GOVERNANCE AND THE EFFICIENCY
OF ECONOMIC SYSTEMS
GESY

Discussion Paper No. 146

Budget Processes: Theory and Experimental Evidence

Karl-Martin Ehrhart*
Roy Gardner**
Jürgen von Hagen***
Claudia Keser****

March 2006

*Karl-Martin Ehrhart, University of Karlsruhe
**Roy Gardner, Indiana University and ZEI, University of Bonn
***Jürgen von Hagen, University of Bonn, Indiana University, and CEPR
****Claudia Keser, IBM T.J. Watson Research Center and CIRANO, Montreal

Budget Processes: Theory and Experimental Evidence

Karl-Martin Ehrhart*, Roy Gardner**, Jürgen von Hagen***, and Claudia Keser****

*University of Karlsruhe,
** Indiana University and ZEI, University of Bonn,
*** University of Bonn, Indiana University, and CEPR,
**** IBM T.J. Watson Research Center and CIRANO, Montreal

March 2006

Abstract: *This paper studies budget processes, both theoretically and experimentally. We compare the outcomes of bottom-up and top-down budget processes. It is often presumed that a top-down budget process leads to a smaller overall budget than a bottom-up budget process. Ferejohn and Krehbiel (1987) showed theoretically that this need not be the case. We test experimentally the theoretical predictions of their work. The evidence from these experiments lends strong support to their theory, both at the aggregate and the individual subject level.*

Keywords: Budget processes, voting equilibrium, experimental economics
JEL codes: H61, C91, C92

Correspondence: Jürgen von Hagen, Institut für Internationale Wirtschaftspolitik, Lennestrasse 37, D-53113 Bonn, vonhagen@uni-bonn.de

1. Introduction

A budget process is a system of rules governing the decision-making that leads to a budget, from its formulation, through its legislative approval, to its execution. Consider the budget process of the United States government. The President formulates a budget proposal as part of his annual obligation to report on the State of the Union. Each house of Congress then reworks the budget proposal, with a final budget being passed by both houses for presidential approval.

In the last quarter century, the details of the budget process, both in the United States and in other countries, have been the object of considerable research (Wildavsky, 1975; Ferejohn and Krehbiel, 1987; Alesina and Perotti, 1995, 1999; von Hagen and Harden, 1995, 1996; see also the contributions in Poterba and von Hagen, 1999). There is a growing body of empirical research, based on international comparative studies, suggesting that the design of budget processes has considerable influence on the fiscal performance of governments. This has also been reflected in political decisions. In the United States, the Budget Act of 1974, the Gramm-Rudman-Hollings Act of 1985, and the Budget Enforcement Act of 1991 all tried to reduce excessive government spending and deficits by changes in the budget process. In the European Union, the Maastricht Treaty on European Union of 1992 mandates reform of budget processes of the member states to enhance fiscal discipline.

One aspect of the budget process that has received considerable attention is the sequence of budgeting decisions. Traditionally, Congress voted on budget items line-by-line, or category-by-category. The sum of all spending approved by Congress emerged as the overall budget—a budget process called *bottom up*. The budget reforms stemming from the Budget Act of 1974 replaced this tradition with a different sequence. First, Congress was to vote on the total size of the budget. Once that was determined, Congress would allocate that total budget among spending categories. A budget process of that type is called a *top-down* process. It was argued at the time, that a top-down budget process would lead to a better outcome, in particular, to a smaller budget, than would a bottom-up budget process (Committee on the Budget, 1987).

A similar presumption is shared by many international organizations, which act as if a top-down budget process is inherently preferable to a bottom-up process. The Organization of Economic Cooperation and Development (OECD, 1987) reported approvingly that several

countries adopted top-down budget processes in quest of greater fiscal discipline. Schick (1986) analyzes this report, explaining (and supporting) the thinking behind it in great detail. Blöndahl (2003) argues that the move to top-down budgeting systems has contributed importantly to strengthening fiscal discipline in many OECD countries in the late 1990s. The International Monetary Fund (IMF) expresses a similar preference for top-down processes (IMF, 1996). At the same time, many member states of the European Union use a bottom-up budgeting process, with negotiation between the finance minister and individual spending ministers (von Hagen and Harden, 1995, 1996).

The presumption in favor of top-down budgeting stands in stark contrast to voting equilibrium theory. Suppose rational agents participate as voters in a budget process. In particular, if voters are sophisticated in the sense of Farquharson (1969) and Kramer (1972), they consider the implications of voting in early stages of the budget process for later stages of the process. Furthermore, assume that voters have convex preferences over the individual dimensions of the budget, and that the budget process divides the decision-making process into a sequence of one-dimensional majority decisions. Based on these assumptions, Ferejohn and Krehbiel (1987) show that the equilibrium of a top-down budget process generally differs from the equilibrium of a bottom-up process. Moreover, there is no unambiguous relation between sequence and the size of the budget. Depending on the voters' preferences, a top-down process can lead to larger or smaller budgets.[1]

This argument depends crucially on the rationality of voters—itself an empirical issue. One way to get at this empirical issue is with controlled laboratory experiments. While laboratory experiments create artificial environments, they have the advantage over international comparisons that the design of an institution and the setting of a decision-making process can be controlled much more precisely.

A number of previous studies have tested voting models in laboratory settings. Holt and Eckel (1989) and Davis and Holt (1993) observe sophisticated voting in experiments on two-stage voting games. Similarly, Gardner and von Hagen (1997) find that equilibrium best

---

[1] The nomenclature surrounding "equilibrium" is complicated. Ferejohn and Krehbiel (1987) refer to the voting equilibrium of a top-down process as a "budget process equilibrium," and the voting equilibrium of a bottom-up process as an "appropriations process equilibrium," nomenclature inspired in part by the US Congress. Following Shepsle (1979) and McKelvey (1979), these are now usually called "structurally induced equilibrium." At the same time, they can be thought of as a Nash equilibrium of the associated voting game, as in Fudenberg and Tirole (1991). To avoid confusion, we will simply use the term "equilibrium" throughout.

describes the outcomes in their experimental trials of bottom-up and top-down budget processes. McKelvey and Ordeshook (1984) find mixed evidence for equilibrium, depending on the degree to which they give subjects the opportunity to communicate and thereby circumvent the procedure. They observe that equilibrium has greater explanatory power when communication between subjects is strictly limited, and especially when a Condorcet winner exists. Apart from Gardner and von Hagen (1997), none of these studies specifically deals with budgeting decisions where predictions about the sum of the voting outcomes in all dimensions are of particular interest. Furthermore, while all of these previous experiments examine two-dimensional category spaces under complete information of all players' preferences, the study on which we report in this paper extends to four dimensions and incomplete information.

Our study is based on a series of 128 independent trials of voting over budgets in which we examine three aspects of voting games: equilibrium, dimensionality, and information. The first aspect is a testable implication of the equilibrium theory of Ferejohn and Krehbiel (1987), where the outcome of a budget process depends in a predictable way on voters' preferences and structure of the process. To test this implication, we vary voters' preferences and the structure of the process (bottom-up or top-down) in a systematic way over these 128 trials.

The second aspect concerns the effect that dimensionality—the number of spending categories—has on the budget process and its outcome. Our experiment includes treatments with two and four dimensions. This leads to a gain in applicability, since budget processes in practice only rarely deal with two dimensions, and serves as a robustness check against the criticism that the experiments support equilibrium theory only because they consider very simple decision making problems.

The third aspect concerns the effect of incomplete information on the budget process and its outcome. Contrasted with complete information (where each voter knows the preferences of all voters), our experiment also include treatments with incomplete information. In these treatments, a voter knows only his or her own preferences, and not the preferences of any other voter. This extension is again made in the interest of realism. Many budgets are processed in situations where a voter has limited knowledge of the preferences of other voters.

Incomplete information requires yet another notion of equilibrium, such as Bayes-Nash equilibrium (BNE) for games with uncertainty. This concept, however, presupposes players'

3

consistent beliefs over the other players' types, i.e. their preferences.[2] Note that our so-called incomplete information game does not meet this criterion, because neither the subjects receive any information about (the distribution of) the other subjects' preferences nor do we control subjects' beliefs. Thus, strictly speaking, we do not face a game with uncertainty but with ambiguity for which the concept of BNE does not apply. We prefer a setting of ambiguity to a setting with uncertainty for two reasons: first, we believe that complex stochastic models would overtax subjects' calculation abilities in the experiment, and second, it is realistic to assume that in anonymous budget processes voters often have very little knowledge of other voters' preferences. In this context, the incomplete information treatments serve as a stress test of the complete information voting equilibrium theory.

We find experimental evidence of impacts of incomplete information, in particular that agenda tend to be significantly longer under incomplete information. Nevertheless, complete information equilibrium theory continues to have substantial predictive power, even though it may not strictly speaking apply.

Our main result is that institutions imbedded in a budget process matter. The data from all treatments correspond closely to the theory of structurally induced equilibrium, and institutions drive those equilibria. The subjects display a high degree of sophistication over all treatments. Both extra dimensionality and incomplete information increase the complexity of the decision problem faced by the subjects, and increase the number of periods needed to reach a final decision. In contrast to popular notions about budgeting, there is no unambiguous relationship between the sequence of budgeting decisions and the size of the budget. Whether or not top-down budgeting leads to smaller budgets than bottom-up budgeting depends on the preferences of the actors involved in the budget process.

The paper is organized as follows. The next section sets out the general model. Section 3 describes the various model specifications and the design of the experiment, as carried out at the economics behavior laboratory of the University of Karlsruhe. Our aggregate results are presented in Section 4; individual results, in Section 5. Section 6 concludes with the policy implications of this set of experiments.

---

[2] Ordeshook and Palfrey (1988) show the impacts incomplete information can have on equilibrium theoretically. In their incomplete information voting game, a voter strategically decides on his vote by maximizing his payoff, conditional on his consistent beliefs over the distribution of other voters' preferences.

2. A model of budgeting

We present a model of budgeting which starts with and extends the two-dimensional model of Ferejohn and Krehbiel (1987). Such a model involves dispersed preferences and an odd number of voters, at least 3.

2.1 The general model

There are n voters, indexed by i, i=1, ..., n. Using majority rule, the voters decide on the size and allocation of a budget. There are m spending categories in the overall budget. Each budget category corresponds to a dimension of R_+^m, the non-negative orthant of the m-dimensional Euclidean space. Let the vector $\mathbf{x} = (x_1, ..., x_m) \in R_+^m$ denote a possible budget, where x_j represents spending in the budget category j. The total spending implied by the budget vector $\mathbf{x}$ is

$$B = \sum_{j=1}^{m} x_j .$$

Each voter i has preferences over budgets $\mathbf{x}$ represented by his or her utility function $u_i(\mathbf{x})$. We assume that each voter i has an ideal budget (or an ideal point) $\mathbf{x^*(i)} \in R_+^m$. The closer the actual budget is to a player's ideal budget the higher is the player's utility, where closeness is measured by the Euclidean distance function:

$$u_i(\mathbf{x}) = K_i - \sqrt{\sum_{j=1}^{m} [x_j - x_j^*(i)]^2} ,$$

where K_i is the utility attached to the ideal point.[3] In general, each voter i has an ideal point **x*(i)** distinct from that of all other voters.

Several interpretations of players and their ideal points are possible. For instance, the players may be spending ministers in a coalition government. In this case, an ideal point represents the budget size and composition a spending minister would most like to see enacted. As another instance, suppose the player is a member of a legislature. Then the ideal point may represent a legislator's campaign promise to get this budget or something close to it enacted.

In a budget process, voting translates preferences into outcomes. Votes are based on majority rule. Suppose the vote is over two budget proposals, x_1 and x_2. If the number of those voting for x_1 is greater than the number of those voting for x_2, x_1 defeats x_2. It is well known that a voting game of this type does not have an equilibrium unless the players' preferences are such that a Condorcet equilibrium exists (Riker, 1962). However, an equilibrium can be induced by imposing more structure on the budget process. In our experiments, we study two types of budget processes, bottom-up and top-down.

In a bottom-up budget process the sequence of votes is taken on a spending category at a time. If there are two dimensions the vote is taken first on one spending category and then on the other. We define $\mathbf{x^{bu}}$ as the vector consisting of the *respective* median voter's ideal value in each spending category. The vector $\mathbf{x^{bu}}$ is the equilibrium induced by a bottom-up budget process.

To see how this works, consider the five voters whose ideal points, $\mathbf{x^*} = (x_1{^*},x_2{^*})$, are portrayed in Figure 1: (6, 13), (7,9), (8,16), (11,12), and (11,14). These ideal points are used in Design I of our experiments. A vote is taken first on category $\mathbf{x_1}$. The median voter on this category is the third voter, with a desired spending level of 8. This median value has a 3-to-2 majority against any other value of spending on $\mathbf{x_1}$, and it is thus the bottom-up equilibrium in the horizontal dimension. Next, a vote is taken on category $\mathbf{x_2}$. The median voter on this category is the first voter, with a desired spending level of 13. This median value has a 3-to-2 majority against any other value of spending on $\mathbf{x_2}$, and it is thus the bottom-up equilibrium in the vertical dimension. Putting the results from both dimensions together, we get the bottom-up equilibrium $\mathbf{x^{bu}} = (8,13)$, shown in Figure 1 at the intersection of the vertical and horizontal median lines *a* and *b*, and the total budget in equilibrium is $B^{bu} = 21$. Notice that we would have gotten the same

---

[3] In the two dimensional case the Euclidean utility function leads to circular indifference curves. More general preferences are studied experimentally in Lao-Araya (1998), whose results suggest that equilibrium theory is robust with regard to elliptical indifference curves.

bottom-up equilibrium if the vote had been taken over the $\mathbf{x_2}$-category first, followed by the $\mathbf{x_1}$-category.

In a top-down budget process, the sequence of votes starts with a vote on the total budget, B. Then votes are taken on the spending in all but one of the spending categories. The amount of spending on the last dimension is determined residually. The vector $\mathbf{x^{td}}$ is the equilibrium induced by a top-down budget process.

To see how this works with two dimensions, again consider the five voters whose ideal points are portrayed in Figure 1. A vote is taken first on total spending, $B = \mathbf{x_1} + \mathbf{x_2}$. The median voter on this category is the fourth voter, whose desired total spending level is $B^*(4) = \mathbf{x_1}^*(4) + \mathbf{x_2}^*(4) = 23$. The other voters want to spend 16, 19, 24 and 25, respectively. The median value of 23 has a 3-to-2 majority against any other value of total spending, and it is thus the top-down equilibrium in the total spending dimension. Next, a vote is taken on spending in category $\mathbf{x_1}$, given that total spending is $B^{td} = 23$. Consider, thus, the $\mathbf{x_1} - \mathbf{x_2}$ dimension, i.e., the difference of the ideal point components. The median voter on this category is the fifth voter with a difference of -3, the differences of the other players being -7, -2, -8, and -1, respectively. The median voter's most preferred level of spending on $\mathbf{x_1}$, given that $\mathbf{x_1} + \mathbf{x_2} = 23$, is $\mathbf{x_1} = 10$. This median value has a 3-to-2 majority against any other value of spending on $\mathbf{x_1}$, and it is thus the bottom-up equilibrium in the spending on the $\mathbf{x_1}$-dimension. Putting the results from both dimensions together, we get the top-down equilibrium $\mathbf{x^{td}} = (10,13)$, shown in Figure 1 at the intersection of the -45-degree and the 45-degree median lines c and d, respectively. While the -45-degree line represents the $\mathbf{x_1} + \mathbf{x_2}$ dimension, i.e., the 45-degree line represents its orthogonal $\mathbf{x_1} - \mathbf{x_2}$ dimension.

In general, bottom-up and top-down equilibria differ: this is the main result of Ferejohn and Krehbiel (1987). Comparing the two equilibria in Figure 1, we see that the top-down equilibrium implies a total spending of 23 that is larger than the total spending of 21 in the bottom-up equilibrium.

To see the potential for reversal in total spending levels, consider Design II in Figure 2, which differs from Figure 1 in exactly one feature, the location of the fourth voter's ideal point, which is now (9,9) rather than (11,12). This move does not change the bottom-up equilibrium, which is still at

7

(8,13) implying a total spending of $B^{bu} = 21$. However, it changes the top-down equilibrium, since the median total spending level is now $B^*(1) = x_1^*(1) + x_2^*(1) = 8 + 11 = 19$. In this case, the top-down equilibrium spends less than the bottom-up equilibrium. Note that both x^{td} and x^{bu} belong to the convex hull of the set of ideal points, and therefore, are Pareto optimal.
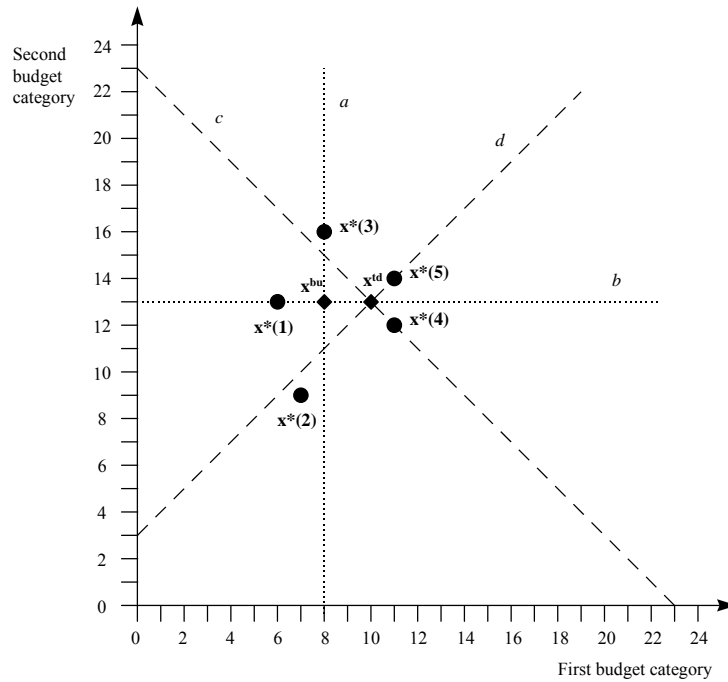

Figure 1:

Ideal points and equilibria in Design I

(*a,b,c,d* are vertical, horizontal, -45°, 45° median lines through the ideal points)

Figure 2:

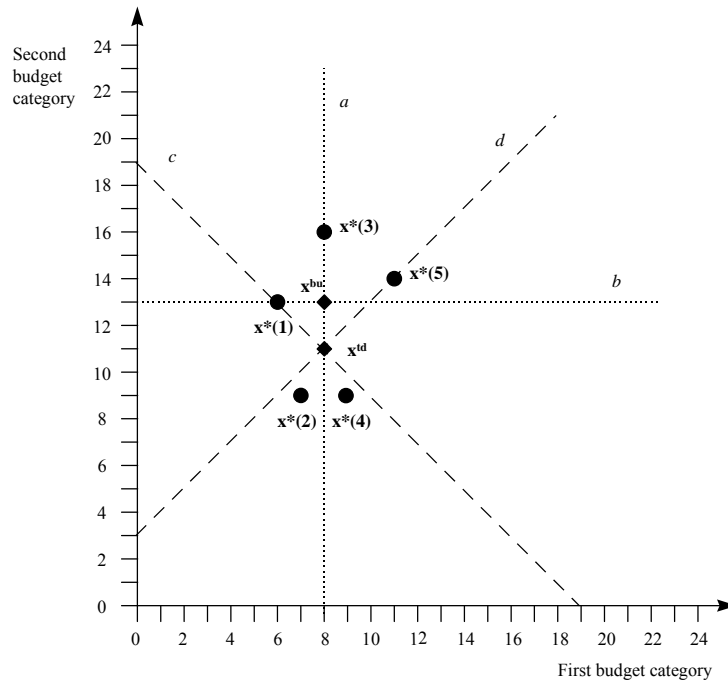Ideal points and equilibria in Design II

(*a*,*b*,*c*,*d* are vertical, horizontal, -45°, 45° median lines through the ideal points)

3. Experimental Design

In all of our experimental treatments the number of voters, n, equals five. The number of spending categories, m, equals either two or four. To specify the voters' utility functions, we have two designs—one design is such that the equilibrium theory of a top-down budget process leads to a larger budget than the equilibrium theory of a bottom-up budget process, and vice versa in the other design.

We discuss first the simpler case m = 2. To specify the voters' utility functions, we have two designs, *Design I* and *Design II*. They are presented in Table 1. Notice that the two designs differ by the fourth voter's ideal point only. Voters 1, 2, 3, and 5 have the same ideal points in both designs. The general intention behind these two designs is to make the difference between the equilibrium induced by a bottom-up process, $\mathbf{x^{td}}$, and the equilibrium induced by a top-down process, $\mathbf{x^{bu}}$, large and in different directions. As can be seen in Table 2, in Design I, the total

budget corresponding to $\mathbf{x^{bu}}$ is smaller than the total budget corresponding to $\mathbf{x^{td}}$, while the opposite is true in Design II. Design I is also shown in Figure 1; Design 2 in Figure 2.

We consider now the case m = 4. The basic principle in getting from two dimensions to four dimensions is projection: (x_1,x_2) maps into (x_1,x_2,x_1,x_2). The ideal points of each player are presented in Table 1. The medians of the ideal points in each dimension are preserved under projection.

For *Design III*, which is the projection of Design I, the medians in dimensions 1 and 3 are 8; in dimensions 2 and 4, 13. Putting the components from the four dimensions together, the bottom-up equilibrium $\mathbf{x^{bu}}$ is given by the vector (8,13,8,13). The total spending under this budget is 42.

The solution $\mathbf{x^{td}}$ induced by the top-down process is the vector (10,13,10,13); this again follows by projection. The total spending under this budget is 46. Notice that $\mathbf{x}^{td}$ is different from $\mathbf{x}^{bu}$, and in particular that $\mathbf{x}^{td}$ spends more than $\mathbf{x}^{bu}$, 46 versus 42.

For *Design IV*, which is the projection of Design II, the medians in dimensions 1 and 3 of the ideal points are 8; in dimensions 2 and 4, 13. Putting the components from the four dimensions together, we get (8,13,8,13) as the bottom-up vector $\mathbf{x^{bu}}$. Total spending under this budget is 42.

The solution $\mathbf{x^{td}}$ induced by the top-down process is the (8,11,8,11). The total spending under this budget is 38. Notice that $\mathbf{x}^{td}$ also differs from $\mathbf{x}^{bu}$. In contrast to Design III, top-down voting leads to a smaller budget, 38, than the budget of size 42 that bottom-up voting adopts.

Table 1:

Individual ideal points and utility function, **x***(i) and $u_i(\mathbf{x})$

Voter i	Two-dimensional				Four-dimensional							
	Design I		Design II		Design III				Design IV			
	$x_1^*(i)$	$x_2^*(i)$	$x_1^*(i)$	$x_2^*(i)$	$x_1^*(i)$	$X_2^*(i)$	$x_3^*(i)$	$x_4^*(i)$	$x_1^*(i)$	$x_2^*(i)$	$x_3^*(i)$	$x_4^*(i)$
1	6	13	6	13	6	13	6	13	6	13	6	13
2	7	9	7	9	7	9	7	9	7	9	7	9
3	8	16	8	16	8	16	8	16	8	16	8	16
4	**11**	**12**	**9**	**9**	**11**	**12**	**11**	**12**	**9**	**9**	**9**	**9**
5	11	14	11	14	11	14	11	14	11	14	11	14
Utility function of voter i $u_i(\mathbf{x})$	$15 - \sqrt{\sum_{j=1}^{2} [x_j - x_j^*(i)]^2}$				$30 - \sqrt{\sum_{j=1}^{4} [x_j - x_j^*(i)]^2}$							

Table 2:

Voting equilibria

Process	Two-dimensional				Four-dimensional							
	Design I		Design II		Design III				Design IV			
	x_1	x_2	x_1	x_2	x_1	x_2	x_3	x_4	x_1	x_2	x_3	X_4
Bottom-up	8	13	8	13	8	13	8	13	8	13	8	13
Σ	21		21		42				42			
Top-down	10	13	8	11	10	13	10	13	8	11	8	11
Σ	23		19		46				38			

The instructions for the experiment are based on those of the classic voting experiment conducted by Fiorina and Plott (1978). Copies of the instructions (in English, translated from the German) are available from the authors upon request.

In the experiment, subjects are told that each of them is member of a group of 5 subjects. In Designs I and II, the group's task is to decide on how many integer-valued tokens to spend on two activities, called A and B. In the instructions for a bottom-up budget process, subjects are told that they first have to decide on the number of tokens to be spent on activity A. Their decision on this number is final. They then have to decide on the number of tokens to be spent on activity B, at which point they have completed their task. In the instructions for a top-down budget process, subjects are told that they first have to decide on the number of tokens to be spent on activities A and B together. Their decision on this number is final. They then have to decide on the number of tokens to be spent on activity A, at which point they have completed their task.

In Designs III and IV, the group's task is to decide on how many tokens to spend on four activities, called A, B, C, and D. In the instructions for a bottom-up budget process, subjects are told that they first have to decide on the number of tokens to be spent for activity A. Their decision on this number is final. They then repeat this process for activities B, C, and D in that order, at which point they have completed their task. In the instructions for a top-down budget process, subjects are told that they first have to decide on the number of tokens to be spent on activities A, B, C, and D together. Their decision on this number is final. They then have to decide on the number of tokens to be spent on activities, A, B, and C in that order, at which point they have completed their task.

At each step, the decision task is to decide on a number of tokens to be spent on some category or combination of categories. The decision process starts with a *proposal on the floor* which equals zero. At any point in time, each subject has the right to propose an amendment. If an amendment is proposed, then the group has to vote on it. If the proposed amendment is accepted, then it becomes the new proposal on the floor. If the proposed amendment is rejected, it has no effect; the proposal on the floor remains unchanged. In that case, each subject is free to propose other amendments, but only one amendment, at a time. At any point of time, a subject may also propose to end the process. If this proposal is accepted, then the proposal on the floor is considered accepted. If the proposal to end deliberations is rejected, then new amendments may be proposed or new proposals for ending the process may be made.

All votes are based on *simple majority rule*. This implies that if three or more members of the group vote in favor of the proposal, then it wins. Otherwise the proposal is rejected.

In the beginning of the experiment, each subject is informed about his personal payoff (or utility) function. The instructions give each subject the exact formula for the payoff function, which is also explained to him in words. In the case of two spending categories (Design I and Design II), the subject is given a table which shows his or her payoff for each combination of numbers in the two spending categories. In all four designs, each subject can, in the final dimension of voting, call up on his or her computer screen to see individual payoff for the proposal on the table and the proposed amendment.

Besides Designs I through IV, which differ with respect to the number of spending categories and the ideal points, we distinguish between two informational treatments. In the complete information treatment each subject knows not only his own ideal point, but also the ideal points of the four other players in his group. In the incomplete information treatment, each player is informed only about his own ideal point. This contrasts with the incomplete information analysis of Ordeshook and Palfrey (1988), where voters do know the distribution of other voters' preferences.

The computerized experiments were organized at the University of Karlsruhe. Subjects were students from various disciplines. Each subject was seated at a computer terminal, which was isolated from other subjects' terminals by wooden screens. The subjects received written instructions that were also read aloud by a research assistant. Before an experiment started, each subject had to answer at his computer terminal a short questionnaire (ten questions) concerning the instructions. Only after all subjects had given the right answers to all questions did decision-making begin. No communication other than through the recognition of proposals and the announcement of the outcomes of votes was permitted.

We organized experimental sessions with 15 or more subjects. Thus, no subject could identify with which of the other participants he or she was grouped. Each subject participated in exactly one trial. Due to the interactive nature of decision making, each group of five subjects yielded one independent observation. For each of the four designs, each of the two budget processes, and each of the two information conditions, we obtained eight independent observations. Table 3 gives an overview of the experimental treatment design. In obtaining these

128 independent observations in total, we also acquired data on 640 subjects, five per experiment.

<p style="text-align:center;">*Table 3*:</p>

<p style="text-align:center;">Treatment design: number of groups (subjects) in each treatment</p>

Information	Process	Two-dimensional		Four-dimensional	
		Design I	Design II	Design III	Design IV
Complete	Bottom-up	8 (40)	8 (40)	8 (40)	8 (40)
	Top-down	8 (40)	8 (40)	8 (40)	8 (40)
Incomplete	Bottom-up	8 (40)	8 (40)	8 (40)	8 (40)
	Top-down	8 (40)	8 (40)	8 (40)	8 (40)

4. Experimental results: aggregate data

This section considers aggregate data from the experiments; the next section individual data. Start with the sizes of the overall budgets we observe in these 128 groups. Tables 4 (for the two-dimensional treatment) and 5 (for the four-dimensional treatment) give group averages of total budget outcomes in all treatments. First consider the direction of the difference in voting outcomes: the Ferejohn and Krehbiel equilibrium prediction is never overturned. In 7 out of 8 cases, when the equilibrium theory predicts a higher budget, a higher budget is observed. In the remaining case (Design I, incomplete information), the result is a tie. Using a binomial test, this result is significant at a 5% level, against an uninformed null hypothesis.

Next, consider the size of the difference in voting outcomes. Here we conduct a series of Mann-Whitney U-tests, against a null hypothesis of no difference in treatment. Take first the complete information treatment. In Design I, the observed outcomes are different at a 10% significance level. In Designs II, III, and IV, the observed outcomes are different at a 5% significance level. Now take the incomplete information treatment. In Design II, the observed outcomes are different at a 10% significance level. In Designs III and IV, the observed outcomes

are different at a 5% significance level. The one case of no difference is again Design I, where a single large outlier in one trial of the bottom-up process is observed.

Table 4:

Average budgets (standard deviation) in the two-dimensional treatments compared to equilibrium

	Design I		Design II	
Information	Bottom-up	Top-down	Bottom-up	Top-down
Complete	21.4	22.5	21.4	19.0
	(1.0)	(1.3)	(1.0)	(1.5)
Incomplete	22.6	22.6	21.5	20.1
	(2.0)	(1.8)	(1.9)	(2.3)
Equilibrium	*21*	*23*	*21*	*19*

Table 5:

Average budgets (standard deviation) in the four-dimensional treatments compared to equilibrium

	Design III		Design IV	
Information	Bottom-up	Top-down	Bottom-up	Top-down
Complete	42.1	46.4	43.0	38.0
	(2.2)	(1.9)	(2.1)	(1.8)
Incomplete	43.4	46.6	43.8	38.6
	(3.7)	(2.1)	(1.6)	(1.6)
Equilibrium	*42*	*46*	*42*	*38*

We summarize these results as follows.

Result 1: Sequence matters. The outcomes observed under bottom-up and top-down budget processes significantly differ from each other in most cases, and they differ in the way predicted by Ferejohn and Krehbiel (1987).

We next argue that equilibrium is a good predictor. To see this, first pool the data from Designs I and II. Next compute the Euclidean distance from an observation to the equilibrium, then average over all observations. The resulting average distance is 1.5. A similar picture emerges for the four-dimensional treatment, where the average Euclidean distance of an observation from the predicted value is 2.6. Pooling over all 128 observations, the average Euclidean distance of the observed budgets from structurally induced equilibrium is 2.1. We summarize this finding as follows:

Result 2: Equilibrium is a good predictor of budget outcome: the average distance of observed outcomes from predicted equilibrium is relatively small.

Note that McKelvey and Ordeshook (1984), who examined committee procedures similar to our procedure but allowed verbal discussion of the proposals, found that results were closer to the equilibrium the more they constrained the degree to which discussion was allowed. A consequence of this is that we provide equilibrium theory its best shot, since we allow no discussion whatsoever.

Let us call an outcome *close* to the equilibrium if it does not deviate from it by more than one unit in any spending category. Table 6 reports the percentages of observations close to equilibrium prediction for all information-dimensionality treatments. Over all treatments, 53.8 percent of the outcomes are close. With complete information, a higher percentage of outcomes is close to the equilibrium than under incomplete information. This is true for each dimensional treatment separately, as well as on average. Performing a Mann-Whitney U-test, these differences

are significant at a 5% level against an uninformed null hypothesis. Second, with lower dimensionality, a higher percentage of outcomes is close to equilibrium than with higher dimensionality. This is true for each information treatment, as well as on average. Again, these differeces are significant at a 5% significance level (Mann-Whitney U-test).

Table 6:

Percentage of budgets close to the equilibrium budget

Information	Two-dimensional	Four-dimensional	Average
Complete	78.1	46.9	62.5
Incomplete	56.3	34.4	45.3
Average	67.2	40.6	53.8

Result 3: Equilibrium is a good predictor of budget outcome: more than half of all observed budgets are close to the predicted equilibrium.

We can use the data in Table 6 to compute Selten's (1991) *measure of predictive success*. Define the *hit rate as* the frequency of outcomes close to the structurally induced equilibrium: the hit rate is given in Table 6. Define the *area rate* as the number of points near equilibrium, divided by the number of points logically encompassed by the experiment. The latter set is the smallest cube containing the origin—the default if voting leads to an impasse—and the maximum ideal point value in any dimension—the highest number we posit any subject to think of logically. Selten's measure is the difference between the hit rate and the area rate.

The area rate is larger in two dimensions than in four dimensions. In two dimensions, the number of points near equilibrium is 9 (the smallest square with equilibrium at its center). The number of points logically encompassed by the experiment is 304, the number of points in the cube with vertices at the origin and (11, 16). The resulting area rate is 9/304, or 3%. Repeating this calculation for four dimensions, the number of points near equilibrium is 81, while the number of points encompassed by the experiment is 41616, the number of points in the

hypercube with vertices at the origin and (11, 16, 11, 16). The resulting area rate is 81/41616, or 0.2%.

Table 7 converts the hit rates of Table 6 into Selten's measure of predictive success. The highest predictive success is in two dimensions with complete information, over 75%. Predictive success decreases with dimensionality, and with loss of information. The lowest predictive success is in four dimensions with incomplete information, barely 34%. To put these predictive successes into context, the predictive success of Nash equilibrium in the experiment of Keser and Gardner (1999) is less than 5%.

Table 7. Predictive Success of Equilibrium

Information	Two-dimensional	Four-dimensional	Average
Complete	75.1	46.7	60.9
Incomplete	64.2	34.2	49.2
Average	69.6	40.4	55.0

Result 4: The predictive success of equilibrium theory is over 75% with complete information and two-dimensions. It falls markedly with incomplete information and higher dimensions.

This result makes intuitive sense. Higher dimensional problems are harder, and subjects do not perform as well at solving them. The same is true for problems where less information is available.

Our final aggregate result concerns how long it takes subjects on average to reach a decision, as measured in moves (1 move = 1 proposal followed by 1 vote). Notice first that in all 128 trials, a decision was reached in every case in fewer than 66 moves. Table 8 shows the average number of moves needed to reach a budget decision in all four information-dimensionality treatments. One sees the same pattern in number of moves as one sees in predictive success. The fewest moves are needed in two dimensions with complete information—the configuration where predictive success is highest. The most moves are needed in four

18

dimensions with incomplete information—the configuration where predictive success is lowest. Using a Mann-Whitney U-test, the difference between average number of moves in 2-dimensions between complete and incomplete information is not significant, while the other three differences in Table 8 are significant at the 5% level (four dimensions: difference between complete and incomplete information; complete information: difference between two and four dimensions; incomplete information: difference between two and four dimensions).

Table 8:

Average (minimum-maximum) number of moves to reach the budget decision

Information	Two-dimensional	Four-dimensional
Complete	11.0 (4-23)	22.6 (10-48)
Incomplete	14.5 (4-29)	28.8 (10-65)

Result 5: Every trial reaches a decision in a finite number of moves. The number of moves needed to reach a budget decision rises with dimensionality and with incomplete information.

This result supports the intuition behind Result 4. In addition, this result is in keeping with the qualitative prediction of Ordeshook and Palfrey (1988), where the agenda is longer under incomplete information than under complete information.

5. Experimental Results: Individual Subject Data

We have seen that equilibrium has considerable predictive success, especially when the budget decision facing subjects is small and when subjects are well-informed. Moreover, subjects never reach a budget impasse. Since all proposals are endogenous, predictive success and avoidance of impasse depend entirely on the quality of proposals and subsequent votes. If

proposals are unattractive, and get voted down repeatedly, the budget process will founder. We now turn to the quality of proposals.

One simple criterion is whether subjects propose their ideal points or not. An ideal point is surely a better proposal than, say "0," and it will represent a Pareto optimum at the very least. Subjects propose their ideal points surprisingly often. Interpret "propose ideal point" as shorthand for "a proposal which, if adopted, would lead to the ideal point value for the subject, when projected onto the dimension determined by that proposal." Table 9 gives the data, drawn from over 10,000 proposals over the course of the 128 trials, excluding proposals to end the process.

Table 9:

Percentage of proposals of a subject's ideal point

Information	Two-dimensional	Four-dimensional
Complete	42.5	40.8
Incomplete	55.9	47.8

We see that ideal points get proposed a lot, over 40% in every treatment. Moreover, in a given dimension, ideal points get proposed more often when information is incomplete. Indeed, employing a χ^2 test against an uninformed null, the difference is significant at 5%. A possible explanation for the high frequency of ideal point values is their safety—a subject can't do better than an ideal point value, at least in a local sense. In the case of incomplete information, proposing one's ideal point also has potential signaling value, as it reveals the proposer's most desired outcome to the other actors.

Result 6: Subjects often propose ideal points, and propose them significantly more often when information is incomplete.

Another related measure of the quality of proposals is whether the proposal, if adopted, would move towards the equilibrium value on that dimension ("equilibrium seeking"). Table 10 gives the relevant percentages.

Table10:

Percentage of equilibrium seeking proposals

Information	Two-dimensional	Four-dimensional
Complete	57.3	49.9
Incomplete	30.8	35.3

Again, these are impressive percentages. In both dimensions, we observe a higher frequency of equilibrium seeking proposals in the case of complete information. Using a χ^2 test against an uninformed null, the difference is significant at 1% in two dimensions, and at 5% in four dimensions.

Result 7: When not proposing their ideal points, subjects strongly prefer equilibrium seeking proposals.

Combining the data of Tables 9 and 10, we conclude that the vast majority of proposals either preserve utility by proposing their ideal points or seek equilibrium even when information is incomplete.

With proposals of such high quality, the subjects have created for themselves quite good material to vote on. Given good proposals to vote on, the final question is how subjects vote, facing such proposals. Here the evidence is again impressive: the vast majority of subjects only vote for proposals that favor them. Table 11 reports the percentages of votes for the more favorable of two proposals, either the one on the table or the one against it. In the event of indifference, we count the vote as for a favorable alternative.

Table11:

Percentage of votes for favorable proposals

Information	Two-dimensional	Four-dimensional
Complete	86.4	72.1
Incomplete	92.4	75.4

Subjects vote for favorable proposals between 72% and 92% of the time. We observe some slippage in the percentage of votes for favorable proposals in the case of four dimensions, compared to two. Using a χ^2 test against an uninformed null, the difference is significant at 5%. The added complexity appears to make it harder for subjects to vote their self-interest, although they still manage to do so much more often than not.

Result 8: Subjects tend strongly to vote for proposals that favor them.

The overall picture that emerges from individual subject behavior is reassuring. Subjects make high quality proposals, and then vote for proposals that are favorable to their individual interests. These twin features of subject behavior drive the predictive success of equilibrium that we observe at the aggregate level.

6. Implications for policy makers

This paper has studied budget processes—the system of rules governing decision-making leading to a budget—experimentally. We conducted a series of 128 experiment trials to study budgeting processes using subjects in a behavior laboratory. At the aggregate level, we find strong support for the Ferejohn and Krehbiel prediction, both qualitatively and quantitatively. We observe no instance of budget impasse, and an overall predictive success of almost 50%. At the individual subject level, subjects tend to make high quality proposals and vote for proposals that favor them. These twin features at the individual level drive the aggregate results towards the predicted equilibrium.

These results have three important policy implications. First and foremost, institutions matter. The kind of budget one gets from a budget process is driven by the voting rule, the

sequence of decisions, and the distribution of ideal points in a predictable fashion. There is no general tendency for top-down budget process to deliver smaller budgets. Although majority voting avoided impasse in all cases, we conjecture that a different rule, such as unanimity voting, would indeed to rather more frequent impasses, such as that involving the EU budget at the EU summit in June 2005.

Second, since sequence matters, policy makers should not presume that a top-down budget process always leads to less spending. Which sequence of decisions leads to smaller total budget depends critically on the preferences of the participants involved in the budget process. Efforts to design budget processes to achieve greater fiscal discipline should focus on other critical aspects, such as the assignment of decision making competences to the participants in the budget process (Hallerberg and von Hagen, 1999).

Third, complexity is costly. If we measure decision-making costs in terms of the number of rounds required to reach closure, then costs go up with more spending categories and with less complete information. To the extent that decision-making costs are important, agenda setters in a budget process, such as finance ministers, are well-advised to keep the overall decision low-dimensional, even if this means relying on local autonomy for more detailed budget allocations. Incomplete information also increases decision-making costs. It does this in several ways: reducing the quality of proposals, reducing the accuracy of votes, and lengthening the number of rounds. This increases the real-world applicability of our results, since complete information, even in a cabinet or legislature of long standing, is rare.

At least one caveat is called for. Our communication condition was quite strict, compared to real world committees and cabinets. As shown by McKelvey and Ordeshook (1984), equilibrium fares less well when communication is open. This is a worthy subject for future investigation.

Acknowledgements

References

Alesina, A., and R. Perotti (1995). "The Political Economy of budget Deficits," *IMF Staff Papers* 42, 1-31.

Alesina, A., and R. Perotti (1999). "Budget Deficits and Budget Institutions," in: J.M. Poterba and J. von Hagen (eds.), *Fiscal Institutions and Fiscal Performance*, Chicago: University of Chicago Press.

Blöndahl, J. R. (2003). "Budgeting Reform in OECD Member Countries: Common Trends." *OECD Journal on Budgeting* 2:4, pp 7-26

Committee on the Budget, United States Senate (1987). *Gramm-Rudman-Hollings and the Congressional Budget Process: An Explanation*, Washington: US Government Printing Office.

Davis, D., and C. Holt (1993). *Experimental Economics*, Princeton: Princeton University Press.

Eckel, C., and C. Holt (1989). "Strategic Voting in Agenda-Controlled Experiments," *American Economic Review* 79, 763-773.

Farquharson, R. (1969). *Theory of Voting*, New Haven: Yale University Press.

Ferejohn, J., and K. Krehbiel (1987). "The Budget Process and the Size of the Budget," *American Journal of Political Science* 31, 296-320.

Fiorina, M. P., and C. R. Plott (1978). "Committee Decisions under Majority Rule: An Experimental Study," *American Political Science Review* 72, 575-598.

Fudenberg, D. and J. Tirole (1991). *Game Theory*, Cambridge: MIT Press.

Gardner, R., and J. von Hagen (1997). "Sequencing and the Size of the Budget: Experimental Evidence," in: W. Albers, W. Güth, P. Hammerstein, B. Moldovanu, and E. van Damme (eds.), *Understanding Strategic Interaction: Essays in Honor of Reinhard Selten*, Berlin: Springer-Verlag, 465-474.

von Hagen, J., and I. Harden (1995). "Budget Processes and Fiscal Discipline," *European Economic Review* 39, 771-779.

von Hagen, J., and I. Harden (1996), "Budget Processes and Commitment to Fiscal Discipline," *IMF Working Paper* WP96/97.

Hallerberg, M. and J. von Hagen (1999). "Electoral Institutions, Cabinet Negotiations, and Budget Deficits within the European Union," in: J. Poterba and J. von Hagen (eds.), *Fiscal Institutions and Fiscal Performance*. Chicago: University of Chicago Press, 209-232.

International Monetary Fund (1996). *World Economic Outlook*, Washington DC, IMF.

Keser, C. and R. Gardner (1999). "Strategic Behavior of Experienced Subjects in a Common Pool Resource Game," *International Journal of Game Theory* 28, 241-252.

Kramer, G. H. (1972). "Sophisticated Voting over Multidimensional Choice Space," *Journal of Mathematical Sociology* 2, 165-180.

Lao-Araya, K. (1998)). "Three Essays on the Fiscal Constitutions of Developing Countries," Ph.D. dissertation, Indiana University.

McKelvey, R. D. (1979). "General Conditions for Global Intransitivities in Formal Voting Models," *Econometrica*, 47(5), 1085-1112.

McKelvey, R. D., and P. Ordeshook (1984). "An Experimental Study of the Effects of Procedural Rules on Committee Behavior," *Journal of Politics* 46(1), 182-205.

Ordeshook, P., and T. Palfrey (1988). "Agendas, Strategic Voting, and Signaling with Incomplete Information," *American Journal of Political Science* 32(2), 441-466.

Organization for Economic Cooperation and Development (1987). *The Control and Management of Government Expenditure*, Paris.

Plott, C., and K. Krehbiel (1979). "Voting in a Committee: an Experimental Analysis," *American Political Science Review*

Poterba, J.M, and von Hagen, J. (1999). *Fiscal Institutions and Fiscal Performance*, Chicago: University of Chicago Press.

Riker, W. H. (1962). *The Theory of Political Coalitions*, New Haven, CT: Yale University Press.

Schick, A. (1986). "Macro-Budgetary Adaptations to Fiscal Stress in Industrialized Democracies," *Public Administration Review* 46, 124-134.

Selten, R. (1991). "Properties of a Measure of Predictive Success," *Mathematical Social Sciences* 21, 153-167.

Shepsle, K. (1979). "Institutional Arrangements and Equilibrium in Multidimensional Voting Models," *American Journal of Political Science* 23, 27-60.

Wildavsky, A. (1975): *Budgeting*, Oxford: Transaction publishers.