Information, polarization and accountability in democracy

Christian Schultz*

January 2005

Abstract

This paper considers a Representative Democracy with uni-dimensional politics. Parties have better information about the economy than voters. With a short term length, the government is accountable, but its incentive to be relected leads to a distorted policy. A longer term length gives less accountability and less distortion. The paper studies this trade off. For high uncertainty about the state of the world and low polarization of parties, voters whose vote depends on the state of the world, favors of short term lengths, otherwise long. The paper furthermore endogenizes party polarization in a citizen-candidate model. A short term length leads to more polarization, which is bad

*Institute of Economics, University of Copenhagen, Studiestraede 6, DK 1455 Copenhagen K, Denmark. e-mail: cs@econ.ku.dk

†This paper was initiated while I visited CES-ifo in the fall 2001. I thank CES-ifo for its hospitality. I thank Torsten Persson, Jean Tirole and Jörgen Weibull as well as seminar audiences at Carlos III, Copenhagen, Stockholm and Yale Universities for very helpful comments.
for voters. Still it is true that if the uncertainty is sufficiently high, swing voters prefer a short term length.

Keywords: Accountability, voting, information, democracy, polarization. JEL: H1, H7, K4

1 Introduction

As Downs (1957) pointed out, the electorate at large has insufficient incentives to become informed about complicated issues in politics, the functioning of the economy etc. Politicians on the other hand are briefed by experts and bureaucrats and it is their job to gather information and take decisions. This asymmetry of information speaks in favor of delegating decision making from the electorate to elected politicians. Such delegation, however, raises problems of congruence of preferences and accountability becomes important. If voters are discontent with the elected politicians, they can vote them out in the next election. On the other hand, governments want to be relected, and this may give them incentives to hide information and distort policies in order to manipulate voters beliefs about the true state of the economy. If a part of the electorate are swing voters, whose vote depends on the state of the economy, governments may want to manipulate voters’ beliefs. A left wing government may be reluctant to cut government expenditures, even though this would be rational under full information, since this may convince voters that rightist policies are better after all so they vote rightist. A short term length improves accountability, but makes reelection a more important concern for governments. This paper studies the trade off involved here.

We consider a society with two parties motivated by ideology and power. When parties’ ideologies are polarized, voters are faced with a non-trivial
choice between left and right. This influences the value of accountability for
the voters. If parties are polarized, one cannot just elect a “better” politician;
as a left wing government will take over if a right wing government is voted
out of office. The paper provides a model with two policy periods. Voters are
uninformed about the state of the economy. There are two types of voters,
partisans, who always vote for the same party, and swing voters, whose vote
depends on the state of the economy.

With a long term length, the government is elected for once at the start
of the first policy period. With a short term length, elections occur at the
start of both policy periods, and the governing party is accountable: It will
be voted out after period one, if the swing voters are not satisfied. Voters are
prospective, they vote for the party they believe chooses a better policy for
them in the following period. Hence they will vote out a right wing party if
they believe that the state of the world favors left wing policies. For instance,
if a voter believes that the government sector is very efficient in providing
the benefits of the welfare state, medicare, schools etc. and that taxation
is not particularly distortory she will want to vote liberal, democrat or
social-democrat, rather than conservative, whereas the opposite beliefs gives
support for the conservatives. This is an advantage of a short term length.
However, there is a flip side: The governing party is interested in reelection.
Since it has information, the voters have not, its policy is a signal about the
state of the world. It may therefore distort its policy in order to manipulate
the beliefs of the electorate and become more popular. To be convincing, a
right wing party has choose an even more rightist policy in some states of
the world, in order to signal to voters that the state of the world really calls
for right wing policies. We show that high uncertainty about the state of the
world tends to make short term lengths preferable for swing voters, since the
option value associated with the possibility of electing another government becomes high. In contrast, high polarization of the parties tends to make long term lengths better. The reason is that the policy distortion associated with short term lengths becomes very large when polarization is high.

Polarization of parties depends on the incentives facing parties and thus on the term length. While there is no generally accepted theory of party formation, Besley and Coate’s (1997) citizen candidate model has attracted considerable attention. In the later part of the paper we endogeneize polarization by embedding the model in a citizen-candidate model. Such models are well-known to have many equilibria and we focus on the least polarized two-candidate equilibrium. We show that a short term length leads to more polarization than a long, and this makes a short term length less attractive than when polarization is exogeneous. However, it is still the case that for sufficiently high uncertainty swing voters prefer a short term length.

Clearly, voters receive information from many sources: newspapers, lobbyists, business, organized labor etc. These many and varied sources of information alleviates the asymmetric information problem. Still empirical assessments show that a large fraction of the electorate typically is poorly informed: As Bartels (1996) puts it “The political ignorance of the American voter is one of the best-documented features of contemporary politics ...”. It is also true that much of the information in media is cheap talk, that “experts” often contradict each other and that the different interest groups provide conflicting information.

The trade off between term lengths has been considered before in the literature. Maskin and Tirole (2001) consider a model, where the salient issue is binary. The electorate’s preferred policy depends on the state of the world, which is only known by politicians. Politicians do not belong to parties and
may or may not have congruent preferences with the electorate at large. If the electorate decides to replace a politician in an election, the preferences of the new politician are chosen at random. So unlike in our framework, the policy space is binary and politicians are not ideologically different. Maskin and Tirole show that the better the electorate is informed about the state of the world, the more attractive is a rudimentary form for direct democracy where the median voter just chooses policy. The choice between representative democracy, where the chosen politician can be replaced, and "juridical power" where there is a very long term length depends on how eager the politician is to be reelected. Unlike our analysis this is an exogenous feature. Like Maskin and Tirole, we also briefly consider the rudimentary Direct Democracy, more a benchmark case, than because of its relevance. It appears hard to believe that the electorate will not find ways of getting information under this institution. In our model such Direct Democracy is optimal when polarization is high and uncertainty low.

Alesina and Tabellini (2003) build on Holmstrom’s (1999) career concern model. Bureaucrats seek to get good reputations for competency in order to increase future pay and career opportunities, politicians do it in order to enhance voters’ perception of his talent so that he be reelected. The different motivations give rise to different effort levels. Alesina and Tabellini then show (among other things) that politicians tend to be best for tasks which are non-technical. They also show that time-consistency problems tends to make bureaucrats more attractive. The driving force behind the results, the career concerns is different from the one explored in the present paper.

Hanssen (2002) studies the strategic choice of mode of governance by an incumbent government. An incumbent government may wish to delegate decision authority to an independent agency in order to raise the cost of
changing a particular policy for a future government. Hansen shows that an incumbent government is more likely to establish an independent judiciary if there is higher probability that it looses the next election and the polarization of politics is larger. In this way the current government ensures that policy also in the future is guided by preferences close to its own. These results are related to the results of Persson and Svensson (1989) and Tabellini and Alesina (1990) on the strategic choice of debt by incumbent governments facing possible defeat in coming elections. Hansson shows that his prediction is confirmed on data from American states. Hanssen’s aim is different from ours. He considers strategic choice of institutions by an incumbent government and not which term lengths are optimal from the point of view of voters.

Following the lead of Barro (1973) and Ferejohn (1986) a large literature considers rent seeking politicians and the disciplining effect of elections in representative democracy, see and Persson and Tabellini (2000) for a recent overview. In this vein Aghinon et al (2002) investigate a model where an elected politician has superior information and can promote and implement reforms. The politician may be good and promote reform but may also be bad and seek to grab rents. Whatever he promotes has to pass a referendum, so a (super) majority can block it. This gives a tradeoff at the constitutional stage, a smaller blocking majority makes it the more difficult for a bad politician to grab rents, but also more difficult to pass reforms. Aghinon et al then study the optimal choice of the size of the supermajority.

The importance of information and polarization for the functioning of representative democracy has been the subject of several papers. Schultz (1996) considers a model where parties are better informed about the state of the world than the electorate and commit to policies before an upcoming
election. If parties’ preferences are sufficiently polarized, the electoral competition will lead to inefficient equilibria, since the parties’ policies do not reflect the state of the world. In Schultz (2002) the electorate is supposed to have inferior information about the state of the world and the preferences of the parties. When parties cannot commit to policies before an election this leads to policy distortions as the incumbent party seeks to manipulate the beliefs of the electorate. The distortion depends on the relative importance of the two kinds of uncertainty: Uncertainty about preferences leads to a bias towards more centrist policies, while uncertainty about the state of the world leads to more extreme policies. Cuikerman and Tommasi (1998) consider a model with two kinds of uncertainty where parties commit to policies before an election. Under some circumstances this leads to situations where a left party most credibly can implement a rightist policy. Harrington (1993) and Letterie and Swank (1998) study a slightly different issue. In their papers the government is unsure about the state of the economy. The policies chosen then act as signals for the governments beliefs.

Kessler (2000) studies the relative merits of representative and direct democracy when voters and politicians initially are uniformed about the state of the world. She studies a one-period citizen-candidate model a la Besley-Coate (1997). In Kessler’s model citizens can exert costly effort in order to become informed. Since ordinary voters have zero chance of being pivotal, they do not invest in information acquisition and the policy decision is uninformed under direct democracy. In contrast, the elected polician in representative democracy has incentives to gather information (for sufficiently low costs) and the policy choice will be informed. In this sense the model endogenizes and explains the asymmetric information Downs (1957) focussed on and I just assume in the present paper. Kessler also assumes that there is
uncertainty about candidates’ preferences. This induces a cost of representative democracy for the voters as the policy choice will be unpredictable. The optimal mode of governance then depends on which kind of uncertainty is the larger. Contrary to me, Kessler does not focus on polarization of parties and the policy distortion it leads to in representative democracies. Furthermore, she does not consider the effects of accountability per se, long term lengths are not considered. The cost of representative democracy in her setting exclusively stems from the uncertainty about the politicians’ preferences.

The remainder of the paper is organized as follows. Section 2 presents the basic model. Direct democracy is treated in section 3. Sections 4 and 5 discuss representative democracy with long and short term length respectively. The optimal mode of governance is derived in section 6. Some extensions and conclusions are provided in section 7.

2 Basics

There are two periods, 1 and 2, where society has to choose a policy $x$, which can be ordered on a left-right dimension, $x \in R$. There are two parties: a left party $L$ and a right party $R$.

There are a continuum of possible states of the world, $s$, uniformly distributed on $[-\sigma, \sigma]$, where $\sigma > 0$. The state of the world is the same in the two periods. Voters are not informed about the true state of the world, but they know the distribution of $s$.

Voters all have quadratic utility functions on the policy chosen, $x$, and have different bliss points. All voters prefer a higher policy, the higher is the state of the world, $s$. There are two types of voters, partisans and swing voters respectively. While a partisan’s vote always will go to the same party,
swing voters’ vote will depend on the state of the world.

Partisan $a$’s bliss point is $a + s$, while swing voter $a$’s bliss point is $a + \phi s$, where $\phi > 1$. The fraction of swing voters is $\mu$. There are a continuum of voters, half of the voters of each type, partisan or swing, have $a \geq 0$, the other half $a \leq 0$, so the median voter is voter 0 (with $a = 0$).

If $x_1$ is chosen in period 1 and $x_2$ in period 2, and the state is $s$, the total utility for a partisan $a$ from the two periods is

$$-(x_1 - (a + s))^2 - \delta (x_2 - (a + s))^2,$$

where the discount factor, $\delta$, fulfills $0 < \delta \leq 1$. A swing voter’s utility is the same except $s$ is multiplied with $\phi$.

As is clear from the utility function, the size of $\sigma$ determines whether uncertainty about the state is important or not. If $\sigma$ is very small, uncertainty is small, and it is not so important to tailor the policy correctly to the state. The opposite holds if $\sigma$ is large.

Each of the parties, $L$ or $R$, is headed by a political leader, who chooses the policy of the party. The leader is a partisan and he is purely ideological and only interested in policy\textsuperscript{1}. The per period utility of the leader of party $R$ is

$$-(x - (r + s))^2$$

where $r + s$ is the bliss point of the leader in state $s$. Just like the voters the party is interested in the sum of discounted utility from the two periods.

As is clear, we assume that the leader is a partisan and not a swing voter. Although one could also analyze the other case, it appears most realistic

\textsuperscript{1}At the expense of longer formulas one could easily add a term reflecting the benefit of power to the utility function, which was positive when the party is in office and zero otherwise.
that parties are not lead by the part of the electorate, who swings the most. If anything members of parties are probably more influenced by political ideology and philosophy than the average citizen and members of parties may also have a longer run perspective. Finally, the platform of a party is usually determined in a rather long process involving conventions. All taken together it appears reasonable to assume that in our model, the party leadership is formed of partisan voters.

Party $L$ has a similar utility function, only difference is that $r$ is replaced with $l < r$. For simplicity, we consider the symmetric case where $l = -r$. Hence, $r$ is a measure of the polarization of parties. The larger is $r$, the more the bliss points of the two parties differ. In the first part of the paper, we take the polarization of the parties as exogenously given, while we endogenize it in subsequent sections.

We assume that voters know the preferences of the parties. Clearly, one could also hold the view that voters may have difficulties in learning the preferences of politicians - in particular if the party has not held power for years. However we will assume that understanding the economy - the state of the world - is the most complicated and important issue and focus on this. For a treatment in representative democracy of the case where voters are uncertain both about the state of the world and the preferences of the parties, see Schultz (2002).

Contrary to the voters, the parties are informed about the state of the world, $s$. As discussed in the Introduction, parties are informed from experts, the governing party has direct access to the bureaucracy, the leaders of the parties are full time politicians whose job it is to gather the relevant information and take decisions. The electorate, on the other hand does not have as strong incentives to gather information.
We focus on the effect of the term length in Representative Democracy, and distinguish between Representative Democracy with a short and a long term length, $RDS$ and $RDL$ respectively. Under Representative Democracy with a long term length, the voters elect a party in the start of period 1. The party governs for both periods and chooses the policy in each period. Under Representative Democracy with a short term length a party is elected in the start of period 1. It chooses the first period policy, which is observed by the voters. A new election occurs in the start of period 2. The newly elected party then chooses policy for period 2. We also briefly comment on a very simple and rudimentary democratic institution: Direct Democracy, where the electorate in each period determines the policy, and we will assume that this implies that the median voter’s preferred policy is chosen. Admittedly, this is perhaps a naive representation of direct democracy. In practice, direct democracy comes in many forms, depending on rules for agenda setting and amendments. The one chosen here could be seen as the archetypical and basic form for direct democracy, where parties and elected politicians have no influence.

In the sequel we will find the expected utility from each mode of governance evaluated at date 0 for an arbitrary voter $a$. As will be clear, the relevant trade-offs are the same for all voters, regardless of their bliss point. They therefore all rank the different modes the same way and it is thus meaningful to speak of an optimal mode of governance.

3 A rudimentary Direct Democracy

In our rudimentary direct democracy, $DD$, voters in each period choose the policy preferred by the median voter without knowledge of the state of nature.
Clearly, this is an abstraction and probably has no bearing on the real world, one may question, whether the electorate would not find ways of providing information. Nevertheless it forms a good starting point for gaining intuition and for later comparisons, and we will therefore consider it. Since there are no informed parties under DD, it portrays the situation where the policy choice is made in the dark. The median (partisan) voter’s expected utility in a period from policy $x$ is

$$- \int_{-\sigma}^{\sigma} (x - s)^2 \frac{1}{2\sigma} ds.$$ 

The optimal policy for the median voter is therefore

$$x = 0.$$ 

and of course this is also the preferred policy of a median swing voter. From the point of view of a swing voter $a$, the expected utility at date 0 from DD is therefore

$$u^D_\phi = (1 + \delta) \int_{-\sigma}^{\sigma} - (0 - (a + \phi s))^2 \frac{1}{2\sigma} ds = - (1 + \delta) \left( \frac{\phi^2 \sigma^2}{3} + a^2 \right).$$ 

(1)

For a partisan voter the expression is the same with $\phi = 1$. We see that the more uncertainty about the state of the world given by the variance$^2$ - the less attractive is Direct Democracy. This effect is even stronger for a swing voter, whose utility is more sensitive to the state of the world.

\[4\] Representative Democracy: long term length

When the term length is long, voters elect a party before date 1, which governs for both periods. The leadership of the party needs not worry about

$^2$Recall that when $s$ is uniform on $[-\sigma, \sigma]$, then the variance equals $\frac{\sigma^2}{3}$. 

12
reelection and chooses its preferred policy, given knowledge of the state, regardless of the views of the voters. The median voter (whether partisan or swing), who does not know the state of the world, is indifferent between the parties, we will assume this implies that each party wins with probability 1/2. The expected utility for a swing voter \( a \) from RDL is therefore

\[
u_L^a = (1 + \delta) \int_{-\sigma}^{\sigma} \left(-\frac{1}{2}(-r + s - (a + \phi s))^2 - \frac{1}{2}(r + s - (a + \phi s))^2\right) \frac{1}{2\sigma} ds,
\]

which yields

\[
u_L^a = - (1 + \delta) \left(\frac{\sigma^2(\phi - 1)^2}{3} + r^2 + a^2\right).
\] (2)

The trade-off associated with a long term length is clearly reflected in these expression. The larger the degree of polarization is, the less attractive is RDL. The advantage associated with RDL is that the chosen policy reflects the state of the world, since it is known to the governing party. A partisan voter, with \( \phi = 1 \), finds that the party prefectly responds to the state of the world and the variance of the state is therefore inessential for this voter. This is not so for a swing voter. Although it is an improvement over Direct Democracy that the party adjusts policy to the state of the world, the governing party will do this too little. The expected utility of a swing voter is therefore decreasing in the variance of the state of the world.

5 Representative Democracy: short term length

With a short term length, the time line is as follows. In the start of period one an election is held. The winning party observes the state of the world and chooses first period policy. Voters observe this policy but it takes time
for the effects of the policy to materialize so its utility consequence is only learned after the second election is held. For instance, the welfare effects of expanding health care comes after some while or similarly the effects of a major tax reform takes time to spell out. Since voters know that the governing party is informed about the state, they may change their belief about the state of the world after observing the policy. In the start of period two a new election is held. The new winner chooses policy for period two.

An equilibrium under Representative Democracy with short term length is a Perfect Bayesian Equilibrium of the game just described. In each period the winning party chooses the policy which maximizes its expected utility given the state of the world, the other party’s strategy and the way voters form beliefs. Voters’ belief formation may affect the probability of winning the next election and parties therefore have an interest in inducing "good" beliefs. Voters vote sincerely given their beliefs. In the first period they observe the policy and form posterior beliefs about the state of the world using Bayes’ rule from the prior and the governing party’s strategy.

Whether voters can infer the true state of the world after period one or not depends on the governing party’s policy strategy, $x(s)$. If voters observe $x$ and know that $x$ is only chosen in state $s$, they can infer that the state is $s$. If $x$ is chosen in many states, they can only infer that it is one of those states. As it turns out the first period strategy will be strictly increasing over some intervals of states, here voters can perfectly infer the state, and we say that the strategy is separating in these intervals. If the strategy is separating on $[-\sigma, \sigma]$, we say the equilibrium is separating. If the policy is constant over an interval of states, we say that the strategy is pooling over this interval.

We find an equilibrium by solving the model backwards and first look at the second period. The second period is the last, so the parties have
no reelection concerns and will choose their bliss points $r + s$ and $-r + s$, respectively. If party $R$ wins its utility is therefore 0, while if $L$ wins it is $-(-r + s - (r + s))^2$. Party $R$’s discounted gain from winning the election is therefore $\delta 4r^2$. Party $L$’s gain from winning is the same. Voters understand that the parties will chose their bliss points in the second period. If voters’ updated beliefs are that the state is uniformly distributed on an interval $[s_l, s_h]$, swing voter $a$ prefers party $R$ if

$$\int_{s_l}^{s_h} \frac{1}{s_h - s_l} (r + s - (a + \phi s))^2 ds - \int_{s_l}^{s_h} \frac{1}{s_h - s_l} (-r + s - (a + \phi s))^2 ds > 0$$

or

$$s^e \equiv \frac{s_h + s_l}{2} > \frac{a}{\phi - 1}$$

where $s^e$ denotes the expected state. A partisan voter prefers party $R$ if $a > 0$ and party $L$ otherwise. This means that the election is determined by the median swing voter and party $R$ wins if the expected state is positive. We will assume that the incumbent wins the election if the voters are indifferent. This tie-breaking rule simplifies a few steps in the proofs below, but has no impact on the results.

Since all agents, voters as well as parties, prefer higher policies in higher states, it is intuitive that in any equilibrium, the first period incumbent will not choose policies which are decreasing in $s$. This is indeed true as Lemma 1 says,

**Lemma 1** In any Perfect Bayesian Equilibrium the first period government’s policy is non-decreasing in $s$.

**Proof.** All proofs are in the Appendix. □
Let $s^e(x)$ denote the voters’ expected state after observing $x$. Lemma 1 implies that if $x_2 > x_1$ both are equilibrium policies then $s^e(x_2) > s^e(x_1)$. For policies $y$, which are not equilibrium policies, expectations are not pinned down by Bayes’ rule. We will assume that the voters’ beliefs are reasonable in the sense that if $x$ is an equilibrium policy and voters observe $y > x$ then $s^e(y) \geq s^e(x)$.

Since the first period incumbent has a strictly positive incentive to win the election, it has an incentive to manipulate the voters’ beliefs, so that they will reflect it. If voters believe that the state is negative party $L$ will win the election. If the state is negative, but close to zero, and party $R$ is incumbent it has a strong incentive to act as if the state is positive so that it will be relected. This is done by choosing a higher policy, which the party finds optimal in higher states. This incentive prevents the existence of a completely separating equilibrium.

**Lemma 2** There is no completely separating equilibrium, where the voters learn $s$ for all $s \in [-\sigma, \sigma]$.

We thus have that in any equilibrium, there will be some $s'$s, where the first period incumbent chooses the same policy. The next Lemma shows that there is only one value of the policy $x$ for which pooling occurs and this occurs over an interval of states, outside this pooling interval, the equilibrium is separating.

**Lemma 3** If $s_1 < s_2$ and $x(s_1) = x(s_2)$ then $x(s) = x(s_1)$ for all $s \in [s_1, s_2]$.

Furthermore if $x(s) = x_1$ for all $s \in [s_1, s_1']$, where $s_1 \neq s_1'$ and $x(s) = x_2$ for all $s \in [s_2, s_2']$, where $s_2 \neq s_2'$, then $x_1 = x_2$.  

16
When the party chooses the pooling policy it inflicts a first period loss on itself in all states (but one) of the pooling interval. The party accepts this loss as there is a gain from pooling, namely that it is going to win the election.

If the pooling interval is long, the difference between the bliss point and the pooling policy is large for some states. This gives a high loss, which can only be justified for the party, if the gain from winning is sufficiently large. The next Lemma exactly shows that the length of the interval is determined by the gain from winning the election. Furthermore it shows that the policy of the pooling interval equals the party’s bliss point at the right (left) end of the interval for party $R$ ($L$).

**Lemma 4** Let $[s_1, s_2]$ be the pooling interval. Then
\[
x(s) = r + s_2 \text{ for } s \in [s_1, s_2] \text{ if } R \text{ is incumbent}
\]
\[
x(s) = -r + s_1 \text{ for } s \in [s_1, s_2] \text{ if } L \text{ is incumbent and}
\]
\[
s_2 = s_1 + 2r\sqrt{\delta}
\]

Outside the pooling interval, the equilibrium is separating. As party $R$ looses for $s < s_1$, it chooses its bliss point in these states. In states $s \geq s_2$, $r + s > r + s_2$, so the party wins the election even if it chooses its bliss point for these $s$. This explains why there is only one pooling interval. Outside the interval pooling is either too expensive or the party wins anyway if it chooses its bliss point, so there is no need to pool.

In the pooling interval party $R$ chooses a higher value of $x$ than its bliss point. The reason is that if it chooses a lower policy, voters will infer that the state is low and then they will want to elect party $L$ instead. In a similar manner party $L$ distorts its policy downwards in order to win.

The pooling interval is situated around zero, but it cannot be pinpointed
exactly. Party $R$ looses for $s < s_1$, where $s_1$ is the left end point of the pooling interval, so $s_1 \leq 0$ and it wins by choosing the pooling policy, so $\frac{s_1 + s_2}{2} \geq 0$, using $s_2 = s_1 + 2r\sqrt{\delta}$, this implies that $s_1 \geq -r\sqrt{\delta}$. When $R$ is incumbent, we therefore have that the left end point $s_1$ fulfills

$$-r\sqrt{\delta} \leq s_1 \leq 0,$$

while if $L$ is incumbent, the right end point $s_2$ fulfills

$$0 \leq s_2 \leq r\sqrt{\delta}.$$

Summing up, we now have

**Proposition 1** Assume party $R$ won the first election.

For any $s_1 \in \left[-r\sqrt{\delta}, 0\right]$ and $s_2 = s_1 + 2r\sqrt{\delta}$ there exists an equilibrium, where its policy $x(s)$ fulfills

$$x(s) = \begin{cases} 
  r + s & \text{for } s < s_1 \\
  r + s_2 & \text{for } s_1 \leq s \leq s_2 \\
  r + s & \text{for } s_2 < s 
\end{cases}$$

These are the only equilibria, when $R$ is incumbent.

If $L$ won the first election, then for any $s_1 \in \left[-2r\sqrt{\delta}, -r\sqrt{\delta}\right]$ and $s_2 = s_1 + 2r\sqrt{\delta}$ there exists an equilibrium, where its policy $x(s)$ fulfills

$$x(s) = \begin{cases} 
  -r + s & \text{for } s < s_1 \\
  -r + s_1 & \text{for } s_1 \leq s \leq s_2 \\
  -r + s & \text{for } s_2 < s 
\end{cases}$$

These are the only equilibria, when $L$ is incumbent.

In an equilibrium party $R$ distorts the policy upwards in the pooling interval in order to make the voters’ expected state is so high that the median
voter votes for party $R$. For very high states this is unnecessary, for low states, the needed distortion in policy is so large that party $R$ does not find it worthwhile, it chooses its bliss point and looses the election.

From the point of view of the median voter, the best equilibrium, when $R$ is incumbent, is when the pooling interval has left endpoint $s_1 = 0$. In this equilibrium, there is a first period distortion, but the voters learn whether $s \geq 0$, so in the second period, the elected government is the most preferred by the median voter given the actual state. From the point of view of party $R$, however, the best equilibrium is the one where $s_1 = -r\sqrt{\delta}$, as it wins the second election in a larger set of states. Figure 1 illustrates the equilibrium, when $R$ is the incumbent.

6 Comparing term lengths

In this section we consider which term length is better for partisan and swing voters by comparing the expected utilities for the voters obtained under two different term lengths. First consider Representative Democracy with a short term length, $RDS$. We focus on the equilibrium, which is best for the median voter, where the pooling interval $[s_1, s_2]$ has $s_1 = 0$ if $R$ wins the first election and $s_2 = 0$ if $L$ wins the first election, i.e. an equilibrium like the one illustrated in Figure 1. This equilibrium represent $RDS$ at its best from the point of view of the median voter. As will be clear, this has no qualitative importance for the results.

At the first election, voters have no further information about the state so each party wins with probability 1/2. At the second election, voters elect party $L$ if $s < 0$ and party $R$ if $s > 0$, and the incumbent - whoever it is - if
Figure 1: Equilibrium policy when $R$ is incumbent
The expected utility in the second period for swing voter \(a\) evaluated at date 0 is therefore
\[
\delta \left( \int_{-\sigma}^{0} (-r + s - (a + \phi s))^2 \frac{1}{2\sigma} ds - \int_{0}^{\sigma} (r + s - (a + \phi s))^2 \frac{1}{2\sigma} ds \right)
\]
\[
= -\delta \left( \frac{\sigma^2}{3} (\phi - 1)^2 + r^2 + a^2 - (\phi - 1) r \sigma \right)
\]
Comparing with RDL where the second period utility is given by (see 2)
\[
-\delta \left( \frac{\sigma^2}{3} (\phi - 1)^2 + r^2 + a^2 \right)
\]
we see that the second period expected utility in RDS is higher than in RDL for all swing voters, and the difference is
\[
\delta (\phi - 1) r \sigma \quad (3)
\]
This is the gain from accountability and arises from the fact that the voters can elect the government they prefer, when they learn the state under RDS, while this is not possible under RDL. We see that the higher is \(r\) the degree of polarization, \(\sigma\) the uncertainty about the state of the world and \(\phi\) - the "degree of swing", the more valuable is accountability. The partisan voters’ preferred government does not depend on the state, hence from their point of view, there is no difference in the second period utility whether there is a short or a long term length.

When the term length is short, the policy will be distorted in the pooling interval. Outside the pooling interval, the policy will be the same regardless of the term length. Evaluated at date zero the difference in expected utility for swing voter \(a\) from RDL and RDS is therefore

\[s = 0^3.\]
\[
\frac{1}{2} \int_{-x}^{0} -(-r + s - (a + \phi s))^2 \frac{1}{2\sigma} ds + \frac{1}{2} \int_{0}^{x} -(r + s - (a + \phi s))^2 \frac{1}{2\sigma} ds \\
- \left( \frac{1}{2} \int_{-x}^{0} -(-r + (-x) - (a + \phi s))^2 \frac{1}{2\sigma} ds + \frac{1}{2} \int_{0}^{x} -(r + x - (a + \phi s))^2 \frac{1}{2\sigma} ds \right)
\]
where \( x = 2r\sqrt{\delta} \). We therefore have that the utility difference equals

\[
\left( \frac{2}{3} \right) \left( 3 - 2(\phi - 2) \sqrt{\delta} \right) \frac{\delta r^3}{\sigma} \tag{4}
\]

Notice that both the size of the gain and the loss is independent of the identity of voter \( a \). For a partisan voter, where \( \phi = 1 \), the expression is positive. First period utility is unambiguously higher under \( RDL \), since the policy is not distorted under \( RDL \) but is under \( RDS \). For a swing voter, the distortion may actually be beneficial if \( \phi \) is sufficiently large, \( \phi > 2 + \frac{3}{2\sqrt{3}} \). Using (3) and (4) we now directly have

**Proposition 2** \( RDL \) is always better than \( RDS \) for partisan voters. For swing voters \( RDS \) is better than \( RDL \) iff the gain from accountability exceeds the loss from the first period distortion under \( RDS \). This is the case when

\[
\frac{\sigma^2}{3} > \frac{6 - 4(\phi - 2) \sqrt{\delta}}{9(\phi - 1)} r^2 \tag{5}
\]

Large uncertainty makes a short term length preferable for the swing voters. For swing voters, who experience a loss under \( RDS \) due to the distortion, i.e. voters with \( \phi < 2 + \frac{3}{2\sqrt{3}} \), larger polarization makes \( RDS \) less attractive. Larger polarization actually both increases the gain and the loss, but the effect on the loss dominates.

The relative merits of our rudimentary direct democracy, \( DD \), and \( RDL \) are easily found using (1) and (2), which directly give
Proposition 3  

\[ \frac{\sigma^2}{3} < \frac{1}{2\phi - 1}r^2 \]  

Everybody agrees that if the variance of the state of nature is low, then \( DD \) is better than \( RDL \). The reason is simple, not much is gained by delegating to an informed party and the voters avoid delegating to a party with preferences, which are more different to their own than the median voters’ are.

If

\[ \frac{1}{2\phi - 1} < \frac{6 - 4(\phi - 2)\sqrt{\delta}}{9(\phi - 1)} \]

which can be rewritten

\[ \phi < \frac{1}{8\sqrt{\delta}} \left( 10\sqrt{\delta} + \frac{3}{2}\sqrt{16\delta + 24\sqrt{\delta} + 1} + \frac{3}{2} \right) \]  

i.e. that swing voters are not too sensitive to the state of the world, we have

Corollary 1  

If (7) is fulfilled, then swing voters prefer

\[ DD \] if \[ \frac{\sigma^2}{3} < \frac{1}{2\phi - 1}r^2 \]

\[ RDL \] if \[ \frac{1}{2\phi - 1}r^2 < \frac{\sigma^2}{3} < \frac{6-4(\phi - 2)\sqrt{\delta}}{9(\phi - 1)}r^2 \]

\[ RDS \] if \[ \frac{6-4(\phi - 2)\sqrt{\delta}}{9(\phi - 1)}r^2 < \frac{\sigma^2}{3} \]

The Corollary is illustrated in Figure 1

7  

Endogeneous polarization

So far we have treated the ideologies of the parties as exogeneous. Although convenient the assumption is not innocous. While there is no well established theory for party formation, the citizen-candidate model of Besley-Coate (1997) has attracted considerable attention. In this section we will
embed our model in the citizen-candidate framework and in this way endogenize the preferences of the parties.

The preferences of the parties, will be the preferences of citizens which decide to run for elections. This is associated with a cost $c$. The decision to enter the electoral race is a profound one, and we will treat it as a long run decision. We will therefore assume that the time line is as follows. First in period 0, citizens decide whether to become a candidate and participate in the electoral race or not. Regardless of whether we consider a short or a long term length only those citizens who decided to become candidates in period zero are able to run at any of the elections. It is hence not possible for a citizen to enter the electoral race between periods one and two. One could, of course, also conceive of the possibility where citizens can decide to enter the electoral race at any time, but we will leave this for later research.

In principle both swing citizens and partisan citizens can decide to become candidates. In line with the analysis of the previous sections, we will focus at the case where partisan citizens can decide to run.

As is well known from the work of Besley and Coate there are multiple equilibria in citizen-candidate models. This will also be the case here. We will focus the analysis on the symmetric two-candidate equilibrium with least polarization. Hence for a given cost, $c$, of becoming a candidate, we will analyze the symmetric two-candidate equilibrium with least polarization. From the point of view of any voter, this is the best two-candidate equilibrium.

Consider first long term length, $RDL$. In a symmetric two candidate equilibrium, candidates are $-r$ and $r$. When citizen $r$ decides whether to enter the electoral race or not, he realizes that policy will be chosen by $-r$
if \( r \) does not enter the race. The utility to \( r \) then becomes

\[
(1 + \delta) \int_{-\sigma}^{\sigma} -(-r + s - (r + s))^2 \frac{1}{2\sigma} ds = -4 (1 + \delta) r^2
\]

If \( r \) chooses to run, he wins the election with probability one half in which case his utility is 0. Taking into consideration the entry cost, \( c \), the expected utility from entering the race for candidate \( r \) is therefore

\[
\frac{1}{2} 0 + \frac{1}{2} (-4 (1 + \delta) r^2) - c
\]

Entering as a candidate is best if the gains from entering the race are larger than the cost, i.e. if

\[
2 (1 + \delta) r^2 \geq c
\]

We therefore have that in the least polarized symmetric two-candidate equilibrium under RDL the bliss point of candidate \( r \) equals

\[
r_{L}(c) = \sqrt{\frac{c}{2 (1 + \delta)}}
\]

In line with previous results in citizen-candidate models, we see that polarization is increasing in the cost of entering the electoral race.

Then consider a short term length, RDS, and look again at \( r \), given \( -r \) has chosen to enter. If \( r \) does not enter \( -r \) will decide policy in the two coming periods. As she is the only candidate, she does not have to worry about reelection in the second period and will just choose her bliss point in both periods. The utility to \( r \) in this case will be

\[
(1 + \delta) \int_{-\sigma}^{\sigma} -(-r + s - (r + s))^2 \frac{1}{2\sigma} ds = -4 (1 + \delta) r^2
\]

If, on the other hand, \( r \) enters the race, he will win with probability \( \frac{1}{2} \) in the first period and he will win in the second period if and only if \( s \geq 0 \).

\[\text{We are as in the previous section focussing attention on the equilibrum, where the pooling interval ends in 0.}\]
expected utility from entering as a candidate is therefore

\[\frac{1}{2} \left( \int_{-\sigma}^{0} (r + s - (r + s))^2 \frac{1}{2\sigma} ds + \int_{0}^{x} (r + x - (r + s))^2 \frac{1}{2\sigma} ds \right) \]

\[+ \frac{1}{2} \left( \int_{-\sigma}^{-x} (-r + s - (r + s))^2 \frac{1}{2\sigma} ds + \int_{-x}^{0} (-r + (-x) - (r + s))^2 \frac{1}{2\sigma} ds \right) \]

\[+ \delta \left( \int_{-\sigma}^{0} (-r + s - (r + s))^2 \frac{1}{2\sigma} ds + \int_{0}^{\sigma} -(r + s - (r + s))^2 \frac{1}{2\sigma} ds \right) - c\]

where \( x = 2r\sqrt{\delta} \). This equals

\[-2 \left( 1 + \delta + \left( 1 + \frac{2}{3}\delta \right) \frac{r}{\sigma} \right) r^2 - c\]

Again entering is best if the gains from entering are larger than the cost, i.e. if

\[-2 \left( (1 + \delta) + \left( 1 + \frac{2}{3}\sqrt{\delta} \right) \frac{r}{\sigma} \right) r^2 - \left( -4 (1 + \delta) r^2 \right) \geq c\]

We therefore have that the least polarized symmetric two-candidate equilibrium under \( RDS \) is given by \( r_S(c) \) which solves

\[\left( 2 (1 + \delta) - 2 \left( 1 + \frac{2}{3}\sqrt{\delta} \right) \frac{r}{\sigma} \right) r^2 = c\]  \hspace{1cm} (11)

The left hand side -the gain from entering - is 0 for \( r = 0 \), positive for small \( r \), has a unique maximum for positive \( r \), and tends to \(-\infty\) as \( r \to \infty \). The maximal value of the left hand side equals \( 8 \frac{(3\delta+2\delta^2)\delta+1)^3}{3} \frac{\sigma^2}{3} \). It follows that

the equation only has solutions for \( c \in \left[ 0, \frac{8((3\delta+2\delta^2)(\delta+1)^3}{3(3\delta+2\delta^2)^3} \frac{\sigma^2}{3} \right] \). For higher values of \( c \), the gains from entering the electoral race is not sufficiently large, and a symmetric two-candidate equilibrium does not exist. In the relevant interval, the solution \( r_S(c) \) is increasing in \( c \) as is easily shown by implicit differentiation.
Comparing (8) and (11) we see that for all $r$, the gain from entering the race is higher under RDL.

It follows that for a given $c$, $r_S(c) > r_L(c)$, so we have

**Proposition 4** The least polarized symmetric two-candidate equilibrium is less polarized under RDL than under RDS.

Partisan voters prefer RDL to RDS for given degree of polarization, cf Proposition 2, due to the policy distortions associated with a short term length. The increased polarization under RDS, when polarization is endogenous, just makes RDS an even worse option for these voters. We therefore directly have

**Corollary 2** Partisan voters all prefer RDL to RDS when polarization is endogeneous.

For swing voters, the situation is not so simple. Under some conditions they prefer a short term length for a given degree of polarization, since it gives the option to exchange government to a type, which the swing voters find better match the state. However, the swing voters also dislike an increase in polarization. Since the formulas become a bit longwinded, we now focus on the case where $\delta = 1$ and $\phi = 2$. We then have that (11) simplifies to

$$r^2_S \left(4 - \frac{10 r_S}{3 \sigma} \right) = c$$

(12)

and solutions exist for $c \in \left[0, \frac{64 \sigma^2}{25 \delta} \right]$. Inserting into the relevant indirect utility functions gives (see appendix)
Proposition 5 Consider the case where \( \delta = 1 \) and \( \phi = 2 \). and \( c \in \left[ 0, \frac{64 \sigma^2}{3} \right] \). When polarization is endogeneous, swing voters prefer RDS to RDL if

\[
c \leq \sigma^2 \left( \frac{12}{11} - \frac{10}{121} \sqrt{33} \right)
\]

and they prefer RDL to RDS if

\[
\sigma^2 \left( \frac{12}{11} - \frac{10}{121} \sqrt{33} \right) < c
\]

We see that swing voter’s preferences depend on the variance and the cost of entering as a candidate. High variance tends to make the swing voters prefer a short term length. Even though a short term length leads to more polarization, the advantage, that the government can be exchanged if the state of the world favors this, dominates. If on the other hand the entry cost in politics is high relative to the variance of state, then even the swing voters prefer a long term length.

Since all partisan voters agree on a long term length is preferable, a majority for sure prefer a long term length if the swing voters also prefer this. If the swing voters prefer at short term length, it depends on who forms the majority, the partisans or the swing voters.

8 Concluding remarks

While a short term length has the advantage that voters can replace a government they are discontent with, it has the disadvantage that governments incentive to be relected make them distort policies when the electorate is poorly informed about the state of the economy. A long term lenght has less accountability and less distortion. When uncertainty about the state of the world is large and parties are not too polarized, then swing voters will prefer
a short term length since accountability is most important. When polarization is high and uncertainty low, they prefer a long term length. Polarization of parties depend on the term length. In a citizen-candidate framework we showed, that a short term length leads to more polarization in the least polarized two-candidate symmetric equilibrium. Even then swing voters prefer a short term length if uncertainty is sufficiently high.

9 Appendix

Proof of Lemma 1. Let $s_2 > s_1$ and let e.g. $x_1 = x(s_1)$ and $k_1$ be the discounted period two expected utility consequence for the government from choosing $x_1$ in the first period. If the incumbent is party $R$, the optimality of $x_1$ at $s_1$ and of $x_2$ at $s_2$ yields

$$-(x_1 - r - s_1)^2 + k_1 \geq -(x_2 - r - s_1)^2 + k_2, \text{ and}$$

$$-(x_2 - r - s_2)^2 + k_2 \geq -(x_1 - r - s_2)^2 + k_1$$

(if $L$ is incumbent $r$ is exchanged with $-r$). Adding and rearranging yields

$$(x_2 - x_1)(s_2 - s_1) \geq 0$$

implying that $s_2 > s_1 \Rightarrow x_2 \geq x_1$. □

Proof of Lemma 2. Assume that party $R$ is the incumbent, (if $L$ is the proof is symmetric). From Lemma 1 we have that if the equilibrium is separating then party $R'$s first period policy $x(s)$ is strictly increasing in $s$. Assume that the equilibrium is indeed separating. Party $R$ will then win the second election if $s \geq 0$ and loose otherwise. As it looses for $s < 0$, it will choose its bliss point in period 1 in these states. Consider $s = -\varepsilon$, where $\varepsilon$ is small. If party $R$ chooses $x(-\varepsilon)$ the first period, it looses the election, while
it wins if it chooses $x(\varepsilon)$. As the gain from winning is $\delta 4r^2$, it has to be the case that
\[
\lim_{\varepsilon \to 0} \left[ -(x(-\varepsilon) - r - (-\varepsilon))^2 - (x(\varepsilon) - r - \varepsilon)^2 \right] \geq \delta 4r^2,
\]
otherwise party $R$ will deviate from $x(-\varepsilon)$ to $x(\varepsilon)$. Since it chooses its bliss point for $s < 0$, we get
\[
\lim_{\varepsilon \to 0} x(\varepsilon) \geq \lim_{\varepsilon \to 0} [r + \varepsilon] + 2r\sqrt{\delta} = r \left(1 + 2\sqrt{\delta}\right)
\]
Consider $\varepsilon_2 > \varepsilon_1 > 0$, then party $R$ wins whether it chooses $x(\varepsilon_2)$ or $x(\varepsilon_1)$. Since the equilibrium is separating we have that
\[
x(\varepsilon_2) > x(\varepsilon_1) > r \left(1 + 2\sqrt{\delta}\right)
\]
For $\varepsilon_2 < 2r\sqrt{\delta}$ we have that
\[
r \left(1 + 2\sqrt{\delta}\right) > r + \varepsilon_2
\]
But then party $R$ prefers to choose $x(\varepsilon_1)$ in state $s = \varepsilon_2$ : the first period utility is higher than if it chooses $x(\varepsilon_2)$ and it stills wins the election. This contradicts that the equilibrium is separating for all $s$.□

**Proof of Lemma 3** The first statement follows directly from Lemma 1.

Look at the second statement. Again assume party $R$ is incumbent. If the two intervals are overlapping, then $x_1 = x_2$. Suppose therefore that $s'_1 < s_2$ and that $x_1 < x_2$. For $s \in [s_1, s'_1]$, party $R$ chooses $x_1$. There is at most one $s \in [s_1, s'_1]$ where $x_1 = r + s$. This implies that party $R$ must win the second election after choosing $x_1$, otherwise it would be better to deviate to $r + s$ for $s$ such that $x_1 \neq r + s$. This implies that $s^e(x_1) \geq 0$. By a similar argument $s^e(x_2) \geq 0$. Consider $s \in [s_2, s'_2]$ for which $x_2 \neq r + s$. If $x_2 < r + s$ party $R$ would be better off deviating to $r + s$, since $s^e(r + s) \geq s^e(x_2) \geq 0$,
so the party would still win the election. If $x_2 > r + s$, the party can deviate to $x_2 - \varepsilon > x_1$. For small $\varepsilon$, this is a better first period policy and $s^e(x_2 - \varepsilon) \geq s^e(x_1) \geq 0$, so the party still wins the election. We conclude that $x_1 < x_2$ is not compatible with equilibrium.

**Proof of Lemma 4** Let $x(s) = \bar{x}$ for $s \in [s_1, s_2]$. Assume $R$ is the incumbent, the proof for $L$ is similar. When party $R$ chooses $\bar{x}$, $s^e(\bar{x}) \geq 0$ and it wins the election. As $x$ is non-decreasing

\[ x(s) = r + s > \bar{x} \text{ for } s > s_2 \]

and thus

\[ \bar{x} \leq \lim_{s \downarrow s_2} r + s = r + s_2 \]

On the other hand we also have that

\[ \bar{x} \geq r + s_2 \]

since otherwise the party can gain by deviating to $r + s_2$ at $s_2$. This gives a better first period policy and it still wins the election. In conclusion, $\bar{x} = r + s_2$.

For $s < s_2$ it must be the case that the party loose the election if it deviates to $x < \bar{x}$, otherwise it would do so. The lowest $s$ for which the party will not deviate down to its bliss point, when this implies that it loose rather than wins the election, is given by the solution to

\[ - (\bar{x} - r - s)^2 = \delta 4r^2 \]

This implies that

\[ s_1 = s_2 - 2r\sqrt{\delta} \]

\[ \square \]
Proof of Proposition 5

The expected utility to swing voter \( a \), from \( RDL \) is (using (2))

\[
    u^L_\phi = -2 \left( \frac{\sigma^2}{3} + r_L^2 + a^2 \right)
\]

while one finds that the expected utility from \( RDS \) is (using (10) with \((a + \phi s)\) replacing the bliss point \((r + s)\), and integrating and simplifying)

\[
    u^S = \frac{1}{3\sigma} \left( 3rs\sigma^2 - 2\sigma^3 - 6r_S^3 - 6r_S^2\sigma - 6a^2\sigma \right)
\]

Hence we get

\[
    u^S - u^L = \frac{r\sigma^2 - 2r^3 - 2r^2\sigma + 2\sigma r_L^2}{\sigma}
\]

Using (9) and (12) we can rewrite this as

\[
    u^S - u^L = \frac{r_S\sigma^2 - 2r^3 - 2r^2\sigma + 2\sigma r_L^2}{\sigma} = \frac{1}{3} \sigma^{-1} (3\sigma^2 - 11r_S^2) r_S
\]

Hence

\[
    u^S > u^L \iff r_S < \sqrt{\frac{3}{11}} \sigma
\]

Inserting \( r_S = \sqrt{\frac{3}{11}} \sigma \) into the left hand side of (12), we get that a swing voter is indifferent between \( RDS \) and \( RDL \) if

\[
    \left( \sqrt{\frac{3}{11}} \sigma \right)^2 \left( 4 - \frac{10}{3} \sqrt{\frac{3}{11}} \sigma \right) = \left( \frac{12}{11} \right) - \frac{10}{121} \sqrt{33} \sigma^2 = c
\]

As \( \frac{\partial r}{\partial c} > 0 \), we conclude that

\[
    u^S > u^L \iff \frac{c}{\left( \frac{12}{11} - \frac{10}{121} \sqrt{33} \right)} \leq \sigma^2
\]
References


[8] Hanssen, A. (2002), Is there an optimal level of judicial independence?, *mimeo, Montana State University*


