Abstract

There is wide theoretical coverage of sophisticated voting despite a clear empirical gap. There are only a handful empirical examples of sophisticated voting which are documented. There is also no systematic searching approach to recover such cases. This paper offers the first systematic approach to identify votes where sophisticated voting seems likely. The algorithm developed here cannot state if sophisticated voting occurred or not, but it can guide scholars to cases where it is promising to focus on.

The proposed algorithm relies heavily on the Bayesian roll call analysis and on a probabilistic choice model. First, ideal points are estimated on a restricted sample, ensuring that the assumptions are not violated. In a second step, a subsample of ideological groups is used to estimate the ideological parameters of a proposal. Finally, based on ideal points of legislators and ideological parameters of proposals it is possible to identify highly atypical voting behavior. The algorithm flags fifteen cases of potential sophisticated voting out of 1281 votes.
"Impressive as this pattern of advances is, the literature on sophisticated voting exhibits a proclivity toward theoretical possibilities rather than empirical realities." (Krehbiel and Rivers, 1990: 549)

1 Introduction

Although there is wide theoretical coverage of the possibility of sophisticated voting, there is a lack of convincing empirical facts. Advocates of the argument that strategic voting occurs on a regular basis have brought some cases to our attention. So far these cases are a tiny subset of all votes. Poole and Rosenthal (1997: 162) only find three cases in the entire literature which they consider to be successful killer amendments.

For the Swiss case two incidents of sophisticated voting were found. One is described by Senti (1998) and the other one by Büttikofer and Hug (2008).

This project intends to develop an algorithm which allows researchers to process large amounts of data. The goal is to narrow down the set of all votes to a much smaller subset of cases where sophisticated voting seems likely to occur. The advantage is that one does not have to rely on interviews with MPs or journalists to be pointed towards cases where sophisticated voting might have appeared. Therefore, this approach is not limited to cases in which all actors realized that sophisticated voting occurred and are willing to share their knowledge.

Given that one receives a subset of votes where sophisticated voting might have occurred, researchers can then focus on those cases and study speeches, party announcements and other sources of ideological positioning.

The paper is structured as follows; section 2 offers a brief overview of the literature on sophisticated voting. Cases which are theoretically interesting are discussed and examined in terms of their compatibility with the Swiss case. Section 3 and 4 present the basic set-up of the algorithm in two different ways. Section 3 tries to explain and justify the proposed procedure in a non-technical fashion and gives an intuitive explanation. Section 4's content is the same as section 3's but presents the steps of the procedure in a more technical way. In section 5 the estimated ideal points and the estimated policy space are presented. Section 6 focusses on the cases which are likely to be incidents of sophisticated voting.
2 Strategic Behavior and Sophisticated Voting

Since the early 1950s it is a well known fact that the parliamentary process offers many opportunities to MPs to engage in strategic behavior. One particular way of acting strategically has caught a lot of attention. What happens if MPs do not always vote according to their preferences?

Farquharson (1969) proposed a model of voting which suggests that MPs vote contrary to their preferences in certain instances. Imagine a situation in which three alternatives exist, where alternative one is pitched against alternative two, and the winner is eventually pitched against alternative three. This situation often occurs with the three alternatives being a bill, an amendment to the bill, and the status quo.

2.1 An Illustrating Example

Imagine a bill proposing to spend a certain amount of money ($b$), an amendment to spend even more money ($ab$), and the status quo ($\phi$) not to spend any money at all. Further, there are three legislators which have the following preference orderings $\phi \succ b \succ a$ (legislator 1), $b \succ \phi \succ a$ (legislator 2), and $a \succ b \succ \phi$ (legislator 3). Here, legislator 1 is opposed to any spending, legislator 2 would like to increase the spending but just moderately. If faced with the decision between spending a lot more ($ab$) and no spending, legislator 2 prefers no spending. Finally, legislator 3 wants to spend as much as possible.

Figure 1: Typical Game Tree illustrating Sophisticated Voting

First, the three legislators have to decide whether the amendment is adopted or not. In a second step they decide on the final vote to adopt the bill, or
if amended, the amended bill or to keep the status quo. This situation is portrayed in figure 1.

If all legislators vote sincere according to their preferences, the amended will be defeated as legislator 1 and 2 vote nay. In the second stage the bill \((b)\) will be pitted against the status quo \((\phi)\). The bill will pass as legislators 2 and 3 will vote yea. This can be understood as the baseline case.

Given the case that legislator 1 realizes at the first voting instance that the situation could be an opportunity to manipulate the outcome in her favor, the outcome will differ. Instead of voting nay legislator 1 can vote yea despite her distaste for more spending. Consequently the amended bill passes. In the final vote legislators 1 and 2 vote nay and the final outcome is \(\phi\). Overall, legislator 1 gets an outcome closer to her preferences.

There are several noteworthy elements of this example. First, by using strategic foresight in the first vote, legislator 1 is able to change the final outcome in her favor. Second, two crucial assumptions are that at least legislator 1 knows the preferences of all legislators and the path of the game. Third, legislator 2 and 3 do reveal their preferences sincerely.

### 2.2 Practical and Theoretical Relevance

There are two different ways to assess the relevance of killer and saving amendments. First, there is a theoretical level which asks if such situations in a legislature can occur and if it is likely for them to occur. Second, at a practical level the question is whether such situations do occur often in everyday legislative politics. Subsequently, one might ask if this is relevant in any terms for describing and understanding preference aggregation.

Social choice theory yielded as an early insight that the possibility of cycles is nontrivial (see e.g. McLean and Urken, 1993). Arrow (1951) proves that there is no method of aggregating preferences which are not prone to cyclical majorities. Black (1958) shows that under certain conditions (single-peaked preferences) in a uni-dimensional space no cyclical majorities arise. McKelvey (1976) shows that in multi-dimensional policy space only under very rare conditions a Condorcet winner can exist.

In a situation with three or more alternatives it is possible that a Condorcet winner exists. But, it is not true in general that a Condorcet winner has to exist and could be recovered by aggregating preferences. Actually, as the number of alternatives and the number of voters increase so does the
Recovering Sophisticated Voting

probability of a cycle (e.g. Black, 1958: 50-51; DeMeyer and Plot, 1970; Riker, 1982: 122).

Moreover, in case a Condorcet winner existed, there is no guarantee that it would be recovered by certainty. This is the main insight of the Gibbard-Satterwhaite theorem. The Gibbard-Satterwhaite theorem shows that there is no voting rule which is strategy-proof and does not offer opportunities to manipulate the outcome (Gibbard, 1973; Satterwhaite, 1975). Hinich and Munger summarize the meaning of the theorem as follows:

"The Gibbard-Satterwhaite theorem is a classical social choice good news, bad news result. On the one hand, it means that all voting procedures (at least, all those that are not trivial or useless) are manipulable under many circumstances. Further, we cannot trust particular votes or messages that voters deliver to represent their true preferences. This is bad news, because it means that voters often don’t vote honestly, and the reasons have to do with the voting procedure itself, not the character of voters" (1997: 166).

The example described in section 2.1 illustrates this fact. The Condorcet winner is the bill ($b$), which wins both pairwise comparisons (with $\phi$ and $a$). But if legislator 1 votes sophisticated, she can change the outcome and the Condorcet winner will not prevail. Manipulated outcomes are theoretically relevant. Situations as described in the example can occur in a legislative setting. We cannot rely on certain special voting procedures which preclude manipulation.¹

This leads to the question if such situations arise frequently in every day parliamentary life. Reconsidering the introductory example some questions arise.

First, as noted in section 2.1 such a situation requires that not only the sequence of votes (game tree) is known to the legislators but also that they know each others preferences (Krehbiel and Rivers, 1990). This is true for the model as presented by Farquharson (1969), but not for the extension of this model brought forward by Enelow (1981). Enelow builds a model in which the legislators do have expectations for the probability of passing for each bill and amended bill. He also provides empirical evidence in form of two cases; the Mathias amendment on a civil rights bill in 1966 and the in the literature well-known Powell amendment to a bill on school construction funding in 1956.

¹This finding holds for all non-random voting methods. Borda-count for example is as well vulnerable to sophisticated voting (Riker, 1982: 143).
Krehbiel and Rivers (1990) doubt that the 1956 case (Powell amendment) is supporting the theory of sophisticated voting. Especially, they disagree with some of the assumed preference orderings by Riker and others. A more recent paper by Calvert and Fenno (1994) disagrees in turn with Krehbiel and Rivers (1990). They provide an additional model for sophisticated voting in which the agenda does not have to be fixed and the preferences do not have to be known with certainty. They also provide empirical support and give evidence of ”[...] clear consciousness of sophisticated voting and agenda control that senators have when such occasions arise.” (Calvert and Fenno, 1994: 372).

Second, if a group of legislators votes in a sophisticated manner, why do others fail to anticipate and prevent a change of the outcome? The literature often focuses on cases in which only one group exercises sophisticated voting. This is no coincidence. Turning back to the example in section 2.1, one can see what happens if all groups engage in sophisticated voting. In such a situation all actors would realize that a vote between a and b is actually a vote between φ (vote for a) and b (vote for b). Therefore legislator 3 would vote for b (although she prefers a to b) and b would win against a. The outcome is b which is the same as if all legislators vote sincerely. If all cast a sophisticated vote, the outcome is the same as if all vote sincerely (e.g. Denzau, Riker, and Shepsle, 1985: 1124). The idea that the outcome changes, which can only occur if not all actors engage in sophisticated voting, is disturbing. Riker summarizes the reason for this discomfort:

"Strategic voting, if successful produces an outcome different from the 'true amalgamation,' however defined, of the values of the members of the group. Of course, just what the 'true amalgamation' is depends [...] on the voting system in use. But, given an agreed-upon constitution for voting, there is some specific outcome produced by voting in accord with one’s true tastes. That specific outcome – the 'true amalgamation' – is precisely what strategic voting displaces” (1982: 156).

Returning to the outlined questions earlier, it was shown that sophis-

---

2 It is often assumed and also empirically shown that sophisticated voting is a coordinative action. This is most clearly stated in Riker (1982: 155): "This suggests that they were influenced to vote as part of a unified maneuver.” (emphasis added). Empirical cases in which only one group behaves sophisticatedly are presented by Riker (1982: 153), Shepsle and Bonchev (1997: 155-157), and Bjurulf and Niemi (1978: 10). Theoretical examples of only one group voting sophisticated are provided by Hinich and Munger (1997: 164), McKelvey and Niemi (1978: 4), and Bjurulf and Niemi (1978:6). There are very few reports where sophisticated voting is widespread or indistinguishable from sincere voting (Enelow and Koehler, 1980).
ticated voting is theoretically relevant. A minority can potentially change the majority outcome. Sophisticated voting can always occur independent of the exact voting rule used. If this occurs frequently is debated. Despite a number of empirical examples (some of which are also doubted to truly reflect sophisticated voting) there is scepticism whether these are very rare incidents or regularly occurring events.

Nevertheless, if sophisticated voting occurs, rarely or more frequently, it is disturbing to most readers from a normative perspective. Even if a Condorcet winner does exist, a potentially minor group can change this outcome by strategic behavior. This seems to tinker with the claim one man, one vote.

3 How to Find the Needles in the Haystack

The leading idea of this study is to narrow down the set of all votes to a subset of votes which consists of cases that are likely to show sophisticated voting. This subset can then be inspected in more detail to see if sophisticated voting actually occurred. To see the advantage of this approach, one has to compare it to other strategies of finding cases which could show sophisticated voting. Normally, this first step of how researchers find sophisticated voting is not documented. One can assume that there is no systematic approach to finding such cases but more that researchers stumble across these cases. This may happen by working closely on a certain case or by interviewing legislators. However, both strategies have considerable downsides. The first cannot be efficient and relies on coincidences, while the latter requires interview partners who know such cases and who are also willing to share this knowledge. The lack of any systematic "finding-algorithm" could partly account for the lack of a considerable amount of empirical evidence.

To my knowledge, this is the first procedure proposed which offers a systematic approach. The main idea is that sophisticated votes are distinguishable from sincere votes. This is to say, that if a group of legislators engages in sophisticated voting, the behavior is atypical compared to their usual voting. In the example in section 2.1 it is legislator 1 who votes contrary to her

---

3 An exception is Wilkerson who first narrows down the set of all votes to those in which there is any report of a legislator calling an amendment a "killer amendment," "bill killers," or using the phrase "kill the bill."
preferences in favor of more spending. The question is then, how to detect such atypical voting patterns.

How would we proceed in a perfect world? We would know the true preferences of a legislator and legislators would vote perfectly according to their preferences without any error. We would also know the ideological positions of any bill and any amendment. Hence, detecting sophisticated voting would be a purely easy task. Observing a wrong vote by a legislator, we would know that this legislator cast a sophisticated vote and maximized her utility in the long run.

In the real world, however, we neither know the true preferences nor can we assume perfect voting, nor do we know the ideological position of any amendment. Working our way back from the ideal world into the real world requires several assumptions and some of these assumptions limit the kind of cases which are recoverable. The next few subsections describe the procedure and try to illustrate what can be estimated, what assumptions are needed to do so, and what implications those assumptions have for the results.

3.1 Three Steps Towards Recovering Sophisticated Voting

The systematic procedure to recovering sophisticated behavior consists of three main steps. First, legislators’ ideal points have to be estimated, then the ideological positions of the amendments are estimated and eventually the residuals are reported. The larger the residuals are, the more likely it is that the case at hand consists of certain sophisticated votes.

There are several challenges to this; first, how can ideal points of legislators be estimated –assuming no strategic behavior– in order to recover sophisticated voting? Second, how can the ideological position of an amendment be determined if people are behaving strategically and therefore their votes do not reflect their preferences? These and related questions will be answered in the following three subsections.

The model that will be used is the spatial voting model. Assuming that ideological preferences can be mapped in a $k$-dimensional space, one can model actors’ utility functions which are decreasing in the euclidean distance.
3.1.1 Preferences, Votes, and Ideal Points

Roll call data is used to estimate legislators ideal points. Several authors forcefully argue that one cannot use roll call data to detect sophisticated voting or, more generally, strategic behavior:

" [...] generally it is inappropriate to use estimates of extant methods (usually generated under assumption of sincere voting) to test models embodying alternate assumptions [...]" (Clinton, Jackman, and Rivers, 2004: 355).

The authors rightfully point to the logical inconsistency of assuming the absence of a phenomenon (here: strategic behavior) to estimate parameters, what eventually leads to interpretations of strategic behavior which was assumed not to be present. This is of course a disappointing conclusion, as especially students of social choice are eager to have a tool allowing them to measure preferences.

But this is not the end of the story. Clinton, Jackman, and Rivers (2004, henceforth CJR) show how one can overcome the problem for a specific setting. They employ a model which assumes that legislators vote independent from their parties conditional on their preferences. It is clear that parties exercise pressure on legislators for certain votes. CJR (2004) extend the model to incorporate this characteristic. In such a situation the legislator has to make a decision if it is worth voting for her preferred outcome and bearing the costs of potential party punishment or if it pays off to vote contrary to her preferences according to her party leaders. The authors decide to incorporate the party pressure into the model. Due to statistical identification it is necessary to make the assumption that there is no party-specific pressure in any vote which eventually gets 65% or more yea or nay votes.\footnote{To be precise, CJR (2004: 365) do not measure party pressure for Democrats and Republicans separately but rather the difference – in terms of direction – between both pressures.}

The cut-off point of 65% may be arbitrary but was also used by others before (Snyder and Groseclose, 2000). The sensitive issue is whether there is no party pressure at all in such cases. One might argue that in such clear decisions there was pressure but it was either not successful or that it would not have been necessary that a bill passes successfully. In either situation there might be some legislators who where influenced by party pressure. But it can be assumed that party pressure was reasonably small. Of course we
are not measuring the *true preferences* in these cases (65% or more yea or nay votes) but it is reasonably near to the legislators preferences.

Turning back to the case at hand, a similar work around will be used. Similar is the idea, that one uses a subset rather than the entire data set. Here, it is necessary to find a subset of the votes which can be used to estimate the ideal points. The goal is to measure not the *true preferences* but rather the typical and legislator-specific voting behavior. This can be done by only relying on roll call data from *final votes*. First, final votes are theoretically appealing as the sophisticated and the sincere vote fall together for all legislators (Enelow and Koehler, 1980: 400; Enelow, 1981: 1066). Therefore it is safe to say that there is no sophisticated voting – in the sense of deviating from a typical pattern – in the final vote. Second, these estimates closely resemble typical voting behavior of a legislator in the lower house in Switzerland. Finally, by relying on all final votes, it is assured that the ideological measures derived are representative for all issues which were debated in any given period.

This workaround is not perfect. In an ideal world the true preferences would be measured directly with a hedonimeter.\(^5\) Again, it is not possible to measure these and so the feasible solution does not capture the preferences with perfect accuracy. As for the described workaround by CJR (2004), it is not perfect but it is a good approximation and reasonably close to what we want to get. After all, the assumption that final votes reflect policy preferences is not a unique claim (Clinton, 2007: 462).

### 3.1.2 Sophisticated Voting as Coordinated Behavior

In the first step, the ideal points are derived on a subset of votes which is representative for all votes. Given these ideal points and the actual voting behavior on amendments, it is possible to derive the ideological coordinates of each amendment. Note that for this second step all votes are used, except the final votes, i.e. all amendments are now in the data set. Given the coordinates of the amendment, it is possible to estimate the expected vote of each legislator and compare this with her actual voting behavior.

The problem lies in the assumptions on which the estimation of the ideological coordinates of the amendment is built on. The standard set-up

---

\(^5\)A hedonimeter is a tool to measure people’s utility, see Colander (2007) for a recent article on early ideas and hopes regarding preference measuring.
assumes sincere voting and does not incorporate sophisticated voting. If a small group of legislators engages in sophisticated voting the ideological coordinates of the amendment will not be heavily biased and the sophisticated voting can be recovered. If this group of legislators is far away of the cutting line, the bias could potentially be large enough and the sophisticated voting could not be recovered. If the group of legislators is of considerable size, the estimates will be biased and the estimated residuals misleading.

Turning back to the arguments and ideas of section 2, remember that only if one group voted sophisticated the outcome changed. These are theoretically interesting cases as a minority can change the majority outcome. The empirical cases presented by Riker (1982) or Bjurulf and Niemi (1978) are examples for sophisticated voting of one group of legislators, and consequently of a minority changing the majority outcome. The interest of this paper lies in recovering exactly such cases. Therefore, only a subset of all cases of sophisticated voting is of interest.

Given that an ideological group votes "[...] as part of a unified maneuver" (Riker, 1982: 155), it is not possible to estimate the ideological coordinates of an amendment. This group has to be excluded from the sample and the ideological coordinates are estimated on the restricted sample. But it is not known a priori which group could engage in sophisticated voting and therefore an iterative process is necessary. Ideological groups have to be defined in advance, coordinates for a given amendment are estimated under exclusion of a certain group. The residuals are thereafter only recorded for that excluded group. This is consistent with the statistical theory underlying the analysis of roll call data.

3.2 Limiting Aspects

After this outline of the proposed procedure it is worth turning the attention to the limitations of this approach. This approach is not intended to recover all kinds of strategic behavior. Actually, this approach focuses on a very special case of strategic behavior, which potentially can change legislative outcomes. Not even all these cases but a subset of all cases of sophisticated voting can be recovered. Only cases in which a very low number (preferably only one) of groups engages in sophisticated voting can be detected.

Readers familiar with the item response theory may think of this, as if some students would on purpose chose a wrong answer category.
These limiting aspects are due to necessary assumptions which have to hold. If these assumptions did not hold, it would not be reasonable to employ methods of roll call voting analysis (ideal points) as presented by Poole (2005), Poole and Rosenthal (1997), or CJR (2004). As it turns out, the limitations are substantially narrower as expected. The interesting cases, those in which a minority can change the outcome, are part of the subset of cases which can be recovered with this approach.

4 The Statistical Set-Up

In this section the statistical model is presented. First, a short repetition of the Bayesian model as proposed by CJR (2004) is given. Second, the estimation of the ideological coordinates of an amendment is derived. Finally, the residuals are defined and an interpretation of them is delivered.

4.1 The Bayesian Analysis of Roll Call Data

To estimate the ideal points of the legislators the Bayesian set up will be used as proposed by CJR (2004). The main idea is based on the spatial voting model. Legislators are more likely to chose points closer to their ideal point as those which are further away. The spatial model assumes a $k$-dimensional ideological space. Ideal points are denoted as $X_i$ and are vectors of length $k$ whereby $i$ is the indicator for the legislator. Therefore, a legislator’s ideal point is defined by $k$ parameters which refer to its coordinates. For any
given vote, the *yea* and the *nay* option are considered to be two points in this ideological space. These points are denoted as $\Theta^y_j$ (for the yea vote) where $j$ is the indicator for the vote. The closer a legislator is to an option, relative to the other option, the more likely she is to vote for it.

The spatial model is also appealing as it interlocks well with probabilistic choice models. Probabilistic choice models assume that decision makers are utility maximizers and that a part of the utility is random. The spatial model directly parametrizes the utility. These models can be estimated by maximum likelihood or within the Bayesian approach.

CJR (2004) model the utility function as quadratic. The utility of voting yea on bill $j$ for legislator $i$ is $U_{ij}(\Theta^y)$. The utility for voting nay is therefore $U_{ij}(\Theta^n)$.

$$U_{ij}(\Theta^y) = -||X_i - \Theta^y_j||^2 + \eta_{ij}$$

$$U_{ij}(\Theta^n) = -||X_i - \Theta^n_j||^2 + \mu_{ij}$$

Assuming that the errors are independent both over legislators and over votes, it is assumed that they have the same expected value ($E(\eta_{ij}) = E(\mu_{ij})$). Further it is assumed that the errors are normally distributed and so is their difference ($\text{var}(\mu_{ij} - \eta_{ij}) = \sigma_{ij}^2$). Using the standard set-up for random utility models:

$$Pr(\text{yea}) = Pr(-||X_i - \Theta^y||^2 + ||X_i - \Theta^n||^2 > \varepsilon_i)$$

and one can formulate the posterior as follows:

$$p(\beta_j, \alpha_j, X_i | y) = \mathcal{L}(\beta_j, \alpha_j, X_i | y) \cdot p(\alpha_j) \cdot p(\beta_j) \cdot p(X_i)$$

whereas the likelihood function is defined as:

$$\mathcal{L}(\beta_j, \alpha_j, X_i | y) = \prod_{i=1}^{n} \prod_{j=1}^{J} \Phi(\beta_j^i - \alpha_j)^{y_{ij}} \times (1 - \Phi(\beta_j^i - \alpha_j)^{y_{ij}})$$

Noteworthy are several things: First, the model is statistically identical
with an item response model. Here, the votes resemble the test questions and the legislators are the same as the test takers. Second, the model is not identified. One has to impose restrictions which make the model identified. This can be illustrated in the unidimensional model where each legislator has a parameter describing her ideal position in the policy space. Imagine a multiplication by the factor $-1$, such a transformation leads to the same spatial model, just the orientation changes (Jackman, 2001: 231). There is nothing in the data which tells us if the conservatives have to be to the right or if they have to be to the left of such a dimension. It is an imposed restriction of the researcher which pins down the model. In general, the number of necessary restriction is $k(k+1)$ (CJR, 2004: 357). In a unidimensional model one has to make 2 restrictions, what is often done by constraining the most liberal legislator to -1 and the most conservative to +1 (see e.g. Rivers, 2003: 17).

The advantage of this model over other models, as proposed by Poole (2005), Poole and Rosenthal (1997) or Snyder and Groseclose (2000), is that it directly incorporates the spatial voting model into the statistical model. By using a Bayesian approach, one can avoid several problems, e.g. in case every legislator votes according to her preferences, the Bayesian choice model does not 'blow up' as an ordinary logit or probit estimated by ML would.

As outlined earlier, only a subset of the votes will be used to estimate legislators’ ideal points. Only final votes will be included to estimate these ideal points. This has several favorable effects, first, it ensures that all relevant questions are considered and second, the assumption of the model – sincere voting – is not violated as in the final vote sincere and sophisticated voting are observably equivalent.\(^7\)

### 4.2 Estimating Amendments’ Coordinates

Given the ideal points of MPs, each vote choice has to be situated in the ideological space. The following assumptions according to the underlying model of CJR (2004) are imposed:

1. The utility function is quadratic.

\(^7\)US American scholars might wonder if this is valid as vote trading would violate the assumptions. Vote trading is not known in the Swiss system. One explanation might be that most of these bargains are not done in a final step (the vote) but in advance by drafting a bill in a committee such that it will gain majority support.
2. MPs choose the option which yields the highest utility to them.

3. The decision process of the MPs is not perfect. The error is assumed to follow a logistic distribution.

Only the last assumption departs from the CJR setup. This is due to computational simplicity. The probability of voting yea is $Pr(U^y - U^n > \zeta)$ where $U^y$ represents the utility of voting yea, $U^n$ represents the utility of voting nay, and $\zeta$ follows the logistic distribution.

For notational simplicity, the following expression describes a two dimensional policy space. This space can be of any dimension and there is nothing special to the two-dimensional space. The two dimensional space is appealing, as it clearly shows that these models are not restricted to an one dimensional policy space. There is also a substantive reason to display the two-dimensional policy space. Statistical analysis of Swiss roll call data implicates that a two-dimensional space is an adequate description for the lower house of Switzerland (e.g. Hug and Schulz, 2007: 313).

For a two-dimensional ideological space ($k = 2$) each choice is defined by two parameters (coordinate on the first dimension and on the second dimension).

$$\Theta^y = (\theta^y_1, \theta^y_2) \tag{6}$$
$$\Theta^n = (\theta^n_1, \theta^n_2) \tag{7}$$

The utility function is quadratic as described above. The ideal position of a given MP is denoted $X$. For a given proposal $j$ the utilities are:

$$U^y_i = -||X_i - \Theta^y||^2 + \eta_i$$
$$U^n_i = -(x_1 - \theta^y_1)^2 - (x_2 - \theta^y_2)^2 + \eta_{i1} + \eta_{i2} \tag{8}$$

The same also holds for the other decision option (voting nay):

$$U^n_i = -||X_i - \Theta^n||^2 + \mu_i$$
$$U^n = -(x_1 - \theta^n_1)^2 - (x_2 - \theta^n_2)^2 + \mu_{i1} + \mu_{i2} \tag{9}$$

\*There is actually one downside of higher dimensionality; as the number of dimensions increases, so does the number of parameters which have to be estimated.
Building on these utilities, one can model the decision process. A possible dependent variable can be assumed to be the degree of acceptance towards a certain bill. Note that the degree of acceptance is not observed but only the final decision (\textit{yea} or \textit{nay} vote). Although the mentioned variable is not observed, one can build a probabilistic choice model by relying on the latent variable idea. By assuming that the utility has a random part, here, $\eta$ and $\mu$ (Steenbergen, 2007).

The probability of voting \textit{yea} is therefore – given the assumptions 1 to 3 (note, that 3 assumes the error $\zeta_i$ to follow a logistic distribution) for a given proposal $j$:

\begin{align}
Pr(\text{yea}) &= Pr(U^y - U^n > 0) \quad \text{(12)} \\
Pr(\text{yea}) &= Pr(-||X_i - \Theta^y||^2 + ||X_i - \Theta^n||^2 > \eta_i - \mu_i) \quad \text{(13)} \\
Pr(\text{yea}) &= Pr(-||X_i - \Theta^y||^2 + ||X_i - \Theta^n||^2 > \zeta_i) \quad \text{(14)}
\end{align}

This expression can also be formulated as:

\begin{equation}
\pi_i = \frac{1}{1 + \exp[-(x_1 - \theta^y_1)^2 - (x_2 - \theta^y_2)^2 + (x_1 - \theta^n_1)^2 + (x_2 - \theta^n_2)^2]} \quad \text{(15)}
\end{equation}

As the probability of a \textit{nay} and \textit{yea} vote is derived, the formulation of the likelihood function collapses for a given vote $j$ to a choice model with two options:

\begin{align}
\mathcal{L} &= \prod_{i=1}^{n} \left[ \pi_i^{y_i} \cdot (1 - \pi_i)^{(1-y_i)} \right] \quad \text{(16)} \\
\ell &= \sum_{i=1}^{n} \left[ y_i \ln(\pi_i) + (1 - y_i) \ln(1 - \pi_i) \right] \quad \text{(17)}
\end{align}

Here $y_i$ is the vote of legislator $i$ on the given proposal $j$. The parameters to estimate in this function are displayed in equation 15.

In this example there are four parameters to estimate: $\theta^y_1$, $\theta^y_2$, $\theta^n_1$, and $\theta^n_2$. This is done by maximizing equation 17; those four values for the $\theta$’s which maximize $\ell$ are the estimates $\hat{\theta}^y_1$, $\hat{\theta}^y_2$, $\hat{\theta}^n_1$, and $\hat{\theta}^n_2$.

There are limitations to this model. It is possible that the maximizing process ’blows up’ if all MPs vote perfectly given their ideal points (Poole, 2005: 37). One way to circumvent this is to use a Bayesian set-up which by incorporating a prior with a non-infinite variance will yield an estimate...
smaller than infinity (Altman, Gill, and McDonald, 2004: 251). Alternatively one can also estimate the model and use the penalized maximum likelihood approach (Firth, 1993; Zorn, 2005).

This estimation will be done for each proposal by excluding consecutively one ideological group. This is necessary because including all groups and afterwards finding a sophisticated vote would be logically inconsistent. It is exactly because of the assumptions of the estimated spatial voting model – the assumption that legislators vote sincere – that one has to exclude a group. When group $t$ is excluded the parameters of the $\text{yea}$ and $\text{nay}$ vote can be estimated. Given these estimates and the ideal points of the group $t$, predictions for the vote choice of each member of $t$ can be made. If the predictions turn out to be wrong for most or all of the legislators this is to be regarded as circumstantial evidence. It is likely then that the group $t$ engaged in sophisticated voting. The next subsection explains what these errors are and what alternative explanations exist.

### 4.3 What Are The Errors

A measure of vote-error (given the prediction) is the difference in distances between the ideal point and the two choices. If the utilities are very similar the difference will be low. If the difference is very large, the voting-error is very unlikely, and it is reasonable to take a closer look at the case. If a legislator votes according to the prediction, the error is set 0. But if a legislator does not vote according to her expected vote, the error is the probability of her expected vote. This ensures that the error grows as the probability of a certain expected action ($\text{yea}$ or $\text{nay}$ vote) increases.

If the model predicts a legislator $i$ to vote on a certain amendment $j$ with $\text{yea}$ by 80% but one actually observes a $\text{nay}$ vote, the error is 0.80. More technically:

$$
\kappa_{ij} = \begin{cases} 
0 & \pi_{ij} \geq 0.5 \text{ and } \text{yea vote} \\
\pi_{ij} & \pi_{ij} \geq 0.5 \text{ and } \text{nay vote} \\
0 & \pi_{ij} < 0.5 \text{ and } \text{nay vote} \\
1 - \pi_{ij} & \pi_{ij} < 0.5 \text{ and } \text{yea vote}
\end{cases}
$$

If, for this given vote $j$, most or all legislators of group $t$ have a positive nonzero $\kappa$, it is an atypical vote. These cases are suspicious of being the result of sophisticated voting as coordinated action. But not all cases in
which most or all of the legislators have an error $\kappa_{ij} \neq 0$ are necessarily the result of sophisticated voting. This leads to the question how the errors can be alternatively interpreted.

Besides the cases in which it is due to sophisticated voting, there are three more possibilities. First, the residuals can be due to an error of the decision maker as in any random utility model. It is very unlikely that a large fraction or even all legislators of an ideological group $t$ would fail to make the ‘right’ decision at the same time.

Second, the effect of all not modeled parameters are pushed into the residual term. As in any other regression model, the effect of a left out covariate influences the residual. If legislators are systematically influenced by factors which are not displayed in the ideological space, it is possible that we observe a supposedly ‘wrong’ decision. What could be such unobserved or unmodeled covariates? Following Kingdon (1981), there are several factors which can influence a legislators decision. Two types of pressure are noteworthy: party pressure and constituency pressure. Party pressure is very likely and was already shown to be relevant for the voting history of a legislator (CJR 2004; Bailer et al., 2007). If a large fraction or all of legislators of an ideological group vote contrary to our expectation, party pressure cannot account for this. How should a party have an ideal point that is different from all its legislators?

Contrary to party pressure, constituency pressure could account for a nonzero error of most or all legislators of an ideological group. This would be the case if a majorly unpopular proposal is on the table, whereas the legislators have different opinions than the people.

Third, if a proposal is on a special policy issue it is possible that it does not fit well in a reduced policy space of only $k$ dimensions. This can be illustrated for the U.S. Congress, which for most periods is well characterized by a one-dimensional policy space reflecting mainly the liberal-conservative axis. If a proposal gets on the floor which touches the north-south divide, this can potentially lead to unexpected results given the one-dimensional model. This possibility cannot be ruled out for this application. It is part of the costs of data reduction and there is no way within the roll call data analysis to avoid these costs.

\footnote{Note, that the error is not due to perceptual error as $\mu_{ij}$ and $\eta_{ij}$ are outside the squared expression (Jackman, 2008).}
5 Empirical Results

The last section described the statistical framework of roll call analysis. In this section, the focus is more on practical issues of applying roll call data analysis. How does one determine the number of dimensions? Which restrictions are imposed to make the model identified? Answers to these two questions and more is provided in the next subsections.

Before turning to these questions some information on the data is provided. The initial data set covers all roll call votes in the 47th Legislature up to December 2005 (2003-2005). There is a total of 1933 roll call votes, whereof 183 are final votes. As some members of the lower house resigned during these two years, there are actually 207 legislators in the data set.\footnote{This data set was put together with a great effort by Daniel Schwarz who uses this data for his PhD thesis. Due to his courtesy this project was made possible.}

5.1 Measuring Preferences – Ideal Points

On the basis of the spatial voting model a Bayesian roll call analysis, as presented by CJR (2004), was performed. This approach requires the researcher to specify \textit{a priori} the number of dimensions of the policy space. How many dimensions should one chose? In extremis one could use as many dimensions as votes, suggesting that each vote is its own dimension. There are at least three downsides to such a proceeding: first, the model is not doing a good job in reduction of complexity. There is no point in roll call data analysis if one has as many dimensions as votes. Second, as the number of dimensions increases so does the number of required restrictions in order to identify the parameters. Restrictions are unpleasant as they often stem from researchers subjective beliefs and may influence the result.\footnote{There are actually only two restrictions which are free in the sense that they do not make any substantive difference. The restiction of an overall mean of 0 and a standard deviation of 1 do not depend on any subjective beliefs as they are pure rescaling restrictions.} Finally, a higher number of dimensions decreases the precision of the estimates. Imagine the extreme case in which the number of dimensions equals the number of votes; in such a situation it is only possible to determine if a voter is above or below the cutting line (between the \textit{nay} and \textit{yea} region) but not the exact position. The disadvantage of a low number of dimensions is that it potentially reduces policy space to a lower number dimensionality space, thereby missing important axes of political life.
There are several contributions for the U.S. Congress where scholars settle with a unidimensional space (e.g. CJR 2004; Wilkerson, 1999). For Switzerland ideological scores are often derived by building a scale where \textit{a priori} votes are classified as contributing to a certain dimension (see e.g. \url{www.smartvote.ch}). All these models set the number of dimensions without any clear criteria. Often a two dimensional model is used where the first dimension is considered to be an economic dimension and the second one is a societal dimension.

Here, two specifications are presented; a unidimensional model and a two dimensional model. The unidimensional model was estimated by imposing a normalizing restriction. This restriction transforms the estimates in a way, that they have a mean of 0 and a standard deviation of 1. This is sufficient for \textit{local} identification.\textsuperscript{12} This model is presented in the appendix.

The two-dimensional model requires more restrictions than the unidimensional model does. Here, the same two restrictions as for the unidimensional model were used; ideal points’ mean is 0 and the standard deviation is 1. There are still additional restrictions needed. One alternative is to use constraints which assign fixed values to two legislators. This would be imposing four constraints (2 parameters, 2 legislators).\textsuperscript{13} There is no straightforward way to impose constraints and it is especially difficult if the meaning of the dimensions is not known \textit{a priori} (Jackman, 2001: 233). How should one determine which legislator is the most liberal one? Note, it is not asked which legislator likes to appear as the most liberal but instead which legislator’s voting behavior is most liberal. Therefore, this question cannot be answered without imposing an \textit{a priori} substantive meaning to the dimensions.

Jackman (2001) describes another way by constraining the items (votes) by imposing a prior on the discrimination parameter. I opted for two votes which can be assigned to two common dimensions; the economic egalitarian and the society dimension. The first vote is on a popular initiative which wanted to force government to keep postal offices whether they were prof-itable or not. As other former public enterprises (e.g. public transportation,

\textsuperscript{12}Obviously, this model is not globally identified as one could multiply the parameters by \(-1\). In this application, left legislators received lower scores and right legislators received higher scores leading to a typical left-right scheme.

\textsuperscript{13}In the Bayesian roll call analysis these constraints are imposed by assigning a prior with a precision of infinity, i.e. a variance of 0.
telecommunication) the postal services were privatized but the government remained the largest shareholder. This question clearly picks up the divide on the economic dimension and the vote was decided by 97 to 84 votes against the initiative. The second vote which should capture the second dimension (society) was on registration of homosexual relationships.

**Figure 3:** Estimated Ideal Points, Two-Dimensional Model

*Note:* Different colors correspond to different parties. Only the five largest factions are labeled. The ellipses are drawn for interpretational reasons and do not correspond to any measure of uncertainty.
The law was intended to provide a legal framework to homosexual couples without giving them the same status as a married heterosexual couple. The vote succeeded with 117 to 50 votes and was mainly opposed by the legislators of the SVP, EDU, and the EVP.

Technically such constraints are imposed by setting a prior on the discrimination parameter. In the two-dimensional model the discrimination parameter is a 2 by 1 vector. Following Jackman (2001: 235), I choose a prior which is modest compared to the estimates of the discrimination parameters in a unidimensional model. The parameter estimates were between -15 and 15. I opted for 5 and imposed a precision of infinity.

Note however, that the two dimensions may not be directly assigned to the two dimensions society and economy just because the votes were picked that way. To recover what each dimension picks up on, one would have to look closely at the discrimination parameters of the single votes. The

<table>
<thead>
<tr>
<th>Model</th>
<th>% Correctly Predicted</th>
<th># Parameters</th>
<th>PRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>79.1</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Unidimensional</td>
<td>93.5</td>
<td>325</td>
<td>0.689</td>
</tr>
<tr>
<td>Two-Dimensional</td>
<td>97.0</td>
<td>765</td>
<td>0.856</td>
</tr>
</tbody>
</table>

*Note: The baseline model assumes every legislator to vote with the majority. Of the 183 final votes, 117 were not unanimous votes and could be used for these models. The number of parameters is \(n \cdot k + j \cdot (k + 1)\) whereby \(n\) is the number of legislators, \(j\) is the number of items (bills), and \(k\) is the number of dimensions. PRE is the proportion of error reduction.*

Baseline in this application is the prediction without a model; that is to say, that every legislator is assumed to vote with the modal category. Assuming that every legislator votes with the modal category yields 79.1% correctly predicted cases. The unidimensional model predicts 93.5% of the cases correctly, while the two dimensional model correctly predicts 97.0% of the cases. Therefore, the unidimensional model has a PRE (proportional reduction of error) of 68.9%. The PRE for the two-dimensional model is 85.6%.
5.2 The Residuals....Where Is The Gold in All This Sand

After obtaining the ideal points, the algorithm, as presented in sections 3 and 4, can flag cases. In doing so, additional settings have to be made.

Table 2: Likely Votes For Sophisticated Behavior

<table>
<thead>
<tr>
<th></th>
<th>Party</th>
<th>Vote Identifier</th>
<th># of Expected</th>
<th># of Votes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Yea Votes</td>
<td>Votes</td>
</tr>
<tr>
<td>1</td>
<td>CVP</td>
<td>208</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>CVP</td>
<td>209</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>CVP</td>
<td>230</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>4</td>
<td>CVP</td>
<td>283</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>CVP</td>
<td>291</td>
<td>21</td>
<td>22</td>
</tr>
<tr>
<td>6</td>
<td>CVP</td>
<td>340</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>CVP</td>
<td>539</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>CVP</td>
<td>555</td>
<td>0</td>
<td>22</td>
</tr>
<tr>
<td>9</td>
<td>CVP</td>
<td>596</td>
<td>3</td>
<td>16</td>
</tr>
<tr>
<td>10</td>
<td>CVP</td>
<td>631</td>
<td>7</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>CVP</td>
<td>727</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>12</td>
<td>CVP</td>
<td>746</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>13</td>
<td>CVP</td>
<td>830</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>14</td>
<td>CVP</td>
<td>898</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>15</td>
<td>CVP</td>
<td>966</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>16</td>
<td>CVP</td>
<td>1109</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>17</td>
<td>CVP</td>
<td>1125</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>18</td>
<td>CVP</td>
<td>1140</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>19</td>
<td>CVP</td>
<td>1173</td>
<td>23</td>
<td>24</td>
</tr>
</tbody>
</table>

*Note:* Entries only for the Christian Democratic Party (CVP). Reading example: On vote 230 thirteen legislators of the CVP voted. All thirteen were expected to vote *yea* but all of them voted *nay*.

In the following section the algorithm was set to flag votes where all legislators of a given party – which participated in a vote – voted contrary to the spatial model’s prediction. The 19 cases in table 2 are only the cases for which the CVP (Christian Democratic Party) was under investigation – the CVP was excluded for estimating the ideological coordinates of the bill. In these 19 votes, all legislators of the CVP (who did not abstain) voted contrary to the model’s prediction.
Overall the algorithm identifies 62 cases for the four potential sophisticated players (SP, CVP, FDP, and SVP). Given that there are 1281 votes it is already an advancement that the set of potential sophisticated votes is narrowed down to 62. Eventually, one will have to look at each bill separately and therefore it is reasonable to start with those bills where it seems more likely to find incidents of sophisticated voting. Here, the focus will be on those bills where all legislators voted the same way. The goal of the algorithm is to find cases where one faction in the parliament engages in sophisticated voting, therefore it is reasonable to focus on cases where the faction acts like a unified player. There are 32 legislators of the CVP in the parliament and there is no vote in which all of them participated. This is already due to the fact that some of legislators were replacements for retiring legislators. If only a low number of legislators participated in the vote, it is unlikely that their behavior makes a difference in the overall outcome. This is why the focus will be on those cases where 70% or more MPs participated in voting.

Note, there are three restrictions imposed here which were set arbitrarily. First, the focus will only be on those cases where every legislator voted contrary to her expected voting behavior. Second, the focus will only be on those cases where all legislators voted the same way. Finally, I focus on those cases in which a large part of the ideological faction participated in a vote (e.g. for the CVP at least 20 legislators). Note, however, that the first and the third restriction use arbitrarily fixed thresholds which can be changed. Following these restrictions, there are three cases which remain; these cases are written in bold typefaces (table 2). It seems most promising to first look at these three cases. One can do this for all four major party factions in the parliament. This yields a total of fifteen especially interesting cases (table 3). These fifteen cases are the core result of this algorithm.

It is possible that some of these cases are the result of a very unusual voting behavior of all other actors. Note, there are more than a dozen different parties in the parliament, if these groups vote in special (and for the two dimensional model atypical) way, the estimated ideological coordinates of the bill may be biased or just wrong. The only way to distinguish between very unusual votes and incidents of sophisticated voting is to look closely at each case and thereby reconstructing the game tree as well as likely preference structures of the parties.
Table 3: Likely Votes For Sophisticated Behavior

<table>
<thead>
<tr>
<th>Party</th>
<th>Vote Identifier</th>
<th># of Expected Yea Votes</th>
<th># of Votes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 SP</td>
<td>715</td>
<td>0</td>
<td>42</td>
</tr>
<tr>
<td>2 SP</td>
<td>1065</td>
<td>0</td>
<td>39</td>
</tr>
<tr>
<td>3 SP</td>
<td>1160</td>
<td>0</td>
<td>48</td>
</tr>
<tr>
<td>4 CVP</td>
<td>555</td>
<td>0</td>
<td>22</td>
</tr>
<tr>
<td>5 CVP</td>
<td>830</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>6 CVP</td>
<td>1109</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>7 FDP</td>
<td>402</td>
<td>0</td>
<td>27</td>
</tr>
<tr>
<td>8 FDP</td>
<td>539</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>9 FDP</td>
<td>555</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>10 FDP</td>
<td>1193</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>11 SVP</td>
<td>275</td>
<td>42</td>
<td>42</td>
</tr>
<tr>
<td>12 SVP</td>
<td>566</td>
<td>0</td>
<td>51</td>
</tr>
<tr>
<td>13 SVP</td>
<td>667</td>
<td>44</td>
<td>44</td>
</tr>
<tr>
<td>14 SVP</td>
<td>754</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>15 SVP</td>
<td>1054</td>
<td>0</td>
<td>41</td>
</tr>
</tbody>
</table>

Note: Entries only for the cases which do not violate the three restrictions from p. 32. Reading example: see table 2.

It was the goal to narrow down the number of potential votes considerably. These cases represent 1.2% of all votes which were investigated. There is no certainty that any of these cases is the product of sophisticated voting but if one wants to look for it, my best advice would be to first look at these 15 cases.

6 Remaining Questions

The algorithm presented here was constructed for an application in Switzerland. The question is whether this approach could also be used in other countries. Put differently, which are the buried assumptions of the algorithm? This section tries to expatiate these assumptions. By stating the necessary assumptions it should be easier to evaluate the opportunities in other settings for this approach.
Another aspect which was not brought up so far is the failure rate. How likely is the algorithm to pick up every case of sophisticated voting? How likely is the algorithm to ignore cases in which sophisticated voting did not occur? The answers to these two questions cannot be conclusively determined.

6.1 Buried Assumption

The approach presented in this paper rests on a number of assumptions. First, the number of ideological groups (party factions) in parliament may neither be too small nor too large. If there are only very few groups, let’s say two, it becomes increasingly difficult to estimate the coordinates of the bill parameters. Especially if the parties vote very cohesively it is likely that the remaining party (one party is always excluded for this step) votes completely \textit{nay} or \textit{yea}. In such a case there is no variance in the voting behavior and the coordinates of the bill are not identified. But the number of ideological groups should also not be too big. If a parliament consists of a high number of parties and they all have similar seat shares it is likely that one group alone will not be able to change the majority outcome by voting sophisticated. If this is the case, only coordination of several ideological groups will be successful. But this algorithm assumes that only one group engages in sophisticated voting in a given vote and that the others vote sincerely.

Second, this approach relies implicitly on a presidential system in which voting behavior is mainly driven by preferences and not by alternative reasons. In presidential systems there is no overly strong party pressure on votes as the governments survival does not depend on approval. If parties and legislators behave according to their preferences this approach should work. Even if legislators occasionally face party pressure, this does not bias the estimation of ideal points. It is important to understand that ideal points are less understood as the legislator’s true preferences but more as her typical voting behavior.

6.2 Type-I-Error and Type-II-Error

This algorithm eventually flags a vote as likely to be the product of sophisticated behavior or not. This raises two possible errors for the algorithm. A type-I-error refers to a vote which is flagged although no sophisticated
voting occurred. A type-II-error is the mirror image, a vote is not flagged although sophisticated voting occurred.

As presented in section 4.3 ("What Are The Errors") it is possible that the measured errors in a voting decision are high (close to 1) although no sophisticated voting happened. Two of three outlined possibilities are very unlikely. The third possibility, a missed dimension, cannot be ruled out. If a vote is mainly decided by preferences on an additional (not included) dimension, it is possible that the voting behavior seems atypical. One way to circumvent this is to use a higher dimensional policy space. This could be an argument for using high dimensions of policy space although the inclusion of an additional dimension would not yield a very high increase in the share of correctly predicted cases. In case not enough dimensions are modelled, it is likely that the type-I-error rate will increase.

The other error type, type-II-errors, is already built-in by the assumption that only one group engages in sophisticated voting. If several groups engage in sophisticated voting, the estimates of a bill’s coordinates will be misleading as the assumption of sincere voting is violated. Excluding a group will not do the job as part of the other groups is engaged in sophisticated voting.

These are only tentative explanations on the likelihood of either error type. As it is not clear how one could determine these error rates analytically or formally, the only way remaining is simulation. A monte carlo analysis could help to identify circumstances which increase or decrease the likelihood of either error type. This aspect remains a task for the future.

7 Conclusion

The idea of sophisticated voting is forty years old and goes back to Farquharson (1969). Despite the wide theoretical attention which this phenomena attracted the empirical literature on sophisticated voting remains scarce. It is not clear if the rare examples of sophisticated voting – which are sometimes doubted by other scholars – are so rare because they do almost never occur or if the examples are so rare because they are hard to identify.

So far there is no systematic approach for how to search for cases of sophisticated voting. This paper develops such a systematic approach and hopes to be a relief at the identification (searching) stage of such projects.
The approach is a three step algorithm which relies on the framework of probabilistic choice models. In a first step ideal points are derived on a restricted sample (Bayesian roll call data). In a second step the ideological coordinates of a proposal are estimated based on a probabilistic choice model on a restricted sample (one party is excluded). Finally, model based voting predictions are compared with actual voting for each party faction. Deviations of expected and actual behavior are likely to be associated with cases of sophisticated voting.

The proposed procedure yields fifteen cases for a two year period of the lower house of parliament in Switzerland from 2003 to 2005. It is now a matter of close case inspections to recover the game tree and to assume a likely preference structure. Only based on these two latter steps a credible argument for empirical evidence of sophisticated voting can be made. The contribution of this paper is that it clearly points out fifteen cases which should be examined first.
References


8 Appendix

8.1 The Unidimensional Model

Figure 4: Estimated Ideal Points and 95% HPD; Unidimensional Model

Note: Points identify ideological positions, horizontal lines represent the 95% HPD. Different colors correspond to different parties. Only the five largest factions are labeled. Not all names of legislators are displayed due to readability.

Readers familiar with Swiss politics will not be surprised by figure 4. At the very left one finds mainly legislators of the Green Party (GP) and the Social Democratic Party (SP). In the central area with a slight leaning towards the right, one sees the legislators of the Christian Democratic Party (CVP) and the Free Democratic Party (FDP). At the right the legislators of the Swiss People’s Party (SVP) can be found. Two aspects are noteworthy. First, there is fairly a lot of party overlap. To the left we see that the SP and GP have a large overlap whereby the SP tends a little bit more towards the center than the GP. The CVP and the FDP share essentially the center area although there is a slight tendency for each of them to one direction (CVP tends a little bit more towards the left, FDP more towards the right). The only party which does not have much overlap is the SVP. Second, there is a clear gap between the SP and the two center parties (CVP, FDP).
Traditionally, parties in Switzerland are divided first between left parties and bourgeois parties. It is exactly here, between the SP and the CVP (and the FDP), where these two distinct groups do not have any ideological overlap.

8.2 Assuring Convergence? (two-dimensional model)

As with any Bayesian model, convergence is a highly important issue. First, the chain has to run long enough in order to converge to its stationary distribution. Second, it has to run long enough within the target distribution to explore the entire target distribution (Gill, 2008: 156). If a chain has converged, it will explore the target distribution and thereby one can learn about this distribution’s moments. Of interest are especially the first two moments, the mean and the variance.

Figure 5: Trace Plots for Three Legislators (1st Dimension)

Note: The three legislators are from three different parties (FDP, SP, SVP) and are located at the left, the center, and at the right of this dimension.

So far there are no implemented convergence diagnostics in the software package¹⁴ apart from trace plots allowing to do visual checks. Given the large number of parameters this is a unfeasible way. Note that one has to check for each parameter as joint convergence of all chains is necessary (Gill, 2008).¹⁵ A small subset of trace plots is presented here and shows clearly that these chains have converged.

¹⁴The model can be estimated in R using the \texttt{pscl} package (Jackman, 2008a).
¹⁵A personal discussion with Simon Jackman at the EITM 2008 (WashU) also provided the same result; one has to inspect all trace plots.
To estimate the model, 2’500’000 iterations of the Gibbs-Sampler were run, the first 100’000 thrown away (burn-in) and the remaining 2’400’000 were thinned by only using every 1000th draw.  

**Figure 6:** Trace Plots for Three Legislators (2nd Dimension)

![Trace Plots](image)

*Note: See note at figure 6.*

It has to be admitted that convergence cannot be guaranteed for all parameters, nor were all trace plots inspected. Again, inspecting all trace plots requires to look at 325 (unidimensional model) or 765 (two-dimensional model) parameters. But to give some confidence to this model, one can compare it with other models. Jackman (2001) presents a model with reported 1072 parameters and settles with 1.5 million iterations.

Another way of approaching this issue is to estimate the model with different starting values. If one uses different starting values and receives eventually the same posteriors, it is very likely that the chains have converged. This is of course not a perfect test, but feasible. The starting values are generated by a command of the `pscl` package and stored as an object. It is possible to manipulate elements of this object and thereby changing the initial values. The ideal points of the legislators are set to 0 as a default. I assigned random numbers (mean 0 and standard deviation 1) to the initial values.

---

16Running the two-dimensional model on a **MacBook Pro** (2.33 GHz Intel Core Duo) takes up to 6.5 hours.
Figure 7: Different Starting Values, Same Posterior Means

Note: First plot for the model with default settings for the initial values of the chain. Second plot for the same model, but with initial values which were assigned randomly.

Subsequently, the posterior means of the legislators ideal points can be compared from both models. Both plots of the ideal points yield the same impression; both plots are in figure 8. When different starting values lead to the same posterior means, it can be taken as an indicator for convergence.
8.3 Code for the Ideal Point Estimation

```r
library(pscl)
library(foreign)

data.ip <- read.dta("/Users/Lucas/Documents/Uni Bern/Project Strategic Behavior/Data/Roll_Call_Matrix_Gesamt.dta")
attach(data.ip)

# Preparing the data elements for the rollcall-command

data.ip<- data.ip[, -189]  # prename1
data.ip<- data.ip[, -187]  # canton
data.ip<- data.ip[, -186]  # prename2
ip.names <- as.matrix(data.ip[, 185], 1)
party <- matrix(data.ip[, 186], length(data.ip[, 186]), 1)
colnames(party) <- "party"
data.ip <- data.ip[, -186]
data.ip <- data.ip[, -185]
data.ip <- data.ip[, -1]

# Generating a roll-call element

rc.ip <- rollcall(data.ip, yea=1, nay=0, missing=9, notInLegis=4,
party.names=ip.names, legis.data=data.frame(party=party),
vote.names=names(data.ip))

# Unidimensional Model

### Unidimensional Model

test_d1 <- ideal(rc.ip,
d=1,
meanzero=TRUE,

```

---

**Recovering Sophisticated Voting**

8.3 Code for the Ideal Point Estimation

# Strategic Voting in the Lower House in Switzerland

# Lucas Leemann

# leemann@ipw.unibe.ch

```

---

```
store.item=TRUE,
maxiter=2000000,
burnin=1e5,
thin=1000)

# Inspecting the results
summary(test_d1)
summary(test_d1,include.beta=TRUE)
plot(test_d1)

# Generating the predicted votes to determine cpc
phat <- predict(test_d1)
# 93.54% correctly predicted cases

# Looking at a couple of trace plots
tracex(test_d1,"Theiler", conf.int=0.95)
tracex(test_d1,"Zisyadis", conf.int=0.95)
tracex(test_d1,"Darbellay", conf.int=0.95)
tracex(test_d1,"Baader", conf.int=0.95)
tracex(test_d1,d=1,legis=c("Leuthard","Cavalli","Theiler"), showAll=TRUE)

####################################################
### Two dimensional model
# (constraints: 2 priors on votes; 2'000'000 iterations)

# Constraints for two bills

c2 <- constrain.items(rc.ip,
x=list("vote_behav33"=c(5,0),
"vote_behav7"=c(0,5)),
d=2)

# Alternativ starting values for convergence check

c2.new <- c2
set.seed(421981)
for (i in 1:2){
  for (j in 1:207){
    c2.new$x[p][j,i]<-rnorm(1,1,1)
  }
}
# Estimating the model (default starting values)

test_d2_cons_def <- ideal(rc.ip, d=2, priors=c2, startvals=c2.new,
store.item=TRUE, maxiter=2000000, burnin=2e5, thin=1000)

# Inspecting the results

plot.ideal(test_d2_cons_def)
summary(test_d2_cons_def)
summary(test_d2_cons,include.beta=TRUE)

# generating predicted votes to determine the cpc

phat_c_2_def <- predict(test_d2_cons_def)
phat_c_2_def
# 97.04 overall correctly predicted

# Trace plots for the first and the second dimension

tracex(test_d2_cons_def,d=1,
legis=c("Leutenegger","Cavalli","Moergeli"),
showAll=TRUE)

tracex(test_d2_cons_def,d=2,
legis=c("Leutenegger","Cavalli","Moergeli"),
showAll=TRUE)
8.4 Code for the Proposed Algorithm

*******************************************************************************
#
####### Strategic Voting in the Lower House in Switzerland #############
#
*******************************************************************************
#
########################## Lucas Leemann ####################################
####################### (leemann@ipw.unibe.ch) #############################
#
*******************************************************************************

# Calling necessary packages

library(pscl)
library(foreign)

# reading in the ideal points
leg.ip<- read.dta("/Users/Lucas/Documents/Uni Bern/Project Strategic Behavior/Data/d2_coord.dta")
attach(leg.ip)

# reading in vote data
vote.data.ip<- read.dta("/Users/Lucas/Documents/Uni Bern/Project Strategic Behavior/Data/Roll_Call_Matrix_allvotes.dta")

*******************************************************************************

### OPTIMIZE FOR TWO-DIMENSIONAL SPATIAL MODEL

# defining a party which will be excluded - here "FDP"
# Note: Only these three letters have to be changed, everything else
# is done automatically

party<="FDP"
partynames <- vote.data.ip[,1286]
num.find <- which(partynames==party)

# defining the legislators of the chosen party
from.mp <- num.find[1]
to.mp <- length(num.find)+from.mp-1
from.mp # First MP of the excluded party
to.mp # Last MP of the excluded party
# Several data matrices which will be filled

dim1 <- leg.ip[-(from.mp:to.mp),2]
dim2 <- leg.ip[-(from.mp:to.mp),3]

brick1 <- rep(NA,2000)
brick2 <- rep(NA,2000)
brick3 <- rep(NA,2000)
brick4 <- rep(NA,2000)

result.sink <- cbind(brick1, brick2, brick3, brick4)
resid.sink <- cbind(brick1, brick2, brick3, brick4, brick2, brick3, brick4, brick2, brick3, brick2, brick3, brick4, brick2, brick3, brick4, brick2, brick3, brick2, brick3, brick2, brick3, brick4, brick2, brick3, brick4, brick2, brick3, brick2)
unity.sink <- rep(NA,2000)
t.sink <- rep(NA,2000)

indicator <- rep(NA,1282)
scanner<-rep(NA,1282)

# Defining a subset of votes, which are not excluded
# Exclusion is due to: all members of a party abstain (algorithm would brake down)
# or only one member of the party votes in a given vote (the start is at i=2 b/c
# the second column of the vote.data.ip matrix is the first vote)

for (i in 2:1282){
  alarm<-which(vote.data.ip[(from.mp:to.mp),i]==9)
  ifelse(length(alarm)==length(num.find),scanner[i]<-0,scanner[i]<-1)
}

reasonable.cases<-which(scanner2>0)

# Actual loop starts here; loops over all votes in the data set (here: 1281)
for (i in reasonable.cases){

# To estimate the coordinates of a y/n-vote, the chosen party is excluded
# Therefore a new matrix has to be built which only contains the other legislators
# Also legislators which abstained have to be excluded (optim cannot handle missings)

abst <- as.vector(vote.data.ip[-(from.mp:to.mp),i])
d2<-cbind(dim1,dim2,abst)

mi.vec<which(abst==9)
d2.new <-d2[-mi.vec,]
st.va.d2 <- rep (1,4)

# creating data and starting values

x1 <- d2[-mi.vec,1]
x2 <- d2[-mi.vec,2]
a <- d2[-mi.vec,3]
datad2<-cbind(x1,x2,a)
st.va.d2 <- rep (.5,4)

#This function estimates the coordinates based on the likelihood on page 20
spatial.d2<-function(theta,d) {
  y1<-theta[1]  # corresponds to the yea in d1
  y2<-theta[2]  # corresponds to the yea in d2
  n1<-theta[3]  # corresponds to the nay in d1
  n2<-theta[4]  # corresponds to the nay in d2
  logl<- sum(a*log(1/(1+exp(-(x1-n1)^2-(x2-n2)^2+(x1-y1)^2+(x2-y2)^2)))
  +(1-a)*log(1/1/(1+exp(-(x1-n1)^2-(x2-n2)^2+(x1-y1)^2+(x2-y2)^2)))))
  # Log-likelihood for Logit, pi.i corresponds to equation 17 (p.19)
  return(-logl)
}

# In case the MLE model 'blows up', the error is surpressed and the loop
# keeps on running

try(result.spd2 <- optim(st.va.d2, spatial.d2, d=datad2, method="BFGS"),
silent = TRUE)
result.sink[i,]<-result.spd2$par

### Generating predicted probabilities
# reading in the estimated coordinates of the n/y positions

\[
y_1 \leftarrow \text{result.sink}[i,1]\\
y_2 \leftarrow \text{result.sink}[i,2]\\
n_1 \leftarrow \text{result.sink}[i,3]\\
n_2 \leftarrow \text{result.sink}[i,4]
\]

# Only ideal points & votes for before excluded faction

\[
dim1.\text{res} \leftarrow \text{leg.ip}[\text{c(from.mp:to.mp)},2]\\
dim2.\text{res} \leftarrow \text{leg.ip}[\text{c(from.mp:to.mp)},3]\\
\text{abst} \leftarrow \text{as.vector(vote.data.ip}[\text{c(from.mp:to.mp),i}])\\
d2<\text{cbind(dim1.\text{res},dim2.\text{res},abst)}
\]

# Generating indicator to drop NA’s

\[
\text{mi.vec}<\text{which(abst==9)}\\
d2.\text{new} \leftarrow \text{d2}[\text{-mi.vec,}]
\]

# This is necessary, as otherwise I run into problems b/c d2.new # might be vector or matrix

\[
\text{ifelse(length(d2.\text{new})==3,}\\
x_1 \leftarrow \text{d2.\text{new}[1],}\\
x_1 \leftarrow \text{d2.\text{new}[,1]})\\
\text{ifelse(length(d2.\text{new})==3,}\\
x_2 \leftarrow \text{d2.\text{new}[2],}\\
x_2 \leftarrow \text{d2.\text{new}[,2]})
\]

# \pi.y Probability of a Yea vote

\[
\pi.y< 1/(1+\exp(-(x_1-n_1)^2-(x_2-n_2)^2+(x_1-y_1)^2+(x_2-y_2)^2))\\
t<\text{length(pi.y)}\\
t.\text{sink[i]} \leftarrow t
\]

# Expected vote decision

\[
\text{pi.y.hat} \leftarrow \text{rep (NA, t)}\\
\text{for (1 in 1:t)}{\\
\text{ifelse(pi.y[l]>0.5, pi.y.hat[l]<-1, pi.y.hat[l]<0) }\\
}
\text{pi.y.hat}
\]

# If unity is equal to 0 or to the number of voting members, it is unified behavior
unity <- sum(pi.y.hat)
unity.sink[i] <- unity

# Error following p. 20/21
resid.d2 <- rep(NA,t)
for (j in 1:t) {
  ifelse(pi.y.hat[j]==1, resid.d2[j] <- ifelse(abst[j]==1,0,pi.y[j]),
         resid.d2[j] <- ifelse(abst[j]==0,0,1-pi.y[j])
  )
}

# If this number is very low, almost all or even all (then it is 0) have voted the
# wrong way
q <- length(which(resid.d2!=0))
w <- t-q
indicator[i] <- w

# For the suspicious cases one might want to know how large the residuals
# actually were
ifelse(indicator[i]==0, for (k in 1:30) {
  resid.sink[i,k]<-resid.d2[k]
},
  resid.sink[i,]<-rep(0,30))

### Suspicious votes?

# (-1) is necessary as the first column of the matrix is no vote
likely.votes <- which(indicator==0) -1
likely.votes

core <- cbind(unity.sink,t.sink,resid.sink)

likely2 <- which(indicator==0)
matrix<-core[likely2,]
rep.num<-length(likely2)
party.vec<-rep(party,rep.num)
end.matrix.svp <-cbind(party.vec,likely.votes, matrix)
matrix.def <- rbind(end.matrix.gps, end.matrix.sp, end.matrix.cvp, end.matrix.fdp, end.matrix.svp)

write.table(matrix.def, file = "/Users/lucas/Documents/Uni Bern/Project Strategic Behavior/Data/error_matrix.txt", sep = ",", col.names = TRUE, qmethod = "double")