

2000

# 2000-12 Endogenous Majority Rules with Changing Preferences

Mattias K. Polborn

Follow this and additional works at: <https://ir.lib.uwo.ca/economicsresrpt>

 Part of the [Economics Commons](#)

---

## Citation of this paper:

Polborn, Mattias K.. "2000-12 Endogenous Majority Rules with Changing Preferences." Department of Economics Research Reports, 2000-12. London, ON: Department of Economics, University of Western Ontario (2000).

**RESEARCH REPORT 2000-12**

**Endogenous Majority Rules with Changing Preferences**

by

**Mattias K. Polborn**

December 2000

Department of Economics  
Social Science Centre  
University of Western Ontario  
London, Ontario, Canada  
N6A 5C2  
econref@julian.uwo.ca

DEC 05 2001

UNIVERSITY OF WESTERN ONTARIO

# Endogenous Majority Rules with Changing Preferences

Mattias K. Polborn\*  
University of Western Ontario

December 5, 2000

## Abstract

This paper provides a new explanation why several US states have implemented supermajority requirements for tax increases. We model a dynamic and stochastic OLG economy where individual preferences depend on age and change over time in a systematic way. In this setting, we show that the first population of voters will choose a supermajority rule in order to influence the outcomes of future elections. We explore the robustness of the basic model and also find some empirical support for predictions derived from the model.

**Keywords:** Supermajority, taxation, constitution, overlapping generations, political economy.

**JEL code:** D72.

---

\*Corresponding author: Mattias Polborn, Department of Economics, University of Western Ontario, London, Ontario, N6A 5C2, Canada; email: mpolborn@julian.uwo.ca . I would like to thank James Davies, Paul Klein and Matthias Messner for helpful comments; the usual disclaimer applies.

# 1 Introduction

In thirteen US states, a supermajority of usually  $2/3$  of the state legislatures is needed for a proposal to increase taxes to pass.<sup>1</sup> Efforts to introduce similar rules are on the way in a number of further states, and there have also been several attempts to introduce a supermajority requirement on the federal level. While it is understandable that anti-tax interest groups lobby for supermajority rules, it is not immediately clear why the median voter (or the median legislator) should agree to a supermajority rule. The standard model of political economy implies that the equilibrium tax rate is that preferred by the median voter, so it would be hard to explain in the standard framework why the median voter in a static world would be willing to transfer power to another voter. This paper develops a model of a dynamic and stochastic OLG economy where individual preferences depend on age and change over time in a systematic way. In this setting, we show that the first population of voters will choose a supermajority rule in order to influence the outcomes of future elections.

Consider an economy in which, as individuals grow older, their salaries, and consequently the taxes they have to pay, increase. In this case, it is plausible that the preferred tax rate is a decreasing function of age.<sup>2</sup> When voting on the tax rate, people should take into account their changing preferences and vote according to some (appropriately defined) “average” preference, provided they expect some stability of today’s result. In a completely deterministic OLG model, individual preferences would change over time, but the distribution of preferences in the population would always remain the same, so a single election on the value of the policy variable would be sufficient to determine the level of the policy variable for the whole future. More generally, preferences in the whole society can change, for example due to shocks to technology (e.g., shocks to the marginal cost of provision of public goods, or new goods becoming available). In this case, a new election is necessary after each shock in order to determine society’s preferences.

For analytical tractability, the basic model assumes a very simple structure: At some ex ante unknown time, a shock to all people’s preferences occurs, and an election on whether the tax rate should be increased or decreased is held. This election is governed by the majority rules chosen in two procedural elections (one for the majority rule for a decrease, and the other one for an increase), held at the beginning of time.

---

<sup>1</sup>These states are Arizona, Arkansas, California, Colorado, Delaware, Florida, Louisiana, Mississippi, Nevada, Oklahoma, Oregon, South Dakota and Washington; see Knight (2000)

<sup>2</sup>Of course, people’s preferences are generally not *only* influenced by their age. We will show the robustness of the model in this respect in section 4.

At that time, people also determine the initial tax rate.

Today's median voter knows that, due to her increased salary in the future, she will be more conservative than average in her future life. From her present point of view, she finds that a future tax increase should be undertaken after the shock only if it is *very* beneficial (i.e., preferred by *more* people than just the future median voter and all people poorer than her). A supermajority requirement guarantees just that. Today's median voter would also like to facilitate tax decreases and so to set the majority necessary for a decrease quite low. However, since changes of laws generally require at least a simple majority, this is the best (for the median voter) which can be achieved.

In section 4, we analyze the robustness of the basic model under more realistic assumptions of recurrent shocks and elections. We also show that states in which the median voter of the median district is only slightly poorer than the median voters in richer districts will employ a higher supermajority rule than states which are more heterogeneous in this regard. The data from US states are consistent with this prediction.

While supermajority rules for tax increases have recently received some interest, the existing literature on this issue is not completely satisfactory. The papers which are most related to the present one are Gradstein (1999), Knight (2000) and Messner and Polborn (1999).<sup>3</sup> Gradstein analyzes a two period model in which a supermajority requirement for an increase of the capital income tax in the second period serves as a commitment device not to increase that tax and thereby to encourage saving. While this model provides a possible explanation for supermajority requirements for capital income tax rate hikes, it does not explain why the states which adopted supermajority requirements typically include *all* taxes in this provision, rather than being constrained to capital income taxes which are the most likely to suffer from time inconsistency problems.

Knight (2000) provides both a theoretical model of supermajority requirements, and an empirical examination of the effects of supermajority rules. At first glance,

---

<sup>3</sup>Other papers relating to supermajorities are Buchanan and Tullock (1962) and Caplin and Nalebuff (1988). Buchanan and Tullock (1962) argue for a unanimity rule as the suitable rule governing social choices because under a simple majority rule, a majority of people could be tempted to change the status quo in a way which is not socially desirable because they can "externalize" part of the cost associated with that decision to the losing minority. However, this argument for rules favoring the status quo cannot explain an asymmetry in rules (e.g., supermajorities for tax increases, simple majorities for tax decreases).

Caplin and Nalebuff (1988) use supermajority rules to prevent cyclical voting behavior in multidimensional elections. However, in our model, voters' preferences are single peaked in a one-dimensional policy space, so a simple majority rule is sufficient to rule out cycles.

supermajority requirements do not seem to have a big impact on tax rates, as average tax rates (in percent of GDP) are the same for states with and without supermajority requirements. However, controlling for observable differences between states and also for differences in political preference, Knight finds that supermajority requirements did significantly reduce tax rates; he estimates that *ceteris paribus*, a supermajority requirement reduces taxes by 8 to 23%.

Knight's theoretical model analyzes a "closed rule" committee system in which the committee chair proposes a new tax rate and no amendments to this proposal are feasible. If the median legislator's preferred tax rate is higher than the status quo tax rate, but lower than that preferred by the committee chair, a supermajority rule can be used to achieve the tax rate preferred by the median legislator. There are some problems with this model. First, a supermajority requirement is not the only institutional arrangement which would work in this setting; more simply, a reorganization of the legislative process such that proposals can be amended, would yield the same result. Second, there is no explanation why we observe supermajority requirements for tax increases, but none for tax decreases; in fact, Knight's model predicts the latter if the median legislator's preferred tax rate is lower than the status quo, but not as low as the committee chair's; there is no a priori reason why this constellation of preferences should be less likely than the reverse. Third, the model predicts that after the introduction of a supermajority rule, we should actually see an *increase* in the tax rate (albeit not as much as without it). Finally, since the model does not appeal to any specific characteristic of decisions on taxation, it cannot explain why we see supermajority rules in taxation, but not in other policy fields.

The modelling technique used in Messner and Polborn (1999) is very related to that used in the present paper. They study an overlapping generations model of voting on investment projects which must be paid by everyone alive at that moment and produce a stream of benefits afterwards. Since old people benefit less from the projects than young ones, preferences in their model change systematically as well, and the median aged voter can use a qualified majority rule (specifically, a 3/4 majority in their basic model) in order to transfer power to people with preferences similar to the ones which today's median voter will have in the future. In contrast to the present paper, which is concerned with sequential decisions on the value of the *same* fundamental policy variable (like a tax rate), Messner and Polborn (1999) analyze the decision on whether to implement investment projects which arise only once, respectively; hence, in their model, it is not possible to meaningfully identify different majority rules for changes in different "directions". In this paper, we will show that the society will vote for a

qualified majority rule for tax *increases* only, while also voting for a simple majority rule for tax decreases.

## 2 The Model

Consider an overlapping generations economy in which people are born on a continuous basis and die after living for one period. There is no population growth, and the population mass is normalized to one. The society has to determine by election the value of a policy variable  $p$ . In addition, the rules which govern a subsequent election on  $p$  (which we call an *issue election*) have to be determined in an initial *procedural election*.

The lifetime utility of an individual is  $\int_0^1 -(p(t) - p^*(t))^2 dt$ , where  $p(t)$  is the policy implemented when the individual is age  $t$  and  $p^*(t)$  is the *instantaneously preferred policy* of an age  $t$  agent; we also refer to  $p^*$  as the agent's "bliss point". We are interested in the case that preferences of individuals over the possible values of  $p$  change depending on age in a monotone way; for concreteness, assume that older people prefer a higher value of  $p$  and that the relationship between the preferred  $p$  and age  $t$  is linear. Without further restriction of generality, we can then assume that the instantaneously preferred policy of an age  $t$  agent equals her age,  $p^*(t) = t$  (this is just a normalization which facilitates notation). Note, however, that in the example of tax rates used in the introduction, the relation between age and preferred tax rate is negative, so that in this case  $p$  should be interpreted as negative linear transformation of the tax rate.<sup>4</sup> Consequently, a *decrease* in  $p$  corresponds to an *increase* in the tax rate, and vice versa.

As an aside, this modelling of preferences is evidently reduced form, but it is easy to construct a richer model which generates bliss points linear in age, and quadratic losses: Suppose that individuals' income is linear and increasing in age and there is a proportional income tax; normalize total income to 1, so that the tax rate  $\tau$  is also equal to the tax revenue. Tax revenue is used to produce  $\tau/c$  units of a public good  $x$ , where  $c$  is the unit cost of a public good. The public good creates benefits according to the marginal benefit function  $\beta - \gamma x$  for all citizens. In this setting, the instantaneously optimal tax rate  $\tau$  for an individual of age  $t$  maximizes  $\int_0^{\tau/c} (\beta - \gamma x) dx - \tau(a + bt)$ ,

<sup>4</sup>Suppose for example that the preferred tax rate  $\tau^*$  equals 40% for the very young and then decreases linearly to 20% for the very old, so  $\tau^*(t) = 0.4 - 0.2t$ . In order to normalize this case, define  $p^* = 2 - 5\tau^*$  as our policy variable; clearly, this function satisfies  $p^*(t) = t$ , and voting on  $p$  is equivalent to directly voting on  $\tau$ .

and hence is

$$\tau = [\beta - (a + bt)] \frac{c}{\gamma}.$$

This is a linear function of  $t$ ; moreover, the loss if the tax rate is different from the instantaneously preferred one is a quadratic function of the difference between the actual and the preferred rate. We refrain from using this richer model since it clutters the notation and does not add much to the central point of this paper. We now return to the simpler, reduced form model.

While individuals' preferences change over time, in the model we specified up to now, the proportion of individuals preferring a specific policy and in particular the median preferred policy remains constant over time, because the age structure of the population does not change. Assume now that there can be a shock which changes all individuals' preferences in the same way. Specifically, the preferred policy after the event has taken place is given by  $p^*(t) = t + \kappa$  where  $\kappa$  is ex ante a random variable which is distributed according to the cumulative distribution function  $F(\cdot)$ ; the corresponding density function is denoted  $f$ . The event is a Poisson event which takes place with frequency  $\lambda$ , so the probability of the event taking place in the next small interval of length  $dt$  (conditional on the event not yet having occurred) is  $\lambda dt$ . Furthermore, we assume that after the event has taken place, no further shocks will occur, i.e. the distribution of preferences then remains fixed forever (of course, *individual* preferences still change with age as described above). The assumption of only one shock enables us to solve the model analytically; in section 4.3, we analyze the more realistic case where shocks continue to occur forever. In the richer model discussed above, the parameter  $\kappa$  could be interpreted either as a shock on preferences (for instance, to  $\beta$ ) or to the costs of public good production,  $c$ .

There are four elections, of which three take place in the beginning. In the *first issue election*, people vote on the present policy,  $p_0$ . In two *procedural elections*, people vote on the "decrease majority"  $M_D$ , which gives the majority that a proposal to decrease the value of the status quo policy must achieve in order to win against the status quo, and on the "increase majority"  $M_I$  (which is defined analogously).<sup>5</sup> We require that the majorities are at least simple majorities:  $M_I, M_D \in [1/2, 1]$ .  $p_0$  remains in effect until the shock has occurred, and then the second issue election is held to determine

---

<sup>5</sup>In order to avoid the problem of multidimensional elections at the beginning of time, we assume that the elections take place sequentially: The first issue election is followed by the procedural election on the increase majority, and last comes the procedural election on the decrease majority. It can be shown that the sequence of the two procedural elections does not matter: The outcome would be the same, independent of whether the people first vote on  $M_I$  and then on  $M_D$  or the other way around.



the future value of the policy variable,  $p_1$ . Figure 1 captures the sequence of elections in this economy.

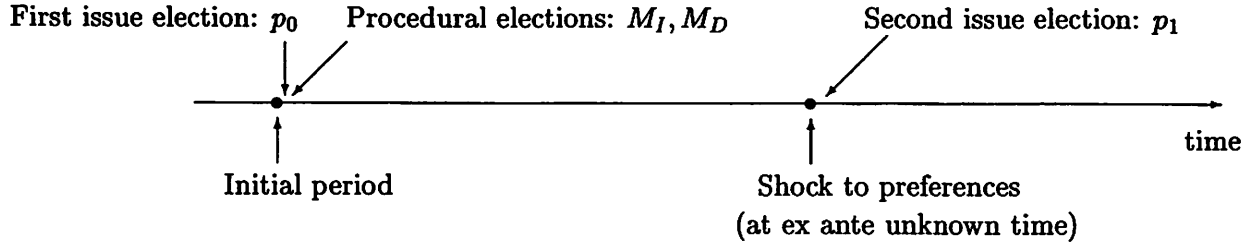


Figure 1: Timing of events

It might seem unusual to have the first issue election *before* the procedural elections; usually, political economy models assume that the choice of the constitution happens in the very beginning, behind a “veil of ignorance”. However, our way of modelling is historically more accurate for modelling real life where constitutions and voting rules are usually drafted by self-interested people with preferences on the issues, not by social planners. Moreover, taxes existed prior to the invention of supermajority requirements for raising taxes, and our modelling sequence also allows us to address an interesting and relevant question: How will the median voter set taxes in the last election in which she is the pivotal voter for increases *and* decreases, given that she knows that the rules for future elections and hence the identity of the pivotal voter for one or the other direction will soon be changed?

Finally, note that in the model, there are no further issue elections between the second procedural and the second issue election and after the second issue election. While this is certainly not realistic, it appears to be a quite innocuous assumption, since in these intervals, neither the preference distribution in the population nor the expected further waiting time for the shock (in the first interval) changes; so even if there were, say, another election some years after the last election considered in the model on whether to change the policy variable  $p_1$ , it appears plausible that there would be no change in  $p_1$  since the distribution of voters’ preferences has not changed. We will discuss what happens if there are more than four elections in more detail in section 4.4.

### 3 Results

We will start this section by analyzing the last election and calculate how the policy chosen,  $p_1$  depends on the majority rules set and the realization of  $\kappa$ . We use this result to get a reduced form expected utility; from that, we determine voting behavior in the procedural elections. Finally, the results of the last issue election and the procedural elections are used to determine the outcome of the initial issue election on the policy variable,  $p_0$ . Results are developed along the way, and Proposition 1 will sum up our results at the end of this section.

Given the realization of  $\kappa$ , the further lifetime utility of an age  $t$  agent if policy  $p_1$  is chosen for the rest of her life is

$$\begin{aligned} \int_t^1 -(p_1 - (\theta + \kappa))^2 d\theta &= \frac{(p_1 - 1 - \kappa)^3 - (p_1 - t - \kappa)^3}{3} \\ &= (1 - t)[p_1(1 + t + 2\kappa) - p_1^2] \end{aligned} \quad (1)$$

Maximizing (1) over  $p_1$  gives the preferred policy

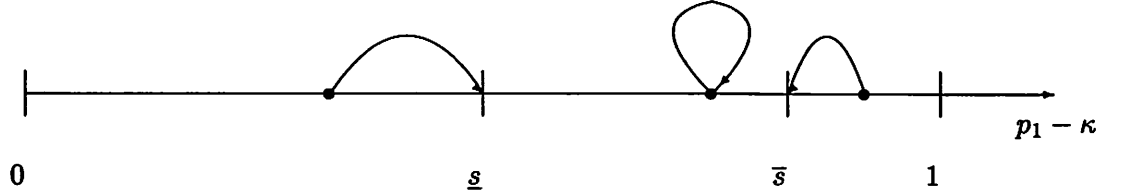
$$p_1^* = \frac{1+t}{2} + \kappa \quad (2)$$

While  $\kappa$  captures a shock to people's preferences, it is actually a more helpful interpretation for deriving results that the shock changed the *effective* policy to  $p_0 - \kappa$ , while preferences over effective policies did not change. The advantage is that *effective* bliss points, as well as the effective policies implemented in equilibrium, all lie in the interval  $[0; 1]$ , independent of  $\kappa$ , and that we will be able to characterize the equilibrium more easily in the space of effective policies.

Writing (2) as  $p_1^* - \kappa = \frac{1+t}{2}$ , it can be interpreted as follows: The best *effective* policy for an age  $t$  individual, given that this policy will remain constant for the rest of her lifetime, is the one which is the instantaneous bliss point of another agent who now has the first agent's average future age,  $(1+t)/2$ . As expected, (2) is increasing in  $t$ ; older people prefer a higher effective level of  $p_1$ . As people are ordered with respect to their preferences according to their age, a majority rule assigns the position of a pivotal voter and through this, the majority rule will influence the result of the second issue election.

In principle, we could now determine how majority rules influence the second issue election, and substitute these results into the expected utility of the median voter at the time of the procedural elections. However, it turns out that an indirect approach is more instructive. Every majority rule for an increase,  $M_I$ , corresponds to a threshold in the space of effective policies, which we call  $\underline{g}$  with the following property: If, after

the shock, the effective policy is below  $\underline{s}$ , it will be increased to  $\underline{s}$ .<sup>6</sup> Similarly, every majority rule for a decrease,  $M_D$ , corresponds to a value  $\bar{s}$  such that, if after the shock the effective policy is above  $\bar{s}$ , it will be decreased to  $\bar{s}$ . Finally, there is a range of inactivity: If the effective policy after the shock is between  $\underline{s}$  and  $\bar{s}$ , there is neither a sufficient majority for an increase nor for a decrease and so the initial policy is kept. All this is illustrated in Figure 2.



**Figure 2:** Mapping effective policies after the shock into effective ex post policies

Formally, we define  $(\underline{s}, \bar{s})$  such that the second issue election yields an effective policy of

$$p_1 - \kappa = \begin{cases} \underline{s} & \text{if } p_0 - \kappa < \underline{s} \\ p_0 - \kappa & \text{if } p_0 - \kappa \in [\underline{s}, \bar{s}] \\ \bar{s} & \text{if } p_0 - \kappa > \bar{s} \end{cases} \quad (3)$$

Which majority rules are necessary to implement  $\underline{s}$  and  $\bar{s}$ ? We assume that  $\underline{s}$  and  $\bar{s}$  are the preferred policies of the pivotal voters for an increase and a decrease of  $p$ , respectively. Specifically,  $\underline{s}$  is the effective policy preferred by the voter aged  $1 - M_I$  ( $\underline{s} = \frac{2 - M_I}{2}$ );<sup>7</sup> and  $\bar{s}$  is the effective policy preferred by the voter aged  $M_D$  ( $\bar{s} = \frac{1 + M_D}{2}$ ). Since the majorities have to be at least simple majorities, we have  $\underline{s} \leq 3/4$  and  $\bar{s} \geq 3/4$ .

As already recognized by Gradstein (1999), the legislative bargaining model due to Baron (1996) provides a possible justification of this assumption.<sup>8</sup> Suppose that after the shock, a voter is chosen at random and proposes a new policy; if approved by a sufficient majority ( $M_I$  in case the proposed policy is an increase in comparison to the status quo, and  $M_D$  for the case of a proposed decrease), the proposal becomes the

<sup>6</sup>As expected,  $\underline{s}$  is determined to be the effective policy preferred by the voter who is pivotal in an election on an increase proposal. We will discuss this in more detail shortly.

<sup>7</sup>If the voter aged  $1 - M_I$  favors an increase of the effective policy from  $p_0 - \kappa$  to  $\underline{s}$  over the status quo, then all voters who are older will also favor this change over the status quo. A symmetric argument applies for the case of a decrease.

<sup>8</sup>Baron (1996) deals with the case of a simple majority rule, but his methods readily extend to the case of a supermajority rule.

new policy; otherwise, after a short time span, a new voter is chosen at random and can propose a new election against the status quo and so on. It follows from Baron's analysis that the outcome of this game is given by (3) with  $\underline{s} = \frac{2-M_I}{2}$  and  $\bar{s} = \frac{1+M_D}{2}$ . An alternative model of the legislative process with a supermajority rule which also yields the result that the pivotal voter's preferred policy is implemented is Gradstein's (1999) own model.

Equation (3) provides us with a reduced form of the election after the shock has occurred; we will use it for calculating the continuation utility of voters, which we will use in the analysis of earlier elections. Let  $Z(t, p_0, \underline{s}, \bar{s})$  denote the expected future utility of an age  $t$  individual if the shock takes place now and the following election results are determined by  $(p_0, \underline{s}, \bar{s})$ . Combining (1) and (3) yields

$$Z(t, p_0, \underline{s}, \bar{s}) = [1 - F(p_0 - \underline{s})] \frac{(t - \underline{s})^3 - (1 - \underline{s})^3}{3} + F(p_0 - \bar{s}) \frac{((t - \bar{s})^3 - (1 - \bar{s})^3)}{3} + \int_{p_0 - \bar{s}}^{p_0 - \underline{s}} [(t - p_0 + \kappa)^3 - (1 - p_0 + \kappa)^3] f(\kappa) d\kappa \quad (4)$$

Define  $W(t; \underline{s}, \bar{s}, p_0)$  as the expected future utility of an individual aged  $t$  and living at the beginning of time, when the elections on the initial policy  $p_0$  and the procedural elections determining  $\underline{s}$  and  $\bar{s}$  take place. Suppressing all but the first argument (time) in  $W(\cdot)$  and  $Z(\cdot)$ ,  $W$  must satisfy

$$\dot{W}(t) = -(p_0 - t)^2 dt + \lambda Z(t + dt) dt + (1 - \lambda dt) W(t + dt) \quad (5)$$

because of the following consideration: In the next short interval of length  $dt$ , an immediate utility of  $-(p_0 - t)^2$  is realized; moreover, with probability  $\lambda dt$ , the shock takes place and the expected further utility after a shock is  $Z(t + dt)$ ; with probability  $(1 - \lambda dt)$ , no shock takes place and then the individual has an expected further utility of  $W(t + dt)$ . Dividing by  $dt$  and letting  $dt$  approach 0, we get the differential equation

$$\dot{W}(t) = \lambda W(t) + (p_0 - t)^2 - \lambda Z(t) \quad (6)$$

where  $\dot{W}(t)$  denotes the time derivative of  $W$ . The boundary condition is  $W(1) = 0$ : An individual who is about to die must have a future utility equal to zero. Solving the differential equation yields

$$W(t, p_0, \underline{s}, \bar{s}) = \int_t^1 [\lambda Z(u; p_0, \underline{s}, \bar{s}) - (p_0 - u)^2] e^{-\lambda(u-t)} du \quad (7)$$

Optimization with respect to  $\bar{s}$  gives as a first order condition

$$\frac{\partial W}{\partial \bar{s}} = \int_t^1 \lambda \frac{\partial Z(u; p_0, \underline{s}, \bar{s})}{\partial \bar{s}} e^{-\lambda(u-t)} du = 0 \quad (8)$$

Using

$$\begin{aligned}
\frac{\partial Z(t)}{\partial \bar{s}} &= -f(p_0 - \bar{s}) \frac{(t - \bar{s})^3 - (1 - \bar{s})^3}{3} + \\
&\quad F(p_0 - \bar{s})[(1 - \bar{s})^2 - (t - \bar{s})^2] + f(p_0 - \bar{s}) \frac{(t - \bar{s})^3 - (1 - \bar{s})^3}{3} \\
&= F(p_0 - \bar{s})(1 - t)[1 + t - 2\bar{s}],
\end{aligned} \tag{9}$$

and solving the first order condition (8) for  $\bar{s}$  we get

$$\bar{s}(t, \lambda) = \frac{\int_t^1 (1 - u^2) e^{-\lambda(u-t)} du}{2 \int_t^1 (1 - u) e^{-\lambda(u-t)} du} \tag{10}$$

as the preferred level of  $\bar{s}$  for an agent aged  $t$ .<sup>9</sup> Equation (10) does not depend on  $p_0$  and  $\underline{s}$ , which is a very useful result as it indicates that the election on  $\bar{s}$  is independent of the two preceding elections.

Next, we will establish that the preferred  $\bar{s}$  is an increasing function of age:

$$\frac{\partial \bar{s}}{\partial t} = \frac{2(1 - t) \int_t^1 e^{-\lambda(u-t)} (1 - u)(u - t) du}{[\int_t^1 2e^{-\lambda(u-t)} (1 - u) du]^2} > 0, \tag{11}$$

as claimed. Consequently, the median voter in the election on  $\bar{s}$  is the person aged  $1/2$  and the winning proposal in the election on  $\bar{s}$  is

$$\bar{s}(1/2, \lambda) = \frac{\int_{1/2}^1 e^{-\lambda(u-1/2)} (1 - u)(1 + u) du}{\int_{1/2}^1 2e^{-\lambda(u-1/2)} (1 - u) du} du \tag{12}$$

This function is decreasing in  $\lambda$ ,<sup>10</sup> and is drawn in Figure 3. Intuitively, the higher is  $\lambda$ , the earlier is the shock expected to take place and the younger is the present median voter at the time of the shock, hence the smaller is the  $\bar{s}$  the present median voter would like to implement. For  $\lambda \rightarrow 0$ ,  $\bar{s}(1/2, \lambda)$  converges to  $5/6$ ; this translates into a  $2/3$  majority, as  $(1 + t_P)/2 = 5/6$  can be solved for the age of the pivotal voter,  $2/3$ . Note that the expected age of the median voter at the time of the shock (conditional on the shock occurring during the present median voter's lifetime) is  $3/4$ , not  $2/3$ . However, since  $Z$  is nonlinear, it is not optimal for the present median voter to transfer power to the voter who is aged  $3/4$ .<sup>11</sup>

<sup>9</sup>One can show that the second order condition is satisfied; see appendix.

<sup>10</sup>See appendix.

<sup>11</sup>Consider an increase of  $\bar{s}$  from  $5/6$ . This change will hurt the present median voter if she happens to be younger than  $2/3$  at the time of the shock, otherwise (i.e. in the majority of cases) she will benefit. However, the decrease in utility if the shock happens while she is still relatively young is quite large, because she still has to live quite long; on the other hand, the increase in utility due to the change if the shock happens late in life is likely small because it applies only for a very short period of time. Of course, since we started in the optimum ( $\bar{s} = 5/6$ ), both effects cancel out exactly.

Why does the median voter transfer the power to be the pivotal voter in the next election to someone who is older than she is? The reason is that when the shock arrives sometime in the future, the present median voter will be older as well, and will then prefer a higher value of the policy variable. The limit of  $\bar{s}(1/2, \lambda)$  for  $\lambda \rightarrow \infty$  is  $3/4$  which corresponds to a simple majority rule for decreases in  $p$  (because  $\frac{1+t}{2} = 3/4$  for  $t = 1/2$ ) and is intuitively clear: If the median aged voter expects the shock to happen very soon, there is no point in transferring the power in the next election to anyone else. However, we will see in section 4.3 that this is an artefact of the assumption that there is only one shock.<sup>12</sup> This assumption is particularly unappealing for  $\lambda$  to infinity, so this result should be interpreted with caution. On the other hand, for  $\lambda$  small, the assumption that there is only one shock is quite innocuous and even if shocks were recurring, the majority rule  $M_D$  chosen would be very close to  $2/3$ .

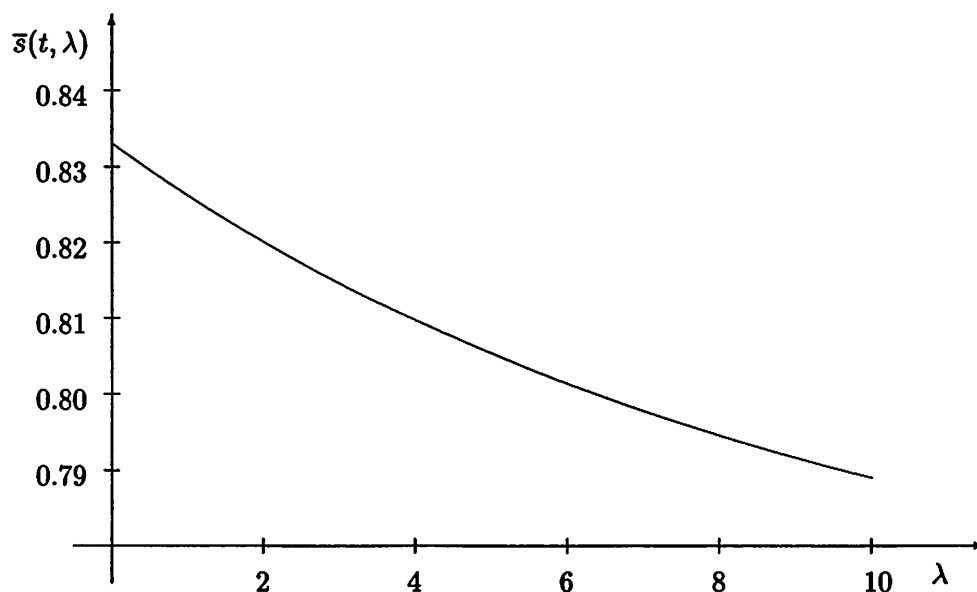


Figure 3

Next, we move to the first procedural election and analyze which value of  $\underline{s}$  is chosen there. Differentiation of  $W$  with respect to  $\underline{s}$  yields

$$\frac{\partial W}{\partial \underline{s}} = \int_t^1 \lambda \frac{\partial Z(u; p_0, \underline{s}, \bar{s})}{\partial \underline{s}} e^{-\lambda(u-t)} du. \quad (13)$$

Since also  $\partial Z / \partial \underline{s} = F(p_0 - \underline{s})(1 - t)[1 + t - 2\underline{s}]$ , the first order condition for  $\underline{s}$  would be the same as the one for  $\bar{s}$ , (8). The intuition for this result is that the median voter would like to transfer *power* in the next issue election to someone who is older than

<sup>12</sup>We will see there that in a model with recurring shocks, an increase in  $\lambda$  actually *increases*  $M_D$ .

she (presently) is, and it does not matter whether the change should be an increase or decrease. However, implementing  $\underline{s} = 5/6$  would mean that only one third of the population would have to vote for an increase of  $p$  in order for it to come into effect. If we assume that at least a simple majority has to vote for any change of the status quo to take place, then the best the median voter in the first procedural election can specify is a simple majority rule for proposals to increase  $p$ .

As mentioned above, an “increase” in our model is the direction of change which favors old people. In the important application of voting on tax rates, the relationship between age and preference is the reverse, and so our results explain why we observe a supermajority requirement for tax *increases* and a simple majority requirement for tax *decreases*.

To complete the analysis of the basic model, let us characterize the initially chosen policy  $p_0$ . Optimization of  $W$  with respect to  $p_0$  yields

$$\frac{\partial W}{\partial p_0} = \int_t^1 \left[ \lambda \frac{\partial Z(u)}{\partial p_0} - 2(p_0 - u) \right] e^{-\lambda(u-t)} du = 0 \quad (14)$$

Substituting  $t = 1/2$  for the age of the median voter and

$$\partial Z(u)/\partial p_0 = (1 - u)[(F(p_0 - \underline{s}) - F(p_0 - \bar{s})) (1 + u - 2p_0) + \int_{p_0 - \bar{s}}^{p_0 - \underline{s}} 2\kappa f(\kappa) d\kappa,$$

we get

$$\begin{aligned} \frac{\partial W}{\partial p_0} = \int_{1/2}^1 \{ & \lambda(1 - u)[(F(p_0 - \underline{s}) - F(p_0 - \bar{s})) (1 + u - 2p_0) + \\ & \int_{p_0 - \bar{s}}^{p_0 - \underline{s}} 2\kappa f(\kappa) d\kappa - 2(p_0 - u) \} e^{-\lambda(u-t)} du = 0 \end{aligned} \quad (15)$$

as an implicit equation which determines the outcome of the first issue election. While (15) cannot be solved for  $p_0$  without making additional assumptions concerning the distribution of  $\kappa$ , we can compare this choice for  $p_0$  with what would have been chosen if the society were constrained to use a simple majority rule in the second issue election. This will show whether there is an immediate effect on *present* policy if a supermajority requirement is introduced. From the implicit function theorem, we know that the sign of  $dp/d\bar{s}$  is the same as the sign of

$$\frac{\partial^2 W}{\partial p_0 \partial \bar{s}} = \int_{1/2}^1 \lambda(1 - u) f(p_0 - \bar{s}) [1 + u - 2\bar{s}] e^{-\lambda(u-t)} du. \quad (16)$$

This expression has the same sign as the derivative of  $W$  with respect to  $\bar{s}$  (cf. (8) and (9)) and hence is positive for all  $\bar{s} \in [3/4; \bar{s}(1/2, \lambda)]$ , where  $\bar{s}(1/2, \lambda)$  is the winning proposal in the procedural election given in (12).

Hence the introduction of a supermajority requirement for decreasing  $p$  after the shock will already influence the choice of  $p_0$  today; in particular,  $p_0$  will be higher than it would have been with a simple majority rule. In the tax application this means that, with a supermajority rule, taxes will *decrease* already prior to that institutional change, in comparison to the equilibrium level of taxes if there were no supermajority requirement for future elections.

Intuitively, from the point of view of the median voter, the policy  $p_0$  chosen has an investment property: Increasing  $p_0$  causes an immediate cost (because  $p_0$  is higher than the present bliss point of the median voter), but there is also an expected benefit in the future when (if  $p_0$  is still in effect) the present median voter will benefit from a higher level of  $p$ . The optimal  $p_0$  balances these two effects. The intuition for the effect of the supermajority rule on  $p_0$  is now as follows: Under a simple majority rule,  $p_0$  is in force only until the shock happens; the value of  $p_1$  chosen afterwards is the level preferred by the then median voter and hence completely independent of  $p_0$ . Therefore, the median voter will just choose the  $p$  which is optimal (in expectation) until the shock occurs. On the other hand, under a supermajority requirement, the effective policy  $p_0 - \kappa$  might fall between  $\underline{p}$  and  $\bar{p}$ ; in this case, the initial choice of  $p_0$  continues to determine the effective policy *even after the shock*. Hence the expected phase during which the median voter will benefit from fixing a high value of the policy variable is *longer*, while the immediate costs of fixing a high value of the policy variable stay the same; consequently, the optimal  $p_0$  increases if there is a supermajority requirement for future elections.

This result is particularly interesting in comparison to Knight (2000). In Knight's model, the median legislator votes for a supermajority requirement in order to secure that the subsequent tax *increase* is positive, but not too big. On the other hand, our model predicts an immediate tax decrease before the supermajority rule is introduced. The enthusiastic support of anti-tax advocacy groups for supermajority rules seems more consistent with our model.<sup>13</sup>

We gather all our results in this section in the following Proposition 1:

**Proposition 1**    1. *In the initial issue election, the winning policy proposal  $p_0$  satisfies (15). This value is higher than what would be chosen if future elections were governed by simple majority rules.*

2. *In the initial procedural elections, the winning proposals specify a simple majority rule for future increases in  $p$  ( $M_I = 1/2$ ), and a supermajority rule for decreases*

<sup>13</sup>See, e.g., the "Americans for Tax Reform" website, [www.atr.org](http://www.atr.org).



in  $p$ :  $M_D = 2\bar{s} - 1$ , where  $\bar{s}$  is given by (12).  $M_D$  is independent of the initial policy value  $p_0$ , a decreasing function of  $\lambda$  and lies between  $1/2$  and  $2/3$ .

3. Depending on the realization of  $\kappa$ , the effective policy chosen in the issue election after the shock is given by (3).

## 4 Extensions

The basic model has the advantage that it is very tractable, but this comes at some expense of realism. It is therefore important to analyze what changes if we alter two quite restrictive assumption of the basic model. First, while “age” is one factor influencing political preferences and hence voting behavior, there are certainly other factors as well. Sections 4.1 and 4.2 address the question of heterogeneity within the electorate, and also study the implementation in a representative democracy in which voters elect a legislature and then the legislature votes on the policy variable. Second, in section 4.3, we analyze the case that shocks occur not only once, but repeatedly. Finally, in section 4.4, we will look at what happens if elections in this society are held on a continuous basis rather than only in the beginning and after the shock occurred.

### 4.1 Representative democracy

Suppose there is a continuum of heterogeneous electoral districts indexed by  $\delta$ , and in each electoral district there is a continuum of agents whose preferences vary with age just as in the basic model. The index  $\delta$  describes a deviation in the preferences of *all* voters of the electoral district<sup>14</sup> from the preferences of voters in the median electoral district, for example because the electoral district is richer than average. The immediate utility of a voter aged  $t$  in electoral district  $\delta$  is  $-(\delta + t - p)^2$ , and the median electoral district (the median  $\delta$ ) is zero (this is just a normalization). Let  $D$  be the cumulative distribution function of  $\delta$ , and let  $D^{-1}$  be its inverse.

Consider the election for a legislature after the shock took place. Within district  $\delta$ , people elect by simple majority a representative who has preferred policy  $\frac{1+1/2}{2} + \delta$ , by the median voter theorem.<sup>15</sup> Consider the behavior of representatives in the legislature in a vote on a proposal to increase  $p$ ; as above, the median legislator is the pivotal voter for a proposed increase. If a proposal to decrease  $p$  needs a supermajority  $M_D$  in the

<sup>14</sup>We will analyze the case of additional, non-age related heterogeneity *within* electoral districts in the next subsection.

<sup>15</sup>It is easy to see that voters have no incentive in this situation to vote strategically for someone who does not represent their true preferences, hence the median voter theorem applies.

legislature to pass, the representative from electoral district  $D^{-1}(M_D)$  is the pivotal voter for that proposal.

We will now turn to the determination of  $M_D$  in the procedural election. Since the preferences of the median voter in the median electoral district have not changed in comparison to the basic model, we know that she would like to implement the same  $(\underline{s}, \bar{s})$  mechanism as in the basic model. Again,  $\underline{s}$  is implemented by a simple majority rule for increases of the policy variable. In order to implement  $\bar{s}$ , the majority rule for a decrease must make the representative who prefers an effective policy of  $\bar{s}$  the pivotal voter:  $\frac{1+1/2}{2} + \delta_{Piv} = \bar{s}$ , so

$$M_D = D\left(\bar{s} - \frac{3}{4}\right). \quad (17)$$

If the electoral districts are very homogeneous so that the maximal value  $\delta$  can take is smaller than  $\bar{s} - \frac{3}{4}$ , then the majority rule chosen will be a unanimity rule. On the other hand, more heterogeneity among electoral districts leads to a decrease of the supermajority requirement:

**Proposition 2** *Suppose  $\delta$  is uniformly distributed on  $[-\delta_{max}; \delta_{max}]$ . In this case, the majority rule chosen in the procedural election is  $M_D = \min\left(1, \frac{1}{2} + \frac{\bar{s} - 3/4}{2\delta_{max}}\right)$ . More heterogeneity among electoral districts decreases the majority rule chosen:  $\frac{\partial M_D}{\partial \delta_{max}} < 0$ . Moreover,  $\lim_{\delta_{max} \rightarrow \infty} M_D = 1/2$ .*

The proof follows immediately from substituting the uniform distribution function  $D(x) = \frac{x + \delta_{max}}{2\delta_{max}}$  in (17).

The intuitive explanation for the relation between homogeneity and the majority rule is as follows. As in the basic model, the median voter in the median electoral district (henceforth, the “median median” voter) would like to transfer power in future elections to some other voter whose preferences correspond to some appropriately defined average over the future preferences of the median median voter. In selecting the pivotal future voter, the median median voter is constrained in representative democracy to choose the median voter of another electoral district.

In a very homogeneous state, the median voters in the other electoral districts are very similar to the median median voter, and in order to have the desired effect, the majority rule must specify a quite high majority. In the other extreme case, if electoral districts are very different from each other, the median median does not expect that his preferences in the future will resemble those of the median voter in any other electoral districts, and hence should prefer not to transfer power.

This result could explain why several US *states* have introduced qualified majority rules for tax increases while the same proposal failed to win support in other states and at the federal level. The states which introduced the supermajority rule could be internally more homogeneous than the United States as a whole, and than those states which rejected supermajority rules.

Differences in preferences between electoral districts are hard to measure directly, but a variable which is probably correlated with the median preferences of an electoral district over tax rates is the median income in that district. If the median incomes in different electoral districts of a state are similar to each other, it is probable that their preferences are also quite similar, and therefore we would expect to see more supermajority rules in such states than in very heterogeneous states.

To test this prediction, we use US Census data for the median income in all US counties.<sup>16</sup> We measure the difference between the states' distribution functions  $D_i$  through what we call the *homogeneity ratio*. This is defined as the ratio between the median income in the "2/3 county" (i.e., the county with the property that 2/3 of the population live in counties with a lower median income), and the median income in the median county.<sup>17</sup> The nearer this ratio is to 1, the more "homogeneous" is the state according to this measure, and the more likely we are to see supermajority rules in homogeneous states, according to proposition 2.

For a few states in which a large fraction of the population lives in one county or where there are only very few counties, our calculation procedure cannot produce reliable results; if the median county and the 2/3 county coincided, we dropped a state.<sup>18</sup> We also exclude from the data set three southern states that introduced supermajority very early (Arkansas in 1934, Louisiana in 1966, Mississippi in 1970),<sup>19</sup> the latter two "following the Voting Rights Act of 1965, presumably to protect the status quo tax

---

<sup>16</sup>Unfortunately, data on median income in state electoral districts (rather than counties) were not available, but county data are probably a good proxy, at least in states which do not have an extreme concentration of the population in one county; for more on the latter problem, see footnote 18.

<sup>17</sup>To be exact, for the median income in the median district, we use a linear interpolation between the median income of the first county such that more than half of the population live in this and poorer counties, and the median income of the county immediately preceding this one, to calculate the median income in the median county. The "2/3 county" (which is the county such that two thirds of the population live in counties with a lower median income) is determined analogously.

<sup>18</sup>This happened in Alaska, Arizona, Delaware, Hawaii, Nevada, New Hampshire, Rhode Island and Utah.

<sup>19</sup>However, we leave Florida which has introduced a first supermajority rule in 1971, applying only to the corporate income tax, in the data set, because it introduced a general supermajority rule for tax increases in 1996.

rates in the face of predictable changes in the electorate” (Knight (2000), p. 43). Since this motivation is different from that analyzed in the present paper, and moreover it is also unclear whether present income data reflect income distribution at the time of the introduction of supermajority rules in these states, we dropped them from our data set.

This leaves us with data from 39 states, and one observation from the US as a whole.<sup>20</sup> Table 1 gives the homogeneity ratio for the six biggest states and for the US as a whole. Observe that the two most homogeneous of these big states have a supermajority rule, while the other ones do not, a result which is consistent with the prediction of Proposition 2.

state	homogeneity ratio	supermajority
CA	1.0397	YES
FL	1.0591	YES
IL	1.1001	NO
NY	1.1363	NO
PA	1.093	NO
TX	1.0613	NO
USA	1.1144	NO

Table 1: Homogeneity ratios for selected states

Does it quantitatively make sense that some states adopt supermajority rules while others do not, given the order of magnitude of these homogeneity ratios? That is hard to tell, for two reasons: First, the “real” median voter and her expectations about future earnings are not easily identifiable in the data.<sup>21</sup> Second, even if we knew today’s median voter’s future earnings profile, we would still have the problem that income heterogeneity is only a proxy for heterogeneity in preferences between electoral districts; if factors other than income also influence preferences over tax rates, the (income) homogeneity ratio is likely to underestimate the heterogeneity in preference between districts.

<sup>20</sup>The US income data were calculated from median income data for house of representatives districts (rather than counties).

<sup>21</sup>If preferences depended only on income and age, we could check consistency as follows: A 2/3 majority rule in the basic model corresponds to a transfer of power from the median voter to a voter who is  $(2/3) - (1/2) = 1/6$  older. If we consider a lifetime of 50 years (as a voter), this means that the median voter would like to transfer power to someone who is now as she will be in about 8 years. If we knew the median voter’s expected income in 8 years, as compared to the median income then, we could compare this ratio with the homogeneity ratio.

Table 2 reports the results which arise when we group states according to whether they were among the 10 most “homogeneous” states (those with the lowest ratio) or not. Four out of the ten most homogeneous states, but only three out of the remaining thirty, had passed supermajority requirements. A  $\chi^2$  test for independence between the homogeneity ratio and the supermajority decision rejects the null hypothesis of independence at the 5% significance level which again can be interpreted as support for our model.

	supermajority	no supermajority
10 most homogeneous states	4	6
other 30 states	3	27

Table 2: Homogeneity ratios and supermajority requirements

## 4.2 Non-age related heterogeneity within electoral districts

In the last subsection, we have introduced heterogeneity among electoral districts, but *within* each electoral district, preferences were only determined by age. Now, we introduce additional non-age related heterogeneity within each electoral district.

Let us assume that, in every electoral district  $\delta$ , there is a continuum of types indexed by  $\gamma$ , such that the preferred policy of type  $(t, \gamma, \delta)$  is  $t + \gamma + \delta$ ; The  $\gamma$ -type of an individual is constant throughout her lifetime, and  $\gamma$  is distributed in all electoral districts according to density  $g$  and distribution function  $G$ . Furthermore, assume that  $\gamma$  is distributed symmetrically around 0.

Proceeding as in the basic model, it is easy to see that in the issue election after the shock, the preferred policy of type  $(t, \gamma, \delta)$  is  $\frac{1+t}{2} + \gamma + \delta$ . Hence electoral district  $\delta$  will elect a representative with preference for  $\frac{1+t}{2} + \delta$ , so in the second issue election, there is no difference to the results derived above. Again, equation (17) will determine how society’s initial choice of  $\bar{s}$  can be implemented by a supermajority rule in the legislature.

How does the initial choice of  $\bar{s}$  change with additional heterogeneity? In the initial procedural election, the median electoral district is decisive, so we analyze that election within the median electoral district and set  $\delta = 0$ . Solving a differential equation which is analogous to (5), we get that if a majority rule implementing  $(\underline{s}, \bar{s})$  and an initial policy  $p_0$  is chosen, the continuation utility of type  $(t, \gamma)$  is given by

$$\int_t^1 [\lambda Z(u; p_0, \underline{s}, \bar{s}, \gamma) - (p_0 - u)^2] e^{-\lambda(u-t)} du, \quad (18)$$

where  $Z(u; p_0, \underline{s}, \bar{s}, \gamma)$  is the expected continuation utility of a type  $\gamma$  individual if the shock happens when the individual has age  $u$  (analogous to  $Z$  as defined in the basic model):

$$\begin{aligned} Z(t, p_0, \underline{s}, \bar{s}, \gamma) = & [1 - F(p_0 - \underline{s})] \frac{(t + \gamma - \underline{s})^3 - (1 + \gamma - \underline{s})^3}{3} + \\ & F(p_0 - \bar{s}) \frac{((t + \gamma - \bar{s})^3 - (1 + \gamma - \bar{s})^3)}{3} + \\ & \int_{p_0 - \bar{s}}^{p_0 - \underline{s}} [(t + \gamma - p_0 + \kappa)^3 - (1 + \gamma - p_0 + \kappa)^3] f(\kappa) d\kappa \end{aligned} \quad (19)$$

Differentiating (18) with respect to  $\bar{s}$  and solving for  $\bar{s}$  yields the preferred level for type  $(t, \gamma)$ ,  $\bar{s}_{het}$ :

$$\bar{s}_{het}(t, \lambda) = \frac{\int_t^1 (1 - u^2) e^{-\lambda(u-t)} du}{2 \int_t^1 (1 - u) e^{-\lambda(u-t)} du} + \gamma = \bar{s}(t, \lambda) + \gamma \quad (20)$$

The number of people who want a level of  $\bar{s}$  which is at least as high as a proposal  $S$  is

$$\int_0^1 \int_{S - \bar{s}(t, \lambda)}^{\infty} g(\gamma) d\gamma dt = \int_0^1 1 - G(S - \bar{s}(t, \lambda)) dt \quad (21)$$

where  $\bar{s}(t, \lambda)(t)$  is the preferred level of  $\bar{s}$  for an age  $t$  individual in the basic model, given in (10). The winning proposal  $S = \bar{s}_{het}$  satisfies that the expression given in (21) equals 1/2.

The following proposition shows conditions under which the equilibrium level of  $\bar{s}$  is smaller in the model with heterogeneous preferences than in the basic model. The condition in part 1 is the more general one (it is satisfied whenever the condition in part 2 of the proposition is), but the condition in part 2 is easier to check.

Part 3 gives an explicit formula for the case of a uniform distribution of  $\gamma$ ; the remarkable property of this formula is that it is the same for *all* uniform distributions on  $[-h; h]$ , independent of  $h$ . In particular, it shows that  $\lim_{h \rightarrow \infty} \bar{s}_{het}(t, \lambda) \neq 3/4$ : Even if age is relatively unimportant as determinant of voting behavior, compared with “other” factors captured as  $\gamma$ , the median voter still has an incentive to change the voting rules in order to implement a different policy after the shock than would arise with a simple majority rule.

**Proposition 3** 1. *If  $1 - G(S - \bar{s}(t, \lambda)(t))$  is strictly concave in  $t$ , then,  $\bar{s}_{het}(t, \lambda) < \bar{s}(t, \lambda)$ .*

2. *If the distribution function  $G$  is a convex function, then  $\bar{s}_{het} < \bar{s}(t, \lambda)$ .*

3. *If  $\gamma$  is uniformly distributed on  $[-h; h]$ , with  $h \geq 1/3$ , then the winning proposal satisfies  $S = \int_0^1 \bar{s}(t, \lambda) dt$ .*

Proof: See appendix.

The most surprising part of this proposition is probably part 3: Even if the  $h$  goes to infinity, the equilibrium value of  $\bar{s}$  does not go to the simple majority equivalent of 3/4. While the equilibrium value of  $\bar{s}$  is lower in the case of intra-constituency heterogeneity, numerical simulation suggests that the difference is not very large,<sup>22</sup> especially taking into account that  $h$  can be arbitrary and hence the analysis remains basically unchanged even if age (and factors correlated with age) is just one factor determining the political preferences and perhaps not even the most important one.

Of course, if we tried to *implement*  $s_{het}$  in a direct democracy (rather than the representative system analyzed here), the necessary majority rule would go to 50% as  $h$  goes to infinity. However, in a representative democracy, the main factor which influences the majority rule chosen is heterogeneity *between the median voters in different electoral districts*, but not so much the heterogeneity of voters *within* electoral districts.

### 4.3 An infinite sequence of shocks

Our basic model assumes that there are no further shocks after the first one; this assumption has the benefit that it makes the model analytically tractable by backwards induction. However, it would be more realistic to have an infinite number of shocks in the following sense: Assume that after the first shock has occurred, the second shock is again a Poisson event with parameter  $\lambda$ , and so on; there is never a “last” shock. In the beginning, people choose the initial policy and the majority rules for future elections. Afterwards, they vote after each shock whether to change the present policy. We will refer to this new model variant as the “infinite shocks model”, and to the basic model as the “one shock model”. In order to limit the complexity, we return in this section to the assumption that only age affects preferences and that policies are determined by referenda.

In a slight abuse of notation, let  $W(t, p, \underline{s}, \bar{s})$  again denote the utility of an individual aged  $t$  if the present policy is  $p$  and majority rules in future elections are such that  $(\underline{s}, \bar{s})$  is implemented. The analogue to (5) is

$$\begin{aligned}
 W(t, p, \underline{s}, \bar{s}) = & -(p - t)^2 dt + (1 - \lambda dt)W(t + dt, p, \underline{s}, \bar{s}) + \\
 & \lambda dt \left\{ [1 - F(p - \underline{s})]W(t, \underline{s}, \underline{s}, \bar{s}) + F(p - \bar{s})W(t, \bar{s}, \underline{s}, \bar{s}) \right. \\
 & \left. + \int_{p-\bar{s}}^{p-\underline{s}} W(t, p - \kappa, \underline{s}, \bar{s}) f(\kappa) d\kappa \right\}
 \end{aligned} \tag{22}$$

---

<sup>22</sup>For example, if  $\lambda = 5$ , without heterogeneity within electoral districts, the  $\bar{s}$  chosen is about 0.805. With heterogeneity (captured by any uniform distribution), it is 0.800.

Note that the (explicit)  $Z$  functions in (5) are replaced by  $W$  functions in (22). This makes the model considerably harder to solve.

As in the basic model,  $\underline{s}$  is determined by the preferred effective policy of a median aged voter after a shock. In the 1-shock model, people know that after the first shock there will be no further one, and hence the median voter in the second issue election votes for an effective policy which minimizes her preference costs over the remainder of her life time:  $\underline{s} = 3/4$ . In the infinite shock model, the median voter knows that the present issue election may not be the last one during her life, and therefore today's choice applies (in expectation) for a shorter span of time, during which the present median voter is on average younger than  $3/4$ ; hence, we would expect that the immediate preferences will receive more weight. Consider the limit case of  $\lambda \rightarrow \infty$ ; there, it is known to the median voter that today's decision will be changed tomorrow, so the best she can do today is to vote for her present bliss point, which yields  $\underline{s} = 1/2$ . In general, we would expect that  $\underline{s}$  is decreasing in  $\lambda$ .

In the basic model,  $\bar{s}$  is the effective policy an individual who has age  $1/2$  at the time of the initial elections would like to see implemented in the second issue election. Very similarly, in the infinite shocks model,  $\bar{s}$  is the effective policy an individual who is median aged at the initial election would like to see chosen in *all* future issue election. For the same  $\lambda$ , the average age of the median voter of the procedural election at an issue election which follows the first one is higher in the infinite shocks model: The average age at the second issue election is the same in both models, but in the infinite shocks model, there can be further elections as well. We would therefore expect that the optimal  $\bar{s}$  is higher than in the one shock model.

Consider now what happens to the majority  $M_D$  which is required to implement  $\bar{s}$ . As we have argued in the penultimate paragraph above, voters will act more "impatiently" in the infinite shocks model as they know that the present election determines the policy only until the next shock happens. Hence we would expect that  $M_D$  is greater in the infinite shocks model, for the following two reasons: The optimal  $\bar{s}$  will be greater, and in order to implement *any given*  $\bar{s}$ , a higher supermajority requirement is needed.

Unfortunately, an analytic, closed form solution could not be derived for the infinite shocks model, so we have to use a Monte Carlo simulation to gain some insights into the properties of the equilibrium. Technical details of the computations can be found in the appendix. As a benchmark case, we take  $\lambda = 2$  (corresponding to an average of 2 shocks within a lifetime) and  $\kappa$  uniformly distributed on  $[-1/2; 1/2]$  (Case 1). In Case 2, we set  $\lambda = 6$  (corresponding to a shock happening roughly once every 8 years) while



keeping the distribution of  $\kappa$ , and in Case 3, we keep  $\lambda = 2$  and change the distribution of  $\kappa$  to a uniform distribution on  $[-2; 2]$ . Results are presented in Table 3, where we also give the results of the basic one shock model for  $\lambda = 2$  and  $\lambda = 6$  for comparison.

	$\underline{s}$	$\bar{s}$	$M_I$	$M_D$
infinite shocks model				
Case 1: $\lambda = 2, \kappa \sim U[-1/2; 1/2]$	0.710	0.824	50%	67.5%
Case 2: $\lambda = 6, \kappa \sim U[-1/2; 1/2]$	0.642	0.820	50%	69.4%
Case 3 $\lambda = 2, \kappa \sim U[-2; 2]$	0.710	0.820	50%	67.2%
for comparison: one shock model				
$\lambda = 2$	0.75	0.8204	50%	64.1%
$\lambda = 6$	0.75	0.80	50%	60%

Table 3

These results are as we would expect them, from our discussion above. First, in the scenarios of the model with infinitely many shocks,  $\underline{s}$  is lower than in the one shock model, and lower for  $\lambda = 6$  than for  $\lambda = 2$ . Second, in the model with infinitely many shocks,  $\bar{s}$  is higher than in the one shock model, and consequently the majority rule for decreases is higher, too.

The distribution of  $\kappa$  does not play a large role in determining the values of  $\underline{s}$  and  $\bar{s}$ . The differences between case 1 and case 3 could be due to random effects in the simulation. Remember that the results of the one shock model were independent of the distribution of  $\kappa$ . However, the intuitive reason for this result there depends on the assumption that there is only one shock, and so there is no reason to conjecture that the same could be true in the infinite shocks model.

#### 4.4 Continuous elections

Up to now, we have assumed that there are exactly two issue and two procedural elections. In this subsection, we return to all other assumptions of the basic model and analyze what changes if society holds issue elections more often.

The intuitive argument for our assumption of a limited number of elections was that there was a new election every time the distribution of preferences in the society changed; between the first election and the shock, or afterwards, this distribution remains constant<sup>23</sup> and it seems fairly reasonable to assume that without changes in

<sup>23</sup>To see this, remember that the shock is a Poisson event; hence the expected further waiting time

“fundamentals”, the outcome of the election should not change.

Assume now that the society holds *issue elections* continuously: At every point in time, there is an election which determines the value of the policy, of course subject to the same majority rules which we have analyzed before. In this case, there is in fact a subgame perfect equilibrium which corresponds to the equilibrium of the basic model: If all individuals expect that the outcome of future elections is the same one as that of the present election as long as fundamentals do not change, it is in their self interest to vote in the same way as they did in the basic model, and so their expectations are correct in equilibrium.

However, there are other equilibria as well. Consider the situation after the shock, and suppose we have a  $2/3$  majority rule for decreasing  $p$ , and the effective policy after the shock (but before any adjustment) is 1. Suppose individuals believe that, whatever the outcome of the present election, in all following elections all individuals will vote for their bliss point (that is, for the policy which minimizes their momentaneous disutility). In this case, it is in the interest of the pivotal voter (aged  $2/3$ ) to vote for a decrease of the effective policy to  $2/3$ , even though she would of course prefer a value of  $p = 5/6$  (as in the basic model), if she would believe that this value were fixed for the rest of her life. It is important to stress that this equilibrium is *not* a result of non-sincere voting. Voters' sincere preferences depend on their expectations about what happens in future elections, and therefore multiple equilibria can arise.<sup>24</sup>

To sum up, if there are issue elections on a continuous basis, the equilibrium we analyze in the basic model is still a subgame perfect equilibrium with sincere voting, but even though it is no longer the unique one, it is still a very plausible equilibrium. On the other hand, the assumption that there are no procedural elections after the shock is crucial for the results; if there were the possibility to change the voting rules at every moment (with a simple majority), then it is evident that a result which is not the one favored by the median voter ex post cannot be an equilibrium. A meaningful is independent of the time which has already elapsed, and so the distribution of voters' preferences is constant before the shock happens.

<sup>24</sup>If another equilibrium than the one we focused on is played after the shock, it might be possible for the society to counteract that behavior by changing the outcome in the first election on the majority rule. In the equilibrium of the basic model, the second period voting results in the preferred policy of the pivotal voter. If we change this assumption and suppose for example that the second period equilibrium play results in some compromise between the (then) median and the “pivotal” voter (in the sense that the outcome is some convex combination of these two people's preferred choices), then, to counteract that tendency to lower policies than desired, the median voter in the procedural election will set a higher supermajority threshold.

procedural election resulting in a supermajority rule requires that the rule cannot be removed by simple majority *ex post*.

## 5 Conclusion

This paper shows how supermajority rules arise endogenously in a dynamic voting model, if voters' preferences change over time in a systematic way. The crucial assumption which yields a supermajority rule for a tax *increase*, but a simple majority rule for a tax decrease, was that when people grow older, their preferred tax rate decreases. This could be the case because of wages which rise with seniority, and the trade-off between publicly supplied goods and taxation becomes less favorable for the individual over time. Another argument is the following: The older people are, the less concerned will they be with the level of state debt, so for a given level of spending, older people should prefer a higher deficit (and less taxation) than young people.

Section 4 has shown that the simple basic model is relatively robust to a variation in some assumptions. In a representative democracy, the supermajority rule needed to implement the equilibrium policies can be higher or lower than in the basic model. If the median voters of different electoral districts in a state are very similar to each other, that state will impose a higher majority rule than a state with very heterogeneous electoral districts. We test this prediction empirically and show that in the US, the most "homogeneous" states are much more likely to adopt a supermajority rule for tax increases, consistent with this prediction. Additional, non-age dependent heterogeneity within the electorate reduces the equilibrium supermajority rule employed, though not much from a quantitative point of view. With infinitely many shocks, the equilibrium supermajority is higher than in the basic model.

A case which fits very well into the predictions of the model is the following. In Missouri, there exists a supermajority rule for schoolboards which want to raise taxes.<sup>25</sup> The remarkable feature of this rule is that it requires a  $2/3$  majority *in a referendum among taxpayers*, something which is usually much more difficult to achieve than a  $2/3$  majority in a legislature. Preferences for schoolboard spending are probably heavily influenced by the number of children in the household who still have to graduate from schools, so the model fits very well here.

More puzzling is why we do in general not observe supermajority rules for the question of how large old age pensions should be. The older people are, the higher

---

<sup>25</sup>See Missouri constitution, Article 10, section 11 c.

is presumably their preferred replacement ratio.<sup>26</sup> We would therefore expect that a simple majority is sufficient to introduce new transfers to old people, and that a supermajority is required for any decrease. While the first prediction is probably true, the second one is more troublesome. I do not know of any such (explicit) supermajority rules in the field of pensions. On the other hand, it also appears that in most countries, lowering the level of pensions is also not an easy exercise, and there are other institutions which can in principle lead to the same results as explicit supermajority rules. See Diermeier and Myerson (1999) for a model in which a bicameral legislature is equivalent to a supermajority rule.

## 6 Appendix

### 6.1 Proof of proposition 1

Most of the arguments are presented in the text. Here, we will the two steps left out there: The second order condition for the first order condition (8) is satisfied, and the function  $s(t, \lambda)$  given in (12) is decreasing in  $\lambda$ .

For the first claim, differentiate (8) again with respect to  $\bar{s}$  and substitute  $t = 1/2$  for the median voter, which yields:

$$\begin{aligned} \frac{\partial^2 W}{\partial \bar{s}^2} &= \int_{1/2}^1 \lambda [-f(p_0 - \bar{s})(1-u)(1+u-2\bar{s}) - 2F(p_0 - \bar{s})(1-u)] e^{-\lambda(u-t)} du = \\ &\int_{1/2}^1 \lambda [-2F(p_0 - \bar{s})(1-u)] e^{-\lambda(u-t)} du < 0 \end{aligned} \quad (23)$$

Hence the second order condition is satisfied.

For the second claim, differentiate  $s(t, \lambda)$  with respect to  $\lambda$ :

$$\frac{\partial \bar{s}(1/2, \lambda)}{\partial \lambda} = \frac{[-\lambda^3 - 2\lambda^2 - 16\lambda + 32]e^{-\lambda/2} + [-2\lambda^2 + 16\lambda - 16 - 16e^{-\lambda}]}{4\lambda^2(2e^{-\lambda/2} + \lambda - 2)^2} \quad (24)$$

A sufficient condition for the first term in square brackets in the numerator to be negative is  $\lambda > 2$ , and a sufficient condition for the second term in square brackets in the numerator to be negative is  $\lambda > 8$ ; hence for  $\lambda > 8$ ,  $\frac{\partial \bar{s}(1/2, \lambda)}{\partial \lambda} < 0$ , and for  $0 \leq \lambda \leq 8$ , a plot of the function shows that  $\frac{\partial \bar{s}(1/2, \lambda)}{\partial \lambda} < 0$  there as well. *QED.*

### 6.2 Proof of proposition 3

**Proof:** The first claim is a consequence of Jensen's inequality.

<sup>26</sup>See Cooley and Soares (1999) for a calibration of a median voter (i.e. simple majority) model of elections on the replacement ratio in the US pension system.

For the second claim, we will first prove that  $\bar{s}(t, \lambda)(t)$  as given in (10) is a concave function: Differentiating  $\bar{s}(t, \lambda)(t)$  twice with respect to  $t$  yields  $\frac{\partial^2 \bar{s}(t, \lambda)}{\partial t^2} =$

$$\lambda \frac{\left[ 2 + e^{-\lambda(1-t)}[-4 - 4\lambda(1-t) + \lambda^2(1-t)^2 - \lambda^3(1-t)^3] + e^{-2\lambda(1-t)}[2 + 4\lambda(1-t) + \lambda^2(1-t)^2] \right]}{[1 - \lambda(1-t) - e^{-\lambda(1-t)}]^3} \quad (25)$$

Since  $1 - x - e^{-x}$  is negative for all  $x > 0$ , the denominator is negative for all  $\lambda(1-t) > 0$ . The numerator is positive for all  $\lambda(1-t) > 0$ : Substituting  $x = \lambda(1-t)$ , the function  $2 + e^{-x}[-4 - 4x + x^2 - x^3] + e^{-2x}[2 + 4x + x^2]$  is an increasing function of  $x$ , for all  $x > 0$ , with range  $[0, 2]$ .

Differentiating  $1 - G(S - \bar{s}(t, \lambda)(t))$  twice with respect to  $t$  yields

$$\frac{d^2}{dt^2}(1 - G(S - \bar{s}(t))) = -g'(S - \bar{s})\left(\frac{d\bar{s}}{dt}\right)^2 + g(S - \bar{s})\frac{d^2\bar{s}}{dt^2}$$

Since the second term is always negative given that  $\bar{s}$  is concave in  $t$ , a sufficient condition for the whole expression to be negative is  $g'(\gamma) \geq 0$  for all  $\gamma$ , or, equivalently,  $G(\cdot)$  convex, as claimed.

For the third claim, substitute  $G(x) = \frac{h+x}{2h}$  in (21). Setting (21) equal to  $1/2$  and simplifying yields the expression given in the proposition. (the condition on  $h$  secures that individuals of all age groups are among those who prefer a higher  $\bar{s}$ ) QED.

### 6.3 Simulation in section 4.3

The simulation results given in Table 3 were calculated as follows. First, 1000 “lives” were drawn randomly according to the parameters specified in Table 3. A life is a sequence of the twelve next shock arrival times,<sup>27</sup> and the respective shock sizes.  $(\underline{s}, \bar{s})$  is then determined to be the optimum for the median voter (subject to the constraints specified in section 4.3), if the median voter knew that his future life were drawn randomly from these 1000 lives (and not the theoretical distribution, which would be very difficult to calculate analytically, but is approximated by the random draw).  $M_D$  is determined as the age of the person whose preferred policy is  $\bar{s}$ , given that the results of future elections are given by  $(\underline{s}, \bar{s})$ , and shocks occur according to the 1000 lives drawn randomly.

This procedure is repeated five times for each of the three cases, and the average of the best  $(\underline{s}, \bar{s}, M_D)$  is reported in Table 3. Differences in the optimal parameter values between different runs are generally quite small. For example, for case 1 ( $\lambda = 2, \kappa \sim$

<sup>27</sup>In all simulations, the minimal time passing until the twelfth shock occurred was greater than 0.5 in every life.

$U[-1/2; 1/2]$ ), the optimal  $\underline{s}$  was between 0.708 and 0.711 in all 5 runs. For  $\bar{s}$ , the range is [0.821; 0.828], and for  $M_D$  it is [0.672; 0.680].

## 7 References

- Baron, D. (1996): "A Dynamic Theory of Public Good Programs", *American Political Science Review*, 90, 316–330.
- Buchanan, J. and Tullock, G. (1962): "The Calculus of Consent.", University of Michigan Press, Ann Arbor.
- Caplin, A. and Nalebuff, B. (1988): "On 64%-majority rule.", *Econometrica*, 56, 787–814.
- Cooley, T. and Soares, J. (1999): "A positive theory of social security based on reputation", *Journal of Political Economy*, 107(1), 135-60.
- Diermeier, D. and Myerson, R. (1999): "Bicameralism and Its Consequences for the Internal Organization of Legislatures" *American Economic Review*, 89, 1182–1196.
- Gradstein, M. (1999): "Optimal Taxation and Fiscal Constitution", *Journal of Public Economics*, 72(3), 471-85.
- Knight, B. (2000) "Supermajority Voting Requirements for Tax Increases: Evidence from the States", *Journal of Public Economics*, 76(1), 41-67.
- Messner, M. and Polborn, M. (1999): "Voting on Majority Rules", University of Western Ontario, mimeo.