 (2.87%)	1,380 (12.11%)	170 (1.49%)	430 (3.77%)	793 (6.96%)	11,393 (99.99%)
Less complex	963 (41.21%)	396 (16.94%)	90 (3.85%)	419 (17.93%)	188 (8.04%)	86 (3.68%)	195 (8.34%)	2,337 (99.99%)
Regulatory DG	7,215 (58.48%)	1,845 (14.96%)	358 (2.90%)	1,569 (12.72%)	265 (2.15%)	285 (2.31%)	800 (6.48%)	12,337 (100%)
Distributive DG	476 (34.17%)	116 (8.33%)	59 (4.24%)	230 (16.51%)	93 (6.68%)	231 (16.58%)	188 (13.50%)	1,393 (100.01%)
Total Register	1,435 (51.90%)	314 (11.36%)	56 (2.03%)	500 (18.08%)	73 (2.64%)	20 (0.72%)	367 (13.27%)	2,765 (100%)

Note: Using a chi-square contingency test shows the distributions are statistically significant from each other.

9 Regression and out-of-sample results

Table 9.1: Year estimates from the OLS, multilevel and fractional logit regressions

	HHI		Transformed HHI	HHI		Transformed HHI	Prop. bus. and occ.		
	OLS		OLS	linear mixed-effects		linear mixed-effects	logistic		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
2002	0.075	0.058	0.004	0.034	0.041	0.004	0.182	0.078	−0.497
	(0.196)	(0.190)	(0.124)	(0.187)	(0.180)	(0.120)	(2.811)	(2.800)	(2.830)
2003	0.087	0.070	0.044	0.008	0.011	0.012	0.786	0.617	0.708
	(0.199)	(0.194)	(0.127)	(0.192)	(0.185)	(0.125)	(2.935)	(2.924)	(3.127)
2004	0.201	0.185	0.104	0.124	0.125	0.067	0.767	0.462	0.102
	(0.189)	(0.182)	(0.118)	(0.185)	(0.177)	(0.116)	(2.780)	(2.754)	(2.764)
2005	0.085	0.065	0.024	0.017	0.007	−0.006	0.304	0.071	−0.320
	(0.188)	(0.183)	(0.118)	(0.181)	(0.175)	(0.115)	(2.761)	(2.744)	(2.750)
2006	0.174	0.159	0.079	0.091	0.090	0.042	0.606	0.343	0.071
	(0.185)	(0.180)	(0.117)	(0.181)	(0.174)	(0.114)	(2.732)	(2.714)	(2.732)
2007	0.105	0.086	0.039	0.043	0.037	0.011	0.211	−0.008	−0.091
	(0.187)	(0.181)	(0.117)	(0.181)	(0.174)	(0.114)	(2.739)	(2.720)	(2.733)
2008	0.091	0.066	0.010	0.010	−0.004	−0.019	0.392	0.064	−0.227
	(0.186)	(0.180)	(0.117)	(0.181)	(0.174)	(0.114)	(2.731)	(2.711)	(2.727)
2009	0.122	0.087	0.029	0.036	0.017	−0.003	0.815	0.466	0.103
	(0.184)	(0.179)	(0.116)	(0.180)	(0.173)	(0.114)	(2.718)	(2.700)	(2.723)
2010	0.113	0.110	0.043	0.039	0.053	0.011	0.453	−0.019	−0.395
	(0.186)	(0.178)	(0.116)	(0.182)	(0.172)	(0.114)	(2.743)	(2.690)	(2.715)
2011		0.121	0.051		0.061	0.020		0.002	−0.508
		(0.179)	(0.116)		(0.173)	(0.114)		(2.696)	(2.717)
N	210	308	302	210	308	302	210	308	302

Note: 2001 is the baseline category. $^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$

Table 9.2: OLS results for 1) open consultations, 2) selective consultations and 3) closed consultations

	Transformed HHI		
	(1)	(2)	(3)
Expenditure policy	-0.052^{*}	-0.033	0.042
	(0.027)	(0.033)	(0.176)
Administrative policy	-0.008	0.038	-0.317
	(0.026)	(0.041)	(0.168)
Concentrated costs	0.048^{**}	0.049^{*}	-0.017
	(0.023)	(0.029)	(0.114)
Actor bias	0.349^{***}	0.053	0.331
	(0.084)	(0.122)	(0.352)
Density	-0.063^{***}	-0.075^{***}	-0.012
	(0.023)	(0.024)	(0.102)
Duration	0.0005	-0.008^{***}	0.015
	(0.002)	(0.003)	(0.015)
Salience	0.00001	0.0002	-0.011
	(0.002)	(0.006)	(0.007)
DG type	-0.118^{***}	-0.064	0.191
	(0.041)	(0.050)	(0.314)
Number of staff	0.00003	0.0001^{**}	-0.0004
	(0.00004)	(0.0001)	(0.0004)
Non-standardized	0.023	0.012	-0.234
	(0.025)	(0.037)	(0.180)
Technical complexity	0.055^{*}	0.092	0.272
	(0.032)	(0.070)	(0.218)
Tech. compl. $\times$ sal.	-0.0002	0.001	
	(0.002)	(0.006)	
N	187	92	23
R^2	0.381	0.567	0.747
Adjusted R^2	0.298	0.437	-0.112
Residual Std. Error	0.119 (df = 164)	0.102 (df = 70)	0.104 (df = 5)
F Statistic	4.586^{***} (df = 22; 164)	4.364^{***} (df = 21; 70)	0.870 (df = 17; 5)

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$

Table 9.3: OLS results without actor bias (1), interest group density (2) and excluding both (3), expanded model

	Transformed HHI		
	(1)	(2)	(3)
Density	-0.075^{***}		
	(0.016)		
Actor bias		0.270^{***}	
		(0.065)	
Expenditure policy	-0.056^{***}	-0.057^{***}	-0.067^{***}
	(0.020)	(0.020)	(0.020)
Administrative policy	-0.001	-0.002	-0.002
	(0.021)	(0.022)	(0.022)
Concentrated costs	0.035^{**}	0.040^{**}	0.036^{**}
	(0.017)	(0.017)	(0.018)
Duration	-0.002	-0.002	-0.003^{**}
	(0.001)	(0.001)	(0.001)
Salience	-0.0005	0.0001	-0.0002
	(0.002)	(0.002)	(0.002)
DG type	-0.077^{***}	-0.096^{***}	-0.084^{***}
	(0.028)	(0.029)	(0.029)
Number of staff	0.00005	0.0001^{*}	0.0001^{*}
	(0.00003)	(0.00003)	(0.00003)
Non-standardized	0.031	0.033^{*}	0.040^{*}
	(0.020)	(0.020)	(0.020)
Technical complexity	0.065^{**}	0.056^{**}	0.065^{**}
	(0.026)	(0.027)	(0.027)
Tech. compl. $\times$ sal.	0.001	-0.0002	0.00004
	(0.002)	(0.002)	(0.002)
Selective	0.025	0.037^{**}	0.033^{*}
	(0.018)	(0.018)	(0.018)
Closed	0.012	0.030	0.025
	(0.028)	(0.028)	(0.029)
N	302	302	302
R^2	0.333	0.322	0.279
Adjusted R^2	0.278	0.265	0.222
Residual Std. Error	0.116 (df = 278)	0.117 (df = 278)	0.121 (df = 279)
F Statistic	6.039^{***} (df = 23; 278)	5.730^{***} (df = 23; 278)	4.908^{***} (df = 22; 279)

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$

Table 9.4: Fractional logit results without actor bias (1), interest group density (2) and excluding both (3), expanded model

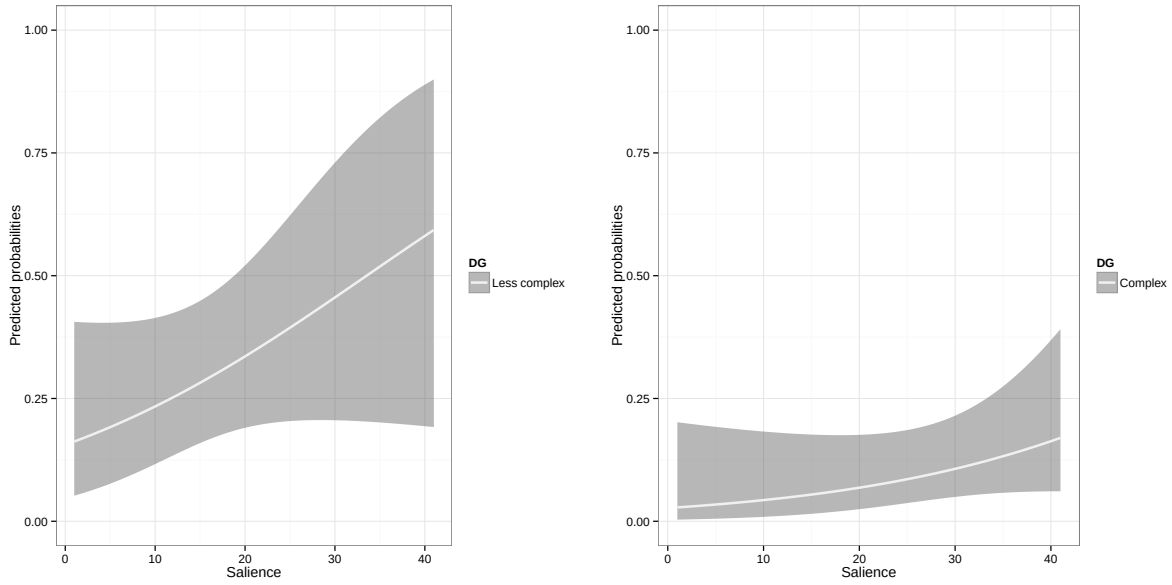
	Proportion of bus. and occ. group participants		
	(1)	(2)	(3)
Density	−0.129		
	(0.311)		
Actor bias		1.130	
		(1.406)	
Expenditure policy	−0.590	−0.566	−0.611
	(0.387)	(0.387)	(0.384)
Administrative policy	−0.046	−0.039	−0.048
	(0.429)	(0.430)	(0.430)
Concentrated costs	0.367	0.392	0.369
	(0.353)	(0.355)	(0.354)
Duration	0.013	0.014	0.010
	(0.028)	(0.028)	(0.027)
Salience	−0.009	−0.007	−0.009
	(0.031)	(0.031)	(0.031)
DG type	−1.142**	−1.207**	−1.152**
	(0.521)	(0.526)	(0.521)
Number of staff	−0.0001	−0.00004	−0.00005
	(0.001)	(0.001)	(0.001)
Non-standardized	0.269	0.255	0.283
	(0.393)	(0.392)	(0.391)
Technical complexity	0.183	0.142	0.179
	(0.505)	(0.507)	(0.506)
Tech. compl. × sal.	0.034	0.032	0.033
	(0.033)	(0.033)	(0.033)
Selective	0.110	0.146	0.123
	(0.374)	(0.374)	(0.373)
Closed	0.035	0.083	0.054
	(0.583)	(0.583)	(0.579)
N	302	302	302
Log Likelihood	−105.789	−105.831	−105.523
AIC	259.578	259.662	257.045

*p < .1; **p < .05; ***p < .01

Table 9.5: OLS and multilevel linear results, arcsine square root transformed proportion of business and occupational group participants as the dependent variable, expanded model

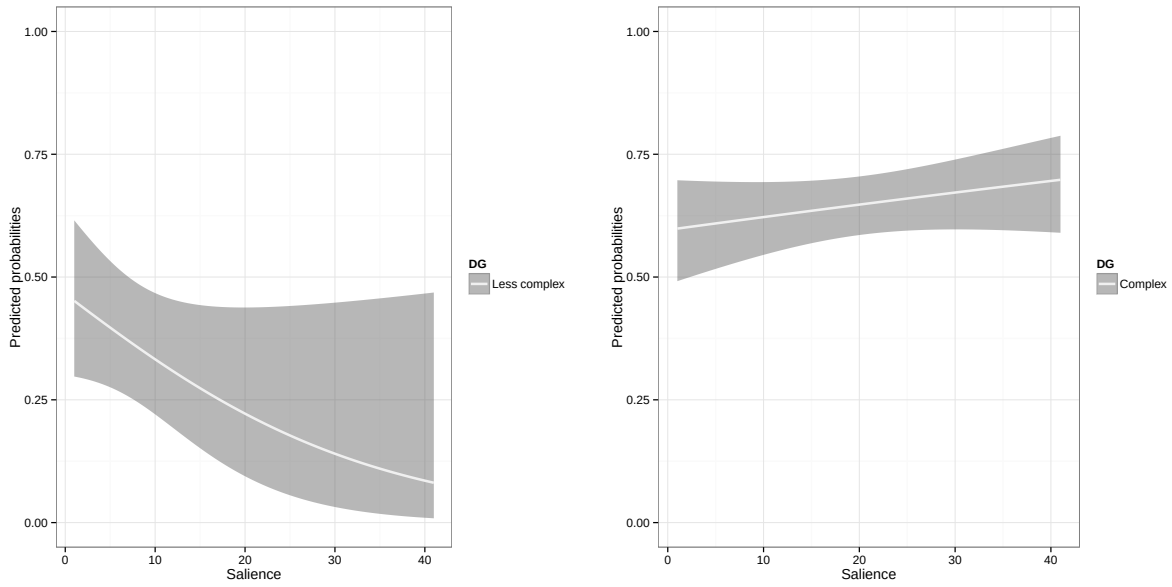
	Arcsine square root transformed proportion of bus. and occ. group participants	
	OLS	linear mixed-effects
	(1)	(2)
Expenditure policy	−0.112***	−0.102**
	(0.039)	(0.040)
Administrative policy	0.011	0.016
	(0.042)	(0.041)
Concentrated costs	0.094***	0.067**
	(0.034)	(0.032)
Actor bias	0.354***	0.221*
	(0.127)	(0.122)
Density	−0.048	−0.047
	(0.031)	(0.030)
Duration	0.004	0.003
	(0.003)	(0.003)
Salience	−0.003	0.0001
	(0.003)	(0.003)
DG type	−0.307***	−0.268***
	(0.056)	(0.078)
Number of staff	−0.00003	−0.0001
	(0.0001)	(0.0001)
Non-standardized	0.058	0.013
	(0.039)	(0.039)
Technical complexity	0.054	0.148*
	(0.052)	(0.084)
Tech. compl. × sal.	0.008**	0.003
	(0.004)	(0.004)
Selective	0.025	0.041
	(0.035)	(0.034)
Closed	−0.001	0.011
	(0.055)	(0.053)
Consultation/policy area level N	302/0	302/30
R^2	0.455	
Adjusted R^2	0.408	
Residual Std. Error	0.230 (df = 277)	
F Statistic	9.626*** (df = 24; 277)	
Log Likelihood		−31.039
AIC		116.079
BIC		216.260

Note: The random intercept variance for model 2 is 0.015, standard deviation 0.121.
*p < .1; **p < .05; ***p < .01



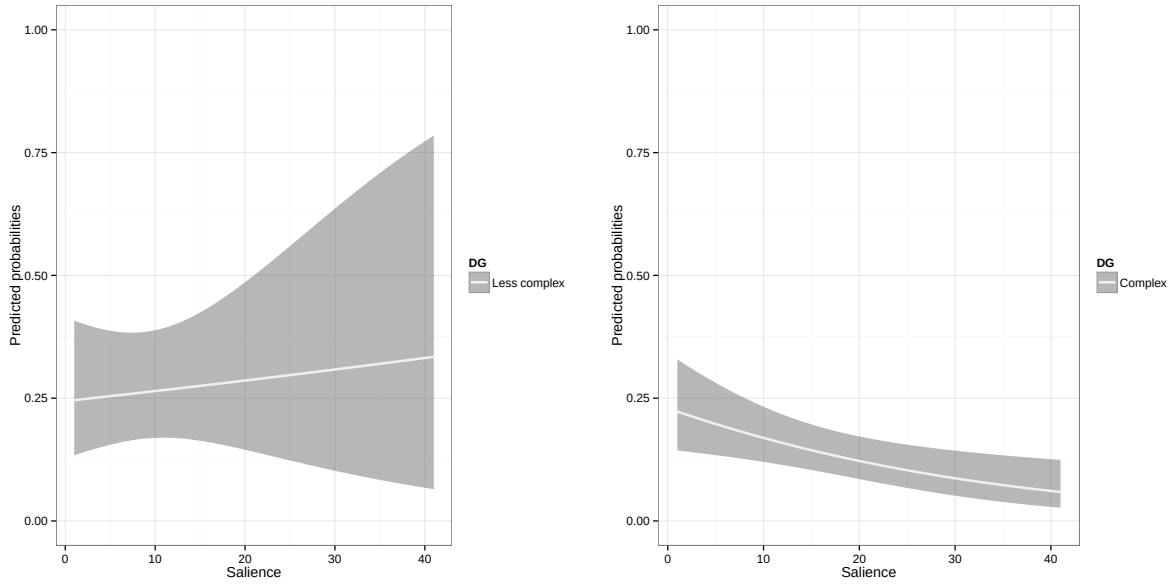
(a) Less complex (b) Complex

Figure 9.1: The effect of salience on the proportion of occupational group participants depending on technical complexity (sparser model), predicted probabilities, 95% confidence intervals



(a) Less complex (b) Complex

Figure 9.2: The effect of salience on the proportion of business group participants depending on technical complexity (sparser model), predicted probabilities, 95% confidence intervals



(a) Less complex

(b) Complex

Figure 9.3: The effect of salience on the proportion of citizen group participants depending on technical complexity (sparser model), predicted probabilities, 95% confidence intervals

Out-of-sample performance results

For the general degree of bias in the expanded model, the error rate is estimated to be between 0.014 and 0.015, depending on the test. Four different ones are used – a validation set, leave-one-out cross-validation, and a five- and ten-fold cross-validation test. Removing activity and salience results in an error rate estimated to be between 0.013 and 0.01. The estimated error rate for the initial model is around 0.015, when using the expanded sample and the square root transformed HHI measure as the dependent variable to make the results comparable.