DISCUSSION PAPER SERIES

No. 9054

TAX CONTRACTS AND ELECTIONS

Hans Gersbach and Maik Schneider

PUBLIC POLICY

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP9054.asp**

ISSN 0265-8003

TAX CONTRACTS AND ELECTIONS

Hans Gersbach, CER-ETH Center of Economic Research and CEPR
Maik Schneider, CER-ETH Center of Economic Research

Discussion Paper No. 9054

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 9054

ABSTRACT

Tax Contracts and Elections*

In this paper we examine the impact of tax contracts, a novel instrument, on elections, policies, and welfare. We consider a political game in which three parties compete to form the government and voters may behave strategically. Parties have policy preferences about the level of public-good provision and benefit from perks when in office. A government raises taxes for both purposes. We show that tax contracts yield moderate policies and lead to lower perks by avoiding the formation of grand coalitions in order to win government. Moreover, in polarized societies they unambiguously improve the welfare of the median voter.

JEL Classification: D72, D82 and H55
Keywords: elections, government formation, political contracts and tax promise

Hans Gersbach
CER-ETH Center of Economic Research at ETH Zurich
Zuerichbergstrasse 18
8092 Zurich
SWITZERLAND

Email: hgersbach@ethz.ch

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=119061

Maik Schneider
CER-ETH Center of Economic Research at ETH Zurich
Zuerichbergstrasse 18
8092 Zurich
SWITZERLAND

Email: schneider@mip.mtec.ethz.ch

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=167162

*We would like to thank Johannes Becker, Christoph Engel, Lars Feld, Kamali Wickramage, Markus Müller, Roger Myerson, and the participants of the Social Choice and Welfare Conference 2008 in Montreal, the Workshop on Political Economy 2008 in Silvaplana, the VfS Meeting 2008 in Graz, and the EEA 2009 in Barcelona for valuable comments

Submitted 05 July 2012

"The coalition is being judged in terms of the promises made in the election campaign. That is unfair!"

Franz Müntefering, head of the Social Democratic Party in Germany, press conference on 29 August, 2006, Berlin, Germany

1 Introduction

The fiscal commons problem is one of the most prominent explanations for excessive spending in democracies. The original reasoning by Buchanan and Tullock (1962) and Weingast et al. (1981) runs as follows: The government budget represents a common pool for political decision-makers. A larger number of decision-makers in connection with the public budget is called higher fragmentation of government. Higher fragmentation tends to increase government spending, the reason being that political decision-makers are concerned with targeting resources from the public budget to projects that benefit the interest groups that support them. The tax scheme spreads out the costs of these special expenditure programs over the whole population. Hence, a governmental politician and the interest group he represents will fully appropriate the benefits of the targeted project, while only bearing a fraction $1/m$ of the costs, where m stands for the appointed decision-makers representing particular constituencies. Accordingly, the higher the number of appointed decision-makers (in other words, the higher the fragmentation of government), the smaller is the share of internalized project costs, thus increasing the government budget, government size, and socially wasteful spending.

In this paper, we propose the fiscal institution of tax contracts as a remedy for the fiscal commons problem. This novel institution works as follows: During election campaigns parties can make promises in the form of limitations on tax rates that they intend to honor when in power. For example, one party may promise not to increase tax rates, while another party may specify a range of value-added taxes. Such tax promises are approved by a public certification body and are called tax contracts. If a party becomes part of the government, at the end of the term of office, the certification body evaluates whether the tax promise has been kept by the government. As tax rates are verifiable, a simple yes/no decision is possible when the contract is up for review. If the answer is yes, there are no further consequences. If the tax promise has been violated, the party is severely punished, e.g. by cutting its public funding or, even more

extremely, depriving it of the right to nominate candidates for the next election. We assume that such punishments are so severe that tax contracts will never be violated. Accordingly, tax contracts enable a party to commit to a range of tax rates during election campaigns.

We consider a political game in which three parties compete to form the government. Parties have policy preferences about the level of public-good provision, and they benefit from perks when in office. A government raises taxes for both purposes. One party is an extremist party, the others are conventional parties. They either form the government on their own, if one of them receives a majority of seats, or they form a grand coalition. The observation that grand coalitions are likelier in the presence of an extremist party is a common feature of democracies with proportional representation. Recent examples include elections in Germany (2005), Austria (2008, 2006, 1995, 1990, 1986), Turkey (1999), and Hungary (1994).[1]

The model produces three insights. First, tax contracts yield moderate policy platforms, the reason being that with tax contracts parties will commit to tax rates close to the ones preferred by the median voter. When parties are in power, they do not have an incentive to choose extreme levels of public-good provision (high or low), as they balance their own perks and benefits from public goods. Thus moderate policy outcomes emerge during legislative periods. However, as parties will always choose some perks when they are in power, the policy outcome will differ from the one preferred by the median voter.

Second, tax contracts lead to lower perks by avoiding the formation of grand coalitions. Without tax contracts, voters on one side of the political spectrum have little incentive to vote strategically for the party on the other side of the political spectrum. Such strategic voting would indeed yield the benefit of avoiding the formation of a grand coalition, which is associated with higher perks. However, the downside is having a party in government that would choose taxes and public-good provision of a kind that is very unattractive to voters on the other side of the political spectrum. If societies are polarized, tax contracts make strategic voting and thus avoidance of grand coalitions attractive, as a governing party will not deviate from the tax promise and will choose moderate public-good provision. This implies smaller utility losses for voters on the other side of the political spectrum than the utility gains from having lower perks for the government.

[1]Source: Social Science Research Center Berlin (WZB).

Third, tax contracts tend to improve the welfare of the median voter due to the moderating effect of tax contracts with respect to policy. By contrast, without tax contracts there may be extremely high or low levels of public-good provision plus (according to the intuition in the previous paragraph) a low incentive to avoid the large amounts of perks associated with grand coalitions. There is one exception: very low levels of polarization, i.e. where the ideal points of parties are very close. Public-good provision is reduced in the presence of tax contracts, whereas without tax contracts taxes are increased. As in terms of utility losses increasing taxes is less expensive than reducing public-good provision, a regime without tax contracts may be favorable in societies with very little polarization.

The paper is organized as follows: The next section relates our paper to the literature. In Section 3 we develop the model. Section 4 examines the formation of coalition governments without contracts. In Section 5 we introduce tax contracts into the legislative game and characterize the equilibria. The effects of tax contracts are identified in Section 6. In Section 7 we examine how tax contracts affect the welfare of the median voter, and in Section 8 we discuss possible extensions of our model. Section 9 concludes.

2 Relation to the Literature

One of the most prominent explanations for excessive spending in democracies is the fiscal commons problem caused by fragmented governments. Several causes for fragmented governments have been put forward in the literature. Buchanan and Tullock (1962) and Weingast et al. (1981) have derived the fiscal commons problem from geographical fragmentation, as representatives at the federal level represent electoral districts. Roubini and Sachs (1989) and Roubini et al. (1989) have focused on coalition size as an indicator of fragmentation. Other interpretations of fragmented government focus on the number of ministers spending money and on cabinet size (Schaltegger and Feld, 2009). Recent theoretical studies of a dynamic version of the fiscal commons problem can be found in Battaglini and Coate (2008).

There exists a large body of empirical literature on the way in which governments with a high degree of fragmentation affect the size of the government and policy outcomes like unemployment and inflation. This literature is surveyed by Schaltegger and Feld (2009), who themselves find that the number of ministers in Swiss cantons is associated

with large government sizes. They also indicate that fiscal referenda tend to alleviate the fiscal commons problem.

A variety of institutional provisions exist that might help remedy the fiscal commons problem. Feld and Kirchgässner (2001) and Feld and Matsusaka (2003) indicate that fiscal referenda help to restrain overspending by governments. Poterba and von Hagen (1999) provide a detailed account of how formal fiscal rules affect fiscal performance, while Persson and Tabellini (2003) relate constitutional rules shaping a particular type of democracy to economic outcomes, including government spending. Tax contracts are a new type of institution that may help tame the fiscal commons problem and foster welfare. A companion paper (Gersbach and Schneider, 2008) examines the impact of tax contracts on government formation in a four-party system with sincere voting.[2]

Our paper is broadly related to the issue of whether political competition will reduce or eliminate the rents of competing political parties. This literature indicates that competition reduces rents, either because there is uncertainty regarding the preferences of parties (Polo, 1998) or because in democracies with proportional representation coordination of voting on non-corrupt parties is difficult to achieve (Myerson, 1993). In our model, all parties will choose some rents when they are in government, as they cannot commit to renouncing perks altogether. However, tax contracts can reduce these rents.

Tax contracts are a particular commitment device for parties in which they can choose the degree of commitment by specifying a tax contract with a large or small range of tax rates they are committed to. The role of commitment to campaign announcements has been a prominent theme ever since the work of Downs (1957). In much of the work on campaigns, it is implicitly or explicitly assumed that candidates could commit to policies they would implement if elected. How candidates can commit has recently been thoroughly studied, and either reputation effects (see e.g. Aragones et al., 2007) or parties (see Levy, 2004) have been considered as commitment devices. Tax contracts are a novel commitment device. In Section 8.2, we discuss how tax contracts differ from reputation effects.

[2]Political Contracts in general are surveyed in Gersbach (2012).

3 The Model

3.1 Voters and policies

We consider a representative democracy with an electorate that consists of a set of voters, $\mathcal{I}$. Voters are identified by their income level y_i, $i \in \mathcal{I}$. The density function $f(y)$ over the interval $[y_{min}, y_{max}]$ describes the income distribution. The median voter's income level is denoted by y_m. Aggregate income Y is given by

$$Y = \int_{y_{min}}^{y_{max}} y f(y) dy.$$

Voters elect a parliament, which in its turn elects a government. The activities of the government are

- choosing the level of a public good $g \in [0, \infty)$ and the perks for the parties in power. Perks comprise targeting expenditures that are welfare-reducing but benefit interest groups and the constituency supporting politicians in power. Perks also include all government expenditures that benefit politicians in power directly, such as administrative empires or prestige projects.[3] Both the public good and the perks are financed by a linear income tax with tax rate t. We assume that taxation is distortionary. Let $\lambda > 0$ denote the shadow cost of public funds. Accordingly, taxation uses $(1 + \lambda)$ of taxpayer resources in order to levy 1 unit of resources for public projects and for perks. The costs of providing the amount g of the public good is $g \cdot c$, where $c > 0$ is the unit cost;

- choosing a binary policy $d \in \{0, \bar{d}\}$, $\bar{d} > 0$, where $d = 0$ represents the status quo, while policy $\bar{d}$ describes large changes, such as waging a war, international agreements on arms control or trade, implementing large-scale regulatory reforms, etc.

The utility of a voter with income y_i is given by

$$U_{y_i}^{\delta_i}(t, g, d) = A \ln g + (1 - t) y_i - \delta_i d, \tag{1}$$

[3] An extreme form of perks is the overly generous opportunity of using public money for private expenditures. The examples revealed in Great Britain in 2009 show how far such attempts can go, see e.g. Economist (2009). The distinction between public goods and perks may not always be clear-cut, however. In particular, investments in public infrastructure can be anything from a pure perk to a pure public good. The infamous bridge to nowhere in Alaska is an example of the former, while investment in national defense is an example of the latter.

where the parameter $A \in \mathbb{R}_+$ expresses the relative strength of the preferences for the public good and $\delta_i \in \{-1, 1\}$ represents the preferences with respect to radical political change. Voters with $\delta = 1$ dislike radical political change, while voters with $\delta = -1$ are in favor of it.[4] We assume that the characteristic $\delta = -1$ is distributed independently of income. Ex ante, this could also be interpreted as every citizen possessing the same probability of being in favor of the extreme policy change.[5]

Before we introduce parties, it is useful to look at voters. Suppose a government (single party or coalition) is elected with s_k seats in parliament. As voters cannot prevent the government from choosing perks, the (second-best) desired policy and financing scheme of a voter characterized by $(y_i, \delta_i = 1)$ is the solution of the following problem:

$$\begin{aligned} &\max_{t,g,d\geq 0} U_{y_i}^{\delta_i}(t, g, d) \\ &\text{s.t. } tY = (1+\lambda)(gc + s_k b), \end{aligned} \tag{2}$$

where b denotes the fixed amount of perks per parliamentary member of the parties in government. Substituting t in $U_{y_i}^{\delta_i}$, the first-order condition with respect to g is

$$\frac{A}{g} - y_i \frac{(1+\lambda)c}{Y} = 0,$$

[4]We have chosen an additively separable utility function to capture the main political conflict over public-good provision. One might consider, for instance, non-separable preferences of the following type: the poor households value the public good (say, public education) more than the rich, and thus A is also a function of the income level. We expect our main results to carry over to such extensions as long as the preferences remain single-peaked in the tax rate. In general, the quasi-linear utility function where utility from the public good is a strictly concave function with infinite marginal utility at $g = 0$ has the following implications for our analysis. First, together with the budget constraint, the voter's maximization over policies for a given choice of d becomes a one-dimensional problem which has a unique interior solution. Second, and related to the first point, the preferences are single peaked for a given value of d. By assuming that $\bar{d}$ is so large that it outweighs the utility that can be obtained from the best tax- and public-good policy, the analysis is based on single-peaked preferences in a single dimension. This is the important property for our analysis that the assumed utility function delivers in an analytically tractable way. A deviation from single-peaked preferences and/or from the implied uni-dimensionality of the preferences may lead to non-existence of equilibria. Preserving the latter properties but allowing for corner solutions by relaxing the assumption of infinite marginal utility or allowing for multiple optima of the voter's problem by relaxing strict concavity to concavity, would substantially increase the complexity of our analysis by creating additional equilibria. However, in the latter case we expect our main results to be robust. To summarize, we expect our qualitative results to hold for utilities that represent unidimensional single-peaked preferences and where perks are costly for the voters but desirable for the parties (and total perks are increasing with the size of the government.)

[5]Let the probability that voter i prefers a radical political change, i.e. that $\delta_i = -1$, be p. Hence with probability $1 - p$ voter i dislikes the extreme policy. As the probability of being in favor of the extreme policy change is independent of income, a suitable version of the law of large numbers for a continuum of random variables implies that a share p of the electorate will be characterized by $\delta_i = -1$ and a share of $1 - p$ of the electorate will be conventional voters with $\delta_i = 1$.

which yields

$$t_i^* = \frac{A}{y_i} + \frac{(1+\lambda)s_k b}{Y}, \tag{3}$$

$$g_i^* = \frac{YA}{y_i(1+\lambda)c}, \tag{4}$$

$$d_i^* = 0. \tag{5}$$

Equations (3) and (4) reveal that voters with higher incomes prefer lower tax rates and lower levels of public-good provision. Note also that if voters were able to determine the amount of perks, they would choose $b = 0$ and $t_i^* = \frac{A}{y_i}$.

3.2 Parties

The total number of seats in the parliament is S. There are two conventional parties denoted by $j, k \in \{L, R\}$ competing for seats in the parliament. We characterize the parties' political orientation by (y_j, δ_j), which we also refer to as the parties' platforms. That is, $(y_L, 1)$ is the platform of party L and $(y_R, 1)$ the platform of party R. The parties' platforms can be interpreted as reflecting the characteristics of their median member. Consequently, the preferred policies of the parties especially favor voters with characteristics close to their platforms. We refer to parties as 'conventional' if their platform involves $\delta = 1$. We assume that $y_L < y_m < y_R$. Accordingly, party L (R) is the left (right) party. The utility of a party k in government is given by

$$V_k = U^1_{y_k}(t, g, d) + \theta s_k b, \tag{6}$$

where θ is a weighting factor and b are the perks in office per seat in parliament of party k. The perks can only be obtained if party k is part of the government. The utility formulation of V_k implies decreasing marginal utility with respect to g. Further, V_k has the property that the first unit of g possesses infinite marginal utility, which induces parties to choose positive levels of public goods. If it is the sole party in power, i.e. if s_k secures a legislative majority, then the party's policy will result from the following utility-maximization problem:

$$\max_{t,g,d \geq 0} V_k$$
$$\text{s.t. } tY = (1+\lambda)(gc + s_k b)$$

Accordingly, party k will choose

$$t_k^* = \frac{A}{y_k} + \frac{(1+\lambda)s_k b}{Y}, \tag{7}$$

$$g_k^* = \frac{YA}{y_k(1+\lambda)c}, \tag{8}$$

$$d_k^* = 0. \tag{9}$$

Comparing (8) with (4), we find that the optimal policy of parties with regard to the public good corresponds to the preferred policy of the respective voter with the characteristic $(y_k, 1)$. The tax rate also corresponds to the voter y_k's preferred tax rate for a given amount of perks. We note that the tax rate t_k^* is higher than voter y_k's preferred tax rate if perks could be avoided. The latter is equal to $\frac{A}{y_k}$.

Finally, there is a protest or extreme party E with platform $(y_E, -1)$ and utility[6]

$$V_E = U_{y_E}^{-1}(t, g, d) + \theta s_E b.$$

According to the preferences of its constituency, the extreme party would like to implement $\bar{d}$. Throughout the paper we assume that policy $\bar{d}$ is sufficiently important in the sense that a voter with $\delta_i = -1$ is better off with the optimal choice of party E than with any choice of the conventional parties that involves $d = 0$.

3.3 Rules and information

We consider a parliamentary democracy with proportional seat allocation, i.e. the seats in parliament are distributed among the parties according to their vote shares. In order to obtain seats in parliament, parties need a certain share of votes denoted by z ($z > 0$). We refer to the event of the extreme party entering parliament as event E. The opposite case is called event ¬E.

The informational environment is as follows:

- At the beginning of the political race and when parties sign the tax contracts, it is common knowledge among all voters and the parties that

[6]The particular value of y_E is not important for our analysis. As the voters of the extreme party are distributed identically as income, the median income in the set of voters supporting the extreme party is y_m. Note that the superscript -1 in $U_{y_E}^{-1}$ refers to $\delta_i = -1$.

- with probability p the platform of party E will find a response from voters, thereby generating a share of $\frac{s_E}{S}$ votes in favor of $\bar{d}$ over $d = 0$. This share is sufficient to enter parliament ($\frac{s_E}{S} \geq z$);
- with probability $1-p$ the platform of party E will not attract sufficient votes for party E to enter parliament.[7]

- Before the election takes place, voters observe their characteristic δ, which will remain private information.

We assume

Assumption 1

$$\frac{s_E}{S} < 1 - \frac{\sqrt{8}}{4}.$$

This upper bound on the seats simplifies the analysis and justifies calling party E an extreme party. The extreme party can, however, be quite large, for example a 25% share of seats is conceivable.

3.4 The political process

The political process involves three main stages:

Stage 1: Tax contracts

Stage 2: Election

Stage 3: Government formation

In the first stage we allow parties to sign tax contracts. A party's tax contract contains a range of tax rates that this party voluntarily commits to during the campaign. The tax contract becomes effective if the party forms a single-party government or is part of the coalition that forms the government. We assume that there are sufficiently severe sanctions if a party violates the tax contract when in government, so that violation will never occur. Such sanctions might be severe fines or drastic reduction or elimination of public financing for the party in question.

[7]The size of the vote share of the extreme party does not matter, as the number of seats for the conventional parties depends only on their relative vote shares.

In the second stage the election takes place, followed by the formation of the government in stage 3. In stage 3 we consider the following subgame: The party with the largest vote share (ties are broken by fair randomization) becomes the "formateur" which can decide to form either a single-party government or a coalition with the other conventional party.

We assume that coalitions are formed only by the conventional parties, i.e. party E is excluded from the government formation process. This is justified by the indivisibility of the policy option d plus two other arguments. First, the defining characteristic of party E is the policy $d = \bar{d}$. For ideological reasons, a compromise on this point is impossible for the extreme party, as its very existence is based on the pursuit of this extreme policy change. If it did not implement its defining policy issue when in government, it would lose its reputation with extreme voters and its entire raison d'être as a party. Ideology acts as a constraint on what the extreme party can do. This is a prominent theme and an important argument in the political economy literature (see Mueller, 2003, Benabou, 2008). Second, a conventional party suffers a large utility loss by accepting $\bar{d}$, as it prefers $d = 0$, which we assume would be larger than the utility it could achieve in a grand coalition. This occurred in Germany when the Social Democrats refused to form a government with an extreme, but small leftist party and favored a grand coalition.[8] We note that even if a conventional party enters into a coalition with the extreme party by accepting $\bar{d}$, it may still suffer indirect losses. For instance, it may face ongoing disputes about its decision to give in to the extreme party. This may even lead to a disruption of the coalition government, which can substantially deteriorate the conventional party's subsequent election prospects.

3.5 Government formation

The party with the highest vote share becomes the formateur which can decide to form a single-party government or to invite the other conventional party to form a coalition government.[9] If the formateur decides to form a single-party government, it proposes a policy (t, g, d) which is implemented if it receives more than 50% of the votes in a vote of confidence. Otherwise a caretaker government takes over. Its functioning is explained below.

[8] This holds if $\bar{d}$ is sufficiently large. Precise conditions are available upon request.

[9] Note that according to our assumptions on the maximal vote share of the extreme party, only a conventional party can become the formateur.

When the formateur invites the other conventional party to establish a coalition government, parties L and R try to form a coalition according to the following stages:

Stage 3.1 Parties maximize

$$\sigma_L V_L + \sigma_R V_R$$

over $(t, g, d = 0)$ subject to the budget constraint

$$tY = (1 + \lambda)(gc + (s_L + s_R)b),$$

where $\sigma_k = \frac{s_k}{s_L + s_R}$ represents the share of seats for party k within the coalition. If tax contracts have been written, they act as further constraints on the maximization problem.

Stage 3.2 Vote of confidence. The proposed coalition is elected if it receives a majority of votes in parliament. If the vote of confidence fails, a "caretaker government" assumes power and implements policy vector $(t, g, d) = (t_{ct}, \frac{t_{ct}Y}{(1+\lambda)c}, 0)$. A caretaker government involves no perks, i.e. $b = 0$ for all parties L, R, and E.

The role of the caretaker government is twofold. First, in parliamentary democracies a caretaker government is supposed to preserve the policy status quo and to assure the fundamental obligations of the state, such as maintaining law and order. It is not supposed to initiate new programs or any projects of political significance (see e.g. Laver and Shepsle, 1994). In our model, the caretaker government finances a certain amount of the public good but without perks for any of the parties. Second, caretaker governments are outside options for parties when they consider either forming a single or a coalition government.

We make two assumptions with respect to the caretaker policy. First, each party prefers being in power, either alone or in a coalition, to a caretaker government. This will always be the case if the value the parties attach to perks is sufficiently high, i.e. if θ is large enough. Second, there does not exist any policy $(\tilde{t}, \tilde{g}, \tilde{d})$ that fulfills the following two conditions simultaneously: (a) a party prefers to implement the policy in a single, possibly minority party government compared to forming a grand coalition, and (b) the other conventional party prefers policy $(\tilde{t}, \tilde{g}, \tilde{d})$ over a caretaker government. In Appendix B, we give the formal conditions for these two assumptions.

The two assumptions imply that the party with the highest vote share will form a single-party government if it possesses the absolute majority of seats in the legislature. Otherwise it will aim at forming a grand coalition rather than a minority government.

3.6 Summary of the game

It is helpful at this point to summarize the game. The players are the parties L, R, and E, and the voters $\mathcal{I}$. Each voter's strategy set is $\mu_i = \{L, R, E\}$. Each party's strategy set is described by $\nu_k = [0,1]^2 \times \{sg, gg\} \times [0,1] \times [0, \frac{Y}{(1+\lambda)c}] \times \{0, \bar{d}\} \times \{yes, no\}^4$. We explain the different parts of the strategy set in more detail. The first component reflects the tax contract choice in the first stage of the game. Choosing a tax contract means selecting an upper and a lower boundary of the interval of tax rates $\tau_k = [\underline{t}_k^c, \bar{t}_k^c]$ to which the party commits itself. This stage is not present in the game without tax contracts.

The party with the highest vote share can decide between forming a single party government denoted by sg and forming a coalition with the other conventional party, gg. The choice of sg is accompanied by selecting a policy vector $(t, g, d) \in [0,1] \times [0, \frac{Y}{1+\lambda}] \times \{0, \bar{d}\}$ which corresponds to the third, fourth and fifth subsets of the strategy set. If a grand coalition is aimed at, the policy proposal is determined via the bargaining process described above. Hence, there is no additional choice necessary for any of the bargaining partners when the formateur has chosen gg.[10] In both cases, sg or gg, the respective policy proposals will be followed by a vote of confidence, where each party either approves, yes, or disapproves, no, the proposed policy vector.[11]

The outcome of the political game is the policy vector (t, g, d) and the number of seats occupied by the government, which determines the amount of perks. The payoffs of the political game are the utilities derived by the players from this outcome according to their utility functions as given in equations (1) and (6).

The government formation processes could be extended. In particular, one could envisage a further subgame where another party becomes the formateur if the vote of confidence fails and the caretaker government is the last option when every party has

[10]Note that in the game with tax contracts, the choice of a tax rate outside of τ_k, as fixed in the first stage, involves harsh punishment. In this way, tax contracts do not restrict the parties' tax rates per se but will typically influence the equilibrium choices of taxes.

[11]Note that there are four possible situations (or information sets) where a vote of confidence can occur. In both states of the world, E and ¬E, the formateur either proposes a single-party government associated with a particular policy proposal or a grand coalition.

once enjoyed the proposal right and no majority support has been obtained in the legislature. However, our shorter government formation process captures the key equilibrium property that a conventional party can only hope to establish a single-party government if it possesses an absolute majority of legislative seats. Otherwise it will form a grand coalition with the other conventional party.

To simplify the analysis, we further assume that the two conventional parties are symmetric, i.e. the median voter is indifferent between the most desired policy/financing scheme of parties L and R if both parties receive half the seats.[12]

4 Election and Government Formation without Tax Contracts

In this section we examine the political process in the absence of tax contracts. This case not only captures what occurs in present-day practice in representative democracies, it also serves as a benchmark for the scenario with tax contracts. If no tax contracts are allowed, only stages 2 and 3 occur, and in particular, the parties are not restricted with respect to tax rates in the government formation process. We solve the game by backward induction and consequently start at the government formation stage.

4.1 Government formation without tax contracts

Recall that after the election, the party with the highest vote share is recognized as a formateur. This party, which on assumption 1 will be a conventional party, can decide to form a single-party government or a coalition government.

Suppose that the formateur has a majority of seats in the legislature. When deciding on a single-party government, it will propose the policy described by equations (7) - (9), which will be approved in the vote of confidence. When inviting the other party to join with them in forming a grand coalition, the two conventional parties bargain and maximize a weighted sum of their utilities. Solving the maximization problem described

[12] Formally, from $A\ln(\frac{YA}{y_L(1+\lambda)c}) - A\frac{y_m}{y_L} - \frac{S(1+\lambda)b\,y_m}{2Y} = A\ln(\frac{YA}{y_R(1+\lambda)c}) - A\frac{y_m}{y_R} - \frac{S(1+\lambda)b\,y_m}{2Y}$ we obtain that the platforms of the parties have to satisfy $\ln(\frac{y_R}{y_L}) = y_m(\frac{1}{y_L} - \frac{1}{y_R})$. In Section 8.6 we argue that our main results are also valid for asymmetric party platforms, as long as the vote-share advantage is not too large.

in Stage 3.1 of the government formation process yields the coalition government's policy as $d = 0$ and

$$t_{LR} = \frac{A}{y_{LR}} + \frac{(1+\lambda)(s_L + s_R)b}{Y}, \tag{10}$$

$$g_{LR} = \frac{YA}{c(1+\lambda)y_{LR}}, \tag{11}$$

where $y_{LR} = \frac{s_L y_L + s_R y_R}{s_L + s_R}$. This policy will be approved in the vote of confidence as it involves higher utility for the conventional parties than the alternative, a caretaker government.

Now consider the case where the formateur does not have the majority of seats. When trying to establish a single-party government as a minority government, the formateur will not be able to find a policy supported by the other conventional party in the vote of confidence due to our assumptions on the caretaker government's policy. Consequently, the formateur knows that aiming at a minority government will result in a caretaker government. By contrast, a grand coalition yields the policy outcome described by equations (10) and (11), which will be approved by the vote of confidence against a caretaker government. We summarize the equilibrium of the government formation subgame in the following lemma:

Lemma 1
If the party with the highest vote share possesses a majority of seats in parliament, it will form a single-party government implementing the policy vector (7) - (9). *Otherwise a grand coalition will be formed implementing the policy outcome given by* (10), (11), *and* $d = 0$.

The proof follows directly from the fact that (a) each conventional party prefers its own ideal policy vector to that of a coalition government, which in turn is preferred to a caretaker government, and (b) a caretaker government will result if a minority government is aimed for by the formateur.

4.2 Election

In the next step we examine the decision of the voters. We assume that voters with $\delta_i = -1$ vote in favor of party E.[13] As we have a continuum of voters, an individual

[13]Sincere voting for the extreme party can also be justified by ideological considerations, see Benabou (2008).

voter has no influence on the outcome, and any voting outcome may be supported as a Nash equilibrium. Accordingly, we use the following selection criteria: A combination of voting strategies, policy, and financing decisions by the party or coalition in power is called an equilibrium if

(i) it is a subgame-perfect Nash equilibrium,

(ii) there is no subset of conventional voters that can do better by changing their voting behavior.

Essentially, we are looking for subgame-perfect equilibria with coalition-proof voting. If there are multiple equilibria, we will introduce a further refinement and require that in any equilibrium the set of citizens who vote strategically be minimal.[14]

If the entire constituency votes sincerely and the extreme party does not enter parliament, there will be a single-party government of either L or R with probability $\frac{1}{2}$. If E enters, only a grand coalition between L and R will be possible.

A grand coalition in event E can be avoided if a sufficiently large subset of the electorate votes strategically. Let the minimal sets of strategic voters needed to prevent a grand coalition in the case of party E entering parliament be $[\hat{y}_R, y_m]$ and $[y_m, \hat{y}_L]$, where $\hat{y}_R < y_m$ and $\hat{y}_L > y_m$. The income levels of the critical voters $\hat{y}_L$ and $\hat{y}_R$ are defined by

$$\begin{aligned} \hat{y}_L &= F^{-1}\left(\frac{S}{2(S-s_E)}\right), \\ \hat{y}_R &= F^{-1}\left(\frac{S-2s_E}{2(S-s_E)}\right). \end{aligned}$$

The notation means that the critical voter $\hat{y}_k$ will vote for party k when voting strategically but will prefer party j when voting sincerely.

There are three types of voting outcomes that can be induced by strategic voting.[15]

(1) A subset of voters of measure $|[y_m, \hat{y}_k]|$ ($k = L$ or R) deviates from sincere voting, thereby allowing party k to form a single-party government with a minimal

[14]The justification is that coordination of voting (e.g. through interest groups) is costly and difficult to achieve if the set is large relative to the electorate. If there are multiple equilibria, they are qualitatively equivalent in the sense that government formation is the same across parties.

[15]In principle, strategic voting on both sides of the political spectrum could cancel out each other. Such constellations cannot occur in an equilibrium and are neglected.

majority in event E and with a supermajority comprising $\frac{S^2}{2(S-s_E)}$ seats in event $\neg E$.[16]

(2) A subset of voters greater than $|[y_m, \hat{y}_k]|$ deviates from sincere voting. This enables party k to form a single-party government with supermajorities in both events E and ¬E. In this case, the number of seats for party k will exceed $\frac{S}{2}$ in event E and $\frac{S^2}{2(S-s_E)}$ in event ¬E.

(3) A subset of voters smaller than $|[y_m, \hat{y}_k]|$ deviates from sincere voting. This implies that in event E a grand coalition will still come about, albeit with a policy tilted towards party k's ideal policy. In event $\neg E$, party k will form a single-party government with a supermajority of seats.

Note that in general there are several possible coalitions of voters for establishing a certain voting outcome, e.g. achieving a single-party government with a minimal majority of seats in event E. We focus on strategic voting by voters with incomes closest to the median income. The reason is as follows:

Suppose an arbitrary subset of voters that would vote for R when voting sincerely switches to party L. This implies that L will form a single-party government in ¬E. In event E, either a grand coalition will be established, where L has larger bargaining power than R, or a single-party government of L will materialize. Due to the single-peaked preferences, the closer a voter's income is to the median income, the less he will suffer from this policy shift toward L. Hence, if no subset of voters closest to the median has an incentive to deviate from sincere voting, then no subset of voters farther away from the median will have such an incentive either.

The next lemma rationalizes why there may be strategic voting.

[16]The size of the supermajority can be derived as follows: As the voters of the extreme party are distributed like the remainder of the electorate with respect to income, party L will just be able to form a single-party government in event E if

$$\int_{y_{min}}^{\hat{y}_L} f(y)\big(1 - \frac{s_E}{S}\big)dy = \frac{1}{2},$$

where $\frac{s_E}{S}$ reflects the vote share of the extreme party in event E. This equation can be transformed into $F(\hat{y}_L) = \frac{S}{2(S-s_E)}$, leading to the definition of $\hat{y}_L$ as given above. Symmetrically, we obtain $\hat{y}_R$ via $F(\hat{y}_R) = 1 - \frac{S}{2(S-s_E)}$. These definitions imply that a party obtains a vote share of 1/2 in event E and consequently $\frac{S}{2(S-s_E)}$ in event ¬E, if voters coordinate on it. Hence, the supermajority achieved by the single-party government in ¬E comprises $\frac{S^2}{2(S-s_E)}$ seats.

Lemma 2

A grand coalition involves higher perks than a single-party government with $\frac{S^2}{2(S-s_E)}$ *seats.*

The proof follows simply from showing that $S-s_E > \frac{S^2}{2(S-s_E)}$ reduces to $s_E < S(1-\frac{\sqrt{8}}{4})$, which is given according to Assumption 1.

A detailed analysis of the incentives for strategic voting is given in Appendix A. Here we provide a brief summary. We start our summary by examining the conditions under which deviation (1) may be profitable. As a grand coalition occupies more seats in the legislature than a single-party government, it is associated with higher perks. A subset of voters may thus consider deviating from sincere voting to avoid a grand coalition in the case of the extreme party becoming part of the legislature. The minimal set of strategic voters needed to avoid a grand coalition is such that in event E a single-party government will occupy a minimal majority of seats in the legislature. To be willing to deviate from sincere voting, this set of voters has to accept a single-party government formed by the less-preferred party in both events – whether E enters the legislature or not – in exchange for lower perks.

For example, suppose that subset $[y_m, \hat{y}_L]$ of the electorate votes strategically for party L. With probability p, the extreme party will enter parliament, and party L will form a single-party government with a minimal majority of seats $\frac{S}{2}$. Without the deviation from sincere voting, a grand coalition would have been formed with $S-s_E$ seats. Hence the perks associated with $\frac{S}{2} - s_E$ seats can be avoided with the help of the strategic voters. However, from the perspective of those who consider voting strategically, the beneficial effect of lower perks comes at the price of a policy tilted toward their less-preferred party rather than the moderate policy of a grand coalition. Moreover, if the extreme party cannot enter the legislature, party L will definitely form the single government with a supermajority of $\frac{S^2}{2(S-s_E)}$ seats. With sincere voting, the strategic voters, who actually prefer party R, would have a chance of $\frac{1}{2}$ that their preferred party will form the government in event ¬E. With sincere voting too, there would be minimal majorities for the single-party governments. Hence, we can say that from the perspective of those who consider voting strategically, the only benefit of avoiding a grand coalition is lower perks in event E. The costs are a worse policy in event E and an expected worse policy and higher perks in event ¬E. Therefore we infer that strategic voting will only occur if the perks associated with a grand coalition are large and the probability of the extreme party entering parliament is high.

Let us now turn to the question whether deviations (2) and (3) can dominate deviation (1). A comparison of (1) and (2) on the basis of the descriptions given in the list above reveals that the policy outcome would be no different. The only difference is with perks. As deviation (2) yields a larger number of governmental seats and thus larger amounts of perks in any of the possible cases, it is strictly dominated by deviation (1). The result of a comparison between (1) and (3) is not so clear. There may be parameter constellations of a kind that make deviation (3) favorable. These are however rather special cases.[17] Accordingly, in the remainder of the paper we concentrate on strategic voting according to deviation (1).

4.3 Subgame-perfect equilibria without tax contracts

In equilibrium we obtain

Proposition 1

There exists a p^{nc} such that

(i) If $p \leq p^{nc}$, there exists a unique equilibrium of the political process with sincere voting

1.

$$\mu_i = \begin{cases} E & \text{if } \delta_i = -1, \\ R & \text{if } \delta_i = 1, y_i > y_m, \\ L & \text{if } \delta_i = 1, y_i < y_m, \\ L \text{ or } R & \text{if } \delta_i = 1, y_i = y_m. \end{cases}$$

2. If E does not enter, L and R have a chance of $\frac{1}{2}$ to form the government. The policy outcome if party k assumes power is given by

$$\begin{aligned} t_k^* &= \frac{A}{y_k} + \frac{(1+\lambda)b}{Y}\frac{S}{2}, \\ g_k^* &= \frac{YA}{y_k(1+\lambda)c}. \end{aligned}$$

3. If E does enter, L and R form a grand coalition with outcome

$$\begin{aligned} t_{LR}^* &= \frac{A}{y_{LR}} + \frac{(1+\lambda)b(S-s_E)}{Y}, \\ g_{LR}^* &= \frac{YA}{y_{LR}(1+\lambda)c}, \end{aligned}$$

where $y_{LR} = \frac{y_L+y_R}{2}$.

[17] A more detailed discussion can be found in Appendix A.

(ii) If $p > p^{nc}$, then there exists an equilibrium with strategic voting

1. $$\mu_i = \begin{cases} E & \text{if } \delta_i = -1 \\ R & \text{if } \delta_i = 1, y_i > \hat{y}_j, \\ L & \text{if } \delta_i = 1, y_i < \hat{y}_j, \\ j & \text{if } \delta_i = 1, y_i = \hat{y}_j, \end{cases}$$

   for at least one party $j \in \{L, R\}$.

2. *Party j will form a single-party government independently of whether the extreme party enters parliament. The policy outcome will be*

$$\begin{aligned} t_j^* &= \begin{cases} \frac{A}{y_j} + \frac{(1+\lambda)b}{Y}\frac{S}{2}, & \text{if } E \text{ enters parliament}, \\ \frac{A}{y_j} + \frac{(1+\lambda)b}{Y}\frac{S}{2}\frac{S}{S-s_E}, & \text{if } E \text{ does not enter parliament}, \end{cases} \\ g_j^* &= \frac{YA}{y_j(1+\lambda)c}. \end{aligned}$$

The precise threshold value p^{nc} is given in Appendix A. A formal proof of Proposition 1 is provided in Appendix D.1.

According to the intuition of the strategic voter's trade-off given above, the benefits of strategic voting are realized in event E when the formation of a grand coalition is avoided. In event ¬E, strategic voting bears utility losses relative to sincere voting. Hence, if the probability p of event E occurring is low, the probability of reaping the benefits of strategic voting is low, but the complementary probability $1-p$ of realizing the costs is high. Thus strategic voting is unattractive for low values of p, but becomes attractive for high values of p.[18]

5 Tax Contracts

We now allow parties to sign tax contracts in which a party k restricts its tax policy to an interval $\tau_k = [\underline{t}_k^c, \bar{t}_k^c]$ before the election takes place. We make the assumption that if a party is indifferent between two sets of tax rates, it will choose the smaller one. Recall that tax contracts enlarge the parties' strategy sets by the possibility of making credible commitments before the election takes place. We solve the political process

[18] In general, the threshold p^{nc} may be one, in which case sincere voting occurs for all feasible values of p. In Appendix A we provide necessary and sufficient conditions for $p^{nc} \in (0, 1)$.

by backward induction. Voters take optimal voting decisions, and conventional parties offer optimally chosen tax contracts anticipating voting decisions and government formation.

5.1 Preliminary analysis

Again we are looking for subgame-perfect equilibria that are coalition-proof among conventional voters and perform some preliminary steps that are important in deriving the equilibria.

When tax contracts are allowed, government formation takes place as described in Section 4, with the additional constraint that the parties' tax contracts be honored. This means that a grand coalition is only feasible if the conventional parties' tax contracts overlap.[19] Otherwise a caretaker government will assume power given that no party occupies an absolute majority of seats. In case of a single-party government being formed, we establish the following result:

Lemma 3
Suppose that party k can form the government and has signed the tax contract τ_k.

1. *If $t_k^* \in \tau_k$, party k will choose t_k^*.*
2. *If $\bar{t}_k^c < t_k^*$, party k will choose $\bar{t}_k^c$.*
3. *If $\underline{t}_k^c > t_k^*$, party k will choose $\underline{t}_k^c$.*

Lemma 3 says that if a party is the sole party in power, it will always implement its most-preferred tax rate, given that it is allowed to do so by its tax contract. If its most-preferred tax rate is not included in the tax contract, then the party will choose the boundary of its tax contract that is closest to its ideal tax rate.

Concerning the election stage of the political process, there is an important difference between tax contracts and standard elections regarding preferences of voters for a particular party. Without tax contracts voters can be unambiguously attributed to preferring a certain party. That is, for any given amount of governmental seats, individual i will prefer one conventional party over the other conventional party as a particular party has a particular level of public good it desires. In general, this is not

[19] Formally this means adding the constraint $t \in \tau_L \cap \tau_R$ to the coalition's maximization problem at stage 3.1.

possible in the setting with tax contracts. As the parties fix a set of tax rates they will implement once in office, the desired level of public goods by a party depends on the election outcome. Accordingly, voters' preferences depend on their (rational) expectations about the result of the election.[20] We adopt the following definition of sincere voting when tax contracts are offered:

Definition 1 (sincere voting)
A voter votes sincerely if he votes for the party whose policy he likes best, given that no coalition is formed and both conventional parties receive the same number of seats.

The assumption that not all voters will behave as strategic players in the usual game-theoretic sense appears to be very realistic and has a long tradition in the literature. The most prominent line is Baron (1994), Grossman and Helpman (1996), McKelvey and Ordeshook (1987), and Ortin and Schultz (2005). However, as we will see at the end of this section, we could dispense with the above definition and assume that all conventional voters will correctly predict the election outcome and vote accordingly. Additionally, our equilibria with and without tax contracts would not change if all conventional voters voted entirely strategically.

Finally, the median voter's most preferred tax rate under sincere voting is important for the parties when choosing their tax contracts at the first stage of the political process. We denote this tax rate by t_{y_m} and derive it formally in Appendix C.

5.2 Equilibria with tax contracts

We now turn to the equilibria of the entire game with tax contracts. We obtain

Proposition 2
There exists a threshold value $\hat{p}$ distinguishing the following equilibria:

(i) If $p \leq \hat{p}$,

[20] More precisely, without tax contracts the parties' desired levels of public good provision do not depend on the election outcome as the tax rate is adjusted to finance the different amounts of perks associated with the election outcomes (see e.g. equations (7) and (8)). Since preferences are linear in the consumption good (or in "money"), this does not affect a voter's preferred public good policy and, hence, its preference for a certain party. This is different with tax contracts. Now the tax rates cannot be adjusted to the particular amount of perks prescribed by the election outcome. Instead, different amounts of perks have to be financed by adjusting public good provision. However, utility from public goods is subject to diminishing returns. Hence, it cannot be precluded that for given tax contracts of the parties a voter prefers R when the total amount of perks is rather small but party L when the amount of perks is large.

1. $\tau_L = \tau_R = \{t_{y_m}\}$.

2. $$\mu_i = \begin{cases} R & y_i > y_m, \\ L & y_i < y_m, \\ R \text{ or } L & y_i = y_m. \end{cases}$$

3. *If E does not enter, each conventional party will win the election with probability $\frac{1}{2}$, form the government, and choose $g(t_{y_m}, \frac{S}{2})$. If E enters, both conventional parties will form a grand coalition and choose $g(t_{y_m}, S - s_E)$.*

(ii) For all $p > \hat{p}$,

1. $\tau_L = \tau_R = \{t_{y_m}\}$.

2. *either*

   (a)

   $$\mu_i = \begin{cases} R & y_i > \hat{y}_L, \\ L & y_i \leq \hat{y}_L. \end{cases}$$

   Party L forms the government independently of whether E enters. Party L chooses $g(t_{y_m}, \frac{S}{2})$ if E enters and $g(t_{y_m}, \frac{S^2}{2(S-s_E)})$ if E does not enter.

   or

   (b)

   $$\mu_i = \begin{cases} R & y_i \geq \hat{y}_R, \\ L & y_i < \hat{y}_R. \end{cases}$$

   Party R forms the government independently of whether E enters. Party R chooses $g(t_{y_m}, \frac{S}{2})$ if E enters parliament and $g(t_{y_m}, \frac{S^2}{2(S-s_E)})$ if E does not enter.

In the proposition we use $g(t, s) = tY/((1 + \lambda)c) - sb/c$ obtained from the budget constraint (2) to denote the amount of the public good provided by a government comprising s seats and implementing a tax rate t. A proof of Proposition 2 can be found in Appendix D.3. The main conclusions from Proposition 2 are that strategic voting to avoid a grand coalition occurs when p is larger than a certain threshold $\hat{p}$, and that tax contracts yield moderate policy outcomes. That is, the conventional parties will commit themselves to a tax rate close to the one preferred by the median voter and, as perks are restricted by $(S - s_E)b$, the amount of public-good provision cannot

be extremely low or excessively high. However, recall from Section 5.1 that the policy outcome in event E with tax contracts is not the policy most preferred by the median voter.

In the equilibrium described in Proposition 2 the parties fix one tax rate in their tax contracts rather than an interval with positive measure. The reason is that a party enlarging the tax set by pushing one boundary of the set closer to its own ideal point means that this boundary tax rate will be implemented when forming a single-party government (see Lemma 3).[21] Under sincere voting, this involves a loss in terms of vote share for the party. It also implies that voters coordinate on the other conventional party when voting strategically due to the minimal set of strategic voters refinement. Consequently, such a deviation is not profitable as the party forgoes the possibility of forming a single-party government in both cases, with and without strategic voting. On the other hand, enlarging the tax set by shifting one boundary further away from its own ideal would not make a difference, given that the other conventional party chooses $\tau = \{t_{y_m}\}$. As a consequence of our minimal tax contracts assumption, the party will choose a tax contract containing t_{y_m} only.[22]

[21] In general, we note that the intersection between the parties' tax sets in equilibrium must comprise a single tax rate. Otherwise, at least one party has an incentive to reduce its tax set to obtain a better policy in a grand coalition without affecting its election prospects. If the parties value being in government (perks) sufficiently strongly, any equilibrium will involve both conventional parties including the same tax rate in their tax contract. Moreover, in all circumstances except one, the unique equilibrium will be $\tau_L = \tau_R = \{t_{y_m}\}$ as otherwise at least one party will find it profitable to expand its tax set towards t_{y_m} as discussed previously for the case that $\tau_L = \tau_R = \{t_{y_m}(S/2, S - s_E)\}$. There is one constellation where the equilibrium $\tau_L = \tau_R = \{t_{y_m}\}$ is not unique. If perks are not very important to parties, one of the parties may choose to forgo the possibility of forming a single-party government in event $\neg E$ for a more favorable policy in a grand coalition in event E. In this case, the party forming the single-party government specifies an interval with positive measure in its tax contract, while the party aiming at the grand coalition policy being closer to its ideal will sign a tax contract containing only one policy that is different from $\{t_{y_m}\}$ (but included in the tax-interval of the other party). As this equilibrium can only occur under special parameter values and never when there is strategic voting, we concentrate on the equilibria characterized in Proposition 2.

[22] The intuition why the parties choose $\{t_{y_m}\}$ even in the limiting cases where $p = 0$ and $p = 1$ can be summarized as follows. If $p = 0$, the extreme party will not enter with certainty thereby eliminating any incentives for strategic voting. Under sincere voting, deviating from $\{t_{y_m}\}$ is not beneficial for any of the parties as it would obtain a lower vote share implying that the opposing party could form the government. Consider now the limiting case where $p = 1$, i.e. the extreme party enters parliament with certainty. In this case, voters will vote strategically to avoid a grand coalition. In particular, due to the refinement requiring a minimal set of strategic voters, the strategic voters coordinate on the party with the policy closer to the one preferred by the median voter, as the set of strategic voters will be smaller than when coordinating on the party with the policy farther from t_{y_m}. Hence, neither party has an incentive to deviate from $\{t_{y_m}\}$ as this would imply that voters coordinated on the opposing party which would then be able to form a single-party government. Without the refinement, multiple equilibria are possible depending on the beliefs of parties and voters on voting behavior of the electorate.

In the next section we provide detailed intuitive explanations for Proposition 2 by comparing the policy outcomes of the political process with and without tax contracts in societies that exhibit different degrees of political polarization. First, however, it is useful to note the following corollary:

Corollary 1

The equilibria in Proposition 2 will remain the same if all voters correctly predict the election outcome.

Corollary 1 follows from the proof of Proposition 2. Given that the parties offer $\tau_L = \tau_R = \{t_{y_m}\}$ and given e.g. the voting decisions of voters in $[\hat{y}_L, y_m]$ in the case of (ii) 2.(a), it is optimal for voters in $[y_{min}, \hat{y}_L)$ to vote for party L and voters in $(\hat{y}_L, y_{max}]$ to support R. Next to voting behavior on the basis of $\tau_L = \tau_R = \{t_{y_m}\}$, the crucial issue is whether parties gain an incentive to deviate from t_{y_m} once all voters possess correct expectations of the voting outcome. To show that this is not the case, the reasoning in the proof of Proposition 2 can be readily applied as long as the deviating party loses vote shares. The latter must be the case for the following reason: First, holding perks fixed, the policy of the deviating party attracts fewer voters. As this is the case, the perks may decline, which will tend to increase the vote share. However, this (positive) effect on vote share from lower perks can never outweigh the (negative) effect on vote shares from the deviation in policy (i.e. tax rates). Our definition of sincere voting neglects this 'perks effect'. However, if voters have correct expectations the proof of Proposition 2 still holds true as the neglected perks effect is dominated by the 'policy effect' on vote shares.

6 Effects of Tax Contracts

Tax contracts have two effects on policy outcomes. First, tax contracts keep policies moderate, and second, they help to avoid the large amounts of perks associated with grand coalitions. The first effect follows directly from Proposition 2. In this section we explain why grand coalitions are less likely to occur in a regime with tax contracts than in a regime without tax contracts.

We start by defining the degree of political polarization as the distance between the two conventional parties' platforms. We write for short: $\Delta_y = y_R - y_L$. When varying the degree of political polarization, we preserve the symmetry assumption, i.e. the

median voter will receive the same utility from both parties' preferred policies.

If, in the regime with tax contracts, the probability of the extreme party entering parliament is higher than $\hat{p}$ defined in Proposition 2, then a subset of voters will deviate from sincere voting to reduce the perks associated with a grand coalition. In the next lemma we state that this already happens if the probability of the extreme party becoming part of the parliament is smaller than $\frac{1}{2}$.

Lemma 4
$\hat{p} < \frac{1}{2}$.

The proof is given in the appendix. Without tax contracts, the costs of deviation from sincere voting in terms of public-good policy clearly depend on the political polarization of the society. It is interesting to ask how the two regimes – with and without tax contracts – compare with respect to preventing a grand coalition by strategic voting. Without tax contracts, this will happen if p exceeds the critical threshold p^{nc} according to Proposition 1. We obtain

Proposition 3
For all degrees of political polarization, $\hat{p} < p^{nc}$.

A formal proof is given in Appendix D.6. Proposition 3 says that for all probabilities p for which grand coalitions are avoided by strategic voting in the regime without tax contracts, they will also be avoided in the regime with tax contracts. Moreover, there is a set of p values for which strategic voting occurs with tax contracts but not without tax contracts. In these situations tax contracts reduce the perks enjoyed by the government.

There are two reasons for this result. First, the utility loss of financing a supermajority is higher in a regime with tax contracts. Second, the costs in terms of public-good policy incurred by avoiding a grand coalition by strategic voting are lower when tax contracts are allowed.

To sharpen the intuition of the first point, consider the case without political polarization, i.e. where both parties have the median voter's utility as their political platform. Then the only difference between a regime with and without tax contracts is the way in which the additional seats of supermajorities are financed. In the regime without tax contracts, an additional governmental seat is financed by a tax increase that involves constant marginal utility losses for the voters. A tax increase in the event of

the extreme party entering parliament is not possible in the regime with tax contracts, and additional governmental seats are financed by lower public-good provision. Due to the strict concavity of voters' utility from the public good, the marginal utility loss of an additional seat increases with the size of the supermajority. As a consequence, the relative benefit of having a smaller supermajority is larger in the regime with tax contracts.

As for the second argument, we can say that when there is political polarization, the attractiveness of voting strategically declines in the regime without tax contracts. The reason is that the costs for the strategic voters increase with the degree of political polarization, as avoiding the perks associated with a grand coalition implies the implementation of the less favorable policy pursued by the party at the other end of the political spectrum. In this way, the weight attached to the utility gain from avoiding a grand coalition, p, must be higher for strategic voting to occur. By contrast, the critical probability for strategic voting in the regime with tax contracts is not affected by the level of political polarization, as parties commit themselves to moderate tax rates. Hence, the cost incurred by the strategic voters in terms of less favorable policies is absent with tax contracts. In this way, the effect of tax contracts on avoiding large amounts of perks in grand coalitions is particularly significant in polarized societies.

7 Winners and Losers of Tax Contracts

In this section we examine how voters and parties are affected by the introduction of tax contracts.

7.1 Utility of the Median Voter

We first examine how tax contracts impact on taxes, public-good provision, and perks, and how this is evaluated by the median voter. In accordance with the findings in the previous section, we need to distinguish three cases with respect to the probability of the extreme party entering parliament:

- $p \leq \hat{p}$: sincere voting equilibrium occurs in both regimes – with and without tax contracts;
- $\hat{p} < p < p^{nc}$: strategic voting to avoid a grand coalition occurs only in the regime

with tax contracts but not in the regime without tax contracts;

- $p > p^{nc}$: strategic voting occurs with and without tax contracts.

We summarize the main insights in the next proposition.

Proposition 4
If the level of polarization is

- *high, the median voter benefits from tax contracts for all values of p;*
- *very low and $\hat{p} < p < p^{nc}$, the median voter may also benefit from tax contracts;*
- *very low and $p \leq \hat{p}$ or $p > p^{nc}$, the median voter will prefer the regime without tax contracts.*

A proof of the proposition can be found in Appendix D.7. The intuition can be summarized as follows: As indicated by the discussion in the previous section, tax contracts yield moderate policies and can induce strategic voting to avoid large perks for grand coalitions. On the other hand, financing supermajorities is more expensive, as it has to be done by cutting back on public-good provision rather than increasing taxes.

In the cases where either sincere voting or strategic voting occurs in both regimes, i.e. if $p \leq \hat{p}$ or $p > p^{nc}$, the trade-off is between the benefit from ensuring moderate policies and the cost of supermajority financing. The former is large in polarized societies and small if political cleavage is negligible. The latter is independent of polarization. Consequently, tax contracts are favorable if polarization is high and less favorable if polarization is low. If $\hat{p} < p < p^{nc}$, tax contracts possess the additional benefit of inducing strategic voting to avoid a grand coalition, so tax contracts can also be beneficial if polarization is low.

From Proposition 4 we can conclude whether voters will favor the introduction of tax contracts. A majority of voters will favor tax contracts if they are supported by the median voter. This would occur in highly polarized societies and also in less polarized societies if tax contracts induce strategic voting to avoid large perks.

7.2 Impact on Parties

We next discuss how parties are affected by tax contracts. It is necessary to distinguish two considerations. First, if parties have the possibility to sign tax contracts,

it is individually rational to use them. Given the opposing party does not sign a tax contract, a party can increase its vote share and thus significantly increase its chances to form a single-party government by signing a tax contract as it allows to commit to a policy that attracts more voters than the ideal policy of its opponent.[23]

Second, an entirely different question is whether parties would like that the possibility to sign tax contracts is introduced. On the one hand, the conventional parties would benefit from the moderating effect on policy outcomes, as utilities are concave with respect to public-good policy. On the other hand, we have seen that tax contracts can lead to fewer perks. Consequently, the parties' appraisal of tax contracts depends on the degree of political polarization and the value they attach to perks. If the former is low and the latter is high, the introduction of tax contracts would not find a majority in the legislature.

7.3 Role of Polarization

One of our main results is that tax contracts are particularly desirable for voters and parties in societies with strong polarization. This may become of elevated importance in the future as political polarization tends to increase in many industrialized countries. McCarty et al. (2006) document increasing polarization in American politics and society. They argue that political polarization is also a consequence of increased income inequality (p. 75). Also in many democracies with proportional representation (such as Germany and the Netherlands) which our analysis focuses on, income inequality has increased over the last decades (OECD, 2011). The increase in income inequality is mainly attributed to factors such as globalization, technological change, and labor market reforms. Hence, it is likely that income inequality will be increasing further potentially leading to an increase in political polarization. In this case, tax contracts can be an increasingly valuable instrument to keep policies moderate and to avoid excessive amounts of perks.

[23] As shown earlier, the equilibrium when tax contracts are allowed involves both parties signing $\{t_{y_m}\}$. This implies that using tax contracts is individually rational. If it were not individually rational, there would exist an equilibrium where each party j is indifferent between signing a tax contract which includes the interval $[t_{LR}, t_j^*]$ and not signing a tax contract.

8 Discussion and Extensions

In this section we discuss several key features of our model and outline possible extensions of the model.

8.1 Aggregate shocks

We have so far abstained from discussing aggregate shocks to income. In the presence of income shocks, tax contracts would tie the hands of the government and would prevent intertemporal smoothing of such shocks through tax policies. There are at least two potential ways of striking a balance between flexibility and commitment to tax contracts in the presence of aggregate shocks.

First, tax contracts could be conditioned on aggregate income shocks, i.e. a party would commit to a tax policy conditional on aggregate shocks. While there are financial securities contingent on GDP (so-called GDP-linked bonds) competition with GDP-linked tax contracts would be highly complex. A milder form would be to allow an adjustment of the tax rates by x% in a contract if GDP growth is above or below a certain threshold.

Second, one could allow for the cancellation of tax contracts in crises, catastrophes, or severe recessions if this is approved by a supermajority of members in parliament. The supermajority has to be sufficiently large. In our model, it would have to be larger than a grand coalition. In such extraordinary circumstances, a vast majority of members of parliament might prefer cancellation of tax contracts. However, such possibilities might also motivate opposition members to reject renegotiation proposals and to force the government to honor its promise in order to hurt them in the next election. Such blockades are, however, risky because they make the opposition accountable for the economic situation.[24]

8.2 Reputation effects and career concerns

In our one-shot election model, the parties' future reputation or candidates' career concerns do not play a direct role. An important branch of the literature on political

[24]The U.S. public debate in 2012 and attempts to affect government revenues and expenditures through pledges (e.g. http://www.atr.org) is an example of such forces in the context of divided governments.

campaigns has investigated whether policy-motivated candidates or parties make (or can make) credible campaign promises. For instance, Aragones et al. (2007) show that with policy-motivated candidates, voters may punish a candidate for reneging on campaign promises by voting him out of office. Once an office-holder reneges, any other promises he makes will not be given credence. If voters and candidates are sufficiently patient, the threats of such punishments can support an equilibrium in which campaign promises are kept. However, only a subset of policy promises are believed by voters, as candidates' credibility is limited by what the candidate stands to lose in the future if he reneges after being elected.

Tax contracts offer a complementary way of making campaign promises credible and can work in circumstances where career concerns are insufficient. We provide three examples. First, even though the deselection threat may motivate candidates to honor their promises, there are many other equilibria in which promises are not kept and this is disregarded by voters. Tax contracts avoid equilibria in which campaign promises are just cheap talk. Second, career concerns only induce candidates to keep promises if discount factors are not too low. With tax contracts, parties are immediately punished as an organization, so penalization does not depend on the threat not to reelect a particular candidate. Third, without tax contracts the punishment threat works because voters will not believe the promises of a candidate once he has reneged. With tax contracts, even candidates who have reneged on non-contracted promises in the past can make credible promises in the next election by signing such a contract.

8.3 Practical implementation

In this paper we suggest that tax contracts have positive effects on welfare. Of course, the practical implementation of such contracts has to be dealt with carefully. We could imagine the following procedure: A new official body called "certification authority" reviews proposals for tax contracts during campaigns. It assesses whether the tax contract, including the tax base and tax rates, is described in such a way that a clear-cut answer is possible as to whether the contract has been fulfilled or not when the contract is up for review. If, and only if, the answer is affirmative will the certification authority approve the proposal. A certified proposal is published and functions as a tax contract. Later, the certification authority can also serve as a an external enforcer. The review and enforcement of political contracts will therefore require a special body,

which could be a separate entity of the constitutional court. In Germany, for example, it could also be one of the tasks of the Federal President.[25]

8.4 Endogenous perks

One important feature of our model is inspired by the empirical literature discussed in the introduction indicating that the amount of perks associated with a single-party government is smaller than with a grand coalition. This is reflected by our assumption that perks per seat are exogenously given. In this section, we discuss two ways of endogenizing perks.

8.4.1 Choice of perks by parties

We could allow that the parties in government choose the level of perks. Suppose the utility of conventional parties is

$$V_k = U^1_{y_k}(t, g, d) + \theta s_k \ln(b_k) \ , \tag{12}$$

with $b_k \in [0, \infty)$. This specification involves decreasing marginal returns on perks. If a party k forms a single-party government with s_k seats and is free to choose the set $\{t, g, b\}$, it will maximize (12) subject to government budget constraint (2).

The solution of the government problem is

$$\begin{aligned} b^*_k &= \frac{\theta Y}{y_k(1+\lambda)}, \\ g^*_k &= \frac{AY}{cy_k(1+\lambda)}, \\ t^*_k &= \frac{A + \theta s_k}{y_k}. \end{aligned}$$

Hence, with this specification the optimal amount of perks for a party is a fixed amount per seat of the party. However, in absolute terms a party with a larger income y_k associated with its platform will spend less on perks per seat. Nevertheless, the relation between perks per seat and public-good provision is constant across parties because

[25] In Germany, a transfer of this task to the Federal President would require a change of the constitution. As a rule, the introduction of tax contracts and the creation of a certification authority in other countries would require a modification of the countries' constitutions. We are grateful to Christoph Engel for extensive discussions on this subject.

$\frac{b_k^*}{g_k^*} = \frac{\theta c}{A}$. This means that, independently of the party in power, a unit of the public good comes with a particular amount of perks. As the total amount of perks increases with the number of governmental seats, our analysis and qualitative results can be extended to this type of endogenous choices of perks.[26]

8.4.2 Tax and perks contracts

As it is difficult or impossible to prevent the office-holder from generating rents, we have assumed so far that perks are not contractible. Suppose as a benchmark that it were possible for parties to contract on both taxes and perks. Then the amount of perks would be determined during campaigns. As voters unanimously dislike perks and preferences regarding public-good provision are single-peaked, competition of parties with tax-cum-perks contracts would yield the median voter result regarding public-good provision. That is, in equilibrium every conventional party will sign a tax-cum-perks contract including zero perks and the median voter's most-preferred tax rate and public-good provision. Moreover, it would not matter any more whether the government is a coalition or whether a single party is in office, as perks would be zero in all cases. As a consequence, voters no longer have any incentive to vote strategically.

8.5 Uncertainty about conventional parties' vote shares

We have focused on the uncertainty surrounding the extreme party's vote share and neglected additional uncertainty about the vote share of the conventional parties. There are several ways of incorporating additional uncertainty about the election outcome for conventional parties. The source of the uncertainty may be shifts of preferences of voters or popularity shocks that are unobservable by parties or voters before the election takes place. Suppose, for example, that with probability $\frac{1}{2}$ a certain measure of voters votes for party j under sincere voting, whereas with probability $\frac{1}{2}$ they vote sincerely for party k.

In this extended setup, the outmost member of the smallest set of strategic voters required to avoid a grand coalition is farther away from the median voter relative to the situation in our basic model. This has two implications. First, if there is strategic voting, the supermajorities of the single-party government may be large and even larger

[26] More general specifications of utilities from perks, captured by $\theta s_k w(b)$ with $w'(b)_{b\to 0} = \infty, w' > 0, w'' < 0$, can be also applied as long as aggregate perks are increasing in s_k.

if the popularity shock turns out to be in favor of the party that the strategic voters have coordinated on. Ceteris paribus, this reduces the incentive to avoid grand coalitions by strategic voting. Second, in the regime without tax contracts, the incentive for the outmost member in the set of strategic voters to vote strategically is lower, as his political position is farther away from the ideal point of his less-favored party.

With additional uncertainty about election outcomes, we need to adjust the refinement. In our basic set-up we used the equilibrium refinement that the set of strategic voters be minimal. In the model with uncertain vote shares for conventional parties, the corresponding refinement is that the *expected* set of strategic voters be minimal. Then the equilibrium with $\tau_j = \{t_{y_m}\}$ can be sustained.[27]

To summarize, additional uncertainty about election outcomes for conventional parties results in lower incentives for strategic voting relative to the basic model, as in expectation supermajorities cannot be avoided so easily. Nevertheless, compared to the regime without tax contracts, the main effects of tax contracts, i.e. that they lead to more moderate policies and help to reduce perks from grand coalitions, are still present, albeit weaker in the extended model.

8.6 Asymmetry of party platforms

In our analysis we have considered the situation where the two conventional parties' platforms are symmetric with respect to the median voter. This may be a plausible situation when there are two large parties in a country, located politically at center left and center right. However, our results are also robust to asymmetric platforms, as long as the advantage of one party under sincere voting without tax contracts is not so large that it will possess a majority of seats independently of whether the extreme party enters parliament or not.

When we allow for asymmetric platforms of the conventional parties, two additional considerations matter. First, the set of strategic voters necessary to avoid the formation of a grand coalition will be smaller and, due to the minimal strategic voters refinement, coordination of voters would always favor the conventional party with the vote share advantage under sincere voting without tax contracts. Second, there is an additional incentive to vote strategically in the regime without tax contracts. The

[27] The reasoning follows the logic of Proposition 2. If one party deviates from its tax contract $\{t_{y_m}\}$ toward its ideal point, the expected vote share will decline.

party with the more favorable platform for the median voter will form a government with a supermajority of seats if voters vote sincerely and E does not enter parliament. If the probability of the extreme party entering is small, voters may have an incentive to avoid the supermajority of a single-party government in event $\neg E$. This avoids excessive perks in event $\neg E$ and shifts the policy of a grand coalition away from the ideal point of the party with the vote share advantage. The latter may be beneficial to voters located around the median. With tax contracts the vote share advantage of one party will vanish, as both parties will make the same tax promise.

8.7 Multidimensional policy spaces

Tax contracts as introduced in this paper are a partial commitment device in the sense that parties can commit to policies on the financial side of policy-making but not on the expenditure side, which is usually much less verifiable. In our model, each tax rate is associated with a certain level of the public good and exogenously given perks. In a framework with endogenous perks, as discussed in Section 8.4, or multiple public goods, each tax rate committed to by a party would be associated with this party's optimal portfolio of public goods and perks. In this way, competition of the parties is uni-dimensional at the tax-rates level, as in our simpler basic model. Even though the parties' portfolio of public goods and perks for a given tax rate may be inefficient, the competition of parties in at least one dimension can be welfare-improving relative to the situation where no binding promises can be made.

However, in the absence of a tight balanced-budget constraint, a tax contract will still allow for the financing of public goods by public debt, which introduces a second policy dimension. To prevent tax contracts from being effectively offset by running budget deficits, one might allow the parties to sign joint tax- and budget-deficit contracts. As public debt may be desirable in downturns, where debt-financing can be used to stabilize macroeconomic shocks, tax contracts would tie the hands of the government in an unattractive way. This would introduce the problem discussed in Section 8.1, where we suggest different approaches to striking a balance between commitment and flexibility.

At a more abstract level, a multidimensional policy space introduces two difficulties for political contracts. First, if contracts do not force parties to commit on every dimension of the policy space, the government may adopt inefficient policies on the dimension on

which it is not committed. This potential source of inefficiency has to be traded off against advantages of contracts in other dimensions. Second, the existence of equilibria may become an issue. When the parties' multidimensional public-good policies and perks are mapped on the dimension in which parties can commit (tax rates), voters' preferences are not necessarily single-peaked in this dimension. Also, the conditions on the non-existence of an equilibrium in multidimensional policy spaces set out by Duggan and Fey (2005) would likely carry over to our model when it is extended by contracts on all dimensions.

To sum up, allowing for contracts on the financing side of government activities coupled with ways of retaining some flexibility to respond to aggregate shocks appears to be the most promising line of future research on tax contracts.

9 Conclusion

We have examined a novel institution called tax contracts and find that, particularly in polarized societies, it exhibits socially desirable properties. In addition, allowing parties to make binding tax promises may have further advantages, giving parties access to political leadership that can be trusted to fulfill their promises and to provide public services. While such leadership can be gained by maintaining a reputation for rewarding agents and citizens appropriately in a society (Myerson, 2008), tax contracts could function as a complementary measure.

The discussion in the last section has identified a variety of features that have to be considered when tax contracts are introduced. In particular, striking a balance between flexibility and commitment, endogenous perks, and a certification procedure involving changes of constitutional norms are important considerations. Nevertheless, it appears that modern liberal democracy can afford to experiment with the use of tax contracts.

Appendix

A Formal Details on Strategic Voting without Tax Contracts

In this section we formally depict the trade-off of strategic voters between avoiding the large amounts of perks associated with a grand coalition and the expense of a less desirable policy. We subsequently argue that of the three possible deviations from sincere voting given in Section 4.2 only deviation (1) is of interest for the paper. Finally, we define the threshold value p^{nc} which distinguishes sincere and strategic voting equilibria.

To formally depict the trade-off that strategic voters face, we can use equation (7) to write the tax rate of a party k with s_k seats in the legislature as

$$t_k(s_k) := \frac{A}{y_k} + \frac{(1+\lambda)b}{Y}\frac{S}{2} + \frac{(1+\lambda)b}{Y}\left(s_k - \frac{S}{2}\right) = t_k(S/2) + \frac{(1+\lambda)b}{Y}\left(s_k - \frac{S}{2}\right). \quad (13)$$

This notation separates the share of the tax rate financing the perks for a supermajority government from the tax rate of this party if it were to form a single-party government with a minimal majority of seats, $t_k(S/2)$. Accordingly, we decompose the tax rate of a grand coalition into

$$t_{LR}(S-s_E) = \frac{A}{y_{LR}} + \frac{(1+\lambda)b}{Y}\frac{S}{2} + \frac{(1+\lambda)b}{Y}\left(\frac{S}{2} - s_E\right) = t_{LR}(S/2) + \frac{(1+\lambda)b}{Y}\left(\frac{S}{2} - s_E\right).$$

The difference in a voter i's utility from the policies of two different parties j and k, when they form a single-party government, can then be expressed in the following way:

$$\begin{aligned} &A\ln g_k - t_k(S/2)y_i - (A\ln g_j - t_j(S/2)y_i) - \frac{y_i(1+\lambda)b}{Y}\left[\left(s_k - \frac{S}{2}\right) - \left(s_j - \frac{S}{2}\right)\right] \\ &= U_{y_i}(t_k(S/2), S/2) - U_{y_i}(t_j(S/2), S/2) - \frac{y_i(1+\lambda)b}{Y}\left[\left(s_k - \frac{S}{2}\right) - \left(s_j - \frac{S}{2}\right)\right], \end{aligned} \quad (14)$$

where $U_{y_i}(t_k(S/2), S/2)$ reflects the utility of a conventional voter with income y_i from a policy $(t_k(S/2), S/2)$. That is, a government with a minimal majority of seats, $S/2$, implements a tax rate $t_k(S/2)$ implying public-good provision $g_k = g(t_k(S/2), S/2)$ via the budget constraint. In the remainder of this section of the Appendix, we shall abbreviate $(t_k(S/2), S/2)$ to $(t_k, S/2)$ if the argument of t_k equals the number of seats

occupied by the government. Equation (14) reveals that utility differences can be divided into two parts: the difference in the political orientation of parties k and j and the different sizes of supermajorities in the legislature. Note that a party's public-good provision is unaffected by the number of governmental seats. The latter merely determine how expensive the respective level of public-good provision will be. Accordingly, (14) illustrates that a voter preferring the political orientation of party k may trade off a better policy with respect to public-good provision, which is more expensive due to a large supermajority, for a worse public-good policy that is cheaper in terms of perks. For this reason, a subset of voters could decide to vote strategically for their less preferred party.

A necessary condition for the subset $[y_m, \hat{y}_j]$ to deviate from sincere voting (this is deviation (1) described in Section 4.2) is that the outmost member of the set, the voter characterized by $\hat{y}_j$, will vote strategically. Thus, coordination on party j is possible if and only if

$$\begin{aligned}\frac{1-p}{2}[U_{\hat{y}_j}(t_k, S/2) + U_{\hat{y}_j}(t_j, S/2)] + p[U_{\hat{y}_j}(t_{RL}, S - s_E)] \\ < (1-p)U_{\hat{y}_j}(t_j, \frac{S^2}{2(S-s_E)}) + p\, U_{\hat{y}_j}(t_j, S/2).\end{aligned}$$

This condition can be rewritten to directly illustrate the trade-off between public-good policy and perks:

$$\begin{aligned}\frac{1-p}{2}[U_{\hat{y}_j}(t_k, S/2) - U_{\hat{y}_j}(t_j, S/2)] + p[U_{\hat{y}_j}(t_{RL}, S/2) - U_{\hat{y}_j}(t_j, S/2)] \\ < \frac{\hat{y}_j}{Y}(1+\lambda)b\left(p(S/2 - s_E) - (1-p)\frac{S}{2}\frac{s_E}{S-s_E}\right).\end{aligned} \tag{15}$$

The left-hand side of the inequality reflects the utility loss of voting strategically for the less preferred party. With respect to $\hat{y}_j$, this is the utility loss from having j's preferred level of public-good provision in both events, E and ¬E. The right-hand side represents the utility gain from lower perks.

In a further step, we now examine possible motives for deviations from sincere voting according to items (2) and (3) of the list in Section 4.2. We start with deviation (2), i.e. subset $[y_m, \tilde{y}_j]$ with measure greater than $[y_m, \hat{y}_j]$ votes strategically. In this case, party j forms a single-party government in events E and ¬E. We can write the participation constraint for voter $\tilde{y}_j$ as follows:

$$\begin{aligned}&\frac{1-p}{2}[U_{\tilde{y}_j}(t_k, S/2) - U_{\tilde{y}_j}(t_j, S/2)] + p[U_{\tilde{y}_j}(t_{RL}, S/2) - U_{\tilde{y}_j}(t_j, S/2)] \\ &< \frac{\tilde{y}_j}{Y}(1+\lambda)b\left(p(S/2 - s_E) - [(1-p)(s_j^{\neg E} - S/2) + p(s_j^E - S/2)]\right),\end{aligned} \tag{16}$$

where s_j^E and $s_j^{\neg E}$ denote the number of seats of party j in events E and ¬E, respectively. Further, we have $s_j^{\neg E} - S/2 > \frac{S}{2}\frac{s_E}{S-s_E}$ and $s_j^E - S/2 > 0$.

If we compare deviation (2) with deviation (1) using equations (15) and (16), we observe that voter $\tilde{y}_j$ would enjoy the same utility from public-good provision with both deviations. However, the reduction in perks from avoiding a grand coalition in event E is smaller with deviation (2) than with deviation (1). In this sense, the set $[\hat{y}_j, \tilde{y}_j]$ has an incentive to return to sincere voting. Hence we conclude that deviation (2) would not occur in equilibrium.

This may be different for deviation (3). Here the subset of strategic voters $[y_m, \tilde{y}_j]$ is smaller than $[y_m, \hat{y}_j]$, which implies that in event E a grand coalition will be formed. The trade-off for voter $\tilde{y}_j$ is to accept a single-party government formed by his less-preferred party with a supermajority of seats for a better policy with respect to public-good provision by a grand coalition in event E. This trade-off can be illustrated by writing the participation constraint of voter $\tilde{y}_j$ as follows:

$$\begin{aligned}&\frac{1-p}{2}[U_{\tilde{y}_j}(t_k, S/2) - U_{\tilde{y}_j}(t_j, S/2)] + (1-p)(1+\lambda)b\frac{\tilde{y}_j}{Y}(s_j^{\neg E} - S/2) < \\ &p\left[A\ln\left[\frac{(S-s_E)AY}{c(1+\lambda)(s_j^E y_j + s_k^E y_k)}\right] - \frac{(S-s_E)A\tilde{y}_j}{s_j^E y_j + s_k^E y_k} - A\ln\left[\frac{2AY}{c(1+\lambda)(y_j+y_k)}\right] + \frac{2A\tilde{y}_j}{y_j+y_k}\right].\end{aligned} \tag{17}$$

with $s_j^E > s_k^E$. The left-hand side of (17) reflects the expected loss of a less desirable policy with respect to public-good provision in $\neg E$, as well as the amount of additional perks for the supermajority of the single-party government of j. On the right-hand side is the expected utility gain from party j having greater weight in coalition bargaining. Two points are worth mentioning. For this deviation to be profitable for the critical voter, his income $\tilde{y}_j$ must be in $(y_m, \frac{y_j+y_k}{2})$. Accordingly, the changes in the policy of a grand coalition due to the additional vote share $(y_m, \tilde{y}_j]$ are not very high. By contrast, depending on the polarization of society, the policy change in $\neg E$ may be strong, as party j can now form a single-party government implementing its preferred policy. Additionally, party j possesses a supermajority involving higher perks than with sincere voting. Taken together, it seems that deviation (3) will be favorable in

those rather special cases where e.g. p is very high.[28] Thus the focus of the paper is on deviation (1) rather than on deviation (3). In principle, we could just neglect a possible equilibrium supported by deviation (3). To establish a clear formal basis for the analysis we make the following assumption:

Assumption 2
There does not exist a $\tilde{y}_j$ where $j \in \{R, L\}$ and $||[y_m, \tilde{y}_j]|| < ||[y_m, \hat{y}_j]||$ such that (17) holds.

As discussed earlier, strategic voting according to deviation (1) will only occur if (15) is satisfied. As we will show in the next lemma, this condition is not satisfied when $p = 0$. Consequently, there only exist positive values of p where strategic voting occurs if the right-hand side of (15) increases more strongly with p than the left-hand side. In this case there exists a $\hat{p}_j > 0$ such that (15) is satisfied for all $p > \hat{p}_j$ but does not hold for $p < \hat{p}_j$. Clearly, as p is a probability, strategic voting can only occur if $\hat{p}_j < 1$. Consequently, we define

$$p_j^{nc} = \begin{cases} \min\{\hat{p}_j, 1\} \ , & \text{if } \hat{p}_j > 0, \\ 1 \ , & \text{else.} \end{cases}$$

We now obtain the following lemma:

Lemma 5
Given Assumption 2,

(i) if $p > 0$ is sufficiently small, no subset of voters will vote strategically;

(ii) if $[U_{\hat{y}_j}(t_{RL}, S/2) - U_{\hat{y}_j}(t_j, S/2)] < \frac{\hat{y}_j}{Y}(1+\lambda)b(S/2 - s_E)$, there exists a unique $p_j^{nc} \in (0, 1)$ such that for all $p \leq p_j^{nc}$ the subset of voters $[y_m, \hat{y}_j]$ will not deviate from sincere voting and for $p > p_j^{nc}$ the subset of voters $[y_m, \hat{y}_j]$ will vote strategically.

The proof can be found in Appendix D.5. Note that in general $p_L^{nc} \neq p_R^{nc}$.[29] Therefore it is possible that, for a given probability p, all individuals in the set $[y_m, \hat{y}_L]$ prefer to vote strategically. But this is not the case for all individuals in the set $[\hat{y}_R, y_m]$ and

[28] More precisely, deviation (3) is favorable if (17) holds.

[29] Note that the preferences of voters are symmetric in the sense that half of the electorate would prefer either of the conventional parties if they voted sincerely, but in general utility derived from policies is not symmetric. As an example, consider two individuals characterized by incomes y_l and y_r that have the same relative position to the median income, i.e. $y_l/y_m = y_m/y_r$. Our model does not possess the feature that both individuals will enjoy the same level of utility from the median voter's most-preferred policy (t_{y_m}, g_{y_m}).

vice versa. However, if $p > p^{nc} := \min_{j\in\{L,R\}}\{p_j^{nc}\}$, there exists an incentive to vote strategically for least one party j.

B Conditions on Caretaker Policy

For the following arguments it is convenient to express the tax rate in terms of public goods and perks via the budget constraint. Then the condition that the formateur party j will prefer a single-party government with policy $(t, g, d) = (t_{sg}, g_{sg}, 0)$ to a grand coalition with policy $(t_{gg}, g_{gg}, 0)$ can be written as[30]

$$C_1(g_{sg}) \equiv A\ln(g_{sg}/g_{gg}) - \frac{y_j(1+\lambda)c}{Y}(g_{sg} - g_{gg}) + \frac{y_j(1+\lambda)}{Y}s_k b > 0.$$

The first two terms reflect the utility gain of the formateur from the public-good policy in the single-party government relative to that in the grand coalition, while the last term represents the amount of resources saved from the fact that a single-party government does not have to finance the perks for the other conventional party. Note that if the formateur chooses $g_{sg} = g_{gg}$, the first two terms will add up to 0. Hence the last term reflects the utility gain over the grand coalition, which could be used to adjust the public-good policy in favor of the other conventional party to motivate it to approve the single-party government in the vote of confidence if the formateur does not possess the absolute majority of seats in the legislature.

The other conventional party will favor a caretaker government to the single-party government of the formateur if

$$C_2(g_{sg}, g_{ct}) \equiv A\ln(g_{sg}/g_{ct}) - \frac{y_k(1+\lambda)c}{Y}(g_{sg} - g_{ct}) - \frac{y_k(1+\lambda)}{Y}s_j b < 0.$$

The first two terms represent the difference in public-good policy between the formateur's single-party government and the caretaker government, while the last term captures the additional costs in terms of perks associated with the formateur's single-party government.

The challenge for the formateur when aiming at a minority government will be to find a g_{sg} such that $C_1(g_{sg}) > 0$ (single-party government with g_{sg} is beneficial for formateur j) and $C_2(g_{sg}, g_{ct}) > 0$ (other conventional party k prefers single-party government with g_{sg} to a caretaker government). Our assumptions stated in Section 3.5 with respect

[30]Throughout this section we refer to j as the formateur and $k \neq j$ as the competing conventional party.

to the parameter values and caretaker policy imply precisely that such a g_{sg} does not exist.[31]

B.1 Detailed Characterization

We will now further characterize the set of caretaker policies such that conditions (a) and (b) in Section 3.5 cannot be simultaneously satisfied. We will proceed in five steps.

Step 1. First, it is useful to establish the following lemma. We use $C_2(\cdot, g_{ct})$ to denote C_2 as a function of g_{sg} for a given value g_{ct}, while $C_2(g_{sg}, \cdot)$ represents C_2 as a function of g_{ct} for a given g_{sg}.

Lemma 6

(i) $C_1(\cdot)$, $C_2(\cdot, g_{ct})$ *are strictly concave and* $C_2(g_{sg}, \cdot)$ *is strictly convex.*

(ii) The function $C_1(\cdot)$ *has a unique maximum at* $g_{sg} = g_j^*$ *and* $C_2(\cdot, g_{ct})$ *has a unique maximum at* $g_{sg} = g_k^*$. $C_2(g_{sg}, \cdot)$ *possesses a unique minimum at* $g_{ct} = g_k^*$.

Both claims in Lemma 6 can be easily verified in standard ways. The lemma implies that the functions $C_1(\cdot)$, $C_2(\cdot, g_{ct})$ and $C_2(g_{sg}, \cdot)$ have at most two roots on the positive real line. In fact, if $b > 0$, both $C_1(\cdot)$ and $C_2(g_{sg}, \cdot)$ possess exactly two real roots since $C_1(g_j^*) > 0$ and $C_2(g_{sg}, g_k^*) < 0$.

Step 2. The formateur j is willing to offer any policy g_{sg} with $C_1(g_{sg}) > 0$. Hence, the root of $C_1(g_{sg})$ that is closest to the ideal point of the opposed party k, reflects the maximal concession the formateur is willing to make to win the support of the other conventional party. Let us denote this policy by $\hat{g}$. If the set of policies for which $C_1(g_{sg}) > 0$ includes both parties' ideal points, then we let $\hat{g}$ assume the value of the non-formateur party's ideal point g_k^*.[32]

Step 3. It follows from Lemma 6 that a sufficient condition that there does not exist a g_{sg} for which both C_1 and C_2 are positive is that C_2 is negative at the point $g_{sg} = \hat{g}$, i.e. $C_2(\hat{g}, g_{ct}) < 0$.

Whether $C_2(\hat{g}, g_{ct}) < 0$, depends on the caretaker policy g_{ct}. According to Lemma 6, $C_2(\hat{g}, g_{ct})$ possesses two real roots $\underline{g}_{ct}$ and $\bar{g}_{ct}$. Since $C_2(\hat{g}, g_{ct})$ is strictly convex in

[31] Note that $C_1(g_{sg}) > 0$ implies that g_{sg} satisfies condition (a) in Section 3.5 and $C_2(g_{sg}, g_{ct}) > 0$ implies that g_{sg} satisfies condition (b) in Section 3.5.

[32] Even though j would be willing to move farther away from its ideal policy than g_k^*, such an offer would not increase k's utility and hence would not increase k's willingness to support a single-party government of j. Hence, j offers g_k^* in this case.

g_{ct}, $C_2(\hat{g}, g_{ct}) < 0$ if and only if $g_{ct} \in (\underline{g}_{ct}, \bar{g}_{ct})$. That is, party k declines any offer g_{sg} satisfying $C_1(g_{sg}) > 0$ if the caretaker government's policy g_{ct} is not too bad for party k, i.e. is in the interval $(\underline{g}_{ct}, \bar{g}_{ct})$. If g_{ct} is not in this interval, i.e. sufficiently bad for party k, then there exists an offer g_{sg} that is beneficial for j and will be accepted by k.

Step 4. The set $(\underline{g}_{ct}, \bar{g}_{ct})$ depends on the identities of parties k and j, i.e. the interval when L is the formateur is different from that when R is the formateur. Furthermore, it is clear that the intervals depend on the specific parameter values, particularly on the amount of perks, b. Thus, for a given amount of perks, we denote the set of caretaker policies for which no policy g_{sg} satisfying $C_1(g_{sg}) > 0$ and $C_2(g_{sg}, g_{ct}) > 0$ exists by $(\underline{g}^R_{ct}(b), \bar{g}^R_{ct}(b))$ if L is the formateur and R opposes a possible minority government, and by $(\underline{g}^L_{ct}(b), \bar{g}^L_{ct}(b))$ in the opposite case where R is the formateur.

The caretaker policies implying that no g_{sg} exists that satisfies conditions (a) and (b) in Section 3.5 are those that lie in the intersection of $(\underline{g}^R_{ct}(b), \bar{g}^R_{ct}(b))$ and $(\underline{g}^L_{ct}(b), \bar{g}^L_{ct}(b))$. Typically the intersection will be the set $(\underline{g}^L_{ct}(b), \bar{g}^R_{ct}(b))$.[33]

Step 5. An important question is whether the intersection of $(\underline{g}^R_{ct}(b), \bar{g}^R_{ct}(b))$ and $(\underline{g}^L_{ct}(b), \bar{g}^L_{ct}(b))$ is non-empty. Due to difficulties to characterize the roots of $C_1(g_{sg})$ and $C_2(g_{sg}, g_{ct})$ analytically, we employ evidence via numerical simulation, deriving the intervals satisfying the assumptions in Section 3.5 for different amounts of perks, b.

For the baseline simulation we use the values $c = 1, A = .01, \lambda = 0.05, Y = 1000, s_L = 1 = s_R, y_L = .85$. Assuming a median income of 1, we derive y_R according to the assumption of the symmetry of party platforms adopted throughout most parts of the paper.[34] We increase the amount of perks b from $b_{min} = 0$ to $b_{max} = 2$ in steps of 0.002 and derive the respective intervals of caretaker government policies described above. Figure 1 shows the results. The red curves depict the interval $(\underline{g}^L_{ct}(b), \bar{g}^L_{ct}(b))$ while the blue curves show the bounds of the interval $(\underline{g}^R_{ct}(b), \bar{g}^R_{ct}(b))$. The caretaker government policies in the shaded area, $(\underline{g}^L_{ct}(b), \bar{g}^R_{ct}(b))$, satisfy the assumptions made in Section 3.5.

We further depicted the ideal policies of the conventional parties and the policy of a grand coalition by gray lines. The graph illustrates that with larger amounts of perks the interval of admissible caretaker polices increases. Our simulation results confirm that the respective interval is non-empty for all values of b. For $b = 0$ the only caretaker

[33] The reasoning behind it is that, according to Lemma 6, the maximum of C_2 with respect to g_{ct} for a given g_{sg} when L is formateur, g^*_R, is lower than the one when R is the formateur, g^*_L.

[34] See Subsection 3.6.

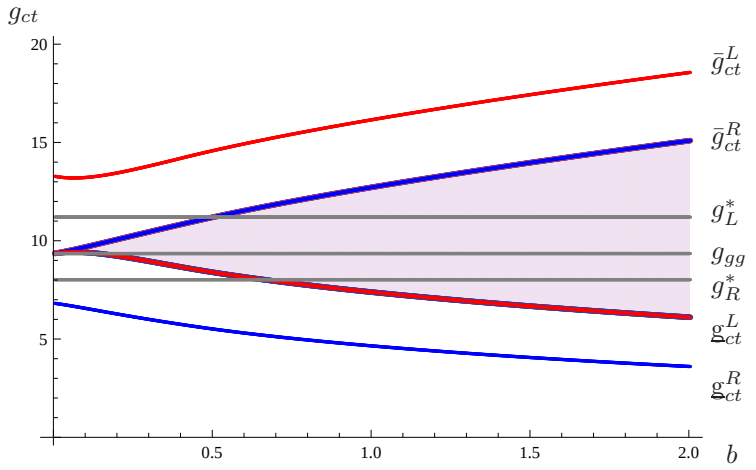

Figure 1: Caretaker policies implying that no g_{sg} satisfies conditions (a) and (b) in Section 3.5

policy for which no g_{sg} satisfies $C_1 > 0$ and $C_2 > 0$ is the policy of a grand coalition[35] as can be derived directly from the definition of the functions $C_1(g_{sg})$ and $C_2(g_{sg}, g_{ct})$. Various robustness checks by changing the different parameter values revealed the same qualitative pattern. In fact, we did not find a specification where the set of admissible caretaker policies is empty.[36]

The graph can also be read in a different way. The figure determines the lowest amount of perks so that a given caretaker policy is in the set of policies satisfying the assumptions in Section 3.5. For example, $g_{ct} = g^*_R$ would be admissible for all b larger than about 0.65.

If a caretaker government policy is given, we can also analytically derive a sufficient condition for b such that no g_{sg} satisfies conditions (a) and (b) in Section 3.5. Since $C_2(g_{gs}, g_{ct})$ is decreasing with the amount of perks, b, $C_1(g_{sg}) > 0$ and $C_2(g_{sg}, g_{ct}) > 0$ cannot be satisfied if $C_2(g_{sg}, g_{ct}) < 0$ for all g_{sg}. As $C_2(\cdot, g_{ct})$ is maximal at $g_{sg} = g^*_k$, $C_2(g_{sg}, g_{ct}) < 0$ for all g_{sg} is satisfied if b is larger than both threshold values (one if L is the formateur and one if R is the formateur) satisfying $C_2(g^*_k, g_{ct}) = 0, k \in L, R$. This sufficient condition would require that perks are sufficiently large that the non-formateur party would decline the formateur's offer even if the latter offered the non-formateur's ideal policy g^*_k. This threshold level of b must be larger or equal to the one depicted in Figure 1 for any given caretaker policy g_{ct}.

Note that we also assume that both conventional parties prefer a grand coalition to a

[35]There are no perks to induce a deviation from the fall back option which is a grand coalition.
[36]We are glad to provide the code of our simulations upon request.

caretaker government. We discuss this additional assumption in the next subsection.

B.2 Both parties prefer grand coalition to caretaker government

Both conventional parties preferring a grand coalition to a caretaker government implies that

$$C_3(g_{ct}) \equiv \frac{AY}{1+\lambda}\ln(g_{gg}/g_{ct}) - cy_j(g_{gg}-g_{ct}) - b\Big[y_j s_k + s_j\big(y_j - \frac{\theta Y}{1+\lambda}\big)\Big] > 0.$$

We observe that this condition is always satisfied if θ is sufficiently large. Note also that θ does not appear in the earlier conditions. Thus a large value of θ has no influence on whether or not C_1 and C_2 are positive.

We next explore the consequences of the condition $C_3(g_{ct}) > 0$. It can be easily verified that $C_3(g_{ct})$ is strictly convex and possesses a unique minimum at $g_{ct} = g_j^*$. Consequently, $C_3(g_{ct})$ possesses at most two real roots, which we denote by $\underline{g}_{ct}^{j3}(b)$ and $\bar{g}_{ct}^{j3}(b)$. The superscript indicates that the roots are derived from condition $C_3(g_{ct}) > 0$ and depend on the particular party under consideration $j \in \{L, R\}$. Furthermore, as the roots depend on the amount of perks, we write them as functions of b. Suppose for a given b, two real roots of $C_3(g_{ct})$ exist. Then $C_3(g_{ct}) < 0$ for all $g_{ct} \in (\underline{g}_{ct}^{j3}(b), \bar{g}_{ct}^{j3}(b))$. That is, the caretaker policy is so close to party j's ideal policy that party j prefers it to a grand coalition. However, for all caretaker policies outside of the interval, j prefers a grand coalition. If no real roots exist, $C_3(g_{ct})$ is always positive, implying that j prefers a grand coalition independently of the value of g_{ct}.

In Figure 2 we add the additional conditions to the graph of Figure 1. That is, we use the same parameter values as in the previous specification and included a very low value of $\theta = 0.004$. The red (blue) curves enclose the values of g_{ct} for a given amount of perks b where party L (R) prefers a caretaker government to a grand coalition.

In the figure, we set the value of both $\underline{g}_{ct}^{j3}(b)$ and $\bar{g}_{ct}^{j3}(b)$ to zero if condition C_3 possesses no real roots, thereby indicating that all g_{ct} satisfy $C_3(g_{ct}) > 0$. We observe in Figure 2 that if the amount of perks is larger than approximately 0.2, a grand coalition is always preferred. While for very low values of b, e.g. at $b = 0.05$, party L would prefer a caretaker government with policy between $g_{ct} = 10$ and $g_{ct} = 12$ while party R would prefer a caretaker government with policy between $g_{ct} = 7.5$ and $g_{ct} = 8.5$. The important insight from the Figure 2 is that the sets of caretaker government policies preferred

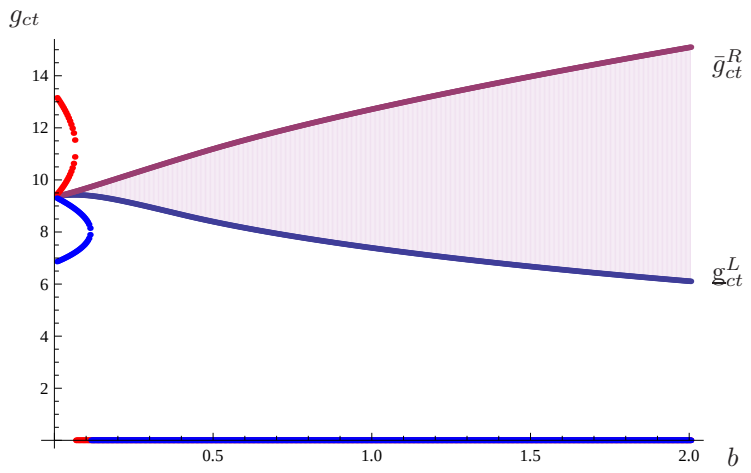

Figure 2: Caretaker policies implying that no g_{sg} satisfies conditions (a) and (b) in Section 3.5 and additionally $C_3(g_{ct}) > 0$ holds.

to a grand coalition do not intersect with the shaded area representing the caretaker government policies for which no g_{sg} satisfies assumptions (a) and (b) in Section 3.5. Hence, the set of caretaker policies satisfying the assumption in Section 3.5 also satisfy the requirement that a grand coalition is preferred to a caretaker government. As C_3 is monotonically increasing in θ, the preceding conclusion is a fortiori true for all values of θ larger than 0.004 chosen in this illustration. However, for extremely low values of θ the set of admissible caretaker policies additionally satisfying $C_3(g_{ct}) > 0$ may be smaller than the set for which no g_{sg} satisfying (a) and (b) of Section 3.5 exists.

C Preferred Tax Rates of Median Voter

Under sincere voting, two situations can occur if the extreme party enters parliament and no party obtains an absolute majority of votes. First, the conventional parties' tax contracts overlap, and a grand coalition accedes to power with $S - s_E$ seats. Second, the tax contracts do not overlap, and there will be a caretaker government. We denote the median voter's preferred tax rates by $t_{y_m}(S/2, S - s_E)$ and $t_{y_m}(S/2, ct)$ when a grand coalition is possible in E and when a caretaker government occurs in E, respectively.[37] We first show that these tax rates are unique, and in general they are not identical.

[37] In other words, $t_{y_m}(S/2, S - s_E)$ is the ex-ante optimal tax rate of the median voter if with probability $1 - p$ a single-party government occupying $S/2$ seats will form and with probability p a grand coalition will be formed. $t_{y_m}(S/2, ct)$ is the ex-ante optimal tax rate of the median voter if with probability $1 - p$ a single-party government occupying $S/2$ seats forms and with probability p a caretaker government takes power.

Finally the constellation $\tau_L = \tau_R = \{t_{y_m}(S/2, S - s_E)\}$ cannot be an equilibrium as we will explain at the end of this section. Therefore we set $t_{y_m} = t_{y_m}(S/2, ct)$.

To derive the preferred tax rate of the median voter, we use $g(t, s)$ as defined in Subsection 5.2 to denote the level of public goods a government provides when it occupies s seats and has committed itself to a tax rate t. If the parties' tax sets are not disjunct, the median voter's preferred (second-best) policy[38] is the solution of the following problem:

$$\max_t \mathbb{E}[U_{y_m}] = p(A \ln g(t, S - s_E) + (1 - t)y_m) + (1 - p)(A \ln g(t, S/2) + (1 - t)y_m).$$

We obtain:

Lemma 7

There exists a unique tax rate $t_{y_m}(S/2, S - s_E)$ that maximizes the median voter's expected utility under sincere voting if $t_{y_m}(S/2, S - s_E)$ is in the intersection of the conventional parties' tax sets. It is given implicitly by

$$\frac{Y}{(1 + \lambda)c}\left(\frac{pA}{g(t, S - s_E)} + \frac{(1 - p)A}{g(t, S/2)}\right) \quad = \quad y_m. \tag{18}$$

The proof is obvious. As $g(t, S - s_E)$ and $g(t, S/2)$ are strictly increasing in t, the left-hand side of (18) is strictly decreasing in t. Hence there is unique $t_{y_m}(S/2, S - s_E)$, which can be determined by solving (18) for t.

Now we turn to the situation where the conventional parties' tax sets do not intersect. In this case there will be a caretaker government in event E whose value to the median voter is independent of the parties' tax-contract choices. So the problem of finding the optimal tax rate boils down to

$$\max_t \mathbb{E}[U_{y_m}] = (1 - p)(A \ln g(t, S/2) + (1 - t)y_m).$$

The solution is given by equation (3) in Section 3, i.e.

$$t_{y_m}(S/2, ct) := \frac{A}{y_m} + \frac{(1 + \lambda)b}{Y}\frac{S}{2}, \tag{19}$$

where *ct* stands for caretaker government.

The tax contract choices $\tau_L = \tau_R = \{t_{y_m}(S/2, S - s_E)\}$ cannot be an equilibrium as at least one party has an incentive to expand its tax set by choosing one boundary

[38] Recall that perks cannot be avoided.

closer to $t_{y_m}(S/2, ct)$. By doing so, it can win more votes under sincere voting.[39] Consequently, if a subset of voters coordinates on one party to avoid a grand coalition, it will be the one that deviated due to the minimal strategic voters refinement. Hence, the two conventional parties signing a tax contract with $t_{y_m}(S/2, S - s_E)$ cannot be an equilibrium. By contrast, the two conventional parties choosing $t_{y_m}(S/2, ct)$ is an equilibrium, as stated in Proposition 2. To simplify notation we abbreviate $t_{y_m}(S/2, ct)$ by t_{y_m}.

D Proofs

D.1 Proof of Proposition 1

We consider $p^{nc} < 1$, as otherwise only sincere voting will occur according to Lemma 5.

Step 1. By construction, the combination of strategies constitutes a subgame perfect equilibrium, as parties respond optimally to the election outcome. Citizens vote according to their preferences.

Step 2. Consider first situation (i), where $p \leq p^{nc}$.

We know from Lemma 5 that neither subset of voters $[y_m, \hat{y}_k]$, $k \in \{R, L\}$, has an incentive to deviate from sincere voting. Hence the equilibrium exists. It is unique, as any other candidate equilibrium must involve strategic voting. However, according to Lemma 5 there is no incentive for this to happen.

(ii) $p > p^{nc}$.

In this case, there exists a set of voters $[y_m, \hat{y}_j]$ that have an incentive to deviate from sincere voting. Therefore we have an equilibrium where a subset of voters of measure $|[y_m, \hat{y}_j]|$ will coordinate on party j. It is generally not unique because if

[39] The reason is that according to the definition of sincere voting in Subsection 5.1, the party whose policy in a single-party government will be closer to $t_{y_m}(S/2, ct)$ obtains more votes. Hence the set of strategic voters when coordinating on the deviating party (to prevent a grand coalition) is smaller than the set of strategic voters necessary when coordinating on the party offering $\{t_{y_m}(S/2, S - s_E)\}$. Hence, the minimal strategic voter requirement implies that voters will coordinate on the deviating party. Thus, the respective deviation is profitable. As another example, consider the deviation of one of the parties to $\{t_{y_m}(S/2, ct)\}$ given the other party offers $\{t_{y_m}(S/2, S - s_E)\}$. In this case, no grand coalition is feasible and the median voter's most preferred policy is $t_{y_m}(S/2, ct)$. Consequently the deviating party will obtain more votes under sincere voting and the set of strategic voters necessary to prevent a caretaker government is lower when coordinating on the deviating party. This implies that such a deviation is beneficial and hence, $\tau_L = \tau_R = \{t_{y_m}(S/2, S - s_E)\}$ cannot be an equilibrium.

$p > \max_{j\in\{L,R\}}\{p_j^{nc}\}$, there also exists an equilibrium where the voters in set $[y_m, \hat{y}_k]$, $k \neq j$, coordinate on party k. □

D.2 Proof of Lemma 3

The first point is obvious, as t_k^* is the utility-maximizing tax rate.

The second and third points follow from the single-peakedness of the parties' preferences, i.e. the strict concavity of $V_k = A\ln(\frac{tY-(1+\lambda)s_k b}{(1+\lambda)c}) + (1-t)y_k + \theta s_k b$. The second derivative of V_k is

$$\frac{d^2V_k}{dt^2} = -A\left(\frac{Y}{tY-(1+\lambda)s_k b}\right)^2 < 0.$$

□

D.3 Proof of Proposition 2

Step 1. First, we show that given $\tau_L = \tau_R = \{t_{y_m}\}$, strategic voting to prevent a grand coalition will occur if and only if $p > \hat{p}$.

For this purpose, we examine whether the subset of voters $[y_m, \hat{y}_j]$, $j \in \{L, R\}$ can improve their utility by voting strategically, given that $\tau_L = \tau_R = \{t_{y_m}\}$. When the subset $[y_m, \hat{y}_j]$ votes strategically, the expected utility of a voter with income $\hat{y}_j$ is given by

$$\begin{aligned}\tilde{U}_{\hat{y}_j} &= p\Big[A\ln\big(g(t_{y_m}, S/2)\big) + (1-t_{y_m})\hat{y}_j\Big] \\ &\quad +(1-p)\Big[A\ln\Big(g\Big(t_{y_m}, \frac{S^2}{2(S-s_E)}\Big)\Big) + (1-t_{y_m})\hat{y}_j\Big]\ .\end{aligned}$$

If the subset of voters in $[y_m, \hat{y}_j]$ votes sincerely, the expected utility of voter $\hat{y}_j$ is written as

$$\bar{U}_{\hat{y}_j} = p[A\ln(g(t_{y_m}, S-s_E)) + (1-t_{y_m})\hat{y}_j] + (1-p)[A\ln(g(t_{y_m}, S/2)) + (1-t_{y_m})\hat{y}_j]\ .$$

The condition $\tilde{U}_{\hat{y}_j} \leq \bar{U}_{\hat{y}_j}$ can be transformed to

$$p \leq \frac{\ln\left(\frac{g(t_{y_m},S/2)}{g(t_{y_m},\frac{S^2}{2(S-s_E)})}\right)}{\ln\left(\frac{g(t_{y_m},S/2)}{g(t_{y_m},S-s_E)}\right) + \ln\left(\frac{g(t_{y_m},S/2)}{g(t_{y_m},\frac{S^2}{2(S-s_E)})}\right)} =: \hat{p}\ .$$

Hence, voters will prefer to vote sincerely if $p \leq \hat{p}$ and to vote strategically if $p > \hat{p}$. Note that this condition is the same for $\hat{y}_L$ and $\hat{y}_R$, as the tax rate does not differ in the cases with and without strategic voting and the valuation of the public good is identical across agents. Hence, in the equilibria with tax contracts there is only one probability threshold for both critical voters $\hat{y}_L$ and $\hat{y}_R$.

Step 2.

We next examine whether parties might deviate in stage 1 of the political game by offering different tax contracts. In general, we proceed in the following way: For each possible deviation of a party, we have to consider how voting behavior changes and what its implications are for government formation. With respect to voting behavior, we first determine the outcome under sincere voting and check whether subsets of voters can improve by voting strategically. This enables us to determine voting behavior in the case of a deviation from the equilibrium tax contracts, and we can compare the deviating party's utility with the utility it would obtain without the deviation.

Without loss of generality, assume that party R offers $\tau_R^d = [\underline{t}_R^d, \bar{t}_R^d]$. (Note that $\tau_L = \{t_{y_m}\}$.) We consider the following deviations:

(d1) $\underline{t}_R^d < t_{y_m}$ and $\bar{t}_R^d = t_{y_m}$. That is, party R expands its set of tax rates toward its ideal policy. With this deviation, R is still able to form a grand coalition with policy $(t_{y_m}, S - s_E)$ but could implement a policy closer to its ideal point if it were the only party in government.

As R would implement a policy with $t_R < t_{y_m}$ (e.g. $(\underline{t}_R^d, s_R)$) in a single-party government, with sincere voting there exists a $\tilde{y} > y_m$ such that all voters with income $y_i > \tilde{y}$ will vote for R, while the other voters will vote for L. If $\tilde{y} < \hat{y}_L$, sincere voting would yield a grand coalition with policy $(t_{y_m}, S - s_E)$ in event E and a single-party government of L in event ¬E involving a policy $(t_{y_m}, S/2 + s^+)$, where s^+ is the number of seats derived from the votes of $[y_m, \tilde{y}]$. In case of a large deviation of $\underline{t}_R^d$ from t_{y_m}, we obtain $\tilde{y} \geq \hat{y}_L$ implying a single-party government of L in both events, E and ¬E, under sincere voting.

If $\tilde{y} < \hat{y}_L$, subsets of voters could now decide to vote strategically in order to either reduce the size of the single-party government of L in event ¬E or to prevent a grand coalition in event E. In the first case, all voters in $[y_m + \varepsilon, \tilde{y}]$, where ε is a positive very small number, will vote for R, so that L will form a single-party government with a minimal majority of seats. In the second case, the subset of

voters $[y_m, \hat{y}_L]$ will coordinate on L rather than R to avoid a grand coalition due to the minimal set of strategic voters refinement. In case that $\tilde{y} \geq \hat{y}_L$, strategic voting would either reduce the size of the single-party government of L to its minimum $S/2$ in event E or $\neg$E, depending on the value of p. A single-party government of L with a minimal majority of seats in event $\neg$E would involve a grand coalition in E with policy t_{y_m}.

The deviating party R will be worse off with the deviation as in the case where a grand coalition occurs in event E, party L will form a single-party government in event $\neg$E, while R will realize the same utility level from a grand coalition in event E as without the deviation. Further, if the voters avoid a grand coalition, R has no chance of forming a single-party government.

(d2) $\underline{t}_R^d < t_{y_m}$ and $\bar{t}_R^d < t_{y_m}$. In this case, party R does not include t_{y_m} in its tax set. Hence, a grand coalition is not possible. In a single-party government, R would choose $(\underline{t}_R^d, s_R)$. Note that we do not preclude $\underline{t}_R^d = \bar{t}_R^d$.

Under sincere voting, there will again exist a critical voter $\tilde{y} > y_m$ such that all voters i with income $y_i > \tilde{y}$ will vote for R, and the other voters will vote for L. The policy outcome would be a caretaker government in event E and a single-party government of L in event $\neg$E with policy $(t_{y_m}, S/2 + s^+)$.

Two improvements by strategic voting are conceivable. First, if a caretaker government yields sufficiently low utility for the voters, it can be avoided by a subset $[\tilde{y}, \hat{y}_k]$ coordinating on party $k \in \{R, L\}$. Due to the minimal strategic voting refinement, the party the voters coordinate on is L. A second situation arises when the caretaker government is acceptable for the voters. Then, as with deviation (d1), the subset $[y_m + \varepsilon, \tilde{y}]$ would coordinate on R in order to have a minimal majority for the single-party government of L in event $\neg$E.

In any of the above cases, party R will have no chance of coming into power. Hence deviation (d2) is not profitable.

(d3) $\underline{t}_R^d \geq t_{y_m}$ and $\bar{t}_R^d > t_{y_m}$. For completeness, we add that a possible deviation by R might also be to expand the set to include policies t farther away from its ideal point, or even to have only policies farther away from the ideal point without including t_{y_m}. The latter deviation leaves R worse off due to the same arguments as above. In the former case, R cannot profit from expanding its tax set in the direction of L's ideal given that $\tau_L = \{t_{y_m}\}$. However, this deviation will incur

no utility loss and is neglected due to the assumption that in case of indifference the smallest tax set will be chosen. Hence, $\tau_L = \tau_R = \{t_{y_m}\}$ is an equilibrium. □

D.4 Proof of Lemma 4

The critical probability $\hat{p}$ as defined in the Proof of Proposition 2 can be written as

$$\hat{p} = \frac{\ln\left[\frac{1}{1-\frac{b}{cg(t_{y_m},S/2)}\frac{S}{2}\frac{s_E}{S-s_E}}\right]}{\ln\left[\left(\frac{1}{1-\frac{b}{cg(t_{y_m},S/2)}(\frac{S}{2}-s_E)}\right)\left(\frac{1}{1-\frac{b}{cg(t_{y_m},S/2)}\frac{S}{2}\frac{s_E}{S-s_E}}\right)\right]}.$$

Multiplying both the denominator and the numerator by (-1) and applying some minor mathematical manipulations yields

$$\hat{p} = \frac{\ln(1-\frac{b}{cg(t_{y_m},S/2)}\frac{S}{2}\frac{s_E}{S-s_E})}{\ln(1-\frac{b}{cg(t_{y_m},S/2)}(\frac{S}{2}-s_E))+\ln(1-\frac{b}{cg(t_{y_m},S/2)}\frac{S}{2}\frac{s_E}{S-s_E})}. \tag{20}$$

We will now show that

$$\left|\ln\left(1-\frac{b}{cg(t_{y_m},S/2)}\left(\frac{S}{2}-s_E\right)\right)\right| > \left|\ln\left(1-\frac{b}{cg(t_{y_m},S/2)}\frac{S}{2}\frac{s_E}{S-s_E}\right)\right|. \tag{21}$$

If this is the case, the denominator can be estimated from below by $2\ln(1-\frac{b}{cg(t_{y_m},S/2)}\frac{S}{2}\frac{s_E}{S-s_E})$, which directly implies that $\hat{p}$ must be smaller than 0.5.

Condition (21) transforms to

$$\frac{S}{2}-s_E > \frac{S}{2}\frac{s_E}{S-s_E}.$$

This inequality is equivalent to the grand coalition occupying more seats than a single-party government with $\frac{S^2}{2(S-s_E)}$ seats. We know from Lemma 2 that this is the case, as long as $s_E < S\left(1-\frac{\sqrt{8}}{4}\right)$, i.e. as long as Assumption 1 holds.

□

D.5 Proof of Lemma 5

(i) follows directly from $\frac{1}{2}[U_{\hat{y}_j}(t_k,S/2)-U_{\hat{y}_j}(t_j,S/2)] > -\frac{\hat{y}_j}{Y}(1+\lambda)b\frac{S}{2}\frac{s_E}{S-s_E}$.

(ii) We rewrite (15) as

$$\frac{1-p}{2}[U_{\hat{y}_j}(t_k, S/2) - U_{\hat{y}_j}(t_j, S/2)] + p[U_{\hat{y}_j}(t_{RL}, S/2) - U_{\hat{y}_j}(t_j, S/2)] \\ -\frac{\hat{y}_j}{Y}(1+\lambda)b\left(p(\frac{S}{2} - s_E) - (1-p)\frac{S}{2}\frac{s_E}{S-s_E}\right) < 0. \tag{22}$$

From (i) we know that for p sufficiently small, (22) is violated. The condition

$$[U_{\hat{y}_j}(t_{RL}, S/2) - U_{\hat{y}_j}(t_j, S/2)] < \frac{\hat{y}_j}{Y}(1+\lambda)b(S/2 - s_E)$$

implies that (22) is satisfied for $p = 1$. As the left-hand side of (22) is either strictly increasing or strictly decreasing with p, we infer that in the case where $[U_{\hat{y}_j}(t_{RL}, S/2) - U_{\hat{y}_j}(t_j, S/2)] < \frac{\hat{y}_j}{Y}(1+\lambda)b(S/2 - s_E)$ holds, it must be increasing with p. Then there is a unique p_j^{nc} where (22), and hence (15), holds with equality. □

D.6 Proof of Proposition 3

The proof has of two parts. In the first (i), we show that without political polarization $\hat{p} < p^{nc}$. In the second (ii), we verify that p^{nc} will increase with the degree of political polarization.

(i)

Using condition (15), p^{nc} can be derived as[40]

$$p^{nc} = \frac{\frac{1}{2}[U_{\hat{y}_j}(t_k, S/2) - U_{\hat{y}_j}(t_j, S/2)] + \frac{\hat{y}_j}{Y}(1+\lambda)b\frac{S}{2}\frac{s_E}{S-s_E}}{\frac{\hat{y}_j}{Y}(1+\lambda)b(\frac{S}{2} - s_E + \frac{S}{2}\frac{s_E}{S-s_E}) - U_{\hat{y}_j}(t_{RL}, S/2) + \frac{1}{2}[U_{\hat{y}_j}(t_k, S/2) + U_{\hat{y}_j}(t_j, S/2)]}. \tag{23}$$

Without political polarization, both conventional parties possess the same political platform, so in terms of policy the costs of voting strategically are zero. With no political costs for deviating from sincere voting, p^{nc} transforms to

$$p^{nc} = \frac{\frac{S}{2}\frac{s_E}{S-s_E}}{\frac{S}{2} - s_E + \frac{S}{2}\frac{s_E}{S-s_E}}$$

Using $\hat{p}$ as given in equation (20), the condition $\hat{p} < p^{nc}$ is equivalent to

$$\left(1 + \frac{\ln(1 - \frac{b}{cg(t_{ym}, S/2)}(\frac{S}{2} - s_E))}{\ln(1 - \frac{b}{cg(t_{ym}, S/2)}\frac{S}{2}\frac{s_E}{S-s_E})}\right)^{-1} < \left(1 + \frac{\frac{S}{2} - s_E}{\frac{S}{2}\frac{s_E}{S-s_E}}\right)^{-1}.$$

[40]Note that as in Appendix A, $U_{\hat{y}_j}(t_j, S/2)$ stands for $U_{\hat{y}_j}(t_j(S/2), S/2)$ and $t_j(S/2)$ is defined by (13).

This condition holds if and only if

$$\frac{\ln(1-\frac{b}{cg(t_{ym},S/2)}\frac{S}{2}\frac{s_E}{S-s_E})}{\frac{S}{2}\frac{s_E}{S-s_E}} > \frac{\ln(1-\frac{b}{cg(t_{ym},S/2)}(\frac{S}{2}-s_E))}{\frac{S}{2}-s_E}. \tag{24}$$

Note that $1-\frac{b}{cg(t_{ym},S/2)}s \in (0,1)$, where $s = \frac{S}{2}\frac{s_E}{S-s_E}, \frac{S}{2}-s_E$. As the function $f(x) = \frac{\ln(1-\gamma x)}{x}$ is strictly declining with x for the constant γ smaller than 1 and $x \in (0, 1/\gamma)$, condition (24) is satisfied, since $\frac{S}{2}\frac{s_E}{S-s_E} < \frac{S}{2}-s_E$.

(ii)

The proof that p^{nc} increases with political polarization proceeds in two steps. First, we prove that $\frac{dp^{nc}}{d\Delta_y} > 0$ under the assumption that

$$\frac{d[U_{\hat{y}_j}(t_k, S/2) - U_{\hat{y}_j}(t_j, S/2)]}{d\Delta_y} > 0.$$

In the second step, we show that this assumption holds.

(1) Recall the definition of p^{nc}:

$$p^{nc} = \frac{\frac{1}{2}[U_{\hat{y}_j}(t_k, S/2) - U_{\hat{y}_j}(t_j, S/2)] + \frac{\hat{y}_j}{Y}(1+\lambda)b\frac{S}{2}\frac{s_E}{S-s_E}}{\frac{\hat{y}_j}{Y}(1+\lambda)b(\frac{S}{2}-s_E+\frac{S}{2}\frac{s_E}{S-s_E}) - U_{\hat{y}_j}(t_{kj}, S/2) + \frac{1}{2}[U_{\hat{y}_j}(t_k, S/2) + U_{\hat{y}_j}(t_j, S/2)]}. \tag{25}$$

The derivative with respect to political polarization is

$$\frac{dp^{nc}}{d\Delta_y} = \frac{\frac{1}{2}\frac{d[U_{\hat{y}_j}(t_k,S/2)-U_{\hat{y}_j}(t_j,S/2)]}{d\Delta_y}D + N\frac{d(U_{\hat{y}_j}(t_{RL},S/2)-\frac{1}{2}[U_{\hat{y}_j}(t_k,S/2)+U_{\hat{y}_j}(t_j,S/2)])}{d\Delta_y}}{D^2},$$

where N and D denote the numerator and the denominator of the fraction in (25), respectively.

We need to show that $\frac{dp^{nc}}{d\Delta_y} > 0$. The corresponding condition can be written as

$$1 + p^{nc}\frac{\frac{d(U_{\hat{y}_j}(t_{RL},S/2)-\frac{1}{2}[U_{\hat{y}_j}(t_k,S/2)+U_{\hat{y}_j}(t_j,S/2)])}{d\Delta_y}}{\frac{1}{2}\frac{d[U_{\hat{y}_j}(t_k,S/2)-U_{\hat{y}_j}(t_j,S/2)]}{d\Delta_y}} > 0. \tag{26}$$

Note that for this transformation, we have assumed that $\frac{d[U_{\hat{y}_j}(t_k,S/2)-U_{\hat{y}_j}(t_j,S/2)]}{d\Delta_y} > 0$. Consider the derivative in the numerator.

$$\frac{dU_{\hat{y}_j}(t_{RL}, S/2)}{d\Delta_y} - \frac{1}{2}\left[\frac{dU_{\hat{y}_j}(t_k, S/2)}{d\Delta_y} + \frac{dU_{\hat{y}_j}(t_j, S/2)}{d\Delta_y}\right].$$

Note that the worst policy for $\hat{y}_j$ is $(t_j, S/2)$. By the concavity of the utility function, it must be the case that

$$\frac{d(U_{\hat{y}_j}(t_{RL}, S/2)}{d\Delta_y} > \frac{dU_{\hat{y}_j}(t_j, S/2)}{d\Delta_y}.$$

Hence we can state that

$$\frac{d(U_{\hat{y}_j}(t_{RL}, S/2) - \frac{1}{2}[U_{\hat{y}_j}(t_k, S/2) + U_{\hat{y}_j}(t_j, S/2)])}{d\Delta_y} > -0.5\Big[\frac{dU_{\hat{y}_j}(t_k, S/2)}{d\Delta_y} - \frac{dU_{\hat{y}_j}(t_j, S/2)}{d\Delta_y}\Big].$$

The denominator of (26) can be written as

$$\frac{1}{2}\Big[\frac{dU_{\hat{y}_j}(t_k, S/2)}{d\Delta_y} - \frac{dU_{\hat{y}_j}(t_j, S/2)}{d\Delta_y}\Big].$$

From this it follows directly that

$$\frac{\frac{d(U_{\hat{y}_j}(t_{RL}, S/2) - 0.5[U_{\hat{y}_j}(t_k, S/2) + U_{\hat{y}_j}(t_j, S/2)])}{d\Delta_y}}{\frac{1}{2}\frac{d[U_{\hat{y}_j}(t_k, S/2) - U_{\hat{y}_j}(t_j, S/2)]}{d\Delta_y}} > -1.$$

Since p^{nc} must be smaller than 1, condition (26) is satisfied for all values of p^{nc}, as long as $\frac{d[U_{\hat{y}_j}(t_k, S/2) - U_{\hat{y}_j}(t_j, S/2)]}{d\Delta_y} > 0$. We will now show that this is the case.

(2) $U_{\hat{y}_j}(t_k, S/2) - U_{\hat{y}_j}(t_j, S/2)$ can be written as

$$A\Big(\ln\Big(\frac{y_j}{y_k}\Big) + \hat{y}_j\Big(\frac{1}{y_j} - \frac{1}{y_k}\Big)\Big).$$

We now examine how the utility difference changes when the degree of political polarization Δ_y marginally increases. As we require symmetry with respect to the median voter's utility from both parties' preferred policies, we will formally take the total derivative with respect to y_j, where the term $\frac{dy_k}{dy_j}$ denotes the change in y_k that is necessary after marginally increasing y_j in order to (re-)establish the symmetry of platforms.

$$\frac{d[U_{\hat{y}_j}(t_k, S/2) - U_{\hat{y}_j}(t_j, S/2)]}{dy_j} = A\left[\frac{y_k}{y_j}\Big(\frac{1}{y_k} - \frac{y_j}{y_k^2}\frac{dy_k}{dy_j}\Big) + \hat{y}_j\Big(\frac{1}{y_k^2}\frac{dy_k}{dy_j} - \frac{1}{y_j^2}\Big)\right].$$

Rearranging terms yields

$$\frac{d[U_{\hat{y}_j}(t_k, S/2) - U_{\hat{y}_j}(t_j, S/2)]}{dy_j} = A\left[\frac{1}{y_j^2}(y_j - \hat{y}_j) + \frac{1}{y_k^2}\frac{dy_k}{dy_j}(\hat{y}_j - y_k)\right].$$

Note that for $j = L$, $\frac{d[U_{\hat{y}_j}(t_k,S/2)-U_{\hat{y}_j}(t_j,S/2)]}{d\Delta_y} = -\frac{d[U_{\hat{y}_j}(t_k,S/2)-U_{\hat{y}_j}(t_j,S/2)]}{dy_j}$, whereas for $j = R$ we would have $\frac{d[U_{\hat{y}_j}(t_k,S/2)-U_{\hat{y}_j}(t_j,S/2)]}{d\Delta_y} = \frac{d[U_{\hat{y}_j}(t_k,S/2)-U_{\hat{y}_j}(t_j,S/2)]}{dy_j}$.

Without loss of generality, consider the case $j = R$. For the utility difference to increase with political polarization, we must have

$$\frac{1}{y_R^2}(y_R - \hat{y}_R) + \frac{1}{y_L^2}\frac{dy_L}{dy_R}(\hat{y}_R - y_L) > 0. \tag{27}$$

The first summand is clearly positive. As $\frac{dy_L}{dy_R} < 0$, the second summand is only positive when $\hat{y}_R < y_L$. Hence if $\hat{y}_R < y_L$, the condition holds.

Suppose now that this is not the case, i.e. $\hat{y}_R > y_L$. According to the platform symmetry assumption, if one of the platforms y_k is given, the position of the other platform is unambiguously determined by the equation given in footnote 10. In this way we can view y_L as a function of y_R and obtain

$$\frac{dy_L}{dy_R} = -\frac{y_L^2}{y_R^2}\frac{y_R - y_m}{y_m - y_L}.$$

Inserting into (27) yields

$$\frac{1}{y_R^2}(y_R - \hat{y}_R) - \frac{1}{y_R^2}\frac{y_R - y_m}{y_m - y_L}(\hat{y}_R - y_L) > 0.$$

As $y_m - y_L > \hat{y}_R - y_L$, we can estimate the second term from below by canceling $y_m - y_L$ and $\hat{y}_R - y_L$. As a consequence, we obtain

$$\frac{1}{y_R^2}(y_m - \hat{y}_R) > 0,$$

which is clearly positive.

Intuitively, this result follows from the strong concavity of the utility function and the platform symmetry with respect to the median voter. More precisely, the argument is that as the ideal point of the critical voter $\hat{y}_R$ is politically to the left of the median voter, the marginal increase of y_R implies a higher utility loss than the corresponding decrease of y_L due to the concavity of the utility function. Hence, the utility difference between the two extremes must increase. □

D.7 Proof of Proposition 4

The following reasoning establishes the proof of the proposition. Consider the first case where $p < \hat{p}$. The condition that tax contracts yield higher welfare for the median voter is

$$\begin{aligned}&(1-p)\, U_{y_m}(t_{y_m}, S/2) + p\, U_{y_m}(t_{y_m}, S - s_E)\\ &> \tfrac{1-p}{2}\, [(U_{y_m}(t_L(S/2), S/2) + U_{y_m}(t_R(S/2), S/2))] + p\, U_{y_m}(t_{RL}(S - s_E), S - s_E).\end{aligned} \tag{28}$$

$U_y(t, s)$ represents the utility a conventional voter ($\delta = 1$) with income y derives from a policy characterized by a tax rate t and implemented by a government with s seats.[41] In condition (28) we explicitly write the tax rates in the regime *without* tax contracts as functions of the number of governmental seats. $t_k(s_k)$ and $t_{LR}(s_L + s_R)$ are defined by (7) and (10), respectively. By contrast, the equilibrium tax rate t_{y_m} in the regime *with* tax contracts as characterized in (19) does not react to different sizes of the government. Condition (28) expresses the utility of the median voter with tax contracts on the left-hand side and without tax contracts on the right-hand side. For the following arguments it is convenient to rewrite (28) in the following way:

$$\begin{aligned}(1-p)\, [U_{y_m}(t_{y_m}, S/2) - 1/2\, (U_{y_m}(t_L(S/2), S/2) + U_{y_m}(t_R(S/2), S/2))]\\ > p\, [U_{y_m}(t_{RL}(S - s_E), S - s_E) - U_{y_m}(t_{y_m}, S - s_E)]\end{aligned} \tag{29}$$

Now, the left-hand side represents the utility difference between the regimes with and without tax contracts in event ¬E and the right-hand side the utility difference in event E. We note again that t_{y_m} represents the most-preferred tax rate of the median voter in the case of a minimal majority. However, in case of a supermajority he would prefer higher taxes. Hence, if the tax rate t_{RL} of the grand coalition without tax contracts is not too far from the median voter's most-preferred tax rate in the case of a government with $S - s_E$ seats, the right-hand side of (29) is clearly positive. As already indicated, this reflects the fact that the median voter prefers to finance the supermajority of the grand coalition via tax increases rather than cuts in public-good provision. In event ¬E, which is represented on the left-hand side of the equation, there will be minimal winning coalitions of single-party governments. The left-hand side is positive since (1)

[41] Note that the tax rate and the number of seats of the government (t, s) are sufficient to characterize the policy with respect to public-good provision via the budget constraint, as each seat carries perks b with it.

the outcome under tax contracts reflects the median voter's most-preferred policy, and (2) the voters' utilities are concave, so the convex combination of two policies yields lower utility than the voters' most-preferred policy. Although the left-hand side of (29) is always positive, it is zero if there is no political polarization, i.e. if party L and party R both have the median voter as their platform voter. Moreover, in this case the right-hand side becomes positive. Consequently, we can infer that with very low levels of polarization the regime without tax contracts gives higher utility for the median voter than the regime with tax contracts. This changes when the degree of polarization becomes sufficiently large. Then the left-hand side is highly positive, compensating for a potentially positive right-hand side of condition (29). This is due to the moderating effect of tax contracts on policy.

In the case where $\hat{p} < p < p^{nc}$, tax contracts may yield higher welfare than the regime without tax contracts, even in societies with a low degree of polarization. The reason is that tax contracts not only lead to moderate tax rates, they also lower perks by avoiding a grand coalition in event E. The condition for tax contracts to be favorable for the median voter is

$$\begin{aligned}(1-p)\left[U_{y_m}(t_{y_m}, \tfrac{S^2}{2(S-s_E)}) - \frac{1}{2}\left(U_{y_m}(t_L(S/2), S/2) + U_{y_m}(t_R(S/2), S/2)\right)\right] \\ > p\left[U_{y_m}(t_{RL}(S-s_E), S-s_E) - U_{y_m}(t_{y_m}, S/2)\right].\end{aligned} \tag{30}$$

Now, the right-hand side is strictly negative, as the grand coalition is avoided in the regime with tax contracts but not in the one without tax contracts. However, there is a utility loss associated with strategic voting in event ¬E, due to the supermajority of the single-party government. Consequently, for small degrees of polarization, the left-hand side becomes negative as well. However, condition (30) may still hold, meaning that tax contracts are also favorable with low polarization.

Finally, when $p > p^{nc}$, voters avoid a grand coalition in both regimes, with and without tax contracts. We assume that voters coordinate on party j. Then the regime with tax contracts will be preferred to that without tax contracts if

$$\begin{aligned}(1-p)\left[U_{y_m}(t_{y_m}, \tfrac{S^2}{2(S-s_E)}) - U_{y_m}(t_j(\tfrac{S^2}{2(S-s_E)}), \tfrac{S^2}{2(S-s_E)})\right] \\ > p\left[U_{y_m}(t_j(S/2), S/2) - U_{y_m}(t_{y_m}, S/2)\right]\end{aligned} \tag{31}$$

Again we see that the median voter prefers supermajorities to be financed by a tax

increase rather than a decrease in the provision of the public good. Consequently, for small degrees of polarization the left-hand side can be negative, whereas the right-hand side is close to zero. However, for a sufficiently high degree of polarization, the political costs of strategic voting in the regime without tax contracts become large, and the moderating effect of tax contracts is dominant.[42] □

[42]In the case where strategic voting equilibria with coordination on L and R are possible, we could also assume that before the game starts there is an equal probability of coordinating votes on L or R, rather than assuming that coordination always favors on party j. We have omitted such an approach as it would not change our qualitative results or add to the result's intuition.

References

Enriqueta Aragones, Thomas R. Palfrey, and Andrew Postlewaite. Political reputations and campaign promises. *Journal of European Economic Association*, 2007.

David P. Baron. Electoral competition with informed and uninformed voters. *American Political Science Review*, 88(1):33–47, 1994.

Marco Battaglini and Stephen Coate. A dynamic theory of public spending, taxation and debt. *American Economic Review*, 98(1):201–236, 2008.

Roland Benabou. Ideology. *Journal of the European Economic Association*, 6:321–352, 2008.

James M. Buchanan and Gordon Tullock. *The Calculus of Consent*. University of Michigan Press, 1962.

Anthony Downs. An economic theory of political action in a democracy. *Journal of Political Economy*, 65(2):135–150, 1957.

John Duggan and Mark Fey. Electoral competition with policy-motivated candidates. *Games and Economic Behavior*, 51:490–522, 2005.

The Economist. Ordered out of office. *The Economist*, May, 21st 2009.

Lars P. Feld and Gebhard Kirchgässner. The political economy of direct legislation: Is there a role of direct democracy in EU decision-making? *Economic Policy*, 33: 329–367, 2001.

Lars P. Feld and John G. Matsusaka. Budget referendums and government spending: evidence from swiss cantons. *Journal of Public Economics*, 87(12):2703 – 2724, 2003.

Hans Gersbach. Contractual democracy. *Review of Law and Economics*, forthcoming, 2012.

Hans Gersbach and Maik T. Schneider. Tax Contracts and Government Formation. *CEPR Discussion Paper, No. 7084*, 2008.

Gene Grossman and Elhanan Helpman. Electoral competition and special interest politics. *Review of Economic Studies*, 63:265–282, 1996.

Michael Laver and Kenneth Shepsle. Cabinet government in theoretical perspective. In Michael Laver and Kenneth Shepsle, editors, *Cabinet Ministers and Parliamentary Government*, chapter 17, pages 285–310. Cambridge University Press, New York, NY, 1994.

Gilat Levy. A model of political parties. *Journal of Economic Theory*, 115:250–277, 2004.

Nolan McCarty, Keith T. Poole, and Howard Rosenthal. *Polarized America: The Dance of Ideology and Unequal Riches*. MIT Press, Cambridge, MA, 2006.

Richard D. McKelvey and Peter C. Ordeshook. Elections with limited information – a multidimensional model. *Mathematical Social Sciences*, 14(1):77–99, 1987.

Dennis C. Mueller. *Public Choice III*. Cambridge University Press, Cambridge, UK, 2003.

Roger B. Myerson. Effectiveness of electoral systems for reducing government corruption: A game-theoretic analysis. *Games and Economic Behavior*, 5(1):118 – 132, 1993.

Roger B. Myerson. Leadership, trust, and power: Dynamic moral hazard in high office. *Working Paper, University of Chicago*, 2008.

OECD. An overview of growing income inequalities in OECD countries: Main findings. *www.oecd.org/els/social/inequality*, 2011.

Ignacio Ortuno Ortin and Christian Schultz. Public funding of political parties. *Journal of Public Economic Theory*, 7(5):781–791, 2005.

Torsten Persson and Guido E. Tabellini. *The Economic Effects of Constitutions*. MIT Press, Cambridge, MA, 2003.

Michele Polo. Electoral competition and political rents. *IGIER Working Paper No. 144*, 1998.

James M. Poterba and Jürgen von Hagen. *Fiscal Institutions and Fiscal Perfomance*. Chicago University Press, Chicago, 1999.

Nouriel Roubini and Jeffrey D. Sachs. Political and economic determinants of budget deficits in the industrial democracies. *European Economic Review*, 33(5):903 – 933, 1989.

Nouriel Roubini, Jeffrey Sachs, Seppo Honkapohja, and Daniel Cohen. Government spending and budget deficits in the industrial countries. *Economic Policy*, 4(8): 100–132, 1989.

Christoph A. Schaltegger and Lars P. Feld. Do large cabinets favor large governments? evidence on the fiscal commons problem for swiss cantons. *Journal of Public Economics*, 93(1-2):35 – 47, 2009.

Social Science Research Center Berlin (WZB). Database "Elections, Parties, Governments" of the Research Unit "Democracy" at the Social Science Research Center Berlin (WZB).

Barry R. Weingast, Kenneth A. Shepsle, and Christopher Johnsen. The political economy of benefits and costs: A neoclassical approach to distributive politics. *The Journal of Political Economy*, 89(4):642–664, 1981.