EUROPEAN UNIVERSITY INSTITUTE, FLORENCE

DEPARTMENT OF POLITICAL AND SOCIAL SCIENCES

The Real World of Deliberation
A Comparative Study of its Favorable Conditions in Legislatures

ANDRÉ BÄCHTIGER and MARCO R. STEENBERGEN

BADIA FIESOLANA, SAN DOMENICO (FI)

André Bächtiger
Marco R. Steenbergen

The Real World of Deliberation. A Comparative Study of its Favorable Conditions in Legislatures[1]

Abstract

While deliberation has moved to the forefront in contemporary democratic theory, it remains an open question whether a deliberative conception of politics could realistically come into operation. In this paper, we ask whether there are certain institutional and issue contexts that are favorable to deliberation, whereby the focus is on legislatures. On the basis of the newly developed "discourse quality index" (DQI), we analyzed the quality of deliberation in 52 parliamentary debates in Switzerland, the United States, Germany, and United Kingdom. We find that consensus institutions (grand coalitions and veto power), presidentialism, second chambers, non-publicity, and low issue polarization further deliberation, particularly by enhancing respectful exchanges among participants. However, we also find intriguing complexity in that public deliberation, while hindering respectful discourse, is conducive to reasoned argument and common good orientations. We conclude that institutions and contexts matter for deliberation, and particularly when the favorable factors combine, we can detect debates that in parts have features of ideal discourses.

Introduction

Deliberation is taking center stage in contemporary democratic theory. Deliberation implies that actors listen to each other, reasonably justify their policy positions, show mutual respect, and reflect upon and evaluate their interests and needs from the point of view of their generalizability. Most importantly, they must be prepared to get persuaded by the "non-coercive coercion of the better argument" (Habermas 1983: 132), whereby the ultimate goal is to find a reasoned consensus about validity claims (Habermas 1981; 1992, 1996; Gutmann/Thompson 1996; Manin 1997; Chambers 1996). From a philosophical point of view, deliberation is seen as a necessary means to arrive at legitimate decisions in modern pluralistic and fragmented societies where a common religious view or a comprehensive moral outlook no longer exists and where the authority of tradition has greatly weakened (Chambers 1995: 244; Benhabib 1996). While normatively (and intuitively) appealing, it is far from certain whether a deliberative conception of politics could realistically come into operation or whether it would remain a principally utopian affair. Scharpf (1997: 166), for instance, sees a practical incompatibility between the consensual ideal of deliberative politics

and the electoral implications of competitive democracies. Why, Scharpf asks, should self-interested political actors enter into a decision making process which redistributes power to the political opponent?

So far, most of the literature on deliberation has a philosophical orientation and discusses in a normative way the strengths and weaknesses of the deliberative model of democracy (see particularly Macedo 1999; Dryzek 2000; Gutmann/Thompson 2000). In recent years, however, a growing body of literature has tried to tackle the question whether arguing or deliberating might also constitute a logic of action in the real world of politics (Keck 1995; Risse 2000; Holzinger 2001a; 2001b; Dryzek/List 2003). However, three major gaps in previous research on real-world deliberation can be detected. First, empirical research on political deliberation is still in its infancy. Only a few concrete studies and projects have been undertaken or are in progress (e.g. Roberts 1997; Holzinger 2001a; Baccaro 2001; Sulkin/Simon 2001; Luskin/Fishkin 2002; Fishkin/Luskin 2004; Ulbert/Risse/Müller 2004). Second, there is also a dearth of measurement instruments that allow researchers to operationalize and quantify the quality of discourse. Third, and most importantly, there is no systematic investigation of *institutions* which might be favorable to a more deliberative mode of policy making. Our argument is that unless more is known about institutions (and other contexts) from a *comparative perspective*, broad assumptions about logics of action, and particularly about the extent and quality of deliberation are quite empty.

This article attempts to narrow the institutional, operational, and empirical gap in the literature on deliberative democracy. The basic research question is whether there are specific institutions and contexts that are favorable to a more deliberative mode of policy making in real world politics, whereby the focus will be on legislatures. Starting from Lijphart's (1984; 1999) seminal distinction between consensus and competitive systems, we propose that it is primarily the properties of consensus institutions (1) – grand coalitions and veto power - that make real world legislators more discursive. While grand coalitions reduce the magnitude of political competition, veto power of actors creates a situation of joint decision-making where elements of discursive rationality might be required to find agreement in policy negotiations. Second, the quality deliberation might also be affected by another classic system variable: presidentialism and parliamentarism (2). We expect that in presidential systems where party discipline is lower, legislators have more leeway to transcend party boundaries and to be open to argument and opinion change. Besides the these two system types, the type of chamber (3) and the type of arena (4) might matter for deliberation too: while second chambers involve a number of organizational provisions (such as longer terms of office and smaller size) as well as norms of civility that favor deliberation, the quality of discourse might be further enhanced when actors deliberate behind closed doors where the pressures of following constituent demands are lowered. And finally, we propose that a non-institutional attribute, the type of policy issue (5), affects deliberation as well, with the expectation that low degrees of issue polarization propel more respectful and more consensual debates. These five hypotheses will be evaluated on the basis of a total of 52 recorded debates (involving over 5.500 speeches) in parliamentary committees as well as plenary sessions of Parliament in Switzerland, the United States, Germany, and Great Britain. The quality of deliberation is assessed by the newly developed "discourse quality index" (DQI) (Steenbergen et al. 2003). The DQI represents an attempt to put the diverse concepts of deliberation in operational terms allowing for a quantitative content analysis of political debates.

Some might question whether parliaments are the right locus for an investigation of institutional effects on deliberation. They might argue that parliaments are not really deliberative bodies where consequential deliberative action takes place, but "rubber stamps" for decisions made by the government and elsewhere with debates being conducted in an adversarial and rhetorical manner between the parties involved. Our argument is that this is a fairly limited view of parliaments. Much in line with the new institutionally-oriented research

program on legislatures (Döring 1995), we will demonstrate that the posited institutional factors (and particularly combinations of them) might indeed be conducive to the achievement of the deliberative ideals within the parliamentary realm. Moreover, from an institutional point of view, parliaments offer a unique opportunity to directly compare the impact of different institutional constellations on actors' behavior while holding the basic organizational setting – legislatures - constant. Put differently, they allow to analyze the effects of a number of key institutional variables in a relatively controlled way, giving important hints what certain institutional mechanisms do to deliberation. And finally, from a research-practical point of view, parliamentary debates are protocoled and accessible, even in the non-public sphere (which, for example, is mostly not true for interactions within the government).

This article proceeds as follows. First, we propose a "weaker" deliberative program that delimits the chances of deliberation in the real world of politics; second we elaborate on our five hypotheses; third, we discuss our research design, present the selection of legislatures and debates, introduce our measurement instrument, the DQI, and discuss the statistical techniques that are used for hypothesis evaluation; fourth, we present the results of our analysis; and fifth, we conclude this article by discussing the relevance of our findings for deliberative democrats and institutional scholars.

A "weaker" deliberative program

When one is concerned with a more deliberative mode of policy making in real world politics, it is fairly problematic to analytically imply the demanding presuppositions of the Habermasian discourse model. As one can expect political actors in democratic systems to primarily focus on the will of their constituents, probably any deliberation in real world politics will fall short of the ideal discourse envisioned by Habermas: neither can one assume a constraint-free environment nor can one posit that only the force of the better argument will prevail or that a "rational consensus" could be found. Put differently, conceptualizing deliberation as process of where political actors exclusively adopt an action orientation toward reaching understanding and abstain from their constituent interests as well as from strategic considerations of how to advance these interests seems to be quite unrealistic indeed. Therefore, to get hold of "deliberation" in the real world of politics, it may be necessary to switch to a "weaker" deliberative program. Such a "weaker" deliberative program, first of all, views deliberation as merely one element of policy making, besides bargaining and voting (see Elster 1998a: 7). From this follows that deliberation is best conceptualized as a continuum ranging from more to less deliberative ways of policy making and not as a categorical conception where "true" deliberation is either present or not. Second, a "weaker" deliberative program does not assume that deliberation will massively change the basic preferences and worldviews of actors in normatively attractive ways (Johnson 1998: 174). But what deliberation can do is to help actors to overcome uncertainty and bounded rationality, to learn about unforeseen consequences and to generate new alternatives, or to structure "messy" preferences; it might, furthermore, enable actors to better understand the intentions, interpretations and expectations of other actors, even if they continue to disagree over substantial matters (Lynch 2002); and finally, deliberation might lead to compromises that rest on greater mutual understanding and respect for diverging standpoints (Knight/Johnson 1994: 285).

André Bächtiger/Marco R. Steenbergen

Favorable Institutional Conditions for Deliberation in Legislatures: Institutions and Issues

The research aim of this article is to identify institutional and contextual factors that are favorable to a more deliberative mode of policy making in the real world of politics, whereby the focus of analysis will be on legislatures. Since our primary focus is on institutional preconditions, it is necessary to define what we mean by institutions and how they can affect deliberation in a general way. By institutions, we understand relatively formal, established "rules of the game" (although not necessarily legal or constitutional rules), which structure the costs and benefits of action, constraining and enabling the behavior of actors (Hall and Taylor 1996; Thelen and Steinmo 1992). We also assume that institutions do not determine actors' deliberative behavior but only comprise action logics that favor or disfavor deliberative action. This means that deliberation may not occur even under the best "engendering" institutional conditions (and vice versa). Therefore, the link between different sets of institutions and the quality of deliberation is a probabilistic and not a deterministic one. And finally, as it may be impossible to fully incorporate the "reflective" and "transformative" aspects of deliberation in a rational incentives' framework (see Müller 1995), we put a special prime on the *enabling* aspects of institutions. Our claim is that there may be certain sets of institutions (and contexts) which are stake-reducing and allow for less calculative and strategic policy-making, while others are stake-enhancing and strongly conducive to such courses of action (Aspingwall/Schneider 2000: 27; March/Olsen 1998: 952-953).

Consensus Versus Competitive Systems. Starting from Lijphart's (1999) seminal distinction between consensus and competitive systems, we propose that two key properties of the consensus model can be considered "lubricants" for a more deliberative mode of policy making: *grand coalitions*, i.e. all major parties are represented in the executive, and *veto power*, i.e. actors possess veto devices to block or overturn decisions.[2] The two criteria form a hierarchical relationship, with grand coalitions and majority-opposition relationships being the main distinguishing feature between consensus and competitive systems and veto power being an element that may or may not be correlated with the two types of party relationships (see also Czada 2000). First, consider the effects of a consensus system on deliberation where all partners in a grand coalition have veto power. In fact, this is the standard form of a consensus system, since by joining a grand coalition, parties obtain veto power, as a political cost (government dissolution or reformation) is involved when their interests are bypassed and they leave the coalition (Tsebelis 1995: 304). Veto power generates, first of all, a logic of joint decision making where argumentative rationality might be necessary for the "problem-solving" or "production" part of policy negotiations in order to successfully cope with different situation definitions, cognitive problems, factual disagreements, as well as standards of appropriateness, fairness and justice (see Risse 2000; Holzinger 2001a; 2001b). Moreover, to maintain the stability of the coalition and to ensure a successful government record, actors will need to adopt a cooperative attitude vis-à-vis their coalition partners comprising a willingness to discuss differences in a constructive and respectful way.[3]

---

[2]   Veto devices can be "institutional", such as supermajorities in constitutional amendment formulas, bicameralism, referenda, or judicial review, or "political", such as resignation threats from a coalition.

[3]   Austen-Smith and Feddersen (2002) contest such reasoning about veto power altogether and even claim that it is actually majoritarian voting rule which promotes deliberation. Although veto power and related unanimity rule create incentives for supporters of the status quo to reveal information, it likewise creates incentives for others to conceal information favoring that status quo. This in turn generates an externality rendering information from all members of the committee suspect. The problem with Austen-Smith and

Two other mechanisms enhance further the deliberative potential in consensus systems. First, when grand coalition arrangements are made permanent and parties obtain relative secure power positions (e.g. obtain a fixed seat), the importance of electoral competition among them is reduced. Second, grand coalitions also obscure policy making clarity, which reduces the possibility that actors can use political successes for partisan electoral advantages (Powell and Whitten 1993). These factors open up an additional "deliberative" space that allows parties to deliberate in a less politicized way, to engage in true argumentative exchanges, as well as to find innovative policy solutions that transcend short-term political goals.

Compare this to a setting with a majority-opposition relationship where the opposition has no veto power (the classical majoritarian setting in Lijphart's scheme). Here, not only can the majority ignore the arguments of the opposition, the fact that the stakes in elections are losing or winning power positions is conducive to competitive interaction orientations among these parties. Moreover, opposition parties who are "cooperative" must always fear that voters tend to assign political successes to the parties in government (Fiorina 1981; Lewis-Beck 1988). Therefore, common problem solving, search for genuine consensus, respect or persuasion may always be a relative loss in the overall competitive game (Scharpf 1997: 166).[4]

It is important to point out that these predictions refer to "ideal" realizations of consensus and competitive systems; they may not always fully hold. First, let us note that it would be misleading to think that grand coalitions simply prevent competitive interaction orientations. Not only must coalition partners face the electorate on their own account, parties might also have an interest in strengthening their position in elections, as this continues to make a difference in the allocation of ministerial posts. This creates incentives to uphold a political identity and defend the interests of their clientele and thus sets a basic constraint on parties' cooperative and deliberative capabilities, which, of course, is further reduced when parties face electoral insecurity or decline. Likewise, it is inappropriate to conceptualize the interaction orientations of opposition parties as merely competitive. For instance, when a policy proposal of the government is highly popular with the broader electorate, the opposition parties may feel pressures to display more cooperative and respectful attitudes.

Second, parties in a grand coalition may not always be able to exert veto power. As grand coalitions are typically oversized, the votes of some parties may not be necessary to pass legislation and when these parties neither have the "opportunity" nor the "motive" to exercise veto (i.e. cannot credibly threaten to resign from the coalition or have a strong office orientation and enjoy the niceties of staying in government) they may be bypassed (Strom 2000: 280).[5] In this case, the majority may forego the careful argumentation of a policy

---

Feddersen's argument is that it involves at least one assumption which is highly problematic in the context of real-world legislatures. They assume that political actors value private interests as much as the common good. This assumption, however, is frequently spoiled in real-world settings: in many cases, real world politicians have clear preferences over different demands and the relative weight they put on private or constituent interests. Under majoritarian decision making, this may create a serious motivational problem for deliberation, as politicians may simply not be willing to listen to minority demands or counterarguments. Hence, only if actors have veto power, other actors are "forced" to listen and deliberate.

[4]   Note that the specified mechanisms favorable to deliberation apply to any coalition setting – be it a minimal-winning coalition or a grand coalition. The reason to focus on grand coalitions is that the quality of deliberation in legislatures is very sensitive to the presence of major opposition parties. Put differently, the more opposition parties are excluded from a coalition arrangement (or the more minimal winning the coalition is), the more adversarial parliamentary deliberation will be.

[5]   Tsebelis (2002: 96) contests such reasoning by arguing that "such simple arithmetic disregards the fact that there are factors than necessitate oversized coalitions. Regardless of what these factors might be, for the coalition to remain intact the will of the different partners must be respected." Yet, he seems to "agree" with Strom that if coalition partners frequently vote against their government, then such a scenario might occur. As Ganghof (2001) rightly notes, Tsebelis' argument would only be true if coalition partners were only

proposal, since persuasion of the minority is not necessary. However, as coalition partners need each other in future interactions, forms of respect may still occur in such situations (Linder 1999).

Conversely, in a majority-opposition setting where the opposition has veto power (e.g. has a majority in the second chamber), higher discourse levels may occur. As the government needs the opposition to pass legislation, this might induce the governing parties to frame their arguments in a more inclusive and respectful way. In addition, if the opposition is willing to enter negotiations. some problem-solving activities with deliberative elements might occur as well. But unlike the situation in grand coalitions where veto power is assumed to be conducive to constructive policy dialogues among coalition partners, we expect a majority-opposition relationship to produce discursive cross-pressures under such circumstances. Party competition will remain strong, and despite pressures for moderation and serious arguing, actors will simultaneously have incentives to attack and degrade the demands and arguments of the political opponent and block respective policy proposals.

In sum, the expectation is neither that a consensus setting automatically leads to a high discourse quality nor that a competitive setting necessarily has a deterministic negative effect on deliberation. Our only claim is that a consensus system enhances the chances that actors engage in deliberation, show respect to demands and arguments of other political actors, and are more constructive.

> HYPOTHESIS 1: Consensus settings lead to a higher level of parliamentary deliberation than competitive settings.

Presidential versus Parliamentary Systems. One of the most important organizational differences between legislatures is the distinction between parliamentary and presidential systems. Using Strom's (2000: 265) minimal definition, a "parliamentary government is a system in which the prime minister and his or her cabinet are accountable to any majority of the members of the parliament and can be voted out of office by the latter, through an ordinary or constructive vote of no confidence". The confidence procedure creates a situation of mutual dependence, creating both incentives and pressures for legislators of the governing parties to vote with "their" government (Diermeier/Feddersen 1998). Disagreement, debate and redrafting of policies always embarrasses the government and ultimately threatens its stability – leading, in general, to strong party discipline. In a presidential regime, to the contrary, the executive is not dependent on legislative confidence, which creates a situation of mutual independence: legislators can vote against the government, initiate, and redraft legislation without threatening governmental stability. Accordingly, party discipline in presidential regimes tends to be lower than in parliamentary regimes. For deliberation, this means that legislators have more leeway to transcend party boundaries and to be open to argument and argumentative change.

However, combining the parliamentary-presidential distinction with competitive and consensus systems leads to complex interactive effects. First, consider the combination of a competitive system and a parliamentary regime. We propose that a parliamentary system strengthens the competitive logic and largely undermines constructive policy discourses. Legislators are not only pressured to defend the proposals of their government, party competition, in general[6], also requires them to steadily attack the political opponent - even if

---

policy-seeking; but as long as one assumes that they are office-seeking too, Strom's argument is corroborated.

[6]   As mentioned before, an exception may be the willingness of the governing party to find agreement with the opposition, when the latter can block a decision.

legislators would privately value or even agree with the political opponent's arguments and demands. If a competitive system is combined with a presidential regime, then there should be more space for deliberation. Not only may legislators be more receptive to arguments across partisan lines, actors from the same party may also find themselves in opposing camps on certain policy issues. This partly blurs the clear-cut majority-opposition pattern as we find it in parliamentary systems with strong party cohesion, and we expect particularly those actors who are on the other side to be moderate or even respectful toward their own party colleagues. Moreover, low party discipline also means that almost every coalition is possible, and this might make legislators more respectful, as today's opponents may be tomorrow's allies. Nonetheless, deliberative quality should be lower than in a consensus system, because presidentialism and lower party discipline do not eliminate strong party competition. As Cox and McCubbins (1993) point out, legislators realize that their electoral fortunes are tied to those of their party, so those legislators who disagree with the political opponent will have a strong incentive to degrade respective arguments. In the combination of a consensus system with parliamentarism (a frequent feature of consensus systems)[7] the management of disagreements – the essence of Habermasian discourse ethics – is mostly "stalled" among coalition MPs. The reason is that deliberation and policy redrafting on part of government MPs always threatens government stability, and party leaders are normally able to preempt it. Rather, disagreements are settled in the executive, which serves as the legislative agenda setter (Tsebelis 2002), or in coalition committees and intra-coalition summits. The resulting compromises will have to be voted on in the legislature, but we should not see much deliberative action, with coalition MPs primarily defending the government's proposals. Therefore, to observe the management of disagreements among coalition MPs, the focus will be on "presidential" consensus systems only[8] – and hence the comparison between presidentialism and parliamentarism will be limited to competitive systems.

> HYPOTHESIS 2: In competitive systems, the quality of deliberation is higher in presidential than in parliamentary regimes.

Second versus First Chambers. From the times of the ancient Greeks onward, there have been staunch advocates, such as Aristotle, Cicero, Mill, and Madison, of a council of elders who could bring their experience and wisdom to the government. In the literature, second chambers are frequently conceived to promote an element of "second" and longer-term thought, and, consequently, more deliberation in the legislative process than first chambers (see Tsebelis/Money 1997: 40). Several factors lubricate this deliberative process: members of the second chamber usually have greater political experience, are usually elected for longer terms, and work in a smaller chamber than their first chamber peers. In addition, second chambers are also said to have "civility" norms which also favor deliberation (Loomis 2002).

> HYPOTHESIS 3: Second chambers have a higher quality of parliamentary deliberation.

Non-public versus Public Arenas. Even though secretive and exclusive arenas tend to horrify most deliberative theorists, many political scientists argue that this very exclusiveness might

---

[7] Prime examples are Austria, Belgium, and the Netherlands.

[8] As Lijphart (2002: 49) correctly notes, the standard form of a presidential regime with a single president and presidential cabinet consisting merely of advisers might be difficult to reconcile with the grand coalition idea where all partners are co-decision-makers. Yet the real world involves more variety in this respect. In the Swiss cantons, for instance, we find directly elected collegial grand coalition executives which are not in a confidence relationship with the respective legislatures.

actually benefit deliberation (e.g. Checkel 1999). Non-public legislative bodies such as parliamentary committees allow legislators to deliberate without external interference, lower the pressures of following constituent demands, and make it easier for politicians to reflect, to show respect for the claims of others, or even to change their opinions. Conversely, in full sessions of parliament, one can expect that most actors will primarily engage in ritualistic rhetoric and strategic arguing. Furthermore, committees are small face-to-face groups that frequently operate over an extended period of time, which may create habits of working together, friendships, as well as knowledge about each other. These outcomes, in turn, may foster trust and, as such, lubricate the deliberative process. In a study on constituent assemblies, Elster (1998b) found that speaking in the public was indeed not conducive to calm and impartial deliberation; behind closed doors, however, actors talked much more seriously although the non-public character simultaneously enhanced bargaining. However, Elster (1998a) has also pointed out that public discourses tend to have a "civilizing effect" on the participants in the sense that explicitly selfish interests can rarely be justified in the public sphere. Moreover, public deliberation also forces participants to adhere to rough standards of reasoned argument. The expectation for discourse quality in the public and non-public sphere is twofold:

> HYPOTHESIS 4: Non-public arenas enhance the quality of parliamentary deliberation in terms of respect and constructive politics, but decrease it in terms of justification rationality and common good orientations.

Low Issue Polarization versus High Issue Polarization. So far, we have only considered institutional determinants of deliberation. But institutional scholars increasingly recognize the importance of the interaction between political institutions and issues (see Scharpf 1997; Tsebelis 2001). Of particular importance is the distance in the policy preferences of actors, what one could call the ideational (or more narrowly, ideological) dimension of policy issues.

To capture the divergence of issue preferences we distinguish between polarized and non-polarized issues, viewing these as endpoints on a continuum that captures the divergence of issue preferences.[9] Non-polarized issues are characterized by a consensus of the elites on key values and goals, while polarized issues are marked by sharp disagreements on these matters. Even if the elites agree that some goal is desirable, an issue may still be polarized because the elites fundamentally disagree about the best means to achieve this goal (see von Beyme 1997). Moreover, "less polarized" issues are, in general, less salient for partisan electoral success.

How does the polarization dimension affect deliberation? If an issue is not very polarized, then we should expect more cooperative interaction orientations among the political actors. And even those actors whose self-interest is in conflict with the values of the discursive consensus might be rhetorically constrained as it is difficult and costly to undermine a prior discursive consensus or a highly valued goal. This should promote more consensual decision making and, as such, should lead to higher levels of discourse.[10] The "issue polarization" hypothesis is, of course, quite a radical departure from the Habermasian discourse model with its emphasis on liberal neutrality, stipulating that all issue types should

---

[9] We prefer this language to the distinction between positional and valence issues introduced by Stokes (1963), since we believe that all issues are valenced in the sense of being affectively charged. However, if the reader prefers the more traditional distinction, positional issues may be substituted for polarized issues and valence issues may be substituted for non-polarized issues, without a loss of generality.

[10] If the issue is perfectly non-polarized, i.e. there is absolute agreement on the goals, policies, and implementation, then we may find little discourse because it is unnecessary to deliberate. Since very few issues fall into this category, we shall not consider it any further.

be equally apt to be submitted to the discourse procedure. Rather, the proposition that it is "valence" or less polarized issues that lead to a higher quality of deliberation harkens back to the republican conception of democracy where people share common values and goals as well as a common conception of the common good.[11]

> HYPOTHESIS 5: Low issue polarization leads to a higher quality of parliamentary deliberation

The hypotheses developed above lay out four institutional factors and one issue characteristic that we believe to affect the level of parliamentary deliberation. In practice, these factors do not work in isolation but will combine to form *constellations* that facilitate or hinder parliamentary deliberation. We should observe the highest level of parliamentary deliberation in a consensus presidential system when all coalition partners possess veto power, when a non-public debate occurs in the second chamber concerning a non-polarized issue. The lowest level of parliamentary deliberation should occur in parliamentary competitive systems where the opposition has no veto power, when a public debate occurs in the first chamber concerning a polarized issue. Other constellations will be located between these two extremes, depending on how many facilitating and how many debilitating factors are present.

Research Design

The foundation of our research design is the idea of typological or theoretical sampling (e.g. Lamnek 1995; McKeown 1999). In contrast to random sampling, the case selection is purposive with respect to the typology. That is, cases are selected to fill out the categories, whereby the goals are to select cases that capture the extremes of the independent variables entailed by the categories and to select cases that are comparable in other regards.

However, we depart from the "qualitative" legacy of typological sampling by conducting statistical analyses on the debates that are selected through this type of sampling. The goal of this statistical approach is not to produce generalizable inferences, since that would require case selection through random sampling. Nevertheless, we think that a typological sampling can serve an important purpose. Its goal is to generate a set of "critical cases" (Eckstein 1975), i.e. cases for which we would expect the theory to hold true, if it is going to hold true at all. If we were to find that the theory does not hold true for those cases, this would cast serious doubt on the theory, regardless of the fact that the cases do not constitute a random sample.

One may wonder why we do not use random sampling. We see two main reasons for this. First, random sampling requires a well-defined population and sampling frame. While it is already difficult to construe what exactly a population of debates is, especially across countries and time, it is near impossible to find a proper sampling frame. There are no lists of debates that permit us to draw a random sample. To be sure, one can lay one's hands on listings of bills, but that sampling frame is insufficiently detailed because a bill can trigger

---

[11] With respect to issue specificity, other hypotheses might be formulated: for instance, one might argue that deliberation is more prevalent in the context of regulative issues than in the context of mere distributive issues (e.g. Elgström/Jönsson 2000: 701). We shall not consider this hypothesis since is quite difficult to properly distinguish between regulative and distributive issues empirically, as most real-world issues are complex mixtures of both. However, the issues we selected always involve a clear regulative dimension (see below). Furthermore, uncertainty might matter for deliberation too in the sense that high levels of uncertainty surrounding an issue could be a strong incentive for actors to deliberate. However, the level of uncertainty surrounding an issue cannot be "objectively" determined but would require interviewing legislators about their respective perceptions, a task which goes beyond the confines of this study.

multiple debates. Second, even if a sampling frame had existed, we would be wary of drawing a random sample because we want to select debates that matter. A random sample might have yielded a selection of debates on relatively minute issues. That might have been correct from a statistical viewpoint, but it would have been of little theoretical interest. A purposive sampling technique is more likely to select such issues than random sampling.

The Selection of Legislatures and Debates
Typological sampling requires that one selects cases—legislatures and debates—that are "ideal" and "most comparable." An assessment of hypotheses 1-5 requires that we select consensus presidential, competitive presidential, and competitive parliamentary systems. For these three basic configurations, we selected Switzerland (a "non-parliamentary" consensus system[12]), the United States (competitive-presidential), Germany and United Kingdom (competitive-parliamentary). With the exception of Britain, the legislatures also allow to investigate debates under strong and weak veto power (which is important in order not to over- or underestimate discourse quality in consensus and competitive systems).[13]

In addition, we need to be able to study debates in two legislative chambers and in two different arenas, open and closed. To shed light on the effect of second versus first chambers, we consider the German, Swiss, and U.S. legislatures. Each of these legislatures is bicameral and in each case the second chamber can be considered a powerful actor.[14] Our analysis of closed versus open arenas of deliberation focuses on the German and Swiss cases. Indeed, the German-Swiss comparison is appropriate due to strong similarities in the institutional workings of committees. In both cases, committee deliberations are non-public and committees have considerable power, for example in terms of rewriting government proposals (Döring 1995). We exclude the United States from the analysis because committee meetings are usually public and there are few other forums in which legislators deliberate behind closed doors. This problem comes into play also with the House of Commons. In addition, committees in the House do not have the prerogative of rewriting legislative proposals, which necessarily curbs the character of their deliberations. This limited diversity in the real world of legislatures involves the problem that we may miss out a real negotiation setting in a competitive system behind closed doors. The only competitive non-public arena in our sample are the *Bundestag* committees; in the latter, however, party discipline is generally strong, which we theoretically expect to seriously undermine negotiation and deliberation. In order to explore the important constellation of a negotiation setting in a competitive system behind closed doors, we use data from the German Conference Committee (*Vermittlungsausschuss*), a body which tries to reconcile conflicts between the

---

12 Analysts frequently denote it "semi-presidential" or "directorial": although parliament elects the government (the Federal Council), the legislature cannot stage a vote of no confidence during the term of the government. If a government proposal is defeated by the Federal Assembly, then it is not necessary for the government to resign. Since the government's existence does not depend on legislative majorities, party discipline tends to be low compared to other European parliamentary systems and MPs have a great deal of autonomy on drawing up legislation.

13 In Switzerland, the distinction is made according to the presence or absence of a referendum threat of one of the coalition partners. The reason for focusing on the institutional veto device of the referendum and not on the political veto device of the resignation threat is that Swiss coalition partners frequently form various coalitions among themselves (Linder 1999) and do not resign from the coalition when they are outvoted. In the United States, it is quite difficult to clearly distinguish between veto and non-veto situations, as low party discipline frequently leads to negotiations among different factions of MPs as well as between the two chambers. Hence, we focus on decision situations where there were direct and credible veto threats (such as a presidential veto or a filibuster threat) and on decision situations where it was "clear" to the actors that a bill would pass with a simple majority. In Germany, a clear veto-situation arises when the opposition possesses a majority in the *Bundesrat* and the consent of the latter is required to pass a bill. Other bills can be generally decided with the simple majority of the governing parties.

14 We do not consider the House of Lords in Britain, since this chamber has relatively little power.

two parliamentary chambers, the *Bundestag* and *Bundesrat*. Although the Conference Committee is located in a system with strong party discipline, its members are partly relieved from partisan directives with the expectation to develop innovative approaches that overcome blockades (Spörndli 2004; Lhotta 2001).

Debates

Comparing debates across different national legislatures raises a fundamental problem. To demonstrate empirically that institutions matter for deliberation we have to control somehow for the impact of different actors' preference profiles that different institutional systems face (Hammond and Butler 2003). Our attempt to control for different preference profiles consists of choosing topics of debate for which actor constellations are similar (for example, social and economic issues with polarization between leftist and rightist parties). This approach is not foolproof: it is generally impossible to find debates on exactly the same set of issues in crossnational research. However, we try to draw comparisons only if the broad parameters of debate are comparable, which means that at least the topic should be comparable.

We selected a total of 52 debates, mostly from the late 1980s and 1990s,[15] from the four countries described earlier. A key selection criterion for the debates is that they should focus on critical decision cases, as opposed to minute issues, and that they should cover topics found in at least two legislative settings. A key differentiating factor is the level of polarization of the issue that is being debated. To draw the distinction between low and high levels of polarization we build on a useful classification developed by Kriesi (2001), which considers two issue dimensions: degree of polarization and salience. Using German data, Kriesi assigns social and economic policies to the polarized pole, a practice that we shall follow here since these issues are tied closely to the left-right ideological divide that has structured the political party systems of Germany, Switzerland, the United Kingdom, and the United States since (at least) the early 20th century (see Lipset and Rokkan 1967).[16]

Identifying issues with low levels of polarization is more difficult, but we have found several such issues. One of these concerns the rights of disabled people and animal welfare (cruelty to animals). Kriesi (2001) argues that crime prevention also fits in the category of low polarization, since most parties agree that crime rates should be lowered even if they do not agree on the details of how this should be done. Compared to rights for the disabled and animal protection, crime is a high salience issue, according to Kriesi. Because of this aspect we shall analyze crime separately from the other, less salient non-polarized issues.[17] The specific debates that we have analyzed can be found in Appendix 3.

One element of the case selection deserves special emphasis. In an ideal world, we would have a fully crossed design in which we would consider the impact of bicameralism, issue type (polarized versus non-polarized), arena (public versus non-public), and veto power for each of the four countries that we selected. Unfortunately that approach is not feasible for a variety of reasons. First, certain cells of the design are empty (for example, non-public arenas in the United States). Second, other cells are not empty but we were unable to detect debates with sufficient numbers of speech acts (at least ten). Third, comparability was also a concern, in that we were sometimes unable to find truly comparable debates. The empty cell problem produces some limitations for the analysis, a problem which we shall address below.

---

15 The exceptions to this time frame are the debates in the German Conference Committee which are drawn from the 1970s.

16 We find supportive evidence by analyzing the Party Manifesto data set (Budge et al. 2001) for differences in party positions on the issues of free market capitalism and the welfare state. On those important economic issues, party differences between the left and right are sizable in all four countries suggesting a high degree of polarization.

17 We have also analyzed other issues that are non-polarized, but those issues are unique to particular countries.

Measures

Predictors. The predictors in our analysis are the various institutional and issue characteristics contained in Hypotheses 1-5. Each characteristic is measured as a dichotomy: either a debate possesses the characteristic or it does not. While one could think of more precise measures, using dichotomies has the advantage of casting the influence of the characteristics in the starkest possible manner.

Dependent Variable: *The Discourse Quality Index (DQI).* To "measure" the quality of deliberation, we have developed a "discourse quality index" (DQI) (Steenbergen et al. 2003), based upon broadly shared concepts of the theoretical literature on deliberative democracy (see Appendix 1). The DQI measures the quality of deliberation with seven indicators that conceptualize deliberation as a continuous variable and allows for a quantitative content analysis at the speech acts' level of recorded debates. The seven indicators are: (1) participation (is it impaired or not?); (2) level of justification (do speakers just forward demands or do they give reasons for their positions, and how sophisticated do they justify their positions?); (3) content of justification (do speakers cast their justifications in terms of a conception of the common good or in terms of narrow group or constituency interests?); (4) respect toward group (do speakers degrade, treat neutrally, or value groups to be helped?); (5) respect toward demand (do speakers degrade, treat neutrally, value, or agree with demands?); (6) respect toward counterarguments (do speakers degrade, ignore, treat neutrally, value, or agree with counterarguments to their position?); and (7) constructive politics (do speakers sit on their positions or submit alternative or mediating proposals?). A detailed description of the seven indicators can be found in Appendix 1.

Habermasian discourse ethics would also require truthfulness (*Wahrhaftigkeit*), which is the absence of deception in expressing intentions. The stated preferences should be sincere rather than strategic so that the discourse can develop in an open and honest manner. While we do acknowledge the importance of truthfulness for deliberative theory, it causes the greatest difficulties from a measurement perspective. To judge whether a speech act is truthful is to make a judgement about a person's true versus stated preferences. This is exceedingly difficult, since the true preferences are not directly observable. The speculative nature of such a judgement is bound to introduce large amounts of (possibly systematic) measurement error and for this reason we shall not further consider truthfulness.

In a series of tests, we could demonstrate that the "inter-coder reliability" of the DQI ranges between good and excellent. Two different coders scored a subset of the debates analysed in this manuscript. The rate of coder agreement ranged from a low of .919 for respect toward counterarguments to a high of 1 for content of justification. Because of the possibility of inter-coder agreement by chance we also computed Cohen's kappa, which controls for this. Kappa ranges from a low of .881 for counterarguments to a high of .954 for constructive politics. These values indicate excellent inter-coder reliability (for more details see Steiner et al. 2004, Chapter 3).

For the empirical analysis the different dimensions of the DQI will be assessed individually (we exclude the participation dimension from the analysis as this did not show sufficient variation[18]). The reason for an individual assessment of dimensions is both theoretical (for some dimensions, there are different theoretical expectations according to the hypotheses) and empirical (the dimensions of the DQI are only weakly (though positively) correlated and thus do not satisfy the criterion of uni-dimensionality (Steenbergen 2000). For the respect category, however, we use combined indices, since this dimension constitutes a theoretical whole. We constructed two such indices. The first is an additive scale of respect

---

[18] There were only seven speakers (out of 5500) whose participation was hampered.

toward demands, counterarguments, and groups, which we dub "overall respect." This index ranges from 0-9. Since respect toward groups may be the product of strategic considerations—some groups may be constituents and may be treated with respect for that reason, not because of discourse ethics—we also computed a second index that simply adds the respect toward demands and counterarguments indicators. We call this indicator "respect toward demands/counterarguments," and it ranges from 0-7. We report means for the two respect indicators, while we report percentages for all of the other indicators.[19]

Statistical Analysis

Our primary analytical vehicle is to present the frequency distributions of the indicators of the DQI for different levels of the predictors; for the respect indicators we report means. To test whether differences are statistically significant, we performed a series of ANOVAs and logistic regressions, depending on the measurement level of a dependent variable. The idea behind these tests is to assess if a predictor exerts a systematic effect on the quality of deliberation, not so much to generalize our results to the universe of all deliberations.

Two specific aspects of our research design need to be addressed, and this will also help to clarify our statistical procedures. First, as mentioned before, there is a problem with empty cells. This means that comparisons can only be made among comparable cases. For instance, if we want to know something about the effect of non-publicity on deliberative quality, we need to compare non-public debates with public ones that only vary with regard to this particular factor; e.g., if a non-public debate is highly polarized and conducted in a second chamber of a presidential competitive setting with veto power of the opposition, then we need a public debate that exactly has the same constellational features. If no such a debate is available, then the respective case cannot be used for the analysis. This procedure simultaneously allows to control for the effects of the other explanatory factors and to calculate a net effect of the factor of interest. As we will always make use of all comparable debates, every analysis is based on a different debate selection.

Second, we are also confronted with an unbalanced design, since the debates entering the analyses have different numbers of speech acts, which frequently translates into different numbers of cases across the levels of the dependent variable. We deal with this issue by computing Type-III sums of squares for the ANOVAs and applying a similar logic to the logistic regression models. In addition, we include a blocking variable into the analyses. This variable divides the speech acts into relatively homogeneous sub-sets, thus decreasing the error variance. Since speech acts come from different debates, blocking them in this way is essential in order to control for idiosyncratic aspects of the debates.[20] The results of the ANOVA and regression significance tests are reported in Appendix 2.

---

[19] For simplicity and transparency reasons we have simply added the different indicators of respect without giving them different weights. However, as the different indicators have unequal numbers of categories, some indicators might distort overall scores and findings. We have rerun all analyses with weighted indices but results do not change.

[20] While we control for the blocking variable in our analysis, we do not report any statistical results for this variable due to space limitations. However, these results are available on request from the authors.

Results

HYPOTHESIS 1: Consensus Versus Competitive Systems
This comparison is performed in two stages: we compare, first of all, Swiss debates to German debates with strong party discipline. Then we compare Swiss debates to US and German debates where party discipline is lower. The second comparison represents the higher confirmation bar, as we theoretically expect that a parliamentary system with strong party discipline such as Germany accentuates the competitive logic, hence underestimating the deliberative potential of competitive settings. The results from this analysis are shown in Table 1:

Table 1
Discourse Quality in Consensus and Competitive Settings: A Swiss-German Comparison

	Setting	
Discourse Quality Dimension	**Consensus (Switzerland)**	**Competitive (Germany)**
Level of Justification:[a]		
No Justification	10.10	20.90
Inferior Justification	13.60	17.80
Qualified Justification	27.30	21.50
Sophisticated Justification I	40.40	34.30
Sophisticated Justification II	8.60	5.50
Overall Respect (Mean):[b]	3.62	3.12
Respect Demands/Counterarg.(Mean):[b]	2.48	2.11
Constructive Politics:[a]		
Positional Politics	81.00	86.20
Alternative Proposal	8.20	6.60
Consensus Appeal	6.20	5.00
Mediating Proposal	4.60	2.20

Notes: Table entries are percentages (which add up to 100 percent in each column). For a description of the cases entering this analysis see http://www.ipw.unibe.ch/discourse/. [a] Heckles excluded (N = 1383 speech acts); [b] heckles included (N = 1896 speech acts).

This table reveals important differences between Switzerland and Germany. For the level of justification and the two respect indicators, the Swiss consensus system scores considerably higher. First, consider the level of justification. Qualified and sophisticated debates were far more common in the Swiss debates (76.3 percent) than in the German debates (61.3 percent). The mean levels of both respect indicators are also higher in Switzerland; the mean difference is .50 for overall respect and .37. for respect toward demands/counterarguments.[21] Regarding constructive politics, there is a tendency toward greater consensus appeals and mediating proposals in the Swiss system (10.8 percent versus 7.2 percent in Germany). In all these dimensions, differences turned out to be statistically significant (at the .01 level) in the logistic regression and ANOVA tests. Let us now look at the second test in Table 2:

---

[21] The significant differences for both respect indicators are robust when heckles are excluded. This also holds true for all other analyses in this article.

Table 2
Discourse Quality in Consensus and Competitive Settings: A Swiss-US/German Comparison

Discourse Quality Dimension	Setting	
	Consensus (Switzerland)	Competitive (USA and Germany)
Level of Justification: [a]		
No Justification	10.60	8.10
Inferior Justification	13.70	15.50
Qualified Justification	26.30	30.00
Sophisticated Justification I	40.30	34.60
Sophisticated Justification II	9.10	11.80
Overall Respect (Mean): [b]	3.67	3.31
Respect Demands/Counterarg.(Mean): [b]	2.53	2.24
Constructive Politics: [a]		
Positional Politics	81.50	81.30
Alternative Proposal	7.70	7.00
Consensus Appeal	5.80	8.40
Mediating Proposal	5.00	3.30

Notes: Table entries are percentages (which add up to 100 percent in each column). For a description of the cases entering this analysis see http://www.ipw.unibe.ch/discourse/. [a] Heckles excluded (N = 2409 speech acts); [b] heckles included (N = 2886 speech acts).

This table again shows important differences in the level of respect between consensus and competitive systems. While these differences are not as large as they were in the Swiss-German comparison, they are still unmistakable. The mean levels of overall respect are also different, at 3.67 for Switzerland and 3.31 for the United States and Germany. ANOVA shows that these differences on the two respect indicators are highly significant (at the .01 level).

In terms of level of justification and constructive politics, Table 2 does not reveal substantive differences between consensus and competitive systems. While for the level of justification, the logistic regression test still indicates a significant difference (with the competitive setting scoring higher), the test fails to show a significant effect for the constructive politics dimension.

Two issues warrant further attention. First, what role does veto power play? In Table 1 and 2 we have only controlled for the effects veto power, but we did not assess its role as a lubricant for deliberation. Testing for this yields mixed results (analysis not shown): while veto power produces only small and statistically insignificant differences for justification and constructive politics, it tends to enhance respect levels – although with the exception of the Germany where veto and non veto situations produce similar respect scores. Apart from this, veto power goes in tandem with the type of party-relationships (grand coalition vs. majority-opposition): it increases respect levels in both the Swiss grand coalition setting and the US majority-minority setting, but does so only in an additive way, i.e. the effects of the different types of party relationships are not wiped out.

Second, the major problem with the comparison between consensus and competitive systems is that one might object that findings are biased since cultural variables have been omitted from the analysis. The argument would be that the more consensual political culture in Switzerland (e.g. Lijphart 1999: 307) would account for the higher respect levels in the Swiss debates, while the more competitive political culture in the US would account for the lower respect levels in the American debates. While the cultural argument may be difficult to refute in the comparison of Swiss and American debates, this is less true for the comparison of Swiss and German debates. Lehmbruch (2003), for instance, stresses the "conflict-averse" German political culture. Thus, the relatively large differences in the respect dimension

between Swiss and German debates are more difficult to attribute to cultural differences solely. Moreover, as soon as institutional situations are changed within the same organizational and cultural setting, DQI scores move in the predicted directions. A particularly striking case in this respect is the comparison between the Swiss and German abortion debate. In both debates, demands, preference profiles, actor constellations, and actors' veto power were almost identical. In addition, the German debate "mimicked" the Swiss institutional setting in two other respects as well: not only was party discipline relaxed for this issue, there were even some claims that opposition demands should be integrated – thus the debate partly approximated a grand coalition setting. While the public debate in the *Bundestag* has lower respect scores than its Swiss counterpart, the differences in the non-public setting of the first chambers tend to wither away: the "overall respect" indicator in the Bundestag committee debate is 3.59 compared to 3.72 in the Swiss first chamber committee debate.[22] These findings underline the validity of our theoretical framework: they show that variation in the institutional parameters can have a profound effect on the quality of deliberation, regardless of the cultural context.

In sum, our results suggest important differences between consensus and competitive systems, in particular with regard to respect. This result takes on additional meaning if one realizes that Swiss coalition parties are not much concerned about coalitional stability and that the coalition parties became increasingly polarized during the 1990s, the time frame of our debates (see Hug and Sciarini 1995). However, it is important to stress that the relationships between the Swiss consensus and the US and German competitive settings are far from being deterministic or "existential": neither are elements of high discourse quality absent from competitive settings nor are consensus settings a safeguard against low discourse quality. The score of demands and counterarguments that are respected and/or agreed with amounts to 9.9 percent (first comparison) and 13.7 percent (second comparison) in the German and American competitive settings (and not tending toward zero as an existential relationship would suggest), while the score of disrespectful speech acts in the Swiss consensus setting achieves a remarkably high score of 23.4 percent (first comparison) and 23.1 percent (second comparison).[23]

HYPOTHESIS 2: Presidential Versus Parliamentary Systems
Our second hypothesis argues that, within the subset of competitive settings, discourse quality should be higher in presidential systems than in parliamentary systems. To test this hypothesis we pit discourse in the United States, a presidential system, against discourse in Germany and the United Kingdom, two examples of parliamentary systems. The results are displayed in Table 3.

---

[22] Similar effects also occur in the British setting: as soon as party discipline is relaxed and pure majoritarian strategies become unavailable, respect levels are enhanced considerably.

[23] This is based on a categorical "respect toward demands/counterarguments" indicator, which collapses respect toward demands and counterarguments in one dimension, involving four categories: disrespect toward demands and/or counterarguments, a neutral category, explicit respect toward demands and/or counterarguments, and acceptance of demands and/or counterarguments.

Table 3
Discourse Quality in the Contexts of Presidential and Parliamentary Competitive Systems

Discourse Quality Dimension	Setting	
	Presidential (USA)	Parliamentary (Germany and UK)
Level of Justification: [a]		
No Justification	11.40	11.20
Inferior Justification	18.60	13.40
Qualified Justification	29.30	8.40
Sophisticated Justification I	34.40	49.60
Sophisticated Justification II	6.30	17.40
Overall Respect (Mean): [b]	3.24	2.62
Respect Demands/Counterarg (Mean): [b]	2.11	1.61
Constructive Politics: [a]		
Positional Politics	89.90	83.00
Alternative Proposal	6.10	8.70
Consensus Appeal	3.70	6.90
Mediating Proposal	.30	1.40

Notes: Table entries are percentages (which add up to 100 percent in each column). For a description of the cases entering this analysis see http://www.ipw.unibe.ch/discourse/. [a] Heckles excluded (N = 497 speech acts); [b] heckles included (N = 1031 speech acts).

As a whole, these results produce some support for our second hypothesis, but they also contradict it in important ways. The contradictions are found on the level of justification and constructive politics dimensions. Sophisticated justifications are more common in parliamentary systems (67 percent) than in presidential systems (40.7 percent), a finding that runs counter to our expectations. In addition, pure positional politics was less common in parliamentary systems (83 percent) than in presidential systems (89.9 percent) and this, too, contradicts our second hypothesis. In the ANOVA and regression tests, these results are statistically significant (at the .05 level or better).

There is support for our third hypothesis when we consider the respect dimension. When heckles are included in the analysis, there are marked differences between the American presidential and the German and British parliamentary system. The respective mean differences are .62 (overall respect) and .49 (respect toward demands/counterarguments). ANOVA and regression tests show these effects to be statistically significant (at the .01 level).

In conclusion, the comparison between parliamentary and presidential competitive systems suggests differences in respect. These differences are in the expected direction and suggest better opportunities for respectful discourse in presidential settings. We also observe some counter-intuitive effects on discourse, whereby parliamentary systems produce higher levels of justification and constructive politics. These results are a puzzle for the moment. Clearly, more theoretical development is required to account for the results, although this should not distract from the finding that respect behaves as we had predicted.

HYPOTHESIS 3: Second versus First Chambers

Table 4
Discourse Quality in First and Second Chambers of the Legislature

Discourse Quality Dimension	Setting	
	Second Chambers	First Chambers
Level of Justification:[a]		
No Justification	15.00	12.70
Inferior Justification	14.70	15.50
Qualified Justification	22.00	26.40
Sophisticated Justification I	33.20	38.20
Sophisticated Justification II	15.10	7.20
Overall Respect (Mean):[b]	3.77	3.40
Respect Demands/Counterarg. (Mean):[b]	2.64	2.29
Constructive Politics:[a]		
Positional Politics	82.60	85.20
Alternative Proposal	7.70	8.60
Consensus Appeal	6.10	3.60
Mediating Proposal	3.60	2.60

Notes: Table entries are percentages (which add up to 100 percent in each column). For a description of the cases entering this analysis see http://www.ipw.unibe.ch/discourse/. [a] Heckles excluded (N = 3694 speech acts); [b] heckles included (N = 4096 speech acts).

As table 4 shows, discourse indeed is of a higher quality in second chambers than in first chambers. Consider first our indicators of respect. The mean difference in the overall respect indicator is .37 and the respect toward demand/counterargument indicator is .36. These effects of chamber on respect are statistically significant in the ANOVA test (at the .01 level). In addition to this powerful effect, we see a slight tendency in first chambers toward greater positional politics compared to second chambers. Logistic regression analysis reveals this effect to be significant as well (at the .05 level). The only difference that is not statistically significant is that on levels of justification; those are comparable in first and in second chambers.

In sum, debates in second chambers tends to be more respectful and also slightly more constructive. This underlines the role of second chambers as "reflexive " bodies, injecting an element of second thought and moderation into the policymaking process.

HYPOTHESIS 4: Non-public versus Public Arenas
Here we also consider the fourth dimension of the DQI: content of justification, as we have developed a specific hypothesis for this dimension in the context of publicity/non-publicity.

Table 5
Discourse Quality in Public and Non-Public Arenas

Discourse Quality Dimension	Setting	
	Non–Public	Public
Level of Justification:[a]		
No Justification	21.10	5.80
Inferior Justification	19.20	6.00
Qualified Justification	30.10	12.10
Sophisticated Justification I	24.60	54.90
Sophisticated Justification II	5.00	21.20
Content of Justification:[a]		
Group Interest	5.00	8.80
Neutral	86.30	60.70
Common Good (Utilitarian)	6.10	17.00
Common Good (Difference Principle)	2.60	13.50
Overall Respect (Mean):[b]	3.67	3.36
Respect Demands/Counterarg (Mean):[b]	2.61	2.21
Constructive Politics:[a]		
Positional Politics	83.20	83.40
Alternative Proposal	7.20	8.10
Consensus Appeal	5.10	6.30
Mediating Proposal	4.50	2.20

Notes: Table entries are percentages (which add up to 100 percent in each column). For a description of the cases entering this analysis see http://www.ipw.unibe.ch/discourse/. [a] Heckles excluded (N = 3086 speech acts, except for content of justification where N = 2166 speech acts); [b] heckles included (N = 4221 speech acts).

As table 5 reveals, there are marked differences in the discourse found in public and non-public arenas. First, consider the aspect of justifications. The level of justifications is significantly lower in non-public than in public arenas. Only about 60 percent of the speech acts in closed debates contained qualified and sophisticated justifications, while the percentage of such speech acts approached 90 percent in open arenas, a statistically significant difference. This clearly supports the argument that publicity strengthens "civility" in that actors want to appear reasonable in public and therefore forward more extended arguments. In non-public arenas, to the contrary, the need to always justify one's positions is clearly reduced. Moreover, non-publicity simultaneously enhances bargaining which does not require justifications, at least not lengthy ones.

Public debates also contained significantly larger numbers of appeals to the common good than non-public debates: 30.5 percent versus 8.7 percent. Apparently, in public, actors seem to have a stronger pressure to make appeals to the common good in order to respond to the moral claims of their constituents.

Where non-public arenas fare much better than public arenas is on the respect dimension. As table 5 shows, overall respect hovers at 3.67 in the non-public sphere and at 3.37 in the public sphere. For respect for demands/counterarguments, differences are even more marked, with 2.61 in the non-public sphere and 2.21 in the public sphere. These effects on respect are statistically significant in an ANOVA (at the .01 level).

There is a major and statistically significant interactive effect (at the .01 level) between non-publicity/publicity and a consensus/competitive system. In consensus systems, difference in respect between public and non-public debates tends to fade. In competitive settings, however, this difference is large. When playing for an audience of citizens, legislators in a competitive system know that there is much to gain by discrediting one's opponents and little to gain from praising them. In a consensus system, conversely, this incentive is reduced.

The only area in which the difference between public and non-public debates is non-significant is constructive politics. Positional politics are the norm, regardless of whether one deliberates in public or behind closed doors.

In the light of the philosophical discussion on deliberation, where publicity is often viewed as a crucial aspect of deliberation, these findings might be troubling and reassuring at the same time. Reassuring should be the fact that public discourse is better in terms of the level and content of justification, although one may wonder about the authenticity of appeals to the public good in this setting. Troublesome should be the fact that public discourse can be quite disrespectful, at least in competitive settings. Respect, of course, plays an important role in forging policy coalitions in the present and in the future and, in this regard, Elster (1998b) appears to have been correct when he argued that deliberation benefits when it is insulated from the public.

HYPOTHESIS 5: Less polarized versus High Polarized Issues
We present these data in two parts. First, we compare low-polarized with high-polarized debates on social issues. In the second analysis, the highly polarized abortion debates in Germany, Switzerland, and the United States are pitted against debates on less polarized issues.

Table 6
Discourse Quality and the Level of Issue Polarization

	Setting	
Discourse Quality Dimension	**Low Polarization**	**High Polarization**
Level of Justification:[a]		
No Justification	10.40	16.10
Inferior Justification	15.80	20.10
Qualified Justification	27.10	26.70
Sophisticated Justification I	33.70	31.20
Sophisticated Justification II	13.00	5.90
Overall Respect (Mean):[b]	3.83	3.43
Respect Demands/Counterarg. (Mean):[b]	2.72	2.33
Constructive Politics:[a]		
Positional Politics	81.60	82.00
Alternative Proposal	8.30	8.70
Consensus Appeal	6.20	7.00
Mediating Proposal	3.90	2.30

Notes: Table entries are percentages (which add up to 100 percent in each column). For a description of the cases entering this analysis see http://www.ipw.unibe.ch/discourse/. [a] Heckles excluded (N = 2995 speech acts); [b] heckles included (N = 3516 speech acts).

The first comparison supports hypothesis 5: there are considerable and statistically significant differences (at the .01 level) between low and highly polarized issues in the levels of justification and in the two respect indicators. Sophisticated justifications are more common when issue polarization is low (45.7 percent) then when issue position is high (37.1 percent). Respect levels are also considerably higher in less polarized issues – the mean differences are .40 (overall respect) and .39 (respect toward demands/counterarguments). For constructive politics, differences are virtually non-existent and also statistically insignificant.

Table 7
Discourse Quality and the Level of Issue Polarization

Discourse Quality Dimension	Setting	
	Low Polarization	High Polarization
Level of Justification:[a]		
No Justification	10.00	11.20
Inferior Justification	13.10	17.60
Qualified Justification	21.90	29.60
Sophisticated Justification I	40.00	28.30
Sophisticated Justification II	15.00	13.30
Overall Respect (Mean):[b]	4.07	3.42
Respect Demands/Counterarg.(Mean):[b]	2.91	2.30
Constructive Politics:[a]		
Positional Politics	83.20	86.00
Alternative Proposal	7.80	7.20
Consensus Appeal	5.40	3.10
Mediating Proposal	3.60	3.70

Notes: Table entries are percentages (which add up to 100 percent in each column). For a description of the cases entering this analysis see http://www.ipw.unibe.ch/discourse/. [a] Heckles excluded (N = 2196 speech acts); [b] heckles included (N = 2219 speech acts).

When we focus on the comparison between abortion and less polarized issues, we obtain a similar pattern of results. Again, levels of justification and respect score significantly higher when issue polarization is low. For constructive politics, however, we find a small but significant difference between the two types of issues (16.8 percent constructive speech acts in the context of less polarized debates and 14 percent in the context of abortion debates).

We should hasten to add, however, that we have observed quite a bit of variance in the debates on polarized issues. Polarization does not preordain poor deliberation. Particularly the German and Swiss abortion debates display an intriguing pattern of disrespectful and respectful speech acts. Horst Eylmann (Christian Democrat) nicely summarizes this ambiguity: "this is […] a debate of personal testimonies, moral appeals, strict positions, and sometimes of big words too, but it also is a debate of critical and skeptical reflexivity" (*Bundestag*, June 25, 1992, p. 8259).[24] Nevertheless, the fact that deliberation has much higher chances when issue polarization is low is a troubling result for deliberative democrats hoping that deliberation can tame or even resolve stark moral disagreements. The findings reveal a close association of deliberation with a republican model of democracy where actors share common goals and common identities.

A constellational view of parliamentary deliberation
After evaluating the five hypotheses individually, we now go back to a "constellational" view and ask what happens when the favorable factors combine in an additive way. From this perspective, we should observe the highest level of parliamentary deliberation in consensus (presidential) systems with strong veto players, when a non-public debate occurs in the second chamber concerning a non-polarized issue. The lowest level of parliamentary deliberation should occur in competitive parliamentary systems where the opposition has not veto power, when a public debate occurs in the first chamber concerning a polarized issue. For the respect dimensions, this expectation is fully corroborated, and under the best

---

[24] "Dies ist […] eine Debatte der persönlichen Bekenntnisse, der moralischen Appelle, der entschiedenen Haltungen, zuweilen auch der grossen Worte, aber auch eine Debatte der kritischen und zweifelnden Nachdenklichkeit."

engendering conditions, we can even detect debates that in parts have features of "ideal discourses". Both the public and the non-public debate on the Swiss language bill in the second chamber achieve respect scores that one might associate with ideal discourses. In the public debate, the score of disrespectful speech acts is almost zero (one single disparaging speech in the public debate (out of 55) and 3 disparaging speeches in the non-public debate (out of 212)) and the higher respect levels (explicit respect or accepting demands and counterarguments) achieve remarkable scores: almost 70 percent in the public debate and 42 percent in the non-public debate. The score for constructive proposals amounts to 25 percent in the non-public debate and to 16.4 percent in the public debate. Compare this to the public debate on the polarized minimum wage debate in the competitive parliamentary and public arena of the British House of Commons: here the score of disrespectful speech acts amounts to 73 percent; there are no instances of speakers accepting the demands or counterarguments of their opponents, while instances of explicit respect are rare (2 percent). And the score of constructive proposals is only 6.5 percent (actually consisting of a total of two alternative proposals).

Conclusion

In this article, we considered the favorable contextual conditions of deliberation, both in terms of institutional characteristics and in terms of issue attributes. What have we learnt? The most important conclusion is that institutional and issue contexts matter for the quality of political discourse. Our data suggest that respect is the most sensitive to institutional and issue attributes. We find that consensus institutions, presidentialism, second chambers, non-publicity, and low issue polarization further respectful exchanges among political actors. In terms of the level of justification, findings are more mixed. Levels of justification do show some changes across institutional settings, but not always in the expected direction and not always robustly; i.e. depending on the case selection we may or may not find the effect. The analysis of levels of justification also reveals that the dimensions of discourse quality do not always travel in the same direction when institutional parameters are changed. As the comparison between public and non-public debates shows, bringing a debate into the public arena greatly improves the level of justification of demands by speakers, but it hinders respectful discourse. For the content of justification, we find that common good orientations are more frequently uttered in the public sphere. The constructive politics dimension, in turn, showed remarkable resilience to institutional variation. Only in the comparison between first and second chambers and the comparison between polarized abortion debates and less polarized issues did we find an institutional effect in the expected direction. In the case of the parliamentary-presidential comparison, we observed an effect but in an unexpected direction. Otherwise, levels of constructive politics looked very similar across institutional settings, with positional politics being the norm. Finally, we can demonstrate that when favorable contextual conditions combine, i.e. when a less polarized issue is debated in a second chamber of a consensus setting where actors possess strong veto power, we find debates that in parts have features of "ideal discourses" with actors being highly respectful, open, and constructive (although only in the sense of "observed" behavior, as the truthfulness claim cannot be substantiated). Recall, however, that our findings cannot be generalized to all parliaments and all parliamentary debates. Yet, as the selected legislatures and debates represent critical cases, the findings give a strong first hint that institutional and contextual factors matter for deliberation.

These are important results both for philosophers and institutional scholars. This study shows that certain types of institutions can build bridges between political philosophy – dealing with the question of what actors *should* do (here: deliberate) – and empirical-analytical political science – dealing with the question of what the polity *can* do

(Rothstein 1998). From this follows the idea of institutional engineering, in that normatively desired outcomes can be achieved through institutional manipulation. Of course, we do not argue that the analyzed institutional mechanisms produce the same results when inserted in radically different cultural contexts. But our results contain a hint that culture is not the entire story: as soon as institutional parameters are changed within one organizational and cultural setting, DQI scores move in the (institutionally) predicted directions. Hence, some tentative advice to institutional designers can be given: if respectful deliberation and argumentative change are the goals, then actor relationships should not be too competitive, actors should not have bound mandates, bodies should be small and explicitly geared toward rational discussion and reflection, and actors should have the opportunity of meeting in secret. This advice will certainly disappoint many deliberative democrats, as deliberation seems to be closely associated with a "pre-modern" and "gentlemen's club" model of politics rather than with modern mass democracies where disciplined parties publicly compete for political power. But this seems to be the price that philosophy has to pay when deliberative politics are to be put in action in the real world of politics.

Our study, furthermore, also sheds new light on the research on legislatures. Many scholars still tend to view legislatures primarily as "rubber stamps" where policies made elsewhere are debated in an adversarial and rhetorical way between the parties involved. While it is fair to say that this describes a substantial part of the debates analyzed in this study, we could demonstrate that under specific institutional and contextual conditions, open, respectful and consensus-oriented deliberation is possible in legislatures. These findings complement the new institutional research program on legislatures by adding to it the element of deliberation, yet showing again that much of it is institutionally driven.

Finally, two major limitations of this study need to be addressed. First, critics might object that one major aspect of deliberation has been blinded out in empirical analysis, namely that we did not establish whether actors have been truthful. Hence, we are not in a position to fully distinguish between genuine deliberation and strategic action. However, one can argue that a clear-cut distinction between deliberation and strategic action is also problematic at a theoretical level, since argumentation may not be free of an action orientation toward success and also strategic actors may – under certain conditions - have an incentive to deliberate (Greve 1999; Johnson 1993; Ulbert/Risse/Müller 2004). Moreover, as soon as we identify actors who are more open, more constructive, who come to respect and agree with the arguments of other actors, and this variation can be explained by institutional factors, then the importance of the truthfulness argument is reduced. It means that we can get the desired outcome through institutional manipulation, regardless whether this matches actors' true preferences or not.

Second, critics might argue that our findings are interesting, but ultimately irrelevant as deliberation was analyzed as a "process" variable only, without tackling the "so what" question, namely whether deliberation had an effect on the outcomes as well. In this regard, let us first note that some coding categories of the DQI – respect and acceptance of demands and counterarguments – involve some outcome dimension as well. However, we do acknowledge that, in general, we analyzed "talk cultures" rather than policy outcomes. So, what is the use of a higher "talk culture", one might ask? Does it really matter if one setting is slightly more respectful than another, or slightly less positional, or slightly better in the justification of demands? In his study on deliberation and outcomes in the German Conference Committee (building on a sample of the debates used in this study), Spörndli (2003; 2004) has addressed exactly this question. He can demonstrate that the general discourse quality (measured by the DQI) has a strong and independent effect on the probability to reach unanimous agreement. However, the deliberation is almost without any effect on the substantive outcome dimension (conceptualized as more and less egalitarian decisions) with the power context being virtually deterministic. Although findings are mixed,

the quality of deliberation is not without consequences. This not only diminishes the truthfulness problematique even further, it also shows that deliberation is more than a procedural "amuse-bouche" that only delights before the real power meal begins.

In sum, our article shows that deliberation is not only a philosophical and utopian affair, but that under certain contextual conditions, it has a chance in the real world of politics too. However, we should not overlook that the large majority of parliamentary debates (even if there were individual actors who followed deliberative standards) are not really deliberative. Thus, while deliberation may not constitute a major logic of action informing many policy making processes, it still captures parts of political reality. Similar to what Wilson (1993: 9) has noted for mankind's moral sense, deliberation is "not a strong beacon light, radiating outward to illuminate in sharp outline all it touches. It is, rather, a small candle flame, casting vague and multiple shadows, flickering and sputtering in the strong winds of power and passion."

References

Aspingwall, Mark and Schneider, Gerald. 2000, "Same menu, separate tables: The institutionalist turn in political science and the study of European integration", *European Journal of Political Research* 38: 1-36.

Austen-Smith, David and Feddersen, Timothy J. 2002, "Deliberation and Voting Rules", Working Paper.

Baccaro, Lucio 2001, "Aggregative and Deliberative Decision-Making Procedures: A Comparison of Two Southern Italian Factories", *Politics & Society* 29: 243-271.

Benhabib, Seyla 1996, "Toward a Deliberative Model of Democratic Legitimacy", in Benhabib, Seyla (ed.), *Democracy and Difference. Contesting the Boundaries of the Political*, Princeton University Press: 67-94.

Beyme, Klaus von 1997, *Der Gesetzgeber. Der Bundestag als Entscheidungszentrum*, Opladen: Westdeutscher Verlag.

Budge, Ian 2001, *Mapping policy preferences: estimates for parties, electors, and governments*, Oxford University Press.

Chambers, Simone 1995, "Discourse and Democratic Practices", in: White, Stephen K. (ed.), The Cambridge Companion to Habermas. Cambridge: Cambridge University Press,

Checkel, Jeffrey T. 1999, "Norms, institutions, and national identity in contemporary Europe", *International Studies Quarterly* 43: 43-83

Cox, Gary W. and McCubbins Mathew D. 1993, *Legislative leviathan: party government in the House*, Berkeley: University of California Press.

Czada, Roland 2000, "Konkordanz, Korporatismus und Politikverflechtung: Dimensionen der Verhandlungsdemokratie", in: Everhard Holtmann und Helmut Voelzkow (Herausgeber), *Zwischen Wettbewerbs- und Verhandlungsdemokratie*, Opladen, Wiesbaden: Westdeutscher Verlag, 23-49.

Diermeier, Daniel and Feddersen, Timothy J. 1998, "Cohesion in Legislatures and the Vote of Confidence Procedure", *American Political Science Review* 92: 611-638.

Döring, Herbert 1995, "Time as a Scarce Resource: Government Control of the Agenda", in: Döring, Herbert (ed.), *Parliaments and Majority Rule in Western Europe*, Frankfurt: Campus Verlag.

Dryzek, John S. 2000, Deliberative Democracy and Beyond, Oxford University Press.

Dryzek, John and List, Christian 2003, "Social Choice Theory and Deliberative Democracy: A Reconciliation", *British Journal of Political Science* 33: 1-28.

Eckstein, Harry 1975, "Case Study and Theory in Political Science", in: Greenstein, Fred I. and Polsby, Nelson W. (eds.), *Handbook of Political Science*, Vol. 7, Strategies of Inquiry, Reading, Mass.: Addison-Wesley.

Elgström, Ole and Jönsson, Christer 2000, "Negotiation in the European Union: Bargaining or Problem-Solving?", in: Journal of European Public Policy, Special Issue: Negotiation and Policy-Making in the European Union – Processes, System and Order, 7: 684-704.

Elster, Jon 1998a, "Introduction", in: Elster, Jon (ed.), *Deliberative Democracy*, Cambridge University Press: 1-18.

Elster, Jon 1998b, "Deliberation and Constitution Making", in: Elster (ed.), *Deliberative Democracy*, Cambridge University Press: 97-122.

Fiorina, Morris P. 1981, *Retrospective Voting in American National Elections*, New Haven, Conn.: Yale University Press.

Ganghof, Steffen 2002, "Veto Points and Veto Players. A Skeptical View", Working Paper, Max Planck Institute for the Study of Societies, Cologne, Germany.

Garner, Robert 1998, "Animal Welfare", in: Cowley, Philip (ed.). *Conscience and Parliament*. London, Frank Cass, 117-131.

George, Alexander L. and McKeown, Timothy J. 1985, "Case Studies and Theories of Organizational Decision Making", *Advances in Information Processing in Organizations* 2: 21-58.

Greve, Jens 1999, "Sprache, Kommunikation und Strategie in der Theorie von Jürgen Habermas", *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 51: 232-59.

Gutmann, Amy and Thompson, Dennis 1996, *Democracy and Disagreement*, Cambridge, Mass.: Belknap Press of Harvard University Press.

Habermas, Jürgen 1981, *Theorie des kommunikativen Handelns*, 2 Bände, Frankfurt a. M.: Suhrkamp.

Habermas, Jürgen 1983, *Moralbewusstsein und kommunikatives Handeln*, Frankfurt a. M.: Suhrkamp.

Habermas, Jürgen 1991, Erläuterungen zur Diskursethik, Frankfurt a. M.: Suhrkamp.

Habermas, Jürgen 1992, Faktizität und Geltung: Beiträge zur Diskurstheorie des Rechts und des demokratischen Rechtsstaats, Frankfurt a. M.: Suhrkamp.

Habermas, Jürgen 1996, Die Einbeziehung des Anderen. Studien zur politischen Theorie, Frankfurt a. M.: Suhrkamp.

Hall, Peter 1986, *Governing the Economy: The politics of state intervention in Britain and France*, Cambridge: Polity Press.

Holzinger, Katharina 2001a, "Verhandeln statt Argumentieren oder Verhandeln durch Argumentieren? Eine empirische Analyse auf der Basis der Sprechakttheorie", *Politische Vierteljahresschrift* 42: 414-446.

Holzinger, Katharina 2001b, „Kommunikationsmodi und Handlungstypen in den Internationalen Beziehungen", *Zeitschrift für Internationale Beziehungen* 2: 243-286.

Hammond, Thomas H. and Butler, Christopher K. 2003, "Some Complex Answers to the Simple Question 'Do Institutions matter?': Policy Choice and Policy Change in Presidential and Parliamentary Systems", *Journal of Theoretical Politics* 15: 145-200.

Hug, Simon and Sciarini, Pascal 1995. "Switzerland - Still a Paradigmatic Case?", in: Schneider, Gerald, Weitsman, Patricia A. and Bernauer, Thomas (ed.), *Towards a New Europe. Stops and Starts in Regional Integration*, Westport, Connecticut and London: Praeger, 55-74.

Johnson, James 1993, "Is Talk Really Cheap? Prompting Conversation Between Critical Theory and Rational Choice", *American Political Science Review* 87: 74-85.

Johnson, James 1998, "Arguing for Deliberation: Some Skeptical Considerations", in: Elster Jon (ed.), *Deliberative Democracy*, Cambridge University Press.

Knight, Jack and James Johnson 1994, "Aggregation and Deliberation: On the Possibility of Democratic Legitimacy", *Political Theory* 22: 277-96.

Kriesi, Hanspeter 2001, "Die Rolle der Öffentlichkeit im politischen Entscheidungsprozess. Ein konzeptueller Rahmen für ein international vergleichendes Forschungsprojekt", Veröffentlichungsreihe der Arbeitsgruppe Politische Öffentlichkeit und Mobilisierung, Wissenschaftszentrums Berlin für Sozialforschung.

Lamnek, Siegfried 1995, *Qualitative Sozialforschung*. Band 2: Methoden und Techniken, Weinheim: Pychologie Verlags-Union.

Lewis-Beck, Michael 1988, *Economics and Elections: The Major Western Democracies*, Ann Arbor: University of Michigan Press.

Lhotta, Roland. 2000, "Konsens und Konkurrenz in der konstitutionellen Ökonomie bikameraler Verhandlungsdemokratie: Der Vermittlungsausschuss als effiziente Institution politischer Deliberation", in: Holtmann, Everhard und Voelzkow, Helmut (Hg.), *Zwischen Wettbewerbs- und Verhandlungsdemokratie*, Opladen/Wiesbaden: Westdeutscher Verlag.

Lijphart, Arend 1984, *Democracies. Patterns of Majoritarian and Consensus Government in Twenty-One Countries*, New Haven, Conn.: Yale University Press.

Lijphart, Arend 1999, *Patterns of Democracy. Government Forms and Performance in Thirty-Six Countries*, New Haven, Conn.: Yale University Press.

Lijphart, Arend 2002, "The Wave of Power-Sharing Democracy", in: Reynold, Andrew (ed.), *The Architecture of Democracy. Constitutional Design, Conflict Management, and Democracy*, Oxford: Oxford University Press, 37-54.

Linder, Wolf, 1999, *Schweizerische Demokratie: Institutionen – Prozesse - Perspektiven*, Bern, Stuttgart, Wien: Haupt.

Lipset, Seymour Martin and Rokkan, Stein (eds.) 1967, *Party Systems and Voter Aligments*, New York: Free Press.

Lynch, Marc 2002, "Why Engage? China and the Logic of Communicative Engagement", *European Journal of International Relations* 8: 187-230.

Loomis Burdett A. 1990, *Dear colleagues, civility and deliberation in the U.S. Senate*, Washington, D.C.: Brookings Institution.

Luskin, Robert C. and Fishkin James S. 2002, "Deliberation and 'Better Citizens", Paper Presented at the ECPR Joint Sessions, Turin, March 23-27.

Macedo, Stephen 1999, "Introduction", in: Macedo, Stephen (ed.), *Deliberative Politics. Essays on 'Democracy and Disagreement'*, Oxford University Press: 3-14.

Manin, Bernard 1987, "On Legitimacy and Deliberation", *Political Theory* 15: 338-68.

McKeown, Timothy J. 1999, "Case Studies and the Statistical Worldview: Review of King, Keohane, and Verba's Designing Social Inquiry: Scientific Inference in Qualitative Research", *International Organization* 53: 161-190.

Mill, John Stuart 1998, *Utilitarianism*, Oxford University Press.

Müller, Harald 1995, "Spielen hilft nicht immer. Die Grenzen des Rational-Choice Ansatzes und der Platz der Theorie kommunikativen Handelns in der Analyse internationaler Beziehungen", *Zeitschrift für Internationale Beziehungen* 2: 367-79.

Powell, Bingham G., Jr. and Whitten Guy D. 1993, "A Cross-National Analysis of Economic Voting: Taking Account of the Political Context", *American Journal of Political Science* 37: 391-411.

Rawls, John 1971, *A Theory of Justice*, Cambridge, Mass.: Belknap Press of Harvard University Press.

Risse, Thomas 2000, "Let's argue! Communicative Action in World Politics", *International Organization* 54: 1-39.

Roberts, Nancy 1997, "Public Deliberation: An Alternative Approach to Crafting Policy and Setting Direction", *Public Administration* 57: 124-132.

Scharpf, Fritz W. 1997, *Games Real Actors Play. Actor-Centered Institutionalism in Policy Research*, Boulder, CO: Westview Press.

Spörndli, Markus 2003, "Discourse Quality and Political Decisions: An Empirical Analysis of Debates in the German Conference Committee", WZB Discussion Paper FS IV 03-101, Berlin: Wissenschaftszentrum Berlin für Sozialforschung.

Spörndli, Markus 2004, *Diskurs und Entscheidung. Eine empirische Analyse kommunikativen Handelns im deutschen Vermittlungsausschuss*, PhD dissertation, Institute of Political Science, University of Bern.

Steenbergen, Marco R. 2000, "Item Similarity in scale analysis", *Political Analysis* 8(3): 261-283.

Steenbergen, Marco R., Bächtiger, André, Spörndli, Markus and Steiner, Jürg 2003, "Measuring Political Deliberation", *Comparative European Politics* 1: 21-48.

Stokes, Donald E. 1963, "Spatial Models of Party Competition", *American Political Science Review* 57: 368-377.

Strom, Kaare 2000, "Delegation and accountability in parliamentary democracies", *European Journal of Political Research* 37: 261-289.

Sulkin, Tracy and Simon, Adam F. 2001, "Habermas in the Lab: A Study of Deliberation in an Experimental Setting", *Political Psychology* 22: 809-826.

Tsebelis, George 1995, "Decision Making in Political Systems: Veto Players in Presidentialism, Paliamentarism, Multicameralism and Multipartyism", *British Journal of Political Science* 25: 289-325.

Tsebelis, George and Money, Jeannette 1997, *Bicameralism*, Cambridge University Press.

Tsebelis, George 2002, *Veto Players. How Political Institutions Work*, Princeton University Press.

Ulbert, Cornelia, Risse, Thomas and Mueller, Harald. 2004, "Arguing and Bargaining in Multilateral Negotiations", Paper presented at the EUI's Swiss Chair Conference on "Empirical Approaches to Deliberative Politics", European University Institute, Florence, May 21-22, 2004.

Wilson, James Q. 1993, "The Moral Sense: Presidential Address, American Political Science Association, 1992", *American Political Science Review* 87: 1-11.

Appendix 1

Discourse Quality Index (DQI)

The DQI attempts to measure political deliberation in a general, valid and reliable way (Steenbergen et al. 2003). It mainly draws on Habermasian discourse ethics, but also incorporates elements of other deliberative models. The unit of analysis of the DQI is a *speech act*, i.e. the discourse by a particular individual delivered at a particular point in a debate. For each speech, we distinguish between relevant and irrelevant parts, and only the relevant parts are coded. A relevant part is one that contains a *demand*, i.e. a proposal on what decision should or should not be made. Our emphasis on demands stems from the fact that they constitute the heart of the deliberation. That is, demands stipulate what ought to be and what ought not to be, and this normative character puts them at the center of discourse ethics. The DQI is composed of seven indicators. Despite the considerable complexity of parliamentary debates, we attempt to keep the coding categories relatively easy, so as to ensure a high level of reliability. The following is an elaboration of these indicators, followed by an overview of the indicators and their codes. We discuss seven the seven indicators under four headings.

1. *Participation*. Participation constitutes a fundamental precondition for deliberation. In parliamentary settings of western democracies, this type of basic participation can usually be seen as given for the elected representatives. Normal participation is only assumed to be impaired if a speaker is cut off by a formal decision, or if she or he feels explicitly disturbed in the case of a verbal interruption by other actors.

2. *Justification*. Fruitful discourse requires the justification of assertions and validity claims. That is, assertions should be introduced and critically assessed through "the orderly exchange of information and grounds between parties" (Habermas 1992: 370). The major problem here is that since the ideal speech situation itself has no content, one cannot apply external standards to what constitutes a good reason; otherwise one would also predetermine the result of a discourse. Hence, we only judge to what extent a speech gives complete justifications for demands and thus makes it accessible to rational critique. The first indicator of the concept of justification is the *level of the justification*, whereby we distinguish between five levels. No justification forms the lowest level of the indicator. Next comes inferior justification, i.e. a reason Y is given why X should or should not be done, but no linkage is made between X and Y (this code also applies if a conclusion is merely supported with illustrations). This is followed by qualified justification where one such a linkage is made. Then come two categories of sophisticated justification: the first category of sophisticated justification is given when at least two complete justifications are given. The second category of sophisticated justification is given when at least two complete justifications are given for a demand and one justification is also embedded in at least tow complete inferences. The second indicator of the concept of justification is the *content of the justification*. The importance of referring to the common good is mainly stressed by deliberative theorists drawing on Rawls (e.g. Gutmann/Thompson 1996). This indicator measures whether the argumentation within a statement is cast in terms of a conception of the common good or in terms of narrow group or constituency interests. With regard to the common good, we distinguish between the common good stated in utilitarian terms, i.e. as the best solution for the greatest number of people (Mill 1998) and the common good expressed through the difference principle, i.e. the common good is served when the least advantaged in a society are helped (Rawls 1971). However, one has always to be aware that a given reference to the

common good may be not much more than "cheap talk", being purely rhetorical and intended to attract potential voters.

3. *Respect*. Respect is a prerequisite for serious listening, which in turn is essential for deliberation. Even more importantly, respect also captures the idea of preference change. Several dimension of respect play a role in discourse. First, it is important to see if speakers show *respect toward the groups which would be helped* with the demand under discussion. This indicator is a translation of Habermas's (1991: 73) postulate of empathy into the parliamentary context. The lowest category is reserved for speeches in which there are only negative statements about the groups. The next category is implicit respect, i.e. there are neither explicitly negative nor positive statements about the groups. The highest category is assigned if there is at least one explicitly positive statement about the groups. In a similar vein, deliberation requires, secondly, that both sides show *respect toward demands* under discussion. Macedo (1999: 10) regards the recognition of the "merit in [the] opponents' claims" as being one of the principal purposes of deliberation. This indicator uses the same categories as the group respect indicator, with one exception: the highest category here is agreeing with a demand. This means willful acceptance of a demand, i.e. the demand needs to be both accepted *and* positively valued. Third, discursive participants need to show *respect toward counterarguments* to their own position. Respect toward counterarguments is a necessary condition for the weighing of alternatives, which some view as an essential element of deliberation (Luskin/Fishkin 2002). The lowest coding category applies when counterarguments are degraded. The next category is reserved for speakers who ignore counterarguments. [25] This is followed by a neutral category where speakers include counterarguments but neither make positive nor negative statements about them. The next category applies for speakers who explicitly value counterarguments. The highest coding category is agreeing with a counterargument (even if this does not cause a transformation of his or her preferences vis-à-vis a demand). Under the rubric of counterarguments we also include personal or partisan attacks and praises.

4. Constructive politics. The last indicator *constructive politic*s is based upon the principal goal of Habermasian discourse ethics to reach a genuine consensus. Here we distinguish between four levels. Positional politics form the lowest level of the indicator. This is followed by alternative proposals, i.e. proposals that attempt to mediate but that do not fit the current agenda. The next category is "consensus appeals". Consensus appeals are calls for compromise and consensus that are pertinent to the current agenda but that are unspecific. For example, a call to "bridge our differences" would be a consensus appeal if it did not contain a specific mediating proposal. The highest level of constructive politics are mediating proposals that fit the current agenda.

---

[25] The original version of the DQI assigned "ignoring" counterarguments to the lowest coding category. The rationale for the above recoding is that, in practice, ignoring counterarguments is not only a feature of bargaining, but frequently also a feature of calm debates, an essential characteristic of Habermasian discourse ethics. Assigning it to the lowest code might give a distorted view of deliberative quality. This is further validated by the fact that there are numerous instances where actors "weighed" the merits of demands without directly going into counterarguments. Hence, it seemed more appropriate to assign a higher code to ignoring counterarguments, albeit it remains a low discourse quality category in the recoded version.

Table 8: Indicators of Discourse Quality

1 Participation

0: interruption of a speaker: a speaker feels explicitly disturbed, or interruption occurs through a formal decision

1: normal participation is possible

2 Justification

2.1 Level

0: no justification

1: inferior justification: conclusion(s) embedded in (an) *in*complete inference(s)

2: qualified justification: one conclusion embedded in a *complete* inference; additional conclusions embedded in *in*complete inferences may be present

3: sophisticated justification (broad): more than one conclusion, each embedded in a complete inference

4: sophisticated justification (in depth): one or several conclusions, each embedded in several complete inferences

2.2 Content

0: explicit statement concerning constituency or group interests

1: neutral statement: either no reference to constituency or group interests or reference to several group interests

2: explicit statement in terms of a conception of the common good in utilitarian or collective terms

3: explicit statement in terms of the difference principle

3 Respect

3.1 toward group in general ("empathy")

0: no respect: explicitly negative statement concerning the group which would be helped

1: implicit respect: no explicitly negative statement concerning the group which would be helped

2: respect: explicitly positive statement concerning the group which would be helped

3.2 toward demands

0: no respect: explicitly negative statement concerning the demand

1: implicit respect: no explicitly negative statement concerning the demand

2: explicit respect: explicitly positive statement concerning the demand

3: demand accepted: speaker agrees with the demands of other actors and values them

3.3 toward counterarguments

0: counterarguments are explicitly degraded

1: counterarguments are ignored

2: counterarguments are neither explicitly degraded nor valued

3: counterarguments are explicitly valued

4: counterarguments are agreed with

4 Constructive Politics

0: positional politics: speaker sits on his/her position

1: alternative proposal: proposal for a different agenda

2: consensus appeal: speakers make an unspecific appeal for compromise or consensus

3: mediating proposal: proposal within the same agenda

Appendix 2: Significance Tests

	Level of Justifi-cation [a] (Ordered Logit)	Content of Justifi-cation [a] (Ordered Logit)	Respect Overall [b] (Anova)	Overall Respect[a] (Anova)	Respect Demands/ Counter-arg. [b] (Anova)	Respect Demands/ Counter-arg. [a] (Anova)	Construc-tive Politics [a] (Ordered Logit)
Cons. System (lenient test)	.693 (.000) (N=1383)		38.100 (.000) (N=1896)	15.586 (.000) (N=1383)	22.958 (0.000) (N=1896)	8.410 (0.004) (N=1383)	. 848 (.000) (N=1383)
Cons. System (strict test)	-.219 (.008) (N=2409)		25.651 (.000) (N=2886)	15.715 (.000) (N=2409)	19.263 (.000) (N=2886)	13.235 (0.000) (N=2409)	-.052 (.687) (N=2409)
Veto Power	.218 (.097) (N=1352)		5.807 (.016) (N=1860)	4.962 (.026) (N=1352)	5.406 (.020) (N=1860)	6.090 (.014) (N=1352)	.233 (.246) (N=1352)
Presid. System	-.859 (.000) (N=497)		48.312 (.000) (N=1031)	8.350 (.004) (N=497)	32.893 (.000) (N=1031)	4.590 (.033) (N=497)	-.699 (.028) (N=497)
Chamber	.011 (.863) (N=3694)		56.735 (.000) (N=4096)	43.800 (.000) (N=3694)	57.841 (.047) (N=4096)	44.096 (.000) (N=3694)	.209 (.031) (N=3694)
Nonpubl. Arena	-1.956 (.000) (N=3086)	-1.415 (.000) (N=2166)	31.404 (0.000) (N=4221)	7.170 (.007) (N=3086)	58.401 (.000) (N=4221)	22.956 (.000) (N=3086)	.223 (.067) (N=3086)
Nonpubl* Cons			59.428 (0.000) (N=4221)	13.634 (0.000) (N=3086)	39.320 (0.000) (N=4221)	9.177 (0.002) (N=3086)	
Issue1	.545 (.000) (N=2995)		54.732 (.000) (N=3516)	44.635 (.000) (N=2995)	56.787 (.000) (N=3516)	44.351 (.000) (N=2995)	.196 (.053) (N=2995)
Issue2	.254 (.002) (N=2196)		113.146 (.000) (N=2219)	109.664 (.000) (N=2196)	110.404 (.000) (N=2219)	107.155 (.000) (N=2196)	.318 (.008) (N=2196)

Notes: For the logistic regression models, entries are ordered logistic coefficients (significance levels in parentheses). Coefficients for block dummies and thresholds are not shown. For the ANOVA models, entries are F-values (significance levels in parentheses). The F-tests are based on a type III error. F-values for block variables and block*main variable interactions are not shown.
[a] heckles excluded; [b] heckles included

Appendix 3

Debates Analyzed

Codings of all debates are available at http://www.ipw.unibe.ch/discourse/

Codes:
lowpol/highpol: less polarized/high polarized debate
vphigh/vplow: veto power of actors high/veto power of actors low
chamber1/chamber2: first chamber/second chamber
nonpub/pub: non-public debate/public debate

Switzerland
- "Bundesverfassung: Sprachenartikel" (91.019)
  Code: lowpol, vphigh; chamber1 and 2, nonpub and pub (4 debates)
- "Tier keine Sache" (parlamentarische Initiative Francois Loeb) (92.437)/ "Wirbeltiere. Gesetzliche Bestimmungen" (parlamentarische Initiative Suzette Sandoz) (93.459)
  Code: lowpol, vplow; chamber1, nonpub (1 debate)
- "Arbeitsgesetz (Revision)" (94.013)
  Code: highpol, vphigh[26]; chamber1 and 2, nonpub and pub (4 debates)
- "Bilaterale Verträge Schweiz/EU: Minimale Arbeits- und Lohnbestimmungen" (99.028-6)
  Code: highpol, vphigh; chamber1 and 2, nonpub and pub (4 debates)
- "Schwangerschaftsabbruch. Revision des Strafgesetzbuches" (parlamentarische Initiative Barbara Haering Binder) (93.434)
  Code: highpol (abortion), vphigh; chamber1 and 2, nonpub and pub (4 debates)
- "Massenentlassungen" (Standesinitiative Genf) (96.320)
  Code: highpol, vplow; chamber1, nonpub (1 debate)

Floor Debates: http://www.parlament.ch and "Amtliches Bulletin der Bundesversammlung (Schweizerische Eidgenossenschaft)"
Committee Debates (National- und Ständerat): Parlamentsdienste Bern.

United States
- H.R.3371/S.1241 "To control and prevent crime"/A bill to control and reduce violent crime" (102nd Congress 1991/92)
  Code: lowpol (crime), vphigh; chamber1 and 2, pub (2 debates)
- H.R.2271: "To establish a clear and comprehensive prohibition of discrimination on the basis of disability" (101st Congress 1989/90)
  Code: lowpol, vplow; chamber1, pub (1 debate)
- H.R.2/S.4: "To amend the Fair Labor Standards Act of 1938 to restore the minimum wage to a fair and equitable rate, and for other purposes" (101st Congress 1989/90)
  Code: highpol, vphigh[27]; chamber1 and 2, pub (2 debates)
- H.R. 1833/S.939: "To amend title 18, United States Code, to ban partial-birth abortions" (104th Congress 1995/96)
  Code: highpol (abortion), vphigh; chamber1 and 2, pub (2 debates)
- H.R.1227/H.AMDT.1084: "To amend the Portal-to-Portal Act of 1947 relating to the payment of wages to employees who use employer owned vehicles"/"Amendment increases the minimum wage by 90 cents from $4.25 to $5.15 per hour over the next two years" (104th Congress 1995/96)
  Code: highpol, vplow; chamber1 and 2, pub (2 debates)

Congressional Record, Washington D.C.: http://thomas.loc.gov

---

[26] Parts of this debate (first chamber, public and non-public sphere) were conducted under conditions of low veto power. We have analyzed this part separately.

[27] Parts of this debate (first chamber, public and non-public sphere) were conducted under conditions of low veto power. We have analyzed this part separately.

Germany
- "Gesetz zur Bekämpfung des illegalen Rauschgifthandels und anderer Erscheinungsformen der Organisierten Kriminalität (OrgK)" (12/106)/"Gesetz über das Aufspüren von Gewinnen aus schweren Straftaten" (12/250)
  Code: lowpol (crime), vphigh; chamber1 and 2, nonpub and pub (3 debates)
- "Gesetz zur Verbesserung der Rechtsstellung des Tieres im bürgerlichen Recht" (12/282)
  Code: lowpol, vplow; chamber1, nonpub (1 debate)
- "Gesetz zur Sozialen Absicherung der Pflegebedürftigkeit"/ "Entgeltfortzahlungsgesetz" (12/337)
  Code: highpol, vphigh; chamber1 and 2, nonpub and pub (4 debates)
- "Gesetz zum Schutz des werdenden Lebens, zur Förderung einer kinderfreundlichen Gesellschaft, für Hilfen im Schwangerschaftskonflikt und zur Regelung des Schwangerschaftsabbruchs" (12/551)
  Code: highpol (abortion), vphigh; chamber1 and 2, nonpub and pub (3 debates)[28]
- "Gesetz zur Änderung des Ladenschlussgesetzes" (488/96)
  Code: highpol, vplow; chamber1, nonpub and pub (2 debates)

Debates of the German Conference Committee (vphigh, chamber1 and 2 combined, nonpub)
- "Gesetz über die Entschädigung von Opfern von Gewalttaten (OEG)" (7/62)
  Code: lowpol
- "Gesetz zum Schutz vor Missbrauch personenbezogener Daten bei der Datenverarbeitung" (7/84)
  Code: lowpol
- "Gesetz über die unentgeltliche Beförderung Schwerbehinderter im öffentlichen Personenverkehr" (8/30)
  Code: lowpol
- "Gesetz über die Prozesskostenhilfe" (8/51)
  Code: lowpol
- "Gesetz zur Verbesserung des Mietrechts und zur Begrenzung des Mietanstiegs" (6/21a)
  Code: highpol
- "Gesetz zur Änderung von Vorschriften der gesetzlichen Rentenversicherung" (7/1)
  Code: highpol
- "Gesetz zur Förderung von Wohnungseigentum und Wohnbesitz im sozialen Wohnungsbau" (7/47b)
  Code: highpol
- "Sozialgesetzbuch, allgemeiner Teil" (7/50)
  Code: highpol
- "Gesetz zur zwanzigsten Rentenanpassung und zur Verbesserung der Finanzgrundlagen der gesetzlichen Rentenversicherung" (8/2)
  Code: highpol
- "Entwurf eines Gesetzes zur Änderung des Einkommenssteuergesetzes und anderer Gesetze" (8/21b)
  Code: highpol

Floor Debates: http://dip.bundestag.de
Committee Debates (Ausschuss-Debatten, Bundestag und Bundesrat): Deutsches Parlamentsarchiv Bonn/Berlin
Conference Committee Debates: *Protokolle des Vermittlungsausschusses des Deutschen Bundestages und des Bundesrates*. Sekretariat des Bundesrates (Hrsg.), Munchen: C.H. Beck, K.G. Saur (Mikroform)

Great Britain

- "Disabled Persons"
  Code: lowpol, vpmedium; chamber1, pub (1 debate)[29]
- "Sunday Trading Bill"
  Code: lowpol, vpmedium; chamber1, pub (1 debate)[30]
- "National Minimum Wage Act (1998)"
  Code: highpol, vplow; chamber1, pub (1 debate)

Hansard (House of Commons Daily Debates):
http://www.parliament.the-stationery-office.co.uk/pa/cm/cmhansrd.htm

---

[28] In this debate, party discipline was explicitly relaxed.
[29] In this debate, party discipline was partly relaxed.
[30] In this debate, party discipline was explicitly relaxed.