

Information, Polarization and Term Length in Democracy*

Christian Schultz[†]

December 2007

Abstract

This paper considers term lengths in a representative democracy where the political issue divides the population on the left-right scale. Parties are ideologically different and better informed about the consequences of policies than voters are. A short term length makes the government more accountable, but the re-election incentive leads to policy-distortion as the government seeks to manipulate swing voters' beliefs to make its ideology more popular. This creates a trade-off: A short term length improves accountability but gives distortions. A short term length is best for swing voters when the uncertainty is large and parties are not very polarized. Partisan voters always prefer a long term length. When politicians learn while in office a long term length becomes more attractive for swing voters.

Keywords: Accountability, voting, information, democracy, polarization, term lengths, term limits. JEL: D72, H1, H7, K4

*I am grateful to the editor Antonio Merlo and two referees for very helpful comments. Furthermore, I benefitted from discussions with Klaus Demset, David Dreyer Lassen, Torsten Persson, Jean Tirole and seminar audiences at Bloomington, Carlos III, Copenhagen, Lund, Purdue, Stockholm, Toulouse, Urbana-Campaign, Yale, ESWC 2005, AEA 2006, and CEPR's ESSET meeting in Gerzensee, 2007.

[†]Department of Economics, University of Copenhagen. Studiestraede 6, DK 1455 Copenhagen K, Denmark. E-mail: cs@econ.ku.dk

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, the efficiency of the welfare state, etc. While the problem may be alleviated since voters receive information from many sources such as media, experts, and interest groups, it is also true that much information from media is cheap talk; "experts" often contradict each other; and different interest groups provide conflicting information. 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"*. Politicians, on the other hand, are briefed by experts and bureaucrats and it is their job to gather information and take decisions. The asymmetry of information speaks in favor of delegating decision-making from the electorate to elected politicians. Hamilton and Madison made the point in the 63rd Federalist Paper, when discussing the Senate:

"To a people as little blinded by prejudice or corrupted by flattery as those whom I address, I shall not scruple to add, that such an institution may be sometimes necessary as a defence to the people against their own temporary errors and delusions".

In representative democracies, decision making is delegated to politicians. If voters become discontent with the politicians they can vote them out. In this way, politicians are accountable at the end of their term, and an important institutional feature of representative democracy is therefore the term length. Term lengths vary around the world; in the US the presidential term is 4 years, while house and senatorial terms are 2 and 6 years, respectively. In European democracies term lengths typically vary from 4 to 7 years. In Tanzania, Julius Nyerere stepped voluntarily down after 20 years. An important question, which this paper will focus on, is how term lengths affect policies.

Hamilton and Madison discussed in the 62th and 63rd Federalist Papers the justifications for longer senatorial terms. They focussed on two points:

"Responsibility, in order to be reasonable, must be limited to objects within the power of the responsible party, and in order to be effectual, must relate to operations of that power, of which a ready and proper judgment can be formed by the constituents. it is evident that an assembly elected for so short a term as to be unable to provide more than one or two links in a chain of measures, on which the general welfare may essentially depend, ought not to be answerable for the final result"

In other words, terms should be sufficiently long such that the results of policies unfold before senators are held accountable. While this sounds eminently plausible, it is associated with a cost, namely that accountability is postponed. An important question in this paper is to identify the important trade offs that determine whether the statement is correct. Furthermore, they stressed that legislators need time to acquire expertise, they found

"a want of due acquaintance with the objects and principles of legislation. It is not possible that an assembly of men ..., continued in appointment for a short time, and led by no permanent motive to devote the intervals of public occupation to a study of the laws, the affairs, and the comprehensive interests of their country, should, if left wholly to themselves, escape a variety of important errors in the exercise of their legislative trust."

Hamilton and Madison's first point indicates that the pros and cons of different term lengths depend on what kind of policies are salient. In this paper, I consider a society where the salient issue in politics is the "grand issue", the size of government or the size of the welfare state, which divides the population on the left-right scale. Welfare state policies are "long term"; an evaluation of their consequences takes time. Consider, for instance, the much discussed health reform in the US, or a major change of the tax system with new implications for incentives. Furthermore, the left-right divide has important consequences for accountability: When voters become discontent with a right wing government, the alternative option is to elect a left wing government. It is not an option just to elect "a better" government.

Citizens have preferences on policies and their consequences. The relation between policies and consequences is determined by the state of the world, which could be thought of as the efficiency of the welfare state. A large spending on a welfare programme will have different consequences depending on the efficiency of the programme and tax distortions. As suggested by the quotes above, voters are poorly informed about the complicated mechanisms that determine different policies' consequences and cannot readily assess the consequences. Nevertheless, they may try to infer information about policy consequences from the fact that the informed government chose specific policies.

Voters are divided into partisans and swing voters. While partisans are loyal to their party, swing voters' preferred party may depend on the state of the world. If, for instance, it turns out that the welfare state is efficient in providing its benefits and taxation is not so distortionary, swing voters tend to prefer the more leftist party, which opts for a large welfare state.

Voters' poor information implies that swing voters may want to change government when they learn about the state of the world; this speaks in favor of a rather short term length. But with a short term length, the consequences

of the enacted policies are not yet known to the voters, and they have to infer information from the government's policies. Since governments do not want to be voted out, they have incentives to manipulate voters' beliefs about the state of the world by distorting the policy choice. A left wing government, who learns that the public sector is in fact inefficient and a new welfare programme therefore is not appropriate, may be reluctant to scale down the programme, since this would reveal the truth about public efficiency and cause swing voters to vote right. The policy distortion will be in direction of the party's ideology - in fact to more extreme policies - in order to persuade swing voters that the state is such that they prefer the party's ideology over the other party's. So, while a short term length improves accountability, it makes re-election a more imminent concern for governments and hence induces policy distortion.

High uncertainty about the state of the world tends to make a short term length preferable for swing voters, since the option of electing another government becomes more valuable. Hence, it may be optimal for swing voters that elections come before all consequences of policies are realized and understood. In contrast, high polarization of parties' ideologies tends to make a long term length better since the policy distortion associated with a short term length becomes large as parties become more eager to win an upcoming election. Partisans stay with their party regardless of the state of the world, so they have no option value of elections. Since a long term length does not lead to policy distortion it is unambiguously better for partisans.

The results are derived in a simple two period model, where an election is held at the start of period one. Under a long term length, this is the only election. Under a short term length a new election is held at the start of period two. At this time voters do not know the state of the world, but they can observe the policy chosen by the first period government.

We also briefly consider a term limit. A term limit precludes the re-election of the leadership of the incumbent party but not the party itself. Assuming that the leaders are genuinely ideologically motivated, their interest in future policies does not depend on whether they are personally elected. However, if they cannot be re-elected, the power motive - the rents from office - becomes less pregnant, and this reduces the re-election motive and thus the distortions from a short term length.

In a brief extension, we consider the acquisition of expertise, which Hamilton and Madison also referred to. The model confirms their insight: Acquisition of expertise, per se, speaks in favor of longer term lengths. Learning makes first period incumbents more competent. Under a short term length this has two effects: First, they become more popular in the second election, so they are re-elected in more states. This makes the fit of the government's ideology and the state of the world slightly worse for swing voters. This effect is affected by polarization. When polarization is high, the "ideological fit" is very important and the effect of learning less important.

Secondly, it becomes more important for the incumbent to win the election, as the opponent is less competent and this also hurts the incumbent. The increased re-election motive distorts the policy further, and this is bad for voters. In sum, learning tilts the trade off towards a long term length.

Term lengths have been considered in a number of papers recently. In Maskin and Tirole (2004) politicians are also better informed than voters, but unlike in my framework, the electorate is not divided on the left-right scale, the policy space is binary (build a bridge or not) and politicians are not ideologically different. Instead, their policy preferences may or may not be congruent with the electorates'. If a politician is voted out, the preference of the new politician is chosen at random. The optimal choice between a short term length and "juridical power" (a long term length) depends on how eager the politician is to be re-elected, which is an exogenous feature, unlike in the present paper, where it is influenced by the polarization of parties. Smart and Sturm (2004) consider Maskin and Tirole's framework and show that term limits reduce the incumbent's incentive to win an election and thus policy distortions, just like it is the case in my framework.

A long term length can be interpreted as a case where the decision maker has life long tenure, like Supreme Court judges or an independent central banker with policy preferences. Hansen (2004) shows that an incumbent government is more likely to establish an independent judiciary if there is higher probability that it loses the next election and if the polarization of politics is larger. Alesina and Tabellini (2006, 2008) compare politicians with bureaucrats and assume that bureaucrats are motivated by career concerns as in Holmstrom (1999).

Schultz (1996) considers policy distortions when politicians are better informed than the electorate and can commit to policies before an upcoming election. If parties' preferences are sufficiently polarized, the electoral competition will lead to inefficient policies which do not reflect the state of the world. This distortion occurs regardless of the term length because the parties commit before the election.

The remainder of the paper is organized as follows. Section 2 presents the basic model. Section 3 discusses a long term length, while Section 4 discusses a short term length and the effect of a term limit. Section 5 considers the choice of term length, while Section 6 considers learning in office. Section 7 concludes and discusses some extensions. Some proofs are in the Appendix.

2 Basics

Society has to choose a policy $x \in \mathbf{R}$ in each of two periods. The policy is chosen by parties: either the left party L or the right party R . Agents preferences over policies depend on the state of the world, s , which is uncertain and uniformly distributed on $[-z, z]$, where $z > 0$. The state of the world represents a fundamental feature of the economy and does not change rapidly. For simplicity, we assume it is the same in the two periods¹. Voters are not informed about the state of the world, but they know its distribution.

Each party is headed by a political leader, whose utility depends on the policies in the two periods, x_1 and x_2 , the party's ideology, and the state of the world. Furthermore, the leader is also motivated by the power of office: he gets office rents b in periods where the party holds office. Party R 's ideology is r , and the leader's utility is²

$$u_R = -(x_1 - (r + s))^2 + O_1 b - (x_2 - (r + s))^2 + O_2 b, \quad (1)$$

where O_t is an indicator variable taking value one, if party R is in office in period t , and zero otherwise. Party L 's utility function is the same except its ideology is $l = -r$, and that the terms $O_t b$ are replaced by $(1 - O_t) b$. For simplicity, we consider the symmetric case where $l = -r$. Hence, r is a measure of the polarization of parties.

Voters have similar utility functions; a voter's ideology is reflected in his idiosyncratic parameter, a . There are two kinds of voters distinguished by how much their bliss points depend on the state of the world as given by the parameter ϕ_i , $i \in \{p, sw\}$: A fraction $1 - \mu > 0$ are partisans, p , who always - as will become clear - vote for the same party. Partisans are characterized by $\phi_p = 1$. A fraction $\mu > 0$ are swing voters, sw , they all have $\phi_{sw} = \phi > 1$. Voter a, ϕ_i 's utility is

$$u = -(x_1 - (a + \phi_i s))^2 - (x_2 - (a + \phi_i s))^2. \quad (2)$$

Everybody, voters and parties, prefer higher policies in higher states.

Partisans agree with the parties on how much policy should depend on the state (they both have $\phi_p = 1$) and their preferred party does not depend on the state. Swing voters, on the other hand, have ideal policies which change more with the state. When the state is lowered, their bliss points swing more to the left than the parties', and they swing more to the right when the state is increased. Swing voters with a close to zero prefer

¹The important feature for the results is that the state in period 2 depends on the state in period 1. One could also assume a MA(1) process for the state. This would not change the qualitative results but result in longer formulas.

²It is straightforward to introduce discounting of second period utilities, but as no qualitative results depend on this, we do not do that.

the left party in low states and the right party in high states. There is a continuum of voters and the median a is 0, both among partisans and swing voters. We thus assume that parties have settled on ideological positions, which are ex ante equally popular in the electorate. One interpretation of the model is that party r 's bliss point is the median of the right leaning partisan voters' bliss points, see Roemer (2001) or Gøromberg et al (2004). With this interpretation a larger polarization of parties stem from a larger polarization of the electorate.

It appears realistic that the electorate is divided between swing voters and partisans, where the latter tend to be more ideological and have more stubborn views: As an example, leftists tend to opt for a large welfare state - even facing proof that it is inefficient and taxes are highly distortionary - while rightists tend to opt for a small welfare state - even in face of evidence of its efficiency. Swing voters' opinions, on the other hand, depend more on the effectiveness - or the quality - of the welfare state.

Voters know the preferences of the parties. Clearly, one could imagine this was not the case - in particular if the party has not held power for years. However, we assume that understanding how policies work - the state of the world - is the most complicated and important issue.

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, and the governing party has direct access to the bureaucracy. Furthermore, 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. In order to simplify the exposition, we assume that voters have no information about the state of the world. This is not necessary for the qualitative results, the important feature is that there is some part of the state of the world which is observed by parties only. One could alternatively assume that $s = s^v + s^g$, and voters only know s^v , while the parties know both s^v and s^g .

As is discussed in the literature on the Condorcet Jury Theorem, see for instance Austen-Smith and Banks (1996), the majority vote may end up as if the electorate is informed if the electorate is sufficiently large, even though each individual only receives a noisy signal on the state. This is relevant if the information is dispersed in society. However, there are also many instances where the government has privileged access to information, such as complicated analyses of the tax system or the details of various budgets made in the bureaucracy, where the Jury Theorem appears less relevant. We focus on this kind of information. As briefly touched upon in the Introduction, there is much evidence suggesting that the electorate is rather poorly informed about economic issues.

3 A 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 winning party needs not worry about re-election 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 voter (a, ϕ_i) from a long term length is therefore³

$$u^{LT} = 2 \int_{-z}^z \left(-\frac{1}{2} (-r + s - (a + \phi_i s))^2 - \frac{1}{2} (r + s - (a + \phi_i s))^2 \right) \frac{1}{2z} ds.$$

Letting $\sigma^2 = z^2/3$ denote the variance of the state of the world, s , we get

$$u^{LT} = -2 \left(\sigma^2 (\phi_i - 1)^2 + r^2 + a^2 \right). \quad (3)$$

The larger the degree of polarization is, the less attractive is a long term length. The advantage of a long term length is that the policy reflects the state of the world, since the governing party knows the state and have no incentives to distort the policy. A partisan voter, with $\phi_p = 1$, finds that the party perfectly responds to the state of the world and the variance of the state is therefore inessential for such a voter. This is not so for a swing voter, who finds that the party's policy responds too little to the state of the world. The expected utility of a swing voter is therefore decreasing in the variance of the state of the world.

4 A short term length

Under a short term length, the first election is held in the start of period one. The winning party observes the state of the world and chooses the first period policy, which is observed by the voters. In the start of period two, a new election takes place, and the new winner chooses the second period policy.

We follow Hamilton and Madison and consider policies where it takes time for the effects to materialize, so their utility consequences are only learned after the second election is held. As discussed in the Introduction, the welfare effects of expanding health care come after some while and major tax reforms take time to spell out. The voters can therefore not infer the

³Notice, that for the median voter, who is indifferent between party L and party R , the utility from a long term length will be the same whether or not the winning probability is $1/2$ for each party in case of a tie. For other voters, however, the expression in (3) would differ if the winning probabilities were different.

state directly from the knowledge of the policy and the experienced utility and have to rely on whatever information is contained in the policy choice of the governing party.⁴ The party knows that the voters form beliefs in this way and it may therefore have an incentive to distort policy in order to induce beliefs, which will help it win the election.

An equilibrium under a short term length is a pure strategy Perfect Bayesian Equilibrium of the game just described. In the first period, the winning party chooses the policy⁵, $x(s)$, which maximizes its expected utility for both periods given the state of the world, s , taking into account its own and the other party's strategies for the next period and the way voters form beliefs. Since voters know that the governing party is informed about the state, they update their belief about the state of the world after observing the policy, using Bayes' rule and the governing party's strategy for the first period, $x(\cdot)$, whenever this is possible. A voter votes for the party, whose second period policy gives him highest expected utility, given his beliefs⁶. Finally, the winner of the second election chooses a policy which maximizes its second period utility given the state of the world.

In the second period, the parties have no re-election concerns and choose their bliss points $r + s$ and $-r + s$, respectively. Party R 's second period utility is $0 + b$ if it wins, and $-(-r + s - (r + s))^2$ if L wins, and the gain from winning is

$$G(b, r) \equiv b + 4r^2. \quad (4)$$

Party L 's gain from winning is the same. Voters foresee the parties' second period choices. When voters observe a first period equilibrium policy x , they update beliefs using Bayes' rule. 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. Let $s^e(x)$ be the expected state after observing x . Since all agents, voters as well as parties, prefer higher policies in higher states, the first period incumbent's policy is non-decreasing in s . This is shown in Lemma 1 in the Appendix. If x' and x both are equilibrium policies and $x' > x$ it therefore follows that $s^e(x') \geq s^e(x)$. This fact does not depend on out of equilibrium beliefs.

If voters observe an out of equilibrium policy, y , which should not be chosen in any state according to the incumbent's strategy, Bayes' rule is not applicable. As is well known, this indeterminacy of beliefs can give rise to many equilibria. We will assume that beliefs are monotone in the sense that

⁴Nothing qualitatively would change if it were assumed that only part of the total utility consequence materializes with a lag. Although perhaps more realistic, this would also make the exposition more cumbersome.

⁵In order to keep the notation as light handed as possible and since no confusion should be possible, we refer to the first period strategy as $x(s)$ and not $x_1(s)$.

⁶This is an optimal strategy, whether voters are inclined to vote strategically or sincerely.

if x is an equilibrium policy and voters observe an out of equilibrium policy $y > x$ then $s^e(y) \geq s^e(x)$ and conversely that $s^e(y) \leq s^e(x)$ if $y < x$. Since all agents' (voters' as well as parties') preferred policies are increasing in the state, this seems a reasonable restriction on beliefs

As the equilibrium policy is non-decreasing, and the state is uniformly distributed, voters' updated belief after observing an equilibrium policy $x(s)$ is either a particular state for sure, or that the state is uniformly distributed on some interval $[s_l, s_h]$. In the latter case, $s^e(x) \equiv \frac{s_h + s_l}{2}$ and voter (a, ϕ_i) prefers party R if

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

which is equivalent to

$$\begin{aligned} a &> -(\phi - 1) s^e(x) \text{ if the voter is a swing voter } (\phi_{sw} = \phi > 1) \quad (5) \\ a &> 0 \text{ if the voter is partisan } (\phi_p = 1). \end{aligned}$$

The higher is the expected state, the more swing voters prefer party R , while partisans with $a > 0$ prefer party R and those with $a < 0$ prefer party L regardless of the state. Hence, the election is determined by the median swing voter and party R wins if the expected state is positive and loses otherwise⁷.

The first period incumbent has an incentive to manipulate the voters' beliefs in order to be re-elected. If, for instance, the state is negative, but close to zero, party L will win if voters learn the truth. An incumbent party R therefore has an incentive to act as it would if the state were positive, so that it will be re-elected. In equilibrium, such mimicking should not pay off, so when the state is just positive, party R has to choose a policy so high, that it will not prefer to mimic it when the state is slightly negative. In equilibrium there is therefore an interval of states $[s_1, s_2]$, situated around $s = 0$, where the incumbent distorts the policy. An R incumbent distorts upwards and an L incumbent distorts downwards (since it is most popular in the low states). This is shown formally in Lemmas 2 and 3 in the Appendix⁸.

⁷If the expected state is zero, the median swing voter is indifferent. We assume that the incumbent wins in this case. This tie-breaking rule simplifies a few steps in the proofs, but has no impact on the results.

⁸The restriction on out of equilibrium beliefs is used in Lemma 3, where it is shown that there is only one pooling interval $[s_1, s_2]$ with policy \bar{x} and it is situated around $s = 0$. Say R is incumbent. The restriction excludes the existence of another pooling interval for states $s > s_2$. Without the restriction, such an interval, with pooling policy \bar{x}' , could exist, since voters' beliefs could be that the state is negative if they do not observe either \bar{x} or \bar{x}' . Belief monotonicity ensures that for any policy $x > \bar{x}$, $s^e(x) \geq s^e(\bar{x})$, so the R incumbent will not be punished by beliefs if he chooses his bliss point when $r + s > \bar{x}$. A similar argument is used in Lemma 4.

When the gain from winning the election, $G(b, r)$, becomes larger, the distortion has to be larger in order to prevent mimicking. The policy becomes more distorted in a given state and the interval of states with distortion becomes longer. Lemma 4 shows this more precisely. As often in signalling games, there is a continuum of equilibria, according to where precisely the pooling interval is located around $s = 0$. Lemma 5 gives the possible locations. We will focus on the most informative equilibrium from the point of view of the median voter. In this equilibrium, voters learn whether the state is positive or negative, and the median swing voter is able to cast her vote based on the information she needs. Hence, in the second period, the elected government is her most preferred given the actual state. Although the formulas would be a bit more complicated, the qualitative results would not change by focussing on one of the other possible equilibria. Proposition 1 is proven in the Appendix.

Proposition 1 *Consider a short term length. If R is incumbent its first period equilibrium policy in the most informative equilibrium is*

$$x_R(s) = \begin{cases} r + s & \text{if } s < 0 \\ r + \sqrt{G(b, r)} & \text{if } 0 \leq s \leq \min[\sqrt{G(b, r)}, z] \\ r + s & \text{if } \sqrt{G(b, r)} < z \text{ and } \sqrt{G(b, r)} \leq s \leq z \end{cases}$$

If L is incumbent its first period policy is

$$x_L(s) = \begin{cases} -r + s & \text{if } -z < -\sqrt{G(b, r)} \text{ and } -z \leq s \leq \sqrt{G(b, r)} \\ -r - \sqrt{G(b, r)} & \text{if } \max[-z, -\sqrt{G(b, r)}] \leq s \leq 0 \\ r + s & \text{if } 0 < s \end{cases}$$

The Proposition is illustrated in Figure 1 for the case where R is incumbent and $\sqrt{G(b, r)} < z$. When s is very high, R chooses its bliss point. This is credible information for the swing voters that the state is high, since party R would not find it worthwhile to deviate to this policy, had the state been negative, even though the gain would be that it won the election. For states between 0 and $\sqrt{G(b, r)}$, party R distorts its policy upwards in order to be re-elected. If it chooses a lower policy, voters will believe that it is a negative state and elect party L for the next period. The distortion is so large that party R does not find it worthwhile to deviate to this policy in a negative state. In negative states, party R chooses its bliss point and faces defeat in the upcoming election. As is clear from the Proposition and Figure 1, the pooling interval is longer the stronger the incentive to win the election is, i.e. the higher is polarization, r , and benefits from power, b .

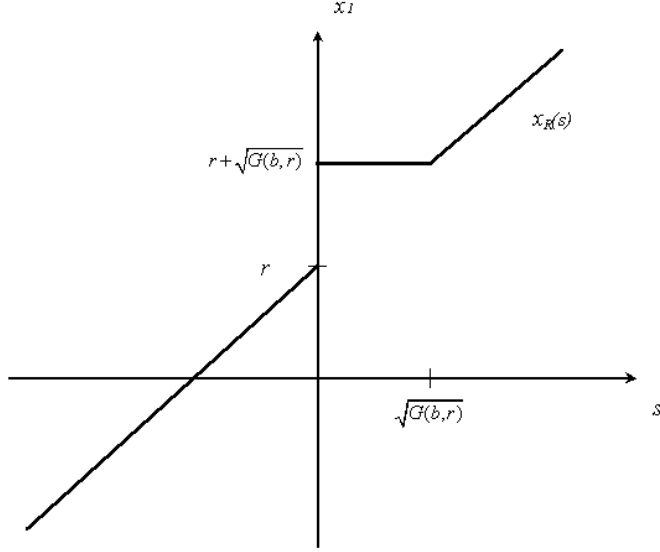


Figure 1. First period policy when R is incumbent

When a term limit is present, the incumbent leadership of the government cannot be re-elected. If the leadership is genuinely interested in politics its interest in future policies remains unchanged. However, since it cannot be re-elected, it is not able to reap the future benefits of office and it seems reasonable to assume that it then puts less weight on the future rents from office, b . Glancing at the Proposition, we see that then the effect of a term limit is that the policy distortion is reduced.

At the first election, voters have no further information about the state, so each party wins with probability $1/2$. The expected utility in the first period for voter a , ϕ_i is therefore

$$u_1^{ST} = \int_{-z}^z \left(-\frac{1}{2} (x_R(s) - (a + \phi_i s))^2 - \frac{1}{2} (x_L(s) - (a + \phi_i s))^2 \right) \frac{1}{2z} ds. \quad (6)$$

At the second election, voters elect party L if $s < 0$ and party R if $s > 0$, and the incumbent - whoever it is - if $s = 0$ ⁹. The expected utility in the second period for voter a , ϕ_i evaluated at date 0 is therefore

$$u_2^{ST} = \int_{-z}^0 -(-r + s - (a + \phi_i s))^2 \frac{1}{2z} ds + \int_0^z -(r + s - (a + \phi_i s))^2 \frac{1}{2z} ds,$$

which simplifies to

$$u_2^{ST} = - \left(\sigma^2 (\phi_i - 1)^2 + r^2 + a^2 - (\phi_i - 1) r z \right). \quad (7)$$

⁹I claimed earlier that the tie-break rule, saying that the incumbent wins if $s = 0$, does not matter for our results. Since $s = 0$ is a measure zero event, the expected utilities, which we now calculate, does not depend on this rule.

5 The choice of term length

Using (3) and (7), and denoting the second period utility under a long term length u_2^{LT} , we can write the difference between the second period expected utilities under a short and a long term length as

$$u_2^{ST} - u_2^{LT} = rz(\phi_i - 1). \quad (8)$$

This is the gain from accountability. It stems from the voters' ability to choose government depending on the state when the term length is short. It is positive for all all swing voters (with $\phi_{sw} = \phi > 1$). The higher is r , the degree of polarization, and z , the uncertainty about the state of the world, the more valuable is accountability for these voters. The partisan ($\phi_p = 1$) voters' preferred government, on the other hand, does not depend on the state, so their expected utility is the same regardless of term length.

When the term length is short, the incumbent distorts policy in some states in order to attract swing voters, while this is not needed when the term length is long. In other states, the incumbent chooses its bliss point under both term lengths. The first period expected utility is therefore higher for all voters when the term length is long. When $\sqrt{G(b, r)} \leq z$, such that incumbent R (L) does not have to distort policy in all positive (negative) states in order to win, the loss from a short term length is (using (3), (6), and Proposition 1)

$$u_1^{LT} - u_1^{ST} = \left(3r + (2 - \phi_i) \sqrt{G(b, r)}\right) G(b, r) \frac{1}{6z}. \quad (9)$$

Notice that both the gain (8) and the loss (9) from a short term length are independent of a voter's bliss point. Using (8) and (9) gives (10) in the following Proposition¹⁰

Proposition 2 *A long term length is always best for partisan voters.*

For swing voters a short term length is best iff the gain from accountability exceeds the loss from the first period distortion.

This is the case if when uncertainty is sufficiently large relative to the gain from winning the election, i.e. when

$$\sigma^2 > \frac{G(b, r)}{3} \text{ and } \sigma^2 \geq \frac{3 + (2 - \phi) \frac{\sqrt{G(b, r)}}{r}}{6(\phi - 1)} \frac{G(b, r)}{3}. \quad (10)$$

It is also the case when uncertainty is not so large but the swing voters' bliss points are sufficiently sensitive to the state, namely when:

$$\frac{G(b, r)}{3} > \sigma^2 \text{ and } \phi \geq \frac{3G(b, r) + 3rz + 6r\sqrt{G(b, r)} - z^2}{6rz + 3\sqrt{G(b, r)}z - 2z^2}. \quad (11)$$

¹⁰Condition (11) is also derived from (3), (6) and Proposition 1 .

In (10) the last inequality is the binding restriction iff

$$\phi \leq \left(9 + 2\frac{\sqrt{G(b,r)}}{r}\right) \left(6 + \frac{\sqrt{G(b,r)}}{r}\right),$$

which is less than 2. Regardless of whether the first or the last inequality binds in (10) it is therefore the case that a larger polarization tends to make it more difficult to fulfil (10), so it is less likely that a short term length is preferred.

When the uncertainty is small relative to polarization, so that the gains from winning the election are so high that the incumbent distorts policy in all states where he is going to be re-elected, i.e. when $G(b,r) > 3\sigma^2 = z^2$ cf. Proposition 1, then the condition for when the swing voter prefers a short term length is given in (11). The condition reveals that if swing voters are sufficiently sensitive to the state, they prefer a short term length even though uncertainty is small relatively to the gain of winning. The reason is that their (expected) ideal policies are so extreme compared with the parties' that the distortion brings policies closer to their ideal policies¹¹. The right hand side is decreasing in z , so the higher the uncertainty, the more types of swing voters prefer a short term length. As the right hand side tends to infinity as z becomes vanishingly small we have that for any given ϕ , if the uncertainty becomes sufficiently small, then the swing voter will prefer a long term length. When polarization r or benefits from office b increases, so that the gain from winning the election increases, the right hand side increases. Hence the larger polarization, the more types of swing voters prefer a long term length. When r tends to infinity all types of swing voters prefer a long term length. Hence, (11) reflects the same trade offs as (10).

Proposition 2 highlights that partisans and swing voters have fundamentally different interests with respect to term lengths. The partisans are not interested in accountability, since they find that the parties adjust policy adequately to the state. They focus on the negative effect of a short term length: The policy distortion coming from the re-election incentive. Swing voters, on the other hand, value the opportunity to oust a government if they find that the state calls for the policies of the other party. This value dominates if the uncertainty of the state of the world is sufficiently large relative to the policy distortion. If the term length is chosen in a popular vote at a constitutional stage, the choice depends on whether partisans or swing voters are in majority.

When a term limit is effective, the leadership of the party governing in period 1 cannot be re-elected. If this implies that office rents b weigh less

¹¹This effect can be visualized in Figure 1, if one adds a steep line through $(0,0)$ representing the median swing voter's bliss point. If the line is sufficiently steep, the voter will prefer larger distortions as this brings the incumbent's policy closer to the swing voter's ideal policies in most of the states where there is distortion.

for the leadership, then $G(b, r)$ decreases and the policy distortion under a short term length is smaller. This makes it more likely that condition (10) is fulfilled and that swing voters prefer a short term length.

6 Learning in office

As discussed in the Introduction, Hamilton and Madison argued that legislators' acquisition of expertise takes time and this speaks for longer term lengths. In this paper we have so far taken the view that legislators are informed agents, but we will now shortly address how the trade off between a short and a long term length is affected by learning. In order to introduce this feature as simple as possible, we assume that the first period is unchanged. The first period state is still s , but there is an additional shock, ε , to the state of the world in period 2, which is uniform on $[-v, v]$, where $v > 0$, so the state in period two is

$$s_2 = s + \varepsilon.$$

The ability to acquire the relevant information increases if the politician has spent time in office: We assume that the first period incumbent observes the new shock, while the opponent does not. More generally, one could assume that the first period incumbent learns the second period shock with a larger probability than the opponent. This would add nothing qualitatively but make formulas longer. The crucial issue is that the incumbent learns more than the opponent. Since both parties are equally inexperienced in the first period, there is no asymmetry here, and for simplicity we assume that they are both capable of observing the first period state.

We will investigate the case, where the expertise of the incumbent is not of overwhelming importance. The variance, $v^2/3$, of ε fulfils

$$\Phi \frac{v^2}{3} \frac{1}{r} < z, \tag{12}$$

where $\Phi \equiv \frac{1}{4} \frac{2\phi-1}{\phi-1}$. As will become clear, if this condition is *not* met, then the incumbent will get the vote of *all* swing voters even in the state where he is most unpopular.

Under a long term length, the government chooses its bliss point in period one. In period two it is experienced and learns ε , and for instance R chooses $r + s + \varepsilon$. The expected utility for voter a , ϕ_i is

$$u^{LT} = -2 \left(\sigma^2 (\phi_i - 1)^2 + r^2 + a^2 \right) - \frac{v^2}{3} (\phi_i - 1)^2. \tag{13}$$

This is as before (compare with (3)), except that the swing voters suffer a utility loss, since they do not find that the policy responds sufficiently much to the new uncertainty, just as it is the case with s .

Then consider a short term length. If, for instance, party R wins the second election, it observes ε and chooses policy $r + s + \varepsilon$ if it was the first period incumbent, while if it were not it does not observe ε and chooses $r + s$. When a voter's belief is s^e , she prefers the incumbent, party R , if

$$-(2\phi_i - 1) \frac{v^2}{3} \frac{1}{4r} - (\phi_i - 1) s^e \leq a, \quad (14)$$

(compare with (5)). Former partisan L voters with $\phi_p = 1$, and $-\frac{v^2}{3} \frac{1}{4r} < a < 0$, will now vote for incumbent R , since it is more experienced. The election therefore depends on the fraction of such voters. This could easily be taken into account but in order not to cloud the formulas, we will assume that there are no such voters. So, a voter with $\phi_p = 1$, has $|a| > \frac{v^2}{3} \frac{1}{4r}$. The election is then determined by the median swing voter, and (14) gives that party R wins if

$$s^e \geq -\Phi \frac{v^2}{3} \frac{1}{r}. \quad (15)$$

The learning gives an incumbency advantage.¹² The swing voters face a trade off: The incumbent has larger expertise, but his ideology may not be the one they prefer in the expected state. Condition (15) is the outcome of this compromise and we see that swing voters re-elect R even in some (not too) negative states. This compromise makes a short term length less attractive for the swing voters. The effect is larger when there is more second period uncertainty, and smaller when there is more polarization. When polarization is large it is important for the swing voters that the government's ideology corresponds to the state. Then differences in expertise weigh less.

The learning of expertise increases the stakes in the election for the first period incumbent, since the second period policy of the inexperienced opponent now becomes even worse for him. If party R is re-elected it will choose $r + s + \varepsilon$ in state $s + \varepsilon$, this gives expected utility 0. If the inexperienced opponent L wins, it will choose $-r + s$ in state $s + \varepsilon$. The gain from winning for incumbent R , $G(b, r, v)$, is therefore

$$G(b, r, v) = b + 4r^2 + \frac{1}{3}v^2. \quad (16)$$

The larger is the second period uncertainty, the larger are the gains from winning the election. The length of the pooling interval is determined by this gain, and it will now be longer and equal to $\sqrt{G(b, r, v)}$.

We focus on the case where the first period uncertainty is so large that the pooling interval does not cover all states where the incumbent wins, which means that $3\sigma^2 > -\Phi \frac{v^2}{3r} + G(b, r, v)$. In the most informative equilibrium

¹²Had we included the former partisans, who now would vote for expertise, this effect would be even larger.

for the median voter, incumbent R wins when (15) is fulfilled. The party chooses the pooling policy $\bar{x} = r - \Phi \frac{v^2}{3r} + \sqrt{G(b, r, v)}$ in states s , fulfilling $-\Phi \frac{v^2}{3r} \leq s \leq -\Phi \frac{v^2}{3r} + \sqrt{G(b, r, v)}$. If L is the first period incumbent, the situation is symmetric, it wins in states $s \leq \Phi \frac{v^2}{3r}$ and the pooling interval ends in $\Phi \frac{v^2}{3r}$.

The expected utility from a short term length, u^{ST} , for voter a , ϕ_i can now be found. The formulas can be found in the Appendix. We then get

Proposition 3 *Assume that $3\sigma^2 > -\Phi \frac{v^2}{3r} + G(b, r, v)$ and that (12) is fulfilled. When a government acquires expertise during its tenure, the swing voters prefer a short term length if*

$$\sigma^2 > \frac{3 + (2 - \phi) \frac{\sqrt{G(b, r, v)}}{r} G(b, r, v)}{6(\phi - 1)} \frac{G(b, r, v)}{3} + \left(2z + \frac{1}{2} \frac{G(b, r, v)}{r} - \Phi \frac{v^2}{3} \frac{1}{r} \right) \Phi \frac{v^2}{3r}. \quad (17)$$

Comparing with Proposition 2, we see that experience changes the trade off between a short and a long term length. If the second period variance, $v^2/3$, equals zero, (17) is the same as (10). Two effects can be identified. First the distortion increases: The pooling interval becomes longer under a short term length as the incumbent is more eager to win the election. If $\phi < 2 \left(1 + \frac{r}{\sqrt{G}} \right)$ so that the swing voters are not very sensitive, and dislike distortions, this makes the first term in (17) larger than the corresponding term in (10), and it also increases the second term. This effect makes a short term length less attractive. The second term is positive under condition (12) and it is increasing in $\Phi \frac{v^2}{3r}$. This effect, which stems from the poorer "fit of ideology and state" also makes a short term length less attractive.

7 Concluding remarks

When voters are poorly informed about the consequences of policies they may wish to replace governments in face of new information, and governments' incentives are influenced by this. While a short term length has the advantage that voters can replace a government faster, it has the disadvantage that governments' incentives to be re-elected make them distort policies. A long term length gives 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. Learning in office twists this trade off, but does not change it fundamentally. Partisans unambiguously prefer a long term length.

I have focussed on the case where it is more difficult for voters to learn about the consequences of policies than the politicians' preferences, as I believe this is most important. However, voters may also be uncertain about parties' preferences. This may in particular be relevant when party's leadership changes. Schultz (2002) considers the case where voters have inferior information both about the state of the world *and* the preferences of the parties at the same time. Uncertainty about preferences in itself leads to a bias towards more centrist policies, i.e. in the opposite direction of the bias stemming from uncertainty about the state. This distortion is beneficial for voters and it would tend to make a short term length more attractive. This case is discussed in a previous working paper version of this paper¹³.

The polarization of parties has been taken as an exogenous variable. However, it is a natural thought that party ideology may depend on the electoral system as such and term lengths in particular. Preliminary investigations of two-candidate equilibria in a citizen-candidate model do point to this. As is well known, the equilibria in citizen-candidate models depend on the incentives to enter at the candidate state. The incentive to enter can be shown to be stronger under a long term length than under a short, since candidates do not need to distort their policy under a long term length. This suggests that polarization will be smaller under a long term length; since the gain from entering is larger, more moderate types will find it worthwhile to pay the entry cost. This will be the subject of further research.

Throughout the paper I assume that both parties are unable to commit to a policy before an election. But one could imagine that by setting up a particular administration, a party in office is capable of committing to a particular set of policies to a degree a party in opposition is not. It is an interesting question how the results would be affected if the first period incumbent can commit to second period policies while the opponent cannot. At first sight, it appears to give the incumbent an advantage, since he has the option to chose more centrist policies and become more popular with the median voter. However, with information problems, the second period policy is a signal for the state, just as the first period policy is, and the incumbent's problem is to signal optimally with two signals. How this would affect results is a non-trivial, interesting question, which will be the subject of further research.

¹³The working paper is available on request.

8 Appendix

Proof of Proposition 1.

We first show a few Lemmas with properties of Perfect Bayesian Equilibria.

Lemma 1 *The first period government's policy is non-decreasing in s .*

Proof. Let $s_2 > s_1$ and let e.g. k_1 be the period two expected utility consequence for the government from choosing $x(s_1)$ in the first period. If the incumbent is party R , the optimally of $x(s_1)$ and of $x(s_2)$ yield

$$\begin{aligned} -(x(s_1) - r - s_1))^2 + k_1 &\geq -(x(s_2) - r - s_1)^2 + k_2, \text{ and} \\ -(x(s_2) - r - s_2))^2 + k_2 &\geq -(x(s_1) - r - s_2)^2 + k_1. \end{aligned}$$

(If L is incumbent r is exchanged with $-r$). Adding and rearranging yield

$$(x(s_2) - x(s_1))(s_2 - s_1) \geq 0,$$

implying that $s_2 > s_1 \Rightarrow x(s_2) \geq x(s_1)$. \square

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

Proof. Assume that party R is the incumbent, (if L is incumbent 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. Consider $s = -\varepsilon$, where $\varepsilon > 0$ is small. By choosing $x(-\varepsilon)$, party R looses the election, while it wins if it chooses $x(\varepsilon)$. As the gain from winning is $b + 4r^2$, it has to be the case that

$$\lim_{\varepsilon \rightarrow 0} \left[-(x(-\varepsilon) - (r - \varepsilon))^2 - \left(-(x(\varepsilon) - (r + \varepsilon))^2 \right) \right] \geq b + 4r^2, \quad (18)$$

otherwise party R will deviate from $x(-\varepsilon)$ to $x(\varepsilon)$. As it looses for $s < 0$, it chooses its bliss point, $x(-\varepsilon) = r - \varepsilon$, so $\lim_{\varepsilon \rightarrow 0} x(-\varepsilon) = r$, and (18) gives

$$\left(\lim_{\varepsilon \rightarrow 0} x(\varepsilon) - r \right)^2 \geq b + 4r^2.$$

Since the policy is non-decreasing and $\lim_{\varepsilon \rightarrow 0} x(-\varepsilon) = r$, we have $\lim_{\varepsilon \rightarrow 0} x(\varepsilon) \geq r$, so the parenthesis above is non-negative and we get

$$\lim_{\varepsilon \rightarrow 0} x(\varepsilon) \geq r + \sqrt{b + 4r^2}.$$

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

$$x(\varepsilon_2) > x(\varepsilon_1) > r + \sqrt{b + 4r^2}.$$

For $\varepsilon_2 < \sqrt{b + 4r^2}$ we have that

$$r + \sqrt{b + 4r^2} > r + \varepsilon_2.$$

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 . \square

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. The proof uses the assumption of monotonicity of out of equilibrium beliefs.

Lemma 3 *In a Perfect Bayesian Equilibrium, 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 $x_1 = x_2$

If $x(s) = \bar{x}$ for all $s \in [s_1, s_2]$, then $s_1 \leq 0$ and $s_2 \geq 0$ and the incumbent wins the election after choosing \bar{x} .

If \bar{x} is the policy of a pooling interval and R is incumbent, then $s^e(\bar{x}) \geq 0$, while if L is incumbent, $s^e(\bar{x}) \leq 0$.

Proof. 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 $s^e(x_1) \geq 0$ (and thus also the last statement in the Lemma). 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, by monotonicity of beliefs, $s^e(r + s) \geq s^e(x_2) \geq 0$, 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 ε , this is a better first period policy and, by monotonicity of beliefs, $s^e(x_2 - \varepsilon) \geq s^e(x_1) \geq 0$, so the party still wins the election. We conclude that $x_1 < x_2$ is not compatible with equilibrium.

Finally, $s^e(\bar{x}) \geq 0$ clearly implies $s_2 \geq 0$. Suppose $s_1 > 0$. Then there is an interval $[-\varepsilon, \varepsilon]$ around zero where $\varepsilon < s_1$, and party R 's policy is separating. As in the proof of Lemma 2 this gives a contradiction.

For all states in the pooling interval except possibly one party R suffers a first period loss by choosing the pooling policy \bar{x} , this can only be optimal if it subsequently wins the election. \square

Lemma 4 *Let R be the incumbent and $[s_1, s_2]$ be the pooling interval and $x(s) = \bar{x}$ for all $s \in [s_1, s_2]$. Then*

$$\bar{x} = r + s_1 + \sqrt{b + 4r^2} \quad (19)$$

Outside the pooling interval, party R chooses $x(s) = r + s$

Proof When party R chooses \bar{x} , then (from Lemma 3) $s^e(\bar{x}) \geq 0$, and R wins the election. By monotonicity of beliefs, if party R chooses $x > \bar{x}$ it also wins. This implies that $\bar{x} \geq r + s_2$, otherwise party R would deviate to $r + s_2$ at s_2 and still win the election. Hence for all $s < s_2$ we have $r + s < \bar{x}$. The largest gain from deviation is then at s_1 . Here the party is indifferent between deviating to $r + s_1$ and loose the election and choose \bar{x} and win, this implies

$$(\bar{x} - r - s_1)^2 = b + 4r^2$$

this gives (19).

Outside the pooling interval the equilibrium is separating (Lemma 3) and the equilibrium policy is increasing (Lemma 1), therefore when $s < s_1$: $s^e(x(s)) = s < s_1 < 0$, and party R loses the election. Hence, it chooses its bliss point $r + s$. When $s > s_2$, $s^e(r + s) = s > s_2 > 0$, and party R wins the election by choosing its bliss point. Clearly, this is then optimal. \square

Lemma 5 *Let R be the incumbent and $[s_1, s_2]$ the pooling interval. Then*

$$\max \left[\frac{-\sqrt{b+4r^2}}{2}, \frac{-z}{2} \right] \leq s_1 \leq 0 \text{ and } \frac{\sqrt{b+4r^2}}{2} \leq s_2 < \min \left[\sqrt{b+4r^2}, z \right]$$

Proof From Lemma 3 $s_1 \leq 0$ and $s^e(\bar{x}) \geq 0$. Together with (19) this gives the Lemma. \square

There is a continuum of equilibria according to where the pooling interval is situated. Lemma 5 gives the possible locations.

We focus on the most informative equilibrium from the point of view of the median swing voter, where she learns whether the state is positive or not. Here the left endpoint of the pooling interval, when R is incumbent is $s_1 = 0$. The first part of Proposition 1, concerning the case where R is incumbent, now follows straightforwardly Lemmas 4 and 5. The proof for L is symmetric. \square

Proof of Proposition 3.

Let $y = \Phi \frac{v^2}{3r}$, and $x = \sqrt{b + 4r^2} + \frac{1}{3}v^2$. The pooling policy of party R is $r + x - y$ and the pooling policy of party L is $-r - x + y$. The expected

utility in the first period from a short term length for swing voter a is

$$\begin{aligned}
u_1^{ST} = & \tag{20} \\
& \int_{-z}^{y-x} \left(-\frac{1}{2} (r+s-(a+\phi s))^2 - \frac{1}{2} (-r+s-(a+\phi s))^2 \right) \frac{1}{2z} ds \\
& + \int_{y-x}^{-y} \left(-\frac{1}{2} (r+s-(a+\phi s))^2 - \frac{1}{2} (-r-x+y-(a+\phi s))^2 \right) \frac{1}{2z} ds \\
& + \int_{-y}^y \left(-\frac{1}{2} (r+x-y-(a+\phi s))^2 - \frac{1}{2} (-r-x+y-(a+\phi s))^2 \right) \frac{1}{2z} ds \\
& + \int_y^{x-y} \left(-\frac{1}{2} (r+x-y-(a+\phi s))^2 - \frac{1}{2} (-r+s-(a+\phi s))^2 \right) \frac{1}{2z} ds \\
& + \int_{x-y}^z \left(-\frac{1}{2} (r+s-(a+\phi s))^2 - \frac{1}{2} (-r+s-(a+\phi s))^2 \right) \frac{1}{2z} ds
\end{aligned}$$

Each party wins the first election with probability $\frac{1}{2}$. In states $s \in [-z, y-x[$ both R and L choose bliss points (R expects to loose and L expects to win anyway). In states $s \in [y-x, -y[$ party R chooses its bliss point (and expects to loose the next election), while party L chooses its poling policy $-r-x+y$. In states $s \in [-y, y[$ both parties pool. In states $s \in [y, y-x[$ only party R pool while in states $s \in [y-x, z]$ none of the parties pool.

In the second period the expected utility is as follows: With probability $\frac{1}{2}$, L is elected in the first period and it is re-elected in states $[-z, y]$. L is then experienced and chooses policy knowing ε . In states $]y, z]$, L loses, R is elected and chooses policy without knowing ε . This explains the first large parenthesis below. With probability $\frac{1}{2}$, R is elected in the first period and the same reasoning explains the second large parenthesis.

$$\begin{aligned}
u_2^{ST} = & \tag{21} \\
& \int_{-v}^v \frac{1}{2} \left(\int_{-z}^y -(-r+s+\varepsilon-(a+\phi(s+\varepsilon)))^2 \frac{1}{2z} ds \right. \\
& \quad \left. + \int_y^z -(r+s-(a+\phi(s+\varepsilon)))^2 \frac{1}{2z} ds \right) \\
& + \frac{1}{2} \left(\int_{-z}^{-y} -(-r+s-(a+\phi(s+\varepsilon)))^2 \frac{1}{2z} ds \right. \\
& \quad \left. + \int_{-y}^z -(r+s+\varepsilon-(a+\phi(s+\varepsilon)))^2 \frac{1}{2z} ds \right) \frac{1}{2v} d\varepsilon
\end{aligned}$$

Doing the integrations in (20) and (21) and comparing with the expected utility from a long term length (13) gives the condition in Proposition 3. \square

References

- [1] Alesina, A and G. Tabellini, 2006, Bureaucrats or Politicians, Part I, A Single Policy Task, *American Economic Review*, 97, March, 169-179
- [2] Alesina, A and G. Tabellini, 2008, Bureaucrats or Politicians: Part II, *Journal of Public Economics*, forthcoming
- [3] Austen-Smith, D. and J. Banks, 1996, Information Aggregation, Rationality and the Condorcet Jury Theorem, *American Political Science Review*, 90 (1), 34-45
- [4] Bartels, L.M., 1996, Uninformed Voters: Information Effects in Presidential Elections, *American Journal of Political Science*, 40 (1), 194-230
- [5] Gomberg, A, F Marhuenda and I. Ortuno-Ortin , 2004, “An endogenous model of political party platforms”, *Economic Theory*.
- [6] Hamilton, A., J. Madison, and J. Jay, 1999,. *The Federalist Papers*. Edited by Clinton Rossiter. New York: Mentor.
- [7] Hanssen, A., 2004, Is there an optimal level of judicial independence?, *American Economic Review*, 94, 3, June, 712,729
- [8] Holmström, B, 1999, Managerial Incentive Problems, A Dynamic Perspective, *Review of Economic Studies*, 66, (1), 169-182
- [9] Maskin, E. and J. Tirole, 2004, The Politician and the Judge, Accountability in Government, *American Economic Review*, 94 (4), September, 1034 - 1054
- [10] Roemer, J., 2001, *Political Competition: Theory and Applications*, Cambridge, MA, Harvard University Press,
- [11] Schultz, C., 1996, Polarization and Inefficient Policies, *Review of Economic Studies*, 63, (2), 331-344
- [12] Schultz, C., 2002, Policy Biases with Voters’ Uncertainty about the Economy and the Government. *European Economic Review*, 46, 487-506
- [13] Smart M. and D. Sturm, 2004, Term Limits and Electoral Accountability, Mimeo University of Munich