Power Fluctuations and Political Economy*

Daron Acemoglu
MIT and CIFAR

Michael Golosov
Yale and NES

Aleh Tsyvinski
Yale and NES

September 2009.

Abstract

We study the constrained Pareto efficient allocations in a dynamic production economy in which the group that holds political power decides the allocation of resources. We show that Pareto efficient allocations take a quasi-Markovian structure and can be represented recursively as a function of the identity of the group in power and updated Pareto weights. For high discount factors, the economy converges to a first-best allocation in which labor supply decisions are not distorted and the levels of labor supply and consumption are constant over time (though there may be transfers from one group to another). For low discount factors, the economy converges to an invariant stochastic distribution in which distortions do not disappear and labor supply and consumption levels fluctuate over time. The labor supply of groups that are not in power are taxed in order to reduce the deviation payoff of the party in power and thus relax the political economy/sustainability constraints.

We also show that the set of sustainable first-best allocations is larger when there is less persistence in the identity of the party in power. This result contradicts a common conjecture that there will be fewer distortions when the political system creates a "stable ruling group". In contrast, political economy distortions are less important when there are frequent changes in power (because this encourages compromise between social groups). Despite this result, it remains true that distortions decrease along sample paths where a particular group remains in power for a longer span of time.

Keywords: commitment, political economy, political stability, power switches, production distortions.

JEL Classification:

*We thank participants at the Canadian Institute of Advanced Research, Econometric Society Meeting 2006, Erasmus, EUI, LAEF conference "Dynamic Political Economy and Optimal Taxation", Minneapolis Fed conference "Economics without Frontiers", MIT, New Economic School, University of Pennsylvania, University of Washington, Society of Economic Dynamics, Stanford, Stockholm, and Yale. This research is supported by the NSF. Golosov and Tsyvinski thank Einaudi Institute of Economics and Finance for hospitality.

1 Introduction

In this paper, we investigate (constrained) Pareto efficient equilibria in an infinite-horizon production economy in which political power fluctuates between different social groups ("parties"). These groups may correspond to social classes with different incomes or to citizens living in different regions. The process for power fluctuation is taken as given. Our objective is to understand the implications of political economy frictions/constraints on the allocation of resources. The key to political economy friction in our model is *lack of commitment*: the group currently in power determines the allocation of resources (the allocation of total production across different groups in the society), and there are no means of making binding commitments to future allocations. This political economy friction leads to an additional sustainability constraint for the group in power, to ensure that it does not expropriate the available resources.

We characterize the (constrained) Pareto efficient allocations in this economy. This focus enables us to understand the implications of political economy frictions on "the best possible" allocations, as clearly identifying the role of political economy in production and consumption distortions.[1] These allocations can be identified as the solution to an optimization problem subject to the participation and sustainability constraints, with different Pareto weights given to the utilities of different groups. We refer to allocations that involve full consumption smoothing and no distortions as "first best". In these allocations, each individual supplies the same amount of labor and receives the same level of consumption at every date, irrespective of which group is in power. The sustainability constraints resulting from political economy imply that first-best allocations may not be supported because the group in power could prefer to deviate from a first-best allocation. In this case, Pareto efficient allocations will involve distortions (in the sense that marginal utility of consumption and disutility of labor are not equalized) and consumption and labor will fluctuate over time. We show that Pareto efficient allocations have

[1] An alternative, complementary strategy is to focus on Pareto dominated equilibria that may emerge either in our game or in some related institutional setting. Much of the political economy literature investigates the role of specific institutions and thus implicitly focuses on such allocations. Such Pareto dominated allocations will naturally induce further distortions relative to the allocations we characterize. While these distortions are often important in practice, from a theoretical point of view they result not from the political economy friction we are focusing on (the commitment problem of the group holding power), but from the additional institutional characteristics leading to Pareto dominate outcomes.

In this light, our exercise can also be interpreted as characterizing the outcomes that would result if the society could introduce specific institutional structures that would ensure the implementation of (constrained) Pareto optimal allocations. However, our analysis in subsection 5.2 already gives some clues as to why we may not always count on different social groups agreeing to institute such arrangements.

a quasi-Markovian structure and can be characterized recursively, conditional on the identity of the group that is in power and Pareto weights. Dynamics are determined by updating the Pareto weights recursively.

We present four sets of results. First, we characterize the structure of political economy frictions as a function of the preference and production structure, the identity of the group in power and the stochastic process regulating power switches. We show that as long as a first-best allocation is not sustainable at the current date, the labor supply (and production) of individuals who belong to groups that are not in power will be distorted downwards—i.e., "taxed". This downward distortion results from the sustainability constraints reflecting the political economy considerations. Intuitively, an increase in production raises the amount that the group in power can allocate to itself for consumption rather than allocating it among the entire population. Reducing aggregate production relaxes the political economy constraints and reduces the rents captured by the group in power. Since starting from an undistorted allocation, the gain to society from rents to the ruling group is first order, while the loss is second-order, (non-first-best) constrained Pareto efficiency allocations involve distortions and underproduction.

The second set of results characterizes the dynamics of distortions. We first show that when discount factors are below some level $\bar{\beta} < 1$, no first best allocation is sustainable. Consequently, distortions always remain, even asymptotically. In particular, we show that in this case all Pareto allocations converge to an invariant *non-degenerate* distribution of consumption and leisure across groups, whereby distortions as well as the levels of consumption and labor supply for each group fluctuate according to an invariant distribution. We then focus on the special case with two social groups (two parties). In this environment we show that there exists a level of the discount factor $\hat{\beta} < 1$, such that when the common discount factor is greater than $\hat{\beta}$, then any Pareto efficient allocation path (meaning an efficient allocation starting with any Pareto weights) eventually reaches a first-best allocation, and both distortions and fluctuations in consumption and labor supply disappear. Finally, we show that regardless of whether first-best allocations are sustainable or not, distortions decrease when the group remains in power for longer. This is because the Pareto weight of a group increases the longer it remains in power, and this, ultimately, translates into fewer distortions.

Our third set of result discusses a central question in political economy—whether a more stable distribution of political power (as opposed to frequent power switches between groups) leads to "better public policies." That is, whether it leads to policies involving lower distortions and generating greater total output. A natural conjecture is that a stable distribution of political power should be preferable because it serves to increase the "effective discount factor" of the group in power, thus making "cooperation" easier. This conjecture receives support from a number of previous political economy analyses. For example, Olson (1993) and McGuire and Olson (1996) contrast an all-encompassing long-lived dictator to a "roving bandit" and conclude that the former will lead to better public policies than the latter. The standard principal-agent models of political economy, such as Barro (1973), Ferejohn (1986), Persson, Roland and Tabellini (1997, 2000), also reach the same conclusion, because it is easier to provide incentives to a politician who is more likely to remain in office.

Our analysis shows that this conjecture is generally not correct (in fact, its opposite is true). The conjecture is based on the presumption that incentives can be given to agents only when they remain in power. Once a politician or a social group leaves power, it can no longer be punished or rewarded for past actions. This naturally leads to the result that there is a direct link between the effective discount factor of a political agent and its likelihood of staying in power. This presumption is not necessarily warranted, however. Members of a social group can be rewarded not only when they are in power, but also *after they have left power*. Consequently, the main role of whether power persists or not is not to affect the effective discount factor of different parties in power, but to determine their deviation payoff. Greater persistence implies better deviation payoffs; in contrast, in the first best, there are no fluctuations in consumption and labor supply, thus along-the-equilibrium-path payoffs are independent of persistence. This reasoning leads to the opposite of the McGuire and Olson conjecture: more frequent power switches tend to reduce political economy distortions and expand the set of sustainable first-best allocations.[2]

Finally, using numerical analysis, we illustrate the relationship between persistence of power and the structure of constrained Pareto efficient allocations. We verify the result that greater persistence reduces the self sustainable first-best allocations. However, we also show that an

[2]This claim does not contradict our above result that along a sample path in which a particular party remains in power for a long time, distortions are decreasing. This is because the current claim concerns the relationship between the set of sustainable first-best allocations and the stochastic process for power switches, while the above result referred to a realization of a sample path for a given stochastic process.

increase in the frequency of power switches does not necessarily benefit all parties. Interestingly, greater persistence might harm—rather than benefit—the party in power. This is because with greater persistence, when power finally switches away from the current incumbent, the sustainability constraint of the new government will be more binding, and this will necessitate a bigger transfer away from the current incumbent in the future.

Our paper is related to the large and growing political economy literature.[3] Several recent papers also study dynamic political economy issues which is the focus of our paper. These include, among others, Acemoglu and Robinson (2000, 2006a), Acemoglu, Egorov and Sonin (2008), Battaglini and Coate (2008), Hassler et al. (2003), Krusell and Rios-Rull (1996), Lagunoff (2005, 2006), Roberts (1999) and Sonin (2003). The major difference of our paper from this literature is our focus on Pareto efficient allocations rather than Markov perfect equilibria. Almost all of the results in the paper are the result of this focus (since Markovian equilibria will involve zero production in this economy).

In this respect, our work is closely related to and builds on previous analyses of constrained efficient allocations in political economy models or in models with limited commitment. These include, among others, the limited-commitment risk sharing models of Thomas and Worrall (1990) and Kocherlakota (1996) and the political economy models of Dixit, Grossman and Gul (2000) and Amador (2003a,b). The main difference between our paper and these previous studies is our focus on the production economy. Several of our key results are derived from the explicit presence of production (labor supply) decisions.[4] In addition, to the best of our knowledge, no existing work has systematically analyzed the impact of the Markov process for power switches on the set of Pareto efficient allocations.[5]

The paper most closely related to our work is a recent and independent contribution by Aguiar and Amador (2009), who consider an international political economy model in which a party that comes to power derives greater utility from current consumption then groups not in

[3] See, among others, Persson and Tabellini, 2000, Persson, Roland and Tabellini, 1997, 2000, Besley and Coate, 1997, 1998, Baron, 1998, Grossman and Helpman, 1994, 1996, Dixit and Londergan, 1995, Bueno de Mesquita et al., 2003.

[4] Acemoglu, Golosov and Tsyvinski, 2008 a,b, 2009 a, b, and Yared, 2009, also consider dynamic political economy models with production, but their models do not feature power switches between different social groups. Battaglini and Coate, 2008, in the context of debt policy consider a model of power switches.

[5] Acemoglu and Robinson 2006b and Robinson 2001 also question the insight that long-lived all-encompassing regimes are growth-promoting. They emphasize the possibility that such regimes may block beneficial technological or institutional changes in order to maintain their political power.

power. Similar to our environment, there is also no commitment and the identity of the power fluctuates over time. Aguiar and Amador characterize a class of tractable equilibria, which lead to fluctuations in taxes on investment (expropriation), slow convergence to steady state due to commitment problems, and potential differential responses to open this depending on the degree of "political economy frictions" parameterized by the difference in the differential utility from consumption for the group in power. In contrast to our model political economy distortions disappear in the long run. In their model the backloading argument similar to Acemoglu, Golosov and Tsyvinski (2008a) applies as despite the current impatience the parties agree on the long term allocations. Battaglini and Coate (2008) is also closely related, since they investigate the implications of dynamic political economy frictions in a model with changes in the identity of the group in power, though focusing on Markovian equilibria and implications for debt and government expenditure.

The rest of the paper is organized as follows. Section 2 introduces the basic environment and characterizes the first-best allocations. Section 3 describes the political economy game and characterizes the level and dynamics of distortions. Section 4 provides a complete characterization of the dynamics of distortions in the case with two parties. Section 5 studies the effect of frequency (persistence) of power switches on political economy distortions. Section 5.2 provides a numerical illustration. Section 6 concludes, while the Appendix contains a number of technical details and proofs omitted from the text.

2 Environment and Benchmark

In this section, we introduce the model and describe efficient allocations without political economy constraints.

2.1 Demographics, Preferences and Technology

We consider an infinite horizon economy in discrete time with a unique final good. The economy consists of N parties (groups). Each party j has utility at time $t = 0$ given by

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u_j(c_{j,t}, l_{j,t}), \tag{1}$$

where $c_{j,t}$ is consumption, $l_{j,t}$ is labor supply (or other types of productive effort), and $\mathbb{E}_0$ denotes the expectations operator at time $t = 0$. To simplify the analysis without loss of any economic insights, we assume that labor supply belongs to the closed interval $\left[0, \bar{l}\right]$ for each party. We also impose the following assumption on utility functions.

Assumption 1 ***(utility function)*** *The instantaneous utility function*

$$u_j : \mathbb{R}_+ \times \left[0, \bar{l}\right] \to \mathbb{R},$$

for $j = 1, ..., N$ is uniformly continuous, twice continuously differentiable in the interior of its domain, strictly increasing in c, strictly decreasing in l and jointly strictly concave in c and l, with $u_j\left(0, 0\right) = 0$ and satisfies the following Inada conditions:

$$\lim_{c \to 0} \frac{\partial u_j(c, l)}{\partial c} = \infty \text{ and } \lim_{c \to \infty} \frac{\partial u_j(c, l)}{\partial c} = 0 \text{ for all } l \in \left[0, \bar{l}\right],$$

$$\frac{\partial u_j(c, 0)}{\partial l} = 0 \text{ and } \lim_{l \to \bar{l}} \frac{\partial u_j(c, l)}{\partial l} = -\infty \text{ for all } c \in \mathbb{R}_+.$$

The differentiability assumptions enable us to work with first-order conditions. The Inada conditions ensure that consumption and labor supply levels are not at corners. The concavity assumptions are also standard, but important for our results, since they create a desire for consumption and labor supply smoothing over time.

The economy also has access to a linear aggregate production function given by

$$Y_t = \sum_{j=1}^{N} l_{j,t}. \tag{2}$$

2.2 Efficient Allocation without Political Economy

As a benchmark, we start with the efficient allocation without political economy constraints. This is an allocation that maximizes a weighted average of different groups' utilities, with *Pareto weights* vector denoted by $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_N)$, where $\alpha_j \geq 0$ for $j = 1, ..., N$ denotes the weight given to party j. We adopt the normalization $\sum_{j=1}^{N} \alpha_j = 1$. The program for the (unconstrained) efficient allocation can be written as:

$$\max_{\left\{[c_{j,t}, l_{j,t}]_{j=1}^{N}\right\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \left[\sum_{j=1}^{N} \alpha_j u_j(c_{j,t}, l_{j,t})\right] \tag{3}$$

subject to the resource constraint

$$\sum_{j=1}^{N} c_{j,t} \leq \sum_{j=1}^{N} l_{j,t} \text{ for all } t. \tag{4}$$

Standard arguments imply that the *first-best allocation*, $\left\{\left[c^{fb}_{j,t}, l^{fb}_{j,t}\right]_{j=1}^{N}\right\}_{t=0}^{\infty}$, which is a solution to the solution to program (3), satisfies the following conditions:

$$\text{no distortions: } \frac{\partial u_j(c^{fb}_{j,t}, l^{fb}_{j,t})}{\partial c} = -\frac{\partial u_j(c^{fb}_{j,t}, l^{fb}_{j,t})}{\partial l} \text{ for } j = 1, ..., N \text{ and all } t, \tag{5}$$

$$\text{perfect smoothing: } c^{fb}_{j,t} = c^{fb}_{j} \text{ and } l^{fb}_{j,t} = l^{fb}_{j} \text{ for } j = 1, ..., N \text{ and all } t. \tag{6}$$

The structure of the first best allocations is standard. Efficiency requires the marginal benefit from additional consumption to be equal to the marginal cost of labor supply for each individual, and also requires perfect consumption and labor supply smoothing.

Note that different parties can be treated differently in the first-best allocation depending on the Pareto weight vector $\boldsymbol{\alpha}$, i.e., receive different consumption and labor allocations.

3 Political Economy

3.1 Basics

We now consider a political environment in which political power fluctuates between the N parties $j \in \mathcal{N} \equiv \{1, ..., N\}$. The game form in this political environment is as follows.

1. In each period t, we start with one party, j', in power.

2. All parties simultaneously make their labor supply decisions $l_{j,t}$. Output $Y_t = \sum_{j=1}^{N} l_{j,t}$ is produced.

3. Party j' chooses consumption allocations $c_{j,t}$ for each party subject to the feasibility constraint
$$\sum_{j=1}^{N} c_{j,t} \leq \sum_{j=1}^{N} l_{j,t}. \tag{7}$$

4. A first-order Markov process m determines who will be in power in the next period. The probability of party j being in power following party j' is $m(j \mid j')$, with $\sum_{j=1}^{N} m(j \mid j') = 1$ for all $j' \in \mathcal{N}$.

A number of features is worth noting about this setup. First, this game form captures the notion that political power fluctuates between groups. Second, it builds in the assumption that the allocation of resources is decided by the group in power (without any prior commitment to what the allocation will be). The assumption of no commitment is standard in political economy models (e.g., Persson and Tabellini, 2000, Acemoglu and Robinson, 2006a), while the presence of power switches is crucial for our focus (see also Dixit, Grossman and Gul, 2000, and Amador, 2003a,b). In addition, we have simplified the analysis by assuming that there are no constraints on the allocation decisions of the group in power and by assuming no capital.

In addition, we impose the following assumption on the Markov process for power switches.

Assumption 2 ***(Markov process)*** *The first-order Markov chain $m(j \mid j')$ is irreducible, aperiodic and ergodic.*

We are interested in *subgame perfect equilibria* of this infinitely-repeated game. More specifically, as discussed in the Introduction, we will look at subgame perfect equilibria that correspond to constrained Pareto efficient allocations, which we refer to as *Pareto efficient perfect equilibria*.[6]

To define these equilibria, we now introduce additional notation. Let $h^t = (h_0, ..., h_t)$, with $h_s \in \mathcal{N}$ be the history of power holdings. Let H^∞ denote the set of all such possible histories of power holding. Let $\mathbf{L}^t = \left(\{l_{j,0}\}_{j=1}^N, ..., \{l_{j,t}\}_{j=1}^N\right)$ be the history of labor supplies, and let $\mathbf{C}^t = \left(\{c_{j,0}\}_{j=1}^N, ..., \{c_{j,t}\}_{j=1}^N\right)$ be the history of allocation rules. A *(complete) history* of this game ("history" for short) at time t is

$$\omega^t = \left(h^t, \mathbf{C}^{t-1}, \mathbf{L}^{t-1}\right),$$

which describes the history of power holdings, all labor supply decisions, and all allocation rules chosen by groups in power. Let the set of all potential date t histories be denoted by Ω^t. In addition, denote an intermediate-stage (complete) history by

$$\hat{\omega}^t = \left(h^t, \mathbf{C}^{t-1}, \mathbf{L}^t\right),$$

and denote the set of intermediate-stage full histories by $\hat{\Omega}^t$. The difference between ω and $\hat{\omega}$ lies in the fact that the former does not contain information on labor supplies at time t,

[6] Throughout, by "Pareto efficient," we mean "constrained Pareto efficient," but we drop the adjective "constrained" to simplify the terminology.

while the latter does. The latter history will be relevant at the intermediate stage where the individual in power chooses the allocation rule.

We can now define strategies as follows. First define the following sequence of mappings $\hat{l} = \left(\hat{l}^0, \hat{l}^1, ..., \hat{l}^t, ...\right)$ and $\hat{C} = \left(\hat{C}^0, \hat{C}^1, ..., \hat{C}^t, ...\right)$, where

$$\hat{l}^t : \Omega^t \to \left[0, \bar{l}\right]$$

determines the level of labor a party will supply for every given history $\omega^t \in \Omega^t$, and

$$\hat{C}^t : \hat{\Omega}^t \to \mathbb{R}^N_+$$

a sequence of allocation rules, which a party would choose, if it were in power, for every given intermediate-stage history $\hat{\omega}^t \in \hat{\Omega}^t$, such that $\hat{C}$ satisfies the feasibility constraint (7). A date t strategy for party j is $\sigma_j^t = \left(\hat{l}^t, \hat{C}^t\right)$. Denote the set of date t strategies by Σ^t. A strategy for party j is $\sigma_j = \left(\left\{\sigma_j^t\right\} : t = 0, 1, ...\right)$ and the set of strategies is denoted by Σ. Denote the expected utility of party j at time t as a function of its own and others' strategies given history ω^t and intermediate-stage history $\hat{\omega}^t$ by

$$U_j\left(\sigma_j, \sigma_{-j} \mid \omega^t, \hat{\omega}^t\right).$$

We next define various concepts of equilibria which we use throughout the paper.

Definition 1 *A* subgame perfect equilibrium (SPE) *is a collection of strategies* $\sigma^* = \left(\left[\sigma_j^t\right] : j = 1, ..., N,\ t = 0, 1, ...\right)$ *such that* σ_j^* *is best response to* σ_{-j}^* *for all* $\left(\omega^t, \hat{\omega}^t\right) \in \Omega^t \times \hat{\Omega}^t$ *and for all* j*, i.e.,* $U_j\left(\sigma_j^*, \sigma_{-j}^* \mid \omega^t, \hat{\omega}^t\right) \geq U_j\left(\sigma_j, \sigma_{-j}^* \mid \omega^t, \hat{\omega}^t\right)$ *for all* $\sigma_j \in \Sigma$*, for all* $\left(\omega^t, \hat{\omega}^t\right) \in \Omega^t \times \hat{\Omega}^t$*, for all* $t = 0, 1, ...$ *and for all* $j \in \mathcal{N}$*.*

Definition 2 *A* Pareto efficient perfect equilibrium *at time* t *(following history* ω^t*),* σ^{**}*, is a collection of strategies that form an SPE such that there does not exist another SPE* σ^{***}*, whereby* $U_j\left(\sigma_j^{***}, \sigma_{-j}^{***} \mid \omega^t, \hat{\omega}^t\right) \geq U_j\left(\sigma_j^{**}, \sigma_{-j}^{**} \mid \omega^t, \hat{\omega}^t\right)$ *for all* $\hat{\omega}^t \in \hat{\Omega}^t$ *and for all* $j \in \mathcal{N}$*, with at least one strict inequality.*

We will also refer to *Pareto efficient allocations* as the equilibrium-path allocations that result from a Pareto efficient perfect equilibrium. To characterize Pareto efficient allocations, we will first determine the worst subgame perfect equilibrium, which will be used as a threat against deviations from equilibrium strategies. These are defined next. We write $j = \mathbf{j}\left(h^t\right)$ or $j = \mathbf{j}\left(\omega^t\right)$ if party j is in power at time t according to history (of power holdings) h^t or according to (complete) history $\omega^t \in \Omega^t$. We also use the notation $h^t \in H_j^t$ whenever $j = \mathbf{j}\left(h^t\right)$. A *worst SPE for party* j at time t following history ω^t where σ^W is a collection of strategies that form a SPE such that there does not exist another SPE σ^{***} such that $U_j\left(\sigma_j^{***}, \sigma_{-j}^{***} \mid \omega^t, \hat{\omega}^t\right) < U_j\left(\sigma_j^W, \sigma_{-j}^W \mid \omega^t, \hat{\omega}^t\right)$ for all $\hat{\omega}^t \in \hat{\Omega}^t$.

3.2 Preliminary Results

The next lemma describes the worst subgame perfect equilibrium. In that equilibrium, all parties that are not in power in any given period supply zero labor and receive zero consumption, while the party in power supplies labor and consumes all output to maximize its per period utility in such a way that marginal utility from consumption is equated with marginal disutility of labor.

Lemma 1 *Suppose Assumption 1 holds. The worst SPE for any party j'' is given by the collection of strategies σ^W such that for all $j \neq \mathbf{j}\left(h^t\right)$: $l_j^t\left(\omega^t\right) = 0$ for all $\omega^t \in \Omega^t$, and for $j' = \mathbf{j}\left(h^t\right)$ $l_{j'}^t\left(\omega^t\right) = \tilde{l}_{j'}$ for all $\omega^t \in \Omega^t$ where $\tilde{l}_{j'}$ is a solution to*

$$\frac{\partial u_{j'}(\tilde{l}_{j'}, \tilde{l}_{j'})}{\partial c} = -\frac{\partial u_{j'}(\tilde{l}_{j'}, \tilde{l}_{j'})}{\partial l} \tag{8}$$

and $c_j^t\left(\hat{\omega}^t\right) = 0$ for $j \neq j'$, $c_{j'}^t\left(\hat{\omega}^t\right) = \sum_{i=1}^N l_i^t\left(\hat{\omega}^t\right)$ for all $\hat{\omega}^t \in \hat{\Omega}^t$.

Proof. We first show that σ^W is a best response for each party in all subgames when other parties are playing σ^W. Consider first party j that is not in power (i.e., suppose that party $j' \neq j$ is in power) at history ω^t. Consider strategy $\sigma_{j,t}$ for party j that deviates from σ_j^W at time t, and then coincides with σ_j^W at all subsequent dates (following all histories). By the one step ahead deviation principle, if σ_j^W is not a best response for party j, then there exists such a strategy $\sigma_{j,t}$ that will give higher utility to this party. Note, first that given σ_{-j}^W, for any $\sigma_{j,t}$, party j will always receive zero consumption (i.e., $c_j\left(\sigma_j, \sigma_{-j}^W \mid \omega^t, \hat{\omega}^t\right) = 0$), and moreover under σ^W, this has no effect on the continuation value of party j. Therefore, at such an history, we have

$$\begin{aligned} U_j\left(\sigma_{j,t}, \sigma_{-j}^W \mid \omega^t, \hat{\omega}^t\right) &= u_j(0, l_{j,t}\left(\sigma_{j,t}\right)) + \beta \mathbb{E}\left[U_j\left(\sigma_{j,t}, \sigma_{-j}^W \mid \omega^{t+1}, \hat{\omega}^{t+1}\right) \mid \omega^t\right] \\ &\leq \beta \mathbb{E}\left[U_j\left(\sigma_j^W, \sigma_{-j}^W \mid \omega^{t+1}, \hat{\omega}^{t+1}\right) \mid \omega^t\right], \end{aligned}$$

for any such $\sigma_{j,t}$, where $l_{j,t}(\sigma_{j,t})$ is the labor supply of party j at this history under the alternative strategy $\sigma_{j,t}$, and $\mathbb{E}\left[U_j\left(\sigma_{j,s}, \sigma_{-j}^W \mid \omega^{t+1}, \hat{\omega}^{t+1}\right) \mid \omega^t\right]$ is the continuation value of this party from date $t+1$ onwards, with the expectation taken over histories determining power switches given current history ω^t. The second line follows in view of the fact that since $u_j(0,0) = 0$, we have $u_j(0, l_{j,t}(\sigma_{j,t})) \leq 0$, and since under σ^Wany change in behavior at t has no effect on future play and $\sigma_{j,t}$ coincides with σ_j^W from time $t+1$ onwards, $\mathbb{E}\left[U_j\left(\sigma_{j,t}, \sigma_{-j}^W \mid \omega^{t+1}, \hat{\omega}^{t+1}\right) \mid \omega^t\right] = \mathbb{E}\left[U_j\left(\sigma_j^W, \sigma_{-j}^W \mid \omega^{t+1}, \hat{\omega}^{t+1}\right) \mid \omega^t\right]$. This establishes that there is no profitable deviations from σ_j^W for any j not in power.

Next consider party j in power at history ω^t. Under σ^W, $l_{j',t} = c_{j',t} = 0$ for all $j' \neq j$, and thus $l_{j,t} = c_{j,t}$. Consider again strategy $\sigma_{j,t}$ for party j that deviates from σ_j^W at time t, and then coincides with σ_j^W at all subsequent dates. Then, using similar notation, we have

$$\begin{aligned} U_j\left(\sigma_{j,t}, \sigma_{-j}^W \mid \omega^t, \hat{\omega}^t\right) &= u_j(c_{j,t}(\sigma_{j,t}), l_{j,t}(\sigma_{j,t})) + \beta\mathbb{E}\left[U_j\left(\sigma_{j,t}, \sigma_{-j}^W \mid \omega^{t+1}, \hat{\omega}^{t+1}\right) \mid \omega^t\right] \\ &\leq u_j(c_{j,t}(\sigma_j^W), l_{j,t}(\sigma_j^W)) + \beta\mathbb{E}\left[U_j\left(\sigma_j^W, \sigma_{-j}^W \mid \omega^{t+1}, \hat{\omega}^{t+1}\right) \mid \omega^t\right], \end{aligned}$$

for any such $\sigma_{j,t}$, where $l_{j,t}(\sigma_{j,t})$ and $c_{j,t}(\sigma_{j,t})$ are the labor supply and consumption of party j at this history under strategy $\sigma_{j,t}$. The second line follows in view of the fact that σ_j^W satisfies (8), and thus $u_j(c_{j,t}\left(\sigma_j^W\right), l_{j,t}\left(\sigma_j^W\right)) = u_j(\tilde{l}, \tilde{l}) \geq u_j(c_{j,t}(\sigma_{j,t}), l_{j,t}(\sigma_{j,t}))$ for any $\sigma_{j,t}$, and again because $\mathbb{E}\left[U_j\left(\sigma_{j,t}, \sigma_{-j}^W \mid \omega^{t+1}, \hat{\omega}^{t+1}\right) \mid \omega^t\right] = \mathbb{E}\left[U_j\left(\sigma_j^W, \sigma_{-j}^W \mid \omega^{t+1}, \hat{\omega}^{t+1}\right) \mid \omega^t\right]$ (from the fact that under σ^W the current deviation by party j has no effect on future play and $\sigma_{j,t}$ coincides with σ_j^W from time $t+1$ onwards). This establishes that there is no profitable deviations from σ_j^W for the party in power. Therefore, σ^W is a SPE. The proof is completed by showing that σ^W is also the worst SPE for any party j. To see this, suppose that all $j' \neq j$ choose strategy σ_{-j}^M to minimize the payoff of j. Since power switches are exogenous, party j can guarantee itself $u_j(\tilde{l}, \tilde{l})$ whenever it is in power and $u_j(c_{j,t}\left(\sigma_j^W, \sigma_{-j}^M\right), l_{j,t}\left(\sigma_j^W, \sigma_{-j}^M\right)) = 0$ whenever it is not in power. Therefore,

$$U_j\left(\sigma_j^W, \sigma_{-j}^M \mid \omega^t, \hat{\omega}^t\right) \geq U_j\left(\sigma_j^W, \sigma_{-j}^W \mid \omega^t, \hat{\omega}^t\right)$$

for any σ_{-j}^M, and thus σ^W is the worst SPE. Intuitively, the worst equilibrium involves all groups other than the one in power supplying zero labor, which minimizes the utility of the group in power. The labor supply decisions constitute a best response, since following a deviation all output is expropriated by whichever group is currently in power. Also note that the same equilibrium is the worst equilibrium for all parties. ■

We denote $V_j^W(h^t)$ to be the expected payoff of party j from period $t+1$ on, conditional on history h^t. Note that generally such utility differs based on the history: $V_j^W(h^t) \neq V_j^W(\hat{h}^t)$ for $h_t \neq \hat{h}_t$. The reason is that the identity of party in power in period t determines the probability of party j being in power in $t+1$. When m satisfies Assumption 2, it can be further simplified, since $V_j^W(h^t)$ depends only on the identify of party in power h_t.

Proposition 1 *Suppose Assumptions 1 and 2 hold. Then, an outcome of any Pareto efficient perfect equilibrium is a solution to the following maximization problem for all $h^t \in \omega^t, \hat{\omega}^t$:*

$$\max_{\{c_j(h^t),l_j(h^t)\}_{j=1,..,N;\ h^t}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\sum_{j=1}^{N} \alpha_j u_j(c_j\left(h^t\right), l_j\left(h^t\right)) \right] \tag{9}$$

subject to, for all h^t,

$$\sum_{j=1}^{N} c_j\left(h^t\right) \leq \sum_{j=1}^{N} l_j\left(h^t\right), \tag{10}$$

$$\mathbb{E}_t \sum_{s=0}^{\infty} \beta^s u_j\left(c_j\left(h^{t+s}\right), l_j\left(h^{t+s}\right)\right) \geq \beta V_j^W(\mathbf{j}(h^t)) \text{ for all } j \neq \mathbf{j}(h^t), \tag{11}$$

and

$$\mathbb{E}_t \sum_{s=0}^{\infty} \beta^s u_{j'}\left(c_{j'}\left(h^{t+s}\right), l_{j'}\left(h^{t+s}\right)\right) \geq \max_{\tilde{l}\geq 0} u_{j'}\left(\sum_{j\neq j'} l_j\left(h^t\right) + \tilde{l}\ , \tilde{l}\right) + \beta V_{j'}^W(j') \text{ for } j' = \mathbf{j}(h^t), \tag{12}$$

for some Pareto weights vector $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_N)$.

Proof. We start by showing that (10)-(12) are necesary and sufficient conditions for any allocation $\{c_j(h), l_j(h)\}_{j=1}^N$ that is an outcome of some SPE. First, we show that any allocation $\{c_j(h), l_j(h)\}_{j=1}^N$ that satisfies (10)-(12) is an outcome of some SPE. For any history ω^t with $h^t \in \omega^t$ let $\sigma^*(\omega^t) = \left\{l_j\left(h^t\right)\right\}_{j=1}^N$ if $\omega^t = \left(h^t, \left\{c_j\left(h^{t-1}\right), l_j\left(h^{t-1}\right)\right\}_{j=1}^N\right)$, and $\sigma^*\left(\omega^t\right) = \sigma^W$ otherwise, and, analogously, $\sigma^*(\hat{\omega}^t) = \left\{c_j\left(h^t\right)\right\}_{j=1}^N$ if $\hat{\omega}^t = \left(h^t, \left\{c_j\left(h^{t-1}\right), l_j\left(h^t\right)\right\}_{j=1}^N\right)$, and $\sigma^*\left(\tilde{\omega}^t\right) = \sigma^W$ otherwise. For any $j \neq \mathbf{j}(h^t)$ if $\sigma_j(\omega^t) \neq \sigma_j^*(\omega^t)$,

$$\begin{aligned} U_j(\sigma_j, \sigma_{-j}^*|\omega^t) &\leq u_j(0,0) + \beta V_j^W(\mathbf{j}(h^t)) \\ &\leq U_j(\sigma_j^*, \sigma_{-j}^*|\omega^t), \end{aligned}$$

where the last inequality follows from (11) and that fact that $u_j(0,0) = 0$. Moreover, for $j' = \mathbf{j}(h^t)$, any $\sigma_{j'} \neq \sigma^*_{j'}$ implies

$$\begin{aligned} U'_{j'}(\sigma_{j'}, \sigma^*_{-j'}|\omega^t, \tilde{\omega}^t) &\leq \max_{\tilde{l}\geq 0} u_{j'}\left(\sum_{j\neq j'} l_j\left(h^t\right) + \tilde{l}\ , \tilde{l}\right) + \beta V^W_{j'}(j') \\ &\leq U_{j'}(\sigma^*_{j'}, \sigma^*_{-j'}|\omega^t, \tilde{\omega}^t). \end{aligned}$$

Therefore, σ^* is an equilibrium.

The necessity of (10)-(12) is straightforward. Condition (10) is feasibility constraint. In any equilibrium σ^* for $j \neq \mathbf{j}(h^t)$ for all $\sigma_j \neq \sigma^*_j$

$$U_j(\sigma^*_j, \sigma^*_{-j}|\omega^t) \geq U_j(\sigma_j, \sigma^*_{-j}|\omega^t)$$

This implies in particular that

$$\begin{aligned} U_j(\sigma^*_j, \sigma^*_{-j}|\omega^t) &\geq u_j\left(c_j\left(\sigma_{j,t}\right), l_j\left(\sigma_{j,t}\right)\right) + \beta \mathbb{E}_t U_j(\sigma_j, \sigma^*_{-j}|\omega^{t+1}) \\ &\geq u_j\left(0, l_j\left(\sigma_{j,t}\right)\right) + \beta V^W_j(\mathbf{j}(h^t)) \\ &= u_j\left(0, 0\right) + \beta V^W_j(\mathbf{j}(h^t)). \end{aligned}$$

Since $u_j\left(0,0\right) = 0$, it implies (11). Similarly, for $j' = \mathbf{j}(h^t)$ best response implies

$$\begin{aligned} U_{j'}(\sigma^*_{j'}, \sigma^*_{-j'}|\omega^t, \tilde{\omega}^t) &\geq \max_{\tilde{l}\geq 0} u_{j'}\left(\sum_{j\neq j'} l_j\left(h^t\right) + \tilde{l}\ , \tilde{l}\right) + \beta \mathbb{E}_t U_j(\sigma_{j'}, \sigma^*_{-j'}|\omega^{t+1}) \\ &\geq \max_{\tilde{l}} u_{j'}\left(\sum_{j\neq j'} l_j\left(h^t\right) + \tilde{l}\ , \tilde{l}\right) + \beta V^W_{j'}(j') \end{aligned}$$

which is condition (12).

To see that σ^* is a Pareto efficient equilibrium, suppose there is any other equilibrium σ^{**} that Pareto dominates σ^*. Since σ^{**} is a SPE,

$$U_j(\sigma^{**}_j, \sigma^{**}_{-j}|\omega^t) \geq \beta V^W_j(h^t)$$

for all $j \neq \mathbf{j}(h^t)$ and

$$U_j(\sigma^{**}_j, \sigma^{**}_{-j}|\omega^t) \geq \max_{\tilde{l}\geq 0} u_{j'}\left(\sum_{j\neq j'} l_j\left(\omega^t\right) + \tilde{l}\ , \tilde{l}\right) + \beta V^W_{j'}(j').$$

Therefore, the outcome of σ^{**} must satisfy (10)-(12). But then the value of (9) would be higher under the outcome of σ^{**} than under σ^{*}, a contradiction. ■

To simplify the notation, we define

$$V_j(h^{t-1}) \equiv \mathbb{E}\left\{\sum_{s=0}^{\infty} \beta^s u_{j'}\left(c_{j'}\left(h^{t+s}\right), l_{j'}\left(h^{t+s}\right)\right)|h^{t-1}\right\}$$

$$V_j[h^{t-1}, i] \equiv \mathbb{E}\left\{\sum_{s=0}^{\infty} \beta^s u_{j'}\left(c_{j'}\left(h^{t+s}\right), l_{j'}\left(h^{t+s}\right)\right)|h^{t-1}, \mathbf{j}(h^t) = i\right\}$$

The difference between $V_j(h^{t-1})$ and $V_j[h^{t-1}, i]$ is that the former denotes expected lifetime utility of party j in period t before the uncertainty which party is in power in that period is realized, while the latter denotes the expected lifetime utility after realization of this uncertainty. From the above definition and Assumption 2,

$$V_j(h^{t-1}) = \sum_{j'=1}^{N} m(j'|h_{t-1}) V_j[h^{t-1}, j'].$$

Proposition 1 implies that in order to characterize the entire set of Pareto efficient perfect equilibria, we can restrict attention to strategies that follow a particular prescribed equilibrium play, with the punishment phase given by σ^W. Notice, however, that this proposition applies to Pareto efficient outcomes, *not* to the strategies that individuals use in order to support these outcomes. These strategies *must be* conditioned on information that is not contained in the history of power holdings, h^t, since individuals need to switch to the worst subgame perfect equilibrium in case there is any deviation from the implicitly-agreed action profile. This information is naturally contained in ω^t. Therefore, to describe the subgame perfect equilibrium strategies we need to condition on the full histories ω^t.

The maximization (9) subject to (10), (11), and (12) is a potentially non-convex optimization problem, because (12) defines a non-convex constraint set. This implies that randomizations may improve the value of the program (see, for example, Prescott and Townsend, 1984 a,b). Randomizations can be allowed by either considering correlated equilibria rather than subgame perfect equilibria, or alternatively, by assuming that there is a commonly observed randomization device on which all individuals can coordinate their actions. In the Appendix, we will formulate an extended problem by introducing a commonly-observed, independently and identically distributed random variable, which all individual strategies can be conditioned upon. We will show that this does not change the basic structure of the problem and in fact there will be randomizations over at most two points at any date, and the history of past randomizations will not play any role in the characterization of Pareto efficient allocations. Since introducing randomizations complicates the notation considerably, in the text we do not consider randomizations (thus implicitly assuming that the problem is convex for the relevant parameters). The equivalents of the main results are stated in the Appendix for the case with randomizations.

We next present our main characterization result, which shows that the solution to the maximization problem in Proposition 1 can be represented recursively.

3.3 Recursive Characterization

Let us first define $M\left(h^{t+s} \mid h^t\right)$ to be the (conditional) probability of history h^{t+s} at time $t+s$ given history h^t at time t according to the Markov process $m\left(j \mid j'\right)$. Moreover, define $P\left(h^t\right)$ be the set of all possible date $t+s$ histories for $s \geq 1$ that can follow history h^t. We write $M\left(h^{t+s} \mid h^{\varnothing}\right)$ for the unconditional probability of history h^{t+s}.

First note that the maximization problem in Proposition 1 can be written in Lagrangian form as follows:

$$\max_{\{c_j(h^t), l_j(h^t)\}_{j=1,..,N}} \quad \mathcal{L}' = \sum_{t=0}^{\infty} \sum_{h^t} \beta^t M\left(h^t \mid h^{\varnothing}\right) \left[\sum_{j=1}^{N} \alpha_j u_j\left(c_j\left(h^t\right), l_j\left(h^t\right)\right)\right] \tag{13}$$
$$+\sum_{t=0}^{\infty} \sum_{h^t} \beta^t M\left(h^t \mid h^{\varnothing}\right) \lambda_j\left(h^t\right) \times$$
$$\left[\sum_{s=t}^{\infty} \beta^{s-t} \sum_{h^s} M\left(h^s \mid h^t\right) u_j\left(c_j\left(h^s\right), l_j\left(h^s\right)\right) - v_j\left(\sum_{i \neq j} l_i\left(h^t\right)\right) - \beta V_j^W(j)\right]$$

subject to (10) and (11), where $\beta^t M\left(h^t \mid h^0\right) \lambda_j\left(h^t\right)$ is the Lagrange multiplier on the sustainability constraint, (12), for party j for history h^t, and $v_j(Y) \equiv \max_{\tilde{l}} u_j\left(Y + \tilde{l}\,, \tilde{l}\right)$. The restriction that $h^t \in P\left(h^{t-1}\right)$ is implicit in this expression.

The proof of Theorem 1 establishes that after history h^{t-1}, this Lagrangian is equivalent to

$$\max_{\{c_j(h^t),l_j(h^t)\}_{j=1,..,N;\ h^t}} \mathcal{L} = \sum_{s=t}^{\infty}\sum_{h^s} \beta^s M\left(h^s \mid h^t\right)\left[\sum_{j=1}^{N}\left(\alpha_j + \mu_j\left(h^{t-1}\right)\right) u_j\left(c_j\left(h^s\right), l_j\left(h^s\right)\right)\right] \tag{14}$$

$$+\sum_{s=t}^{\infty}\sum_{h^s} \beta^s M\left(h^s \mid h^t\right) \times$$

$$\sum_{j=1}^{N} \lambda_j\left(h^s\right)\left(\sum_{s'=s}^{\infty}\sum_{h^{s'}} \beta^{s'-s} M\left(h^{s'} \mid h^s\right) u_j\left(c_j\left(h^{s'}\right), l_j\left(h^{s'}\right)\right) - v_j\left(\sum_{i\neq j} l_i\left(h^s\right)\right) - \beta V_j^W(j)\right),$$

subject to (10) and (11), with μ_j's defined recursively as:

$$\mu_j\left(h^t\right) = \mu_j\left(h^{t-1}\right) + \lambda_j\left(h^t\right)$$

with the normalization $\mu_j\left(h^{\varnothing}\right) = 0$ for all $j \in \mathcal{N}$.

The most important implication of the formulation in (14) is that for any $h^t \in P\left(h^{t-1}\right)$, the numbers

$$\alpha_j\left(h^{t-1}\right) \equiv \frac{\alpha_j + \mu_j\left(h^{t-1}\right)}{\sum_{j'=1}^{N}\left(\alpha_{j'} + \mu_{j'}\left(h^{t-1}\right)\right)} \tag{15}$$

can be interpreted as *updated Pareto weights*. Therefore, after history h^{t-1}, the problem is equivalent to maximizing the sum of utilities with these weights (subject to the relevant constraints). The problem of maximizing (14) is equivalent to choosing current consumption and labor supply levels for each group and also updated Pareto weights $\{\alpha_j\}_{j=1}^{N}$.

In addition, (14) has the attractive feature that $\mu_j\left(h^t\right) - \mu_j\left(h^{t-1}\right) = 0$ whenever $j \neq \mathbf{j}\left(h^t\right)$, i.e., whenever group j is not in power. This is because there is no sustainability constraint for a group that is not in power. This also implies that in what follows, we can drop the subscript j and refer to $\lambda\left(h^t\right)$ rather than $\lambda_j\left(h^t\right)$, since the information on which group is in power is already incorporated in h^t.

This analysis establishes the following characterization result:

Theorem 1 *Suppose Assumptions 1 and 2 hold. Then the constrained efficient allocation has a quasi-Markovian structure whereby consumption and labor allocations* $\left\{c_j\left(h^t\right), l_j\left(h^t\right)\right\}_{j=1,..,N;\ h^t}$ *depend only on* $\mathbf{s} \equiv \left(\{\alpha_j(h)\}_{j=1}^{N}, \mathbf{j}(h)\right)$, *i.e., only on updated weights and the identity of the group in power.*

Proof. The proof of this theorem builds on the representation suggested by Marcet and Marimon (1998). In particular, observe that for any $T \geq 0$, we have

$$\begin{aligned} & \sum_{s=0}^{T} \sum_{h^s} \beta^s M\left(h^s \mid h^{\varnothing}\right) \lambda_j\left(h^s\right) \sum_{s'=s}^{T} \sum_{h^{s'}} \beta^{s'-s} M\left(h^{s'} \mid h^s\right) u_j\left(c_j\left(h^{s'}\right), l_j\left(h^{s'}\right)\right) \quad (16) \\ = & \sum_{s=0}^{T} \beta^s M\left(h^s \mid h^{\varnothing}\right) \mu_j\left(h^s\right) u_j\left(c_j\left(h^s\right), l_j\left(h^s\right)\right) \end{aligned}$$

where $\mu_j(h^s) = \mu_j\left(h^{s-1}\right) + \lambda_j(h^s)$ for $h^s \in P\left(h^{s-1}\right)$ with the initial $\mu_j\left(h^{\varnothing}\right) = 0$ for all j. Substituting (16) in $\mathcal{L}'$ in (13), we obtain that $\mathcal{L}'$ for any h^{t-1}, can be expressed as

$$\begin{aligned} \max_{\{c_j(h), l_j(h)\}_{j=1}^N} \mathcal{L}'' = & \sum_{s=t}^{\infty} \sum_{h^s} \beta^s M\left(h^s \mid h^{t-1}\right) \left[\sum_{j=1}^{N} \left(\alpha_j + \mu_j\left(h^{t-1}\right)\right) u_j\left(c_j\left(h^s\right), l_j\left(h^s\right)\right)\right] \\ & + \sum_{s=t}^{\infty} \sum_{h^s} \beta^s M\left(h^s \mid h^{t-1}\right) \sum_{j=1}^{N} \lambda_j\left(h^s\right) \left(\begin{array}{c} \sum_{s'=s}^{\infty} \sum_{h^{s'}} \beta^{s'-s} M\left(h^{s'} \mid h^s\right) u_j\left(c_j\left(h^s\right), l_j\left(h^s\right)\right) \\ -v_j\left(\sum_{i \neq j} l_i\left(h^s\right)\right) - \beta V_j^W(j) \end{array}\right) \\ & + \sum_{s=0}^{t-1} \sum_{h^s} \beta^s M\left(h^s \mid h^{\varnothing}\right) \sum_{j=1}^{N} \alpha_j u_j\left(c_j\left(h^s\right), l_j\left(h^s\right)\right) \\ & + \sum_{s=0}^{t-1} \sum_{h^s} \beta^s M\left(h^s \mid h^{\varnothing}\right) \sum_{j=1}^{N} \lambda_j\left(h^s\right) \left(\begin{array}{c} \sum_{s'=s}^{t-1} \sum_{h^{s'}} \beta^{s'} M\left(h^{s'} \mid h^s\right) u_j\left(c_j\left(h^{s'}\right), l_j\left(h^{s'}\right)\right) \\ -v_j\left(\sum_{i \neq j} l_i\left(h^s\right)\right) - \beta V_j^W(j) \end{array}\right). \end{aligned}$$

Since after history h^{t-1} has elapsed, all terms in the last two lines are given, maximizing $\mathcal{L}''$ is equivalent to maximizing (14).

Given the structure of problem (14), the result that optimal allocations only depend on $\{\alpha_j(h)\}_{j=1}^N$ and $\mathbf{j}(h)$ then follows immediately. In the Appendix, we prove a generalized version of this theorem, Theorem 6, that allows for randomizations. ■

The result in this theorem is intuitive. When the sustainability constraint for the party in power is binding, the discounted value of this party needs to be increased so as to satisfy this constraint. This is typically done by a combination of increasing current and future utility. The latter takes the form of increasing the Pareto weight of the party in power, corresponding to a move along the constraint Pareto efficient frontier.

It is also worth noting that the existence of a recursive formulation for the problem of characterizing the set of Pareto efficient allocations also has an obvious parallel to the general recursive formulation provided by Abreu, Pearce and Stacchetti (1990) for repeated games with imperfect monitoring. Nevertheless, Theorem 1 is not a direct corollary of their results, since it establishes that this recursive formulation depends on updated Pareto weights and the identity of the group in power, and shows how these weights can be calculated from past realizations of the history h^t.

Theorem 1 allows us to work with a recursive problem, in which we only have to keep track of the identity of the party that is in power and updated Pareto weights. Moreover, the analysis preceding the theorem shows that the Pareto weights are updated following the simple formula (15), which is only a function of the Lagrange multiplier on the sustainability constraint of the party in power at time t. The recursive characterization implies that we can express $V_j(h^{t-1})$ and $V_j[h^{t-1}, i]$ as $V_j(\boldsymbol{\alpha})$ and $V_j[\boldsymbol{\alpha}, i]$.

3.4 Characterization of Distortions

We now characterize the structure of distortions arising from political economy. Our first result shows that as long as sustainability/political economy constraints are binding, the labor supply of parties that are not in power is distorted *downwards*. There is a positive wedge between their marginal utility of consumption and marginal disutility of labor. In contrast, there is no wedge for the party in power. Recall also that without political economy constraints, in the first best allocations, the distortions are equal to zero.

Proposition 2 *Suppose that Assumptions 1 and 2 hold. Then as long as $\lambda\left(h^t\right) > 0$, the labor supply of all groups that are not in power, i.e., $j \neq \mathbf{j}\left(h^t\right)$, is distorted* downwards*, in the sense that*

$$\frac{\partial u_j\left(c_j\left(h^t\right), l_j\left(h^t\right)\right)}{\partial c} > -\frac{\partial u_j\left(c_j\left(h^t\right), l_j\left(h^t\right)\right)}{\partial l}.$$

The labor supply of a party in power, $j' = \mathbf{j}\left(h^t\right)$, is undistorted, i.e.,

$$\frac{\partial u_{j'}\left(c_{j'}\left(h^t\right), l_{j'}\left(h^t\right)\right)}{\partial c} = -\frac{\partial u_j\left(c_{j'}\left(h^t\right), l_{j'}\left(h^t\right)\right)}{\partial l}.$$

Proof. This results follows from the first-order conditions of the maximization problem in Section 7.1 in the Appendix. In particular, combining (22) and (23), we have

$$\frac{\partial u_j\left(c_j\left(h^t\right), l_j\left(h^t\right)\right)}{\partial c} > -\frac{\partial u_j\left(c_j\left(h^t\right), l_j\left(h^t\right)\right)}{\partial l} \text{ whenever } \lambda_j\left(h^t\right) > 0 \text{ for all } j \neq \mathbf{j}\left(h^t\right).$$

A comparison with the equalities in (5) shows that these inequalities correspond to downward distortions. The result about the group in power is obtained analogously. ■

The intuition for why there will be downward distortions in the labor supply of groups that are not in power is similar to that in Acemoglu, Golosov and Tsyvinski (2008a). Positive distortions, which are the equivalent of "taxes," discourage labor supply, reducing the amount of output that the group in power can "expropriate" (i.e., allocate to itself as consumption following a deviation). This relaxes the sustainability constraint (12). In fact, starting from an allocation with no distortions, a small distortion in labor supply creates a second-order loss. In contrast, as long as the multiplier on the sustainability constraint is positive, this small distortion creates a first-order gain in the objective function, because it enables a reduction in the rents captured by the group in power. This intuition also highlights that the extent of distortions will be closely linked to the Pareto weights given to the group in power. In particular, when α_j is close to 1 and group j is in power, there will be little gain in relaxing the sustainability constraint (12). In contrast, the Pareto efficient allocation will attempt to provide fewer rents to group j when α_j is low, and this is only possible by reducing the labor supply of all other groups, thus distorting their labor supplies.

Two immediate but useful corollaries of Proposition 2 are as follows:

Corollary 1 *The (normalized) Lagrange multiplier on the sustainability constraint (12) given history h^t, $\lambda\left(h^t\right)$, is a measure of distortions.*

This corollary follows immediately from Theorem 2, and more explicitly from Section 7.1 in the Appendix, which shows that the wedges between the marginal utility of consumption and the marginal disutility of labor are directly related to $\lambda\left(h^t\right)$. This corollary is useful as it will enable us to link the level and behavior of distortions to the behavior of the Lagrange multiplier $\lambda\left(h^t\right)$.

A related implication of Proposition 2 is that constrained Pareto efficient allocations will be "first-best" if and only if the Lagrange multipliers associated with all sustainability constraints are equal to zero (so that there are no distortions in a first-best allocation).

Corollary 2 *A first-best allocation starting at history h^t involves $\lambda\left(h^{t+s}\right)=0$ for all $h^{t+s}\in P\left(h^t\right)$.*

3.5 Dynamics of Distortions

Proposition 2 states that when the Lagrange multipliers are positive, the allocations are distorted. We now study the evolution of the Largrange multipliers and distortions resulting from the sustainability constraints.

Our first result in this subsection is an immediate implication of the recursive formulation in Theorem 1, but it will play an important role in our results. The lemma that follows shows that if a group is in power today, then in the next period its updated Pareto weight must be weakly higher than today.

Lemma 2 *If $\mathbf{j}\left(h^{t+1}\right)=j$, then $\alpha_j\left(h^{t+1}\right)\geq\alpha_j\left(h^t\right)$.*

Proof. This follows immediately from equation (15) observing that if $\mathbf{j}\left(h^{t+1}\right)=j'$, then $\mu_j\left(h^{t+1}\right)-\mu_j\left(h^t\right)=0$ for all $j\neq j'$. ■

This lemma implies that as long as party j remains in power, its Pareto weight is increasing. We assume the following condition holds.

Condition 1 *For all j, if $\alpha_j=1$, then*

$$V_j[\boldsymbol{\alpha},j]>v_j\left(\sum_{i\neq j}\tilde{l}_i\right)+\beta V_j^W(j)$$

where $\tilde{l}_i$ is a labor supply of party i in equilibrium in the state $(\boldsymbol{\alpha},j)$.

Although Assumption 1 is stated in terms of the endogenous objects, it is easy to see that it will be satisfied unless the discount factor is too low. An implication of Lemma 2 is that when a particular party remains in power for sufficiently long, the Lagrange multipliers on the sustainability constraints and distortions begin declining.

Theorem 2 *Suppose Assumptions 1, 2 and Condition 1 hold. Then for any $\varepsilon>0$ there exists $K\in\mathbb{N}$ such that if $\mathbf{j}\left(h^t\right)=\mathbf{j}\left(h^{t+k}\right)=j$ for $k=1,...,K$, then $\lambda\left(h^{t+m}\right)<\varepsilon$ for all $m\geq K$.*

Proof. Fix j and $h^\infty\in H^\infty$, and suppose that $\mathbf{j}\left(h^t\right)=\mathbf{j}\left(h^{t+k}\right)=j$ for all $k\in\mathbb{N}$. Then from Lemma 2, $\left\{\alpha_j\left(h^{t+k}\right)\right\}_{k=0}^{\infty}$ is a non-decreasing sequence, and moreover, by construction $\alpha_j\left(h^{t+k}\right)\in[0,1]$ for each j and all h^{t+k}. Consequently, $\alpha_j\left(h^{t+k}\right)\to\bar{\alpha}_j$. Next note that $\bar{\alpha}_j<1$. To see this note that the inspection of the maximization problem (14) shows that when $\bar{\alpha}_j = 1$, the constraint (12) is slack. Since the objective function is continuous in the vector of Pareto weights $\boldsymbol{\alpha}$, this implies that for $\alpha_j = 1 - \varepsilon$ with ε sufficiently small, the constraint is also slack and $\lambda_j \left(h^{t+k}\right) = 0$. This implies that there exists $\varepsilon_j > 0$ such that starting with $\alpha_j \left(h^t\right) < 1 - \varepsilon_j$, we cannot have $\alpha_j \left(h^{t+1}\right) = 1$ for any h^t, since from equation (15) this would imply that $\lambda_j \left(h^{t+k}\right) = \infty$, which is not possible. Next equation (15) also implies that if $\lambda \left(h^{t+k}\right) = 0$, then $\alpha_j \left(h^{t+k}\right)$ will remain constant (since $\mathbf{j} \left(h^t\right) = \mathbf{j} \left(h^{t+k}\right) = j$ for all $k \in \mathbb{N}$). Therefore, $\alpha_j \left(h^{t+k}\right) \to \bar{\alpha}_j < 1$. Next, inspection of equation (15) shows that as $\alpha_j \left(h^{t+k}\right) \to \bar{\alpha}_j$, we have $\mu_j \left(h^{t+k}\right) - \mu_j \left(h^{t+k-1}\right) \to 0$ and thus $\lambda \left(h^{t+k}\right) \to 0$ (by virtue of the fact that $\bar{\alpha}_j < 1$). ■

Intuitively, along the path in which a particular group remains in power for a long time, distortions ultimately decline. This is because as a particular group remains in power for a long time, its Pareto weight increases sufficiently and the allocations do not need to be distorted to satisfy the sustainability constraint.

The major result in Theorem 2 is that as a particular group remains in power longer, distortions eventually decline. Intuitively, this follows from the fact that when the group in power has a higher updated Pareto weight, then there is no need to distort allocations as much. In the limit, if the group in power had a weight equal to 1, then the Pareto efficient allocation would give all consumption to individuals from this group, and therefore, there would be no reason to distort the labor supply of other groups in order to relax the sustainability constraint and reduce rents to this group. Put differently, recall that distortions (and inefficiencies) arise because the group in power does not have a sufficiently high Pareto weight and the Pareto efficient allocation reduces its consumption by reducing total output and thus relaxing its sustainability constraint. As a group remains in power for longer, its updated Pareto weight increases and as a result, there is less need for this type of distortions. This reflects itself in a reduction in the Lagrange multiplier associated with the sustainability constraint.

Theorem 2 also suggests a result reminiscent to the conjecture discussed in the Introduction; greater political stability translates into lower inefficiencies and better public policies. This conclusion does not follow from the theorem, however. The theorem is for a given sample path (holding the Markov process regulating power switches fixed). The conjecture linking political stability to efficient public policy refers to a comparison of the extent of distortions for different underlying Markov processes governing power switches. We will be discuss such comparison in greater detail in Section 5.

Theorem 2 does not answer the question of whether distortions will ultimately disappear—i.e., whether we will have $\lambda\left(h^t\right)=0$ for all h^t after some date. More formally, we call any allocation $\{c_j^*, l_j^*\}_j$ a *sustainable first-best allocation* if $\{c_j^*, l_j^*\}_j$ is a first best allocation that satisfies

$$\frac{1}{1-\beta}u_j\left(c_j^*, l_j^*\right) \geq v_j\left(\sum_{i\neq j} l_i^*\right) + \beta V_j^W(j) \text{ for all } j \tag{17}$$

The next theorem addresses this question.

Theorem 3 *Suppose that Assumptions 1 and 2 hold. Then there exists $\bar{\beta}$, with $0<\bar{\beta}<1$ such that*

(a) For all $\beta \geq \bar{\beta}$, there is some first-best allocation that is sustainable;

(b) For all $\beta < \bar{\beta}$, no first-best allocation is sustainable, and $\{c_j(h^t), l_j(h^t)\}_j$ converges to an invariant non-degenerate distribution F.

Proof. See the Appendix. ■

This theorem therefore shows that for high discount factors, i.e., $\beta \geq \bar{\beta}$, first-best allocations will be sustainable. However, when $\beta < \bar{\beta}$, then there will be permanent fluctuations in consumption and labor supply levels as political power fluctuates between different parties. The invariant distribution can be quite complex in general, with the history of power holdings shaping consumption and labor supply levels of each group.

This theorem does not answer the question of whether first-best allocations, when they are sustainable, will be ultimately reached. We address this question for the case of two parties in the next section.

4 The Case of Two Parties

In this section, we focus on an economy with two parties (rather than N parties as we have done so far). We also specialize utility function to be quasi-linear. Under these conditions, we show that when there exists a sustainable first-best allocation (i.e., an undistorted allocations for some Pareto weights), the equilibrium will necessarily converge to a point in the set of first-best allocations. More specifically, starting with any Pareto weights, the allocations ultimately converge to undistorted allocations.

For the rest of this section, we impose the following assumption on the preferences.

Assumption 3 ***(quasi-linearity)*** *The instantaneous utility of each party j satisfies*

$$u_j\left(c_j - h_j\left(l_j\right)\right)$$

with the normalization

$$h_j'\left(1\right) = 1. \tag{18}$$

Assumption 3 implies that there are no income effects in labor supply. Consequently, when there are no distortions, the level of labor supply by each group will be constant, and given the normalization in (18), this labor supply level will be equal to 1.

Under Assumption 3, Theorem 2 implies that as a particular group remains in power for a sufficiently long time, overall output in the economy will increase (since there are no income effects, lower distortions translate into higher labor supply levels). The absence of income effects also simplifies the analysis and dynamics, which is our main focus in this section.

We now state and prove three lemmas, which together will enable us to establish our main result in this section Theorem 4.

We first show that the party with a higher Pareto weight will receive higher value.

Lemma 3 *For any two vectors of Pareto weights α, α', if $\alpha_i > \alpha_i'$ then $V_i\left[\alpha, j\right] \geq V_i\left[\alpha', j\right]$ for $j \in \{1, 2\}$.*

Proof. Without loss of generality, let $i = 1$. Optimality implies

$$\alpha_1 V_1\left[\alpha, j\right] + \alpha_2 V_2\left[\alpha, j\right] \geq \alpha_1 V_1\left[\alpha', j\right] + \alpha_2 V_2\left[\alpha', j\right]$$

and

$$\alpha_1' V_1\left[\alpha', j\right] + \alpha_2' V_2\left[\alpha', j\right] \geq \alpha_1' V_1\left[\alpha, j\right] + \alpha_2' V_2\left[\alpha, j\right].$$

These conditions imply that

$$\left(\alpha_1 - \alpha_1'\right)\left(V_1\left[\alpha, j\right] - V_2\left[\alpha, j\right]\right) \geq \left(\alpha_1 - \alpha_1'\right)\left(V_1\left[\alpha', j\right] - V_2\left[\alpha', j\right]\right)$$

or

$$V_1\left[\alpha, j\right] - V_1\left[\alpha', j\right] \geq V_2\left[\alpha, j\right] - V_2\left[\alpha', j\right]. \tag{19}$$

Suppose that $V_1[\alpha, j] < V_1[\alpha', j]$. Then (19) implies that $V_2[\alpha, j] < V_2[\alpha', j]$. But this is impossible, because then $(V_1[\alpha', j], V_2[\alpha', j])$ would Pareto dominate $(V_1[\alpha, j], V_2[\alpha, j])$. ■

Let $\alpha^* = (\alpha_1^*, \alpha_2^*)$ be a vector of the Pareto weights for which first best allocation is sustainable. Consider any other initial vector $\alpha_0 \neq \alpha^*$and suppose that the first best allocation that corresponds to that vector is not sustainable. This implies that at least for one party the sustainability constraint (12) binds. The next lemma shows that sustainability constraint (12) does not bind if any party has a Pareto weight higher that α_j^*. While the proof of the lemma is somewhat involved, the intuition for the result is straightforward. If a party has a Pareto weight higher than the sustainable first best weight, the planner is treating such agent better than in the sustainable allocation. Therefore, when in power the sustainability constraint does not bind.

Lemma 4 *Suppose Assumptions 1 and 3 hold. Suppose that* $\alpha_j\left(h^{t-1}\right) \geq \alpha_j^*$ *for some* j, h^{t-1} *and* $\mathbf{j}\left(h^t\right) = j$. *Then* $\lambda_j\left(h^t\right) = 0$ *and* $\alpha_j\left(h^t\right) = \alpha_j\left(h^{t-1}\right)$.

Proof. Let us consider the relaxed problem of maximizing (9) without the constraint (12) following history h^t such that $\alpha_j\left(h^{t-1}\right) \geq \alpha_j^*$. We will characterize the solution to this relaxed problem and then show that the solution in fact satisfies (12) establishing that $\lambda_j\left(h^t\right) = 0$ and $\alpha_j\left(h^t\right) = \alpha_j\left(h^{t-1}\right)$.

Without loss of generality assume that $j = 1$. The expected utility of party 1 in history h^t in the relaxed problem is

$$u_1\left(c_1\left(\alpha\left(h^{t-1}\right)\right), l_1\left(\alpha\left(h^{t-1}\right)\right)\right) + \beta\left(m(1|1)V_1\left[\alpha\left(h^{t-1}\right), 1\right] + m(2|1)V_1\left[\alpha\left(h^{t-1}\right), 2\right]\right)$$

where $\left(c_i\left(\alpha\left(h^{t-1}\right)\right), l_i\left(\alpha\left(h^{t-1}\right)\right)\right)$ is a solution to the maximization problem

$$\max_{\{c_i, l_i\}} \alpha_1\left(h^{t-1}\right) u_1\left(c_1, l_1\right) + \alpha_2\left(h^{t-1}\right) u_2\left(c_2, l_2\right)$$

subject to

$$c_1 + c_2 \leq l_1 + l_2.$$

Since there is no sustainability constraint, Assumption 3 immediately implies that $l_j\left(\alpha\left(h^{t-1}\right)\right) = 1$ for all j, and moreover, $u_1\left(c_1\left(\alpha\left(h^{t-1}\right)\right), l_1\left(\alpha\left(h^{t-1}\right)\right)\right)$ is increasing in $\alpha_1\left(h^{t-1}\right)$.

Since Pareto weights α^* correspond to the sustainable allocation,

$$u_1\left(c_1\left(\alpha^*\right), l_1\left(\alpha^*\right)\right)+\beta\left(m(1|1) V_1\left[\alpha^*, 1\right]+m(2|1) V_1\left[\alpha^*, 2\right]\right) \\ \geq v_1\left(l_2\left(\alpha^*\right)\right)+\beta V_1^W(1) \tag{20}$$

Once again, Assumption 3 implies that $l_j\left(\alpha^*\right)=1$ for all j. From Lemma 3, $V_1\left[\alpha\left(h^{t-1}\right), j\right] \geq V_1\left[\alpha^*, j\right]$ for all j. Therefore the solution to the relaxed problem satisfies (12) if

$$u_1\left(c_1\left(\alpha\left(h^{t-1}\right)\right), 1\right)-v_1(1) \geq u_1\left(c_1\left(\alpha^*\right), 1\right)-v_1(1)$$

Since $u_1\left(c_1\left(\alpha\left(h^{t-1}\right)\right), 1\right)$ is increasing in $\alpha_1\left(h^{t-1}\right)$ and $\alpha_1\left(h^{t-1}\right) \geq \alpha_1^*$, this inequality is satisfied. ■

The previous lemma established that if party j is in power and has an updated Pareto weight above α_j^*, its next period updated Pareto weight remains the same. The next key step in our argument is to show that if a party has Pareto weight is below α_j^*, its next period updated Pareto weight is also below α_j^* (even if its current sustainability constraint is binding).

Lemma 5 *Suppose Assumptions 1 and 3 hold. Suppose that* $\alpha_j\left(h^{t-1}\right)<\alpha_j^*$ *for some* j, h^{t-1} *and* $\mathbf{j}\left(h^t\right)=j$. *Then* $\alpha_j\left(h^t\right) \leq \alpha_j^*$ *for all subsequent* h^t.

Proof. See the Appendix. ■

This lemma is the key to the main result in this section. It shows that if the sustainability constraint does not hold for group j that is in power even though its Pareto weight is below α_j^*, then for all subsequent histories its Pareto weight will not exceed α_j^*. The proof utilizes quasi-linearity of preferences to put structure on updated Pareto weights and the corresponding allocations.

Now we are ready to prove the most important result of this section about the convergence to the first best allocations.

Theorem 4 *Suppose that Assumptions 1 and 3 hold. If* $\beta \geq \bar{\beta}$, *then for all* j

$$\left(c_j\left(h^t\right), l_j\left(h^t\right)\right) \rightarrow\left(c_j^*, l_j^*\right)$$

where $\left\{\left(c_j^*, l_j^*\right)\right\}_{j=1}^2$ *is a first best sustainable allocation.*

Proof. Suppose the first best allocation with Pareto weight α^* is sustainable. Without loss of generality, suppose $\alpha_1(h^0) \leq \alpha_1^*$. Lemmas 4 and 5 imply that $\alpha_1(h^t)$ is a monotonically increasing sequence bounded above by α_1^*. Therefore $\alpha_1(h^t)$ must converge. Such convergence is possible only if (12) does not bind for both parties, which is possible only for the first best sustainable allocation. ■

This theorem establishes that if there exist first-best allocations that are sustainable they will be ultimately reached. This implies that the political economy frictions in this situation will disappear in the long run. The resulting long-run allocations will not feature distortions and fluctuations in consumption and labor supply. Note, however, that the theorem does not imply that such first-best allocations will be reached immediately. Sustainability constraints may bind for a while, because the sustainable first-best allocations may involve too high a level of utility for one of the groups. In this case, a first-best allocation will be reached only after a specific path of power switches increases the Pareto weight of this group to a level consistent with a first-best allocation. After this point, sustainability constraints do not bind for either party, and thus Pareto weights are no longer updated and the same allocation is repeated in every period thereafter. Interestingly, however, this first-best allocation may still involve transfer from one group to another.

5 Political Stability and Efficiency

Our framework enables an investigation of the implications of persistence of power on the sustainability of first-best applications. In particular, the "stability" or persistence of power is captured by the underlying Markov process for power switches. If the Markov process $m(j \mid j')$ makes it very likely that one of the groups, say group 1, will be in power all the time, we can think of this as a very "stable distribution of political power".

Such an investigation is important partly because a common conjecture in the political economy literature is that such stable distributions of political power are conducive to better policies. For example Olson (1993) and McGuire and Olson (1996) reach this conclusion by contrasting an all-encompassing long-lived dictator to a "roving bandit". They argue that a dictator with stable political power is superior to a roving bandit and will generate better public policies. This conjecture at first appears plausible, even compelling: what matters for better policies are high "effective discount factors," and frequent switches in the identity of powerholders would reduce these effective discount factors. Hence, stability (persistence) of power should be conducive to better policies and allocations. Similar insights emerge from the standard principal-agent models of political economy, such as Barro (1973), Ferejohn (1986), Persson, Roland and Tabellini (1997, 2000), because, in these models, it is easier to provide incentives to a politician who is more likely to remain in power. We next investigate whether a similar result applies in our context. In particular, we ask which types of Markov processes make it more likely that a large set of first-best allocations are sustainable.

Our main result in this section is that this common conjecture is not generally correct, and that in particular, in our framework, essentially the opposite of this conjecture holds. We show this in two parts. In the next subsection, we show that the opposite of this conjectures is true for the set of sustainable first-best allocations; greater persistence of power encourages deviations and leads to a smaller set of sustainable first-best allocations. We then investigate how changes in persistence of power influences the utility of different players.

5.1 The Set of Sustainable First-Best Allocations

In this subsection, we show that higher persistence of power makes distortions more likely, in the sense that it leads to a smaller set of sustainable first-best allocations. This result is stated in the next theorem. In this subsection, we again consider a general setup (N parties and non quasi-linear utilities).

Theorem 5 *Consider an economy consisting of N groups, with group j having utility functions $u_j(c_j, l_j)$ satisfying Assumption 1. Suppose that $m(j \mid j) = \rho$ and $m(j' \mid j) = (1-\rho)/(N-1)$ for any $j' \neq j$. Then $\bar{\beta}$ is increasing in ρ, i.e., the set of sustainable first-best allocations is smaller when ρ is greater.*

Proof. Recall that a first-best allocation satisfies (17). The left-hand side of this expression is independent of ρ, so is the first term on the right-hand side. Therefore, the desired result follows if the second term on the right-hand side, $V_j^W(j)$, is increasing in ρ, i.e., if (20) holds for any pair (β, ρ), then it holds for any (β, ρ') with $\rho' \leq \rho$, and thus the threshold $\bar{\beta}(\rho)$ about which it holds is increasing in ρ.

We now prove that this is the case. From the specification of the power switching process, we have that group j will remain in power next period with probability ρ, and hence $V_j^W(j)$

satisfies

$$V_j^W(j) = \rho V_j^P + (1-\rho) V_j^{NP}, \tag{21}$$

where V_j^P and V_j^{NP} are respectively the utility of being in power and not in power after a deviation. These are given by

$$V_j^P = u_j\left(\tilde{l}_j, \tilde{l}_j\right) + \beta\rho V_j^P + \beta(1-\rho) V_j^{NP},$$

and

$$V_j^{NP} = \beta\left(1 - \frac{1-\rho}{N-1}\right) V_j^{NP} + \beta\left(\frac{1-\rho}{N-1}\right) V_j^P,$$

where $\tilde{l}_j$ solves (8). Subtracting the second equation from the first, we obtain

$$V_j^P - V_j^{NP} = \frac{u_j\left(\tilde{l}_j, \tilde{l}_j\right)}{1-\beta\rho+\beta\left(\frac{1-\rho}{N-1}\right)},$$

and therefore

$$V_j^P = \frac{1-\beta+\beta\left(\frac{1-\rho}{N-1}\right)}{(1-\beta)\left(1-\beta\rho+\beta\left(\frac{1-\rho}{N-1}\right)\right)} u_j\left(\tilde{l}_j, \tilde{l}_j\right),$$

and substituting this into (21), we obtain

$$V_j^W(j) = \frac{\beta\left(\frac{1-\rho}{N-1}\right)+(1-\beta)\rho}{(1-\beta)\left(1-\beta\rho+\beta\left(\frac{1-\rho}{N-1}\right)\right)} u_j\left(\tilde{l}_j, \tilde{l}_j\right),$$

which is increasing in ρ, establishing the desired result. ■

This theorem implies the converse of the Olson conjecture discussed above: the set of sustainable first-best allocations is maximized when there are *frequent power switches* between different groups. The Olson conjecture is based on the idea that "effective discount factors" are lower with frequent power switches, and this should make "cooperation" more difficult. "Effective discount factors" would be the key factor in shaping cooperation (the willingness of the party in power to refrain from deviating) only if those in power can only be rewarded when in power. This is not necessarily the case, however, in reality or in our model. In particular, in our model deviation incentives are countered by increasing current utility and the Pareto weight of the party in power, and, all else equal, groups with higher Pareto weights will receive greater utility in all future dates. This reasoning demonstrates why "effective discount factor" is not necessarily the appropriate notion in this context. Instead, Theorem 5 has a simple intuition: the value of deviation for a group in power is determined by the persistence of power; when power is highly persistent, deviation becomes more attractive, since the group in power can still obtain relatively high returns following a deviation as it is likely to remain in power. In contrast, with more frequent power switches, the group in power is likely to be out of power tomorrow, effectively reducing deviation value. Since first-best allocations, and thus first-best utilities, are independent of the persistence of power, this implies that greater persistence makes deviation more attractive relative to candidate first-best allocations, and thus first-best allocations become less likely to be sustainable.

5.2 Numerical Illustration: The Form of the Pareto Frontier

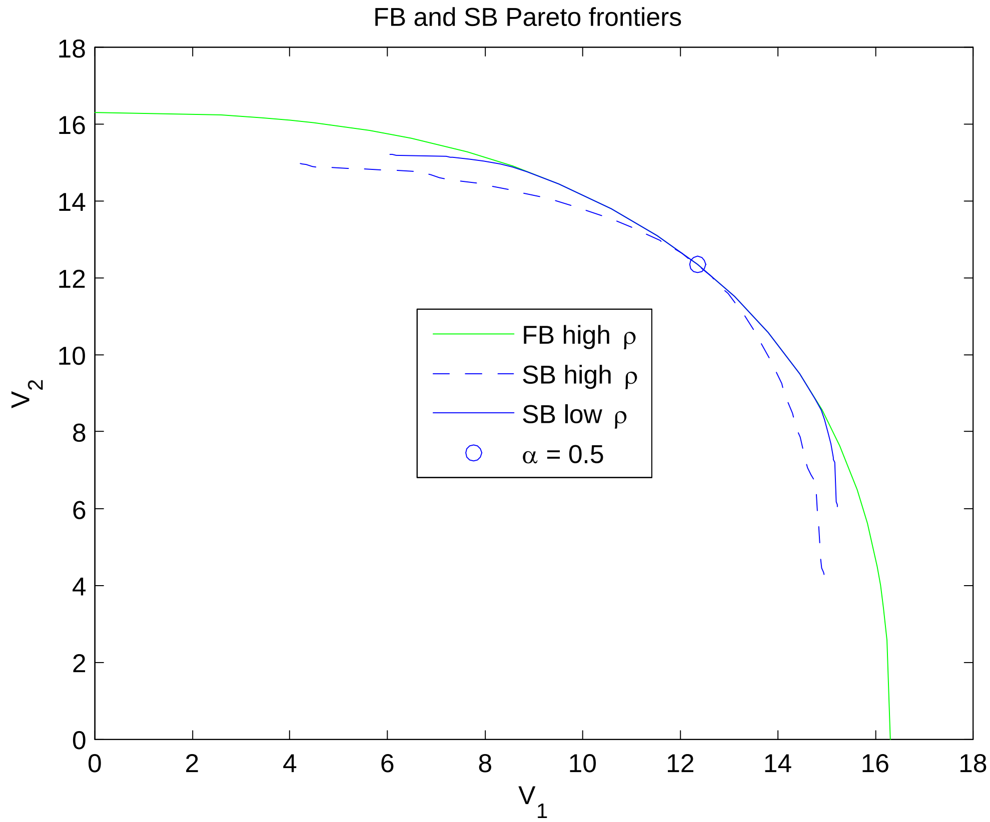
In this subsection, we numerically investigate the effect of persistence and frequency of power switches on the structure of constrained Pareto efficient allocations. In particular, we study how both the ex-ante Pareto frontier, which applies before the identity of the party in power is revealed, and the ex-post Pareto frontier, conditional on the identity of the party in power, vary with the degree of persistence. Our purpose is not to undertake a detailed calibration, but to provide illustrative numerical computations. We focus on an economy with two groups, $j = 1, 2$, and further simplify the discussion by assuming quasi-linear and identical utilities, given by

$$u_j\left(c - h\left(l\right)\right) = \frac{1}{1-\sigma}\left(c - \frac{1}{1+\gamma} l^{1+\gamma}\right)^{1-\sigma}.$$

We set $\gamma = 1$, $\sigma = 0.6$ and choose a symmetric Markov process for power switches with $m(1|1) = m(2|2) = \rho$, so that ρ is the persistence parameter (higher ρ corresponds to greater persistence). In Figure 1, we focus on the ex-ante Pareto frontier. For any given Pareto weight α we define ex-ante utility of party i by

$$V_i^{ea}\left[\alpha\right] = \frac{1}{2} V_i\left[\alpha, 1\right] + \frac{1}{2} V_i\left[\alpha, 2\right].$$

The figure plots $V_2^{ea}\left[\alpha\right]$ and $V_1^{ea}\left[\alpha\right]$ for different values of α and for two different values of levels of the persistence, ρ, $\rho = 0.9$ represented by the inner solid line, and $\rho = 0.6$ shown as the dashed line. We also show the first-best Pareto frontier for comparison (the outer solid line). We chose a discount factor β so that only one first best allocation (that corresponding to the Pareto weights $\alpha_1 = \alpha_2 = 0.5$ is sustainable) when $\rho = 0.9$.



Consistent with Theorem 5, a larger set of first best allocations is sustainable when persistence is lower. This can be seen by observing the common part of the first-best frontier and two other frontiers. For $\rho = 0.6$ this common part is larger than for $\rho = 0.9$ (which is just one point corresponding to $\alpha_1 = \alpha_2 = 0.5$). Also, the whole ex-ante Pareto frontier for low persistence lies above the Pareto frontier for high persistence, which implies that, before uncertainty about the identity of the party in power is realized, both parties are better off, and would prefer to be, in a regime with frequent power switches.

If the institutional characteristics of the society determining the frequency of power switches were chosen "behind the veil of ignorance," then this result would imply that both parties would prefer lower persistence. However, most institutional characteristics in practice are not determined behind a veil of ignorance. Different groups would typically have different amounts of political power, and in the context of our model, one would be "in power". In this case, what would be relevant is the ex-post not be ex-ante Pareto frontier, and we next turn to the ex-post Pareto frontier. This is shown in Figure 2, assuming that party 1 is currently in power. As with Figure 1, this figure also plots $V_2[\alpha, 1]$ against $V_1[\alpha, 1]$ for different values of α, and for high and low levels of persistence ($\rho = 0.9$ and $\rho = 0.6$). Figure 2 first shows that higher persistence imposes a greater "lower bound" on the possible payoff of party 1, which is in power. This can be seen from the fact that the beginning of the right solid line ($\rho = 0.6$) starts lower than the beginning of the left line ($\rho = 0.9$). This implies that greater persistence decreases the highest payoff that party 2 can get. The more surprising pattern in Figure 2 is that for high values of α, the value of party 1 is *lower* with higher persistence. This appears paradoxical at first, since higher persistence improves the deviation value of party 1. We investigate the reason for this pattern in greater detail in Figure 3.

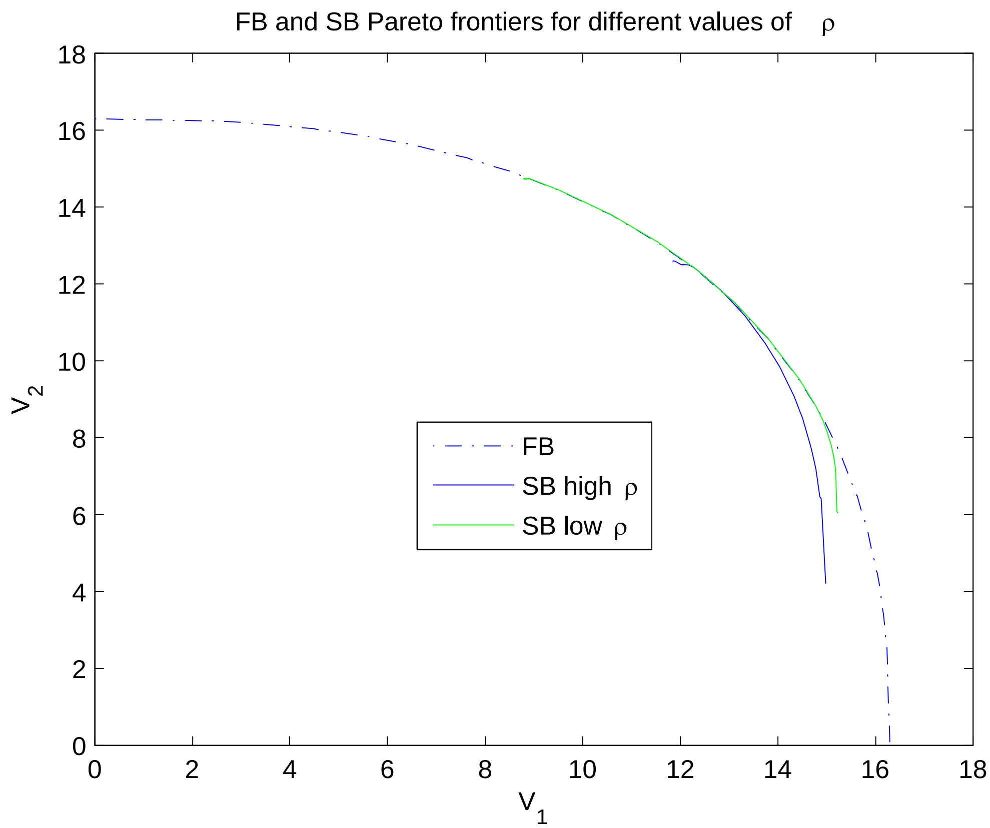
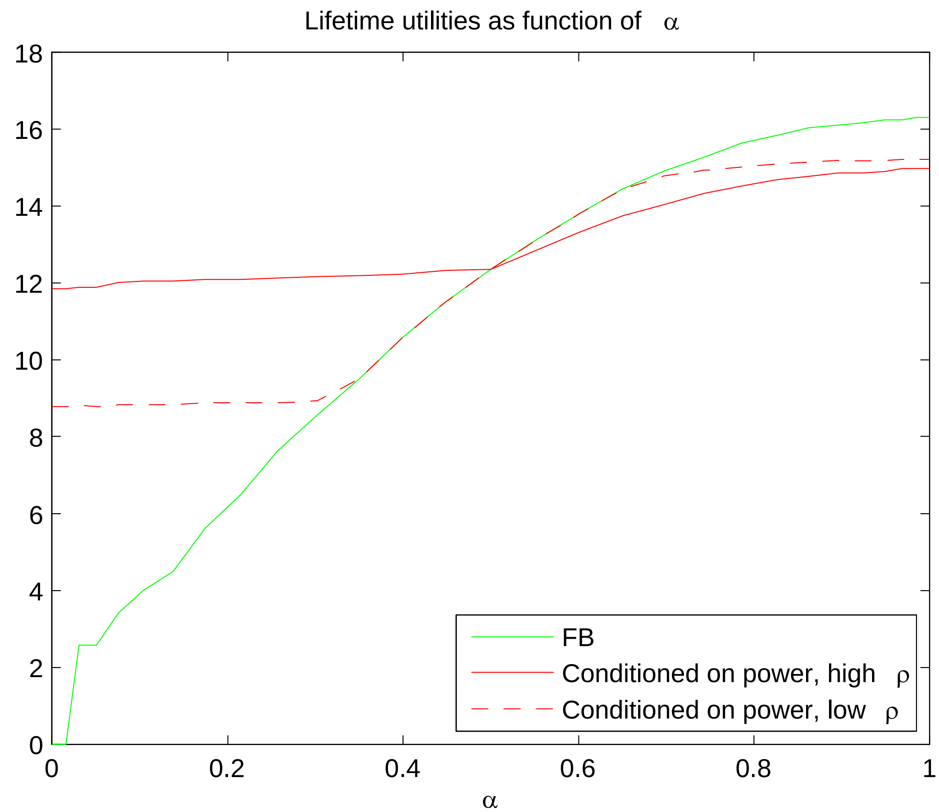


Figure 3 plots the payoff of party 1, $V_1[\alpha, 1]$, for different values values of α_1. The line representing the first best allocation is monotonically increasing with the Pareto weight assigned to party 1. Figure 3 also shows that for α_1 sufficiently high, party 1 obtains higher value with lower persistence (for $\alpha_1 > 1/2$, the line representing $\rho = 0.9$ is below the dashed line representing $\rho = 0.6$). The reason for this is as follows. When α_1 is sufficiently high, the sustainability constraint of party 1 is slack. Thus greater persistence does not necessitate an increase in current consumption or Pareto weight to satisfy its sustainability constraint. But it implies that the deviation value of party 2 will also be higher when it comes to power. When party 2 comes to power, its Pareto weight will be low and thus its sustainability constraint will be binding. A greater deviation value for party 2 at this point therefore translates into higher utility for it and lower utility for party 1. The anticipation of this lower utility in the future is the reason why the value of party 1 is decreasing in the degree of persistence in power switches for α_1 sufficiently high (greater than 1/2 in the figure). The analogue of this argument holds for weights below 1/2. Consistent with Figure 2, for low initial Pareto weights the utility of party 1 is increasing in the degree of persistence ρ (for $\alpha_1 < 1/2$, the line representing $\rho = 0.9$ is above the dashed line representing $\rho = 0.6$). Here, higher persistence of power increases the value of deviation and requires the planner to allocate more utility for this party with a low initial Pareto weight. This reasoning explains why the "lower bound" on the equilibrium payoffs for party 1 is higher with high persistence.

Another important implication of Figure 3 (already visible from Figure 2) is that changes in persistence do not necessarily correspond to Pareto improvements once the identity of the party in power is known. This highlights that even though the set of sustainable first-best allocations expands when the degree of persistence declines, along-the-equilibrium-path utility of both parties (conditional on the identity of the party in power) need not increase. This suggests that we should not necessarily expect a strong tendency for societies to gravitate towards institutional settings that increase the frequency of power switches.



6 Concluding Remarks

In this paper, we studied the constrained Pareto efficient allocations in a dynamic production economy in which the group in power decides the allocation of resources. The environment is a simple model of political economy. In our model, different groups have conflicting preferences and, at any given point in time, one of the groups has the political power to decide (or to influence) the allocation of resources. We made relatively few assumptions on the interactions between the groups; the process of power switches between groups is modeled in a reduced-form way with an exogenous Markov process. Our focus has been on the allocations that can be achieved given the distribution and fluctuations of political power in this society—rather than potential institutional failures leading to Pareto dominated equilibria given the underlying process of power switches. This focus motivated our characterization of (constrained) Pareto efficient equilibria. In the constrained Pareto efficient equilibria, there are well-defined political economy distortions that change over time.

The distortions in constrained Pareto efficient equilibria are a direct consequence of the sustainability constraints, which reflect the political economy interactions in this economy. If these sustainability constraints are not satisfied, the group in power would allocate all production to itself. The results here are driven by the location and shape of the Pareto frontier and by the "power" of different groups, which corresponds to what point the society is located along the Pareto frontier.

We showed how the analysis in the paper is simplified by the fact that these Pareto efficient allocations take a quasi-Markovian structure and can be represented recursively as a function of the identity of the group in power and updated Pareto weights. This recursive formulation allows us to provide a characterization of the level and dynamics of taxes and transfers in the economy.

We demonstrated that for high discount factors the economy converges to a first-best allocation in which there may be transfers between groups, but labor supply decisions are not distorted and the levels of labor supply and consumption do not fluctuate over time. When discount factors are low, the economy converges to an invariant stochastic distribution in which distortions do not disappear and labor supply and consumption levels fluctuate over time, even asymptotically.

We also showed that the set of sustainable first-best allocations is "decreasing" in the degree of persistence of the Markov process for power change. This result directly contradicts a common conjecture that there will be fewer distortions when the political system creates a stable ruling group (see, e.g., Olson, 1993, or McGuire and Olson, 1996, as well as the standard principle-agent models of political economy such as Barro, 1973, Ferejohn, 1986, Persson, Roland and Tabellini, 1997, 2000). The reason why this conjecture is incorrect illustrates an important insight of our approach. In an economy where the key distributional conflict is between different social groups, these groups can be rewarded not only when they hold power, but also when they are out of power (and they engage in consumption and production). Consequently, the probability of power switches does not directly affect "effective discount factors" and potentially invalidating the insight on which this conjecture is based. Because the persistence of the Markov process for power switches reduces deviation payoffs (while first-best payoffs are independent of persistence), greater persistence makes first-best allocations less like you to be sustainable. Nevertheless, it remains true that distortions decrease along sample paths where a particular group remains in power for a longer span of time (holding the underlying stochastic process for power switch is constant). This is because as a particular group remains longer in power, its Pareto weight increases and ultimately makes its sustainability constraint slack, thus removing the labor supply distortions on other groups.

While our analysis focused on the distortions introduced by the political economy friction, it is straightforward to derive implications of these results for tax policy. If the group in power sets taxes and transfers rather than directly deciding allocations, then the constrained Pareto efficient allocation can be decentralized as a competitive equilibrium, but this would necessarily involve the use of distortionary taxes. This observation implies that the fluctuations of distortions, consumption and labor supply levels derived as part of the Pareto efficient allocations in this paper also correspond to fluctuations in taxes—not simply to the presence of and fluctuations in "wedges" between the marginal utility of consumption and the marginal disutility of labor. The result that distortionary taxes must be used to decentralize the Pareto efficient allocation has a simple intuition, for clarifying the source of distortions in our economy: distortionary taxes must be used in order to discourage labor supply, because greater labor supply would increase the amount of output at the group in power can expropriate, tightening its sustainability constraint. Starting from an undistorted allocation, a small increase in taxes (distortions) would have a second-order cost in terms of lost net output, while having a first-order benefit in terms of relaxing the sustainability constraint when the latter is binding (see also Acemoglu, Golosov and Tsyvinski, 2008a).

We believe that the framework studied here is attractive both because we can analyze the effect of political economy distortions without specifying all of the details of interactions between groups and the process of decision-making. Undoubtedly, these institutional details are important in practice, and may lead to the emergence of outcomes inside the constrained efficient Pareto frontier. A natural next step is then to investigate what types of institutional structures can support ("implement") the constrained Pareto efficient allocations. This would give a different perspective on the role of specific institutions, as potential tools regulating the allocation of political power in society and placing constraints on the exercise of such power so as to achieve constrained Pareto efficient allocations. Nevertheless, our results indicating that changes in the frequency of power switches that improve ex-ante welfare do not necessarily improve ex-post welfare for all groups suggest that even when such specific institutions implementing constrained Pareto efficient allocations exist, whether they will emerge in equilibrium needs to be studied in the context of well-specified models. We leave an investigation of these issues to future work. Another important area for future research is to endogenize the Markov process for power switches. In modern societies, fluctuations of political power between different groups arise because of electoral competition, possible political coalitions between different groups lending their support to a specific party or group, or in extreme circumstances, because different groups can use their *de facto* power, such as in revolutions or in civil wars, to gain *de jure* power (see, e.g., Acemoglu and Robinson, 2006a).

7 Appendix

In this Appendix, we provide some of the technical details and results omitted from the text. We start with the first-order conditions that characterize the Pareto efficient allocation. We then discuss how randomizations can be allowed without changing the conclusions in the text.

7.1 First-Order Conditions

Under Assumptions 1 and 2, the constrained efficient allocation satisfies the following first-order conditions for any h^t:

$$\left[\begin{array}{c}\alpha_j\beta^t M\left(h^t \mid h^{\varnothing}\right)+\sum_{s=0}^{t}\sum_{h^{t-s}}\mathbf{I}_{j\neq\mathbf{j}(h^{t-s})}\beta^s M\left(h^t \mid h^{t-s}\right)\eta_j\left(h^{t-s}\right)\\ +\sum_{s=0}^{t}\sum_{h^{t-s}}\lambda\left(h^{t-s}\right)\mathbf{I}_{j=\mathbf{j}(h^{t-s})}\beta^s M\left(h^t \mid h^{t-s}\right)\end{array}\right]\times$$
$$\frac{\partial u_j\left(c_j\left(h^t\right),l_j\left(h^t\right)\right)}{\partial c_j}=\zeta\left(h^t\right)\text{ for }j\neq j', \tag{22}$$

$$\left[\begin{array}{c}\alpha_j\beta^t M\left(h^t \mid h^{\varnothing}\right)+\sum_{s=0}^{t}\sum_{h^{t-s}}\mathbf{I}_{j\neq\mathbf{j}(h^{t-s})}\beta^s M\left(h^t \mid h^{t-s}\right)\eta_j\left(h^{t-s}\right)\\ +\sum_{s=0}^{t}\sum_{h^{t-s}}\lambda\left(h^{t-s}\right)\mathbf{I}_{j=\mathbf{j}(h^{t-s})}\beta^s M\left(h^t \mid h^{t-s}\right)\end{array}\right]\times \tag{23}$$
$$\frac{\partial u_j\left(c_j\left(h^t\right),l_j\left(h^t\right)\right)}{\partial l}=-\zeta\left(h^t\right)+\lambda\left(h^t\right)\frac{\partial\left[\max\limits_{\tilde{l}\geq 0}u_{j'}\left(\sum_{j\neq j'}^{N}l_j\left(h^t\right)+\tilde{l},\tilde{l}\right)\right]}{\partial l_j\left(h_t\right)}\text{ for }j\neq j',$$

$$\left[\begin{array}{c}\alpha_j\beta^t M\left(h^t \mid h^{\varnothing}\right)+\sum_{s=0}^{t}\sum_{h^{t-s}}\mathbf{I}_{j\neq\mathbf{j}(h^{t-s})}\beta^s M\left(h^t \mid h^{t-s}\right)\eta_j\left(h^{t-s}\right)\\ +\sum_{s=0}^{t}\sum_{h^{t-s}}\lambda\left(h^{t-s}\right)\mathbf{I}_{j=\mathbf{j}(h^{t-s})}\beta^s M\left(h^t \mid h^{t-s}\right)\end{array}\right]\times$$
$$\frac{\partial u_j\left(c_j\left(h^t\right),l_j\left(h^t\right)\right)}{\partial c_j}=\zeta\left(h^t\right)\text{ for }j=j', \tag{24}$$

and

$$\begin{aligned}&\left[\begin{array}{c}\alpha_j\beta^t M\left(h^t \mid h^{\varnothing}\right)+\sum_{s=0}^{t}\sum_{h^{t-s}}\mathbf{I}_{j\neq\mathbf{j}(h^{t-s})}\beta^s M\left(h^t \mid h^{t-s}\right)\eta_j\left(h^{t-s}\right)\\ +\sum_{s=0}^{t}\sum_{h^{t-s}}\lambda\left(h^{t-s}\right)\mathbf{I}_{j=\mathbf{j}(h^{t-s})}\beta^s M\left(h^t \mid h^{t-s}\right)\end{array}\right]\\ &\times\frac{\partial u_j\left(c_j\left(h^t\right),l_j\left(h^t\right)\right)}{\partial c_j}\\ =\ &\zeta\left(h^t\right)\text{ for }j=j'.\end{aligned}$$

These readily show that the consumption-labor marginal of the group in power is not distorted relative to first best. In contrast, when the sustainability constraint is binding, so that $\lambda\left(h^t\right)>0$, this margin for groups not in power is distorted downwards (towards lower labor supply).

7.2 Randomizations and Theorem 6

As discussed in the text, the maximization of (9) subject to (10), (11) and (12) is a potentially non-convex one. If non-convexities are important, allowing for public randomizations would improve the achievable value. In this part of the Appendix, we allow for such randomizations. We show that randomizations do not affect any of our major results, mainly because any randomization will be over a finite number of (in fact two) allocations, and each allocation in the support of the stochastic distribution induced by the randomizations will satisfy first-order conditions similar to those analyzed in the text. To establish this result, we will first state a generalized version of Theorem 1, which shows that the characterization of constrained Pareto efficient allocation takes a recursive form even when randomizations are allowed.

Formally, randomizations could be introduced by either considering correlated equilibria or by explicitly introducing a publicly-observed randomization device. We pursue the second strategy, since it allows for a tractable formulation in the context of the problem here.

Let us first formulate a version of Proposition 1 with randomizations. In particular, let $\mathcal{C} = \{\{c_j, l_j\}_{j=1}^N \in \mathbb{R}^{2N} : \sum_{j=1}^N c_j \leq \sum_{j=1}^N l_j \text{ and } 0 \leq l_j \leq \bar{l}\}$ be the set of possible consumption-labor allocations for different groups, and let $\mathcal{P}^\infty$ be the set of probability measures defined over the set $\mathcal{C}^\infty$. Moreover, for each t and $h^t \in H^t$, let $\zeta\left(\cdot \mid h^t\right) \in \mathcal{P}^\infty$ be a probability measure over consumption-labor allocations for different groups given history of power holdings h^t. Then, this randomized-version Proposition 1 can be written as

Problem A :

$$\max_{\zeta(\cdot|h^t)\in\mathcal{P}^\infty} \sum_{t=0}^{\infty} \beta^t \left[\int \sum_{j=1}^N \alpha_j u_j(c_j, l_j) \zeta(d(c_j, l_j) \mid h^t) \right]$$

subject to

$$\sum_{j=1}^N c_j \leq \sum_{j=1}^N l_j \quad \zeta\left(\cdot \mid h^t\right)\text{-almost everywhere, for all } h^t \in H^t \text{ and all } t = 0, 1, ...,$$

$$\begin{aligned} u_j\left(c_j, l_j\right) &\geq 0 \ \zeta\left(\cdot \mid h^t\right)\text{-almost everywhere,} \\ \text{for all } h^t &\in H^t, \text{ all } j = 1, ..., N, \text{ and all } t = 0, 1, ..., \end{aligned}$$

and

$$
\begin{aligned}
\mathbb{E}_t \sum_{s=0}^{\infty} \beta^s u_j \left(c_j, l_j\right) &\geq u_j \left(\sum_{j'=1}^{N} l_{j'} \ , l_j \right) \ \zeta \left(\cdot \mid h^t\right) \text{-almost everywhere,} \\
\text{for all } h^t &\in H_j^t \text{ and all } t = 0, 1, ...
\end{aligned}
$$

We next assume that there exists an independently distributed uniform random variable υ_t publicly observed at each t before actions are taken. Consequently, actions can be conditioned on υ_t. This implies that formally, the history ω^t should now include $\upsilon^t = (\upsilon_0, ..., \upsilon_t)$, and in terms of Proposition 1, we should now condition on $z^t = \left(h^t, \upsilon^t\right)$. Denote the set of z^t's by Z^t and partition this into $Z_1^t, ..., Z_N^t$, depending on which group is in power at time t. Once conditioning on this publicly-observed random variable is allowed, the maximization problem that characterizes the constrained Pareto efficient allocations can be written as

$$
\textbf{Problem B} \ : \quad \max_{\{c_j(z), l_j(z)\}_{j=1}^N} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\sum_{j=1}^{N} \alpha_j u_j (c_j \left(z^t\right), l_j \left(z^t\right)) \right]
$$

subject to

$$
\sum_{j=1}^{N} c_j \left(z^t\right) \leq \sum_{j=1}^{N} l_j \left(z^t\right)) \text{ for all } z^t \in Z^t \text{ and all } t = 0, 1, ..., \tag{25}
$$

$$
u_j \left(c_j \left(z^{t+s}\right), l_j \left(z^{t+s}\right)\right) \geq 0 \text{ for all } z^{t+s} \in Z^{t+s} \text{ and all } j = 1, ..., N. \tag{26}
$$

and

$$
\mathbb{E}_t \sum_{s=0}^{\infty} \beta^s u_j \left(c_j \left(z^{t+s}\right), l_j \left(z^{t+s}\right)\right) \geq u_j \left(\sum_{j'=1}^{N} l_{j'} \left(z^t\right) \ , l_j \left(z^t\right) \right) \text{ for all } z^t \in Z_j^t \text{ and all } t = 0, 1, ...
$$

Problem B allows for randomizations in the same way as Problem A, since allocations can be conditioned on the realization of the random variable today and in all past dates. Any solution to Problem A is a solution to Problem B and vice versa. Given this, our first result is a generalization of Theorem 1.

Theorem 6 *The constrained Pareto efficient allocation with randomizations has a quasi-Markovian structure whereby* $\{c_j(z), l_j(z)\}_{j=1}^N$ *depend only on* $\mathbf{s} \equiv \left(\{\alpha_j(z)\}_{j=1}^N, \mathbf{j}(z)\right)$*, i.e., only on updated weights and the identity of the group in power.*

Proof. The proof follows similar lines to that of Theorem 1. In particular, with a similar manipulation, we can write the maximization problem following history z^{t-1} as:

Problem C :

$$\max_{\{c_j(z),l_j(z)\}_{j=1}^N} \quad \mathcal{L} = \mathbb{E}_t \sum_{s=t}^{\infty} \beta^s \left[\sum_{j=1}^{N} \left(\alpha_j + \mu_j \left(z^{t-1} \right) \right) u_j \left(c_j \left(z^s \right), l_j \left(z^s \right) \right) \right]$$
$$-\mathbb{E}_t \sum_{s=t}^{\infty} \beta^s \sum_{j=1}^{N} \lambda_j \left(z^s \right) \left[\mathbb{E}_s \sum_{s'=s}^{\infty} \beta^{s'-s} u_j \left(c_j \left(z^{s'} \right), l_j \left(z^{s'} \right) \right) - u_j \left(\sum_{i=1}^{N} l_i \left(z^s \right), l_j \left(z^s \right) \right) \right]$$

subject to (25) and (26). By construction, solutions to Problem B and Problem C coincide. The same argument as in the text establishes that Problem C has a recursive structure, where only the updated weights

$$\alpha_j \left(z^{t-1} \right) \equiv \frac{\alpha_j + \mu_j \left(z^{t-1} \right)}{\sum_{i=1}^{N} \left(\alpha_i + \mu_i \left(z^{t-1} \right) \right)}$$

and the identity of the group in power matter for future allocations. ∎

Our next result states that the solution to Problem C will involve randomization using at most two values.

Theorem 7 *To characterize the Pareto efficient allocations, it is sufficient to restrict $\zeta \left(\cdot \mid h^t \right)$ for all $h^t \in H^t$ and for all t to have its support over two vectors of consumption, labor supply and updated Pareto weights.*

Proof. This theorem follows from Lemmas 7 and 8 in Appendix C of Acemoglu, Golosov and Tsyvinski (2006). ∎

In light of Theorem 7, we can consider the following simplified maximization problem, where following history h^{t-1}, the constrained Pareto efficient allocation is a solution to the following Lagrangian:

$$\max_{\left\{c_j^r(h),l_j^r(h),\xi^1(h)\right\}_{j=1}^N} \mathcal{L} = \mathbb{E}_t \sum_{s=t}^{\infty} \beta^s M \left(h^s \right) \left[\sum_{r=1}^{2} \xi^r \left(h^s \right) \sum_{j=1}^{N} \left(\alpha_j + \mu_j^r \left(h^{t-1} \right) \right) u_j \left(c_j^r \left(h^s \right), l_j^r \left(h^s \right) \right) \right]$$
$$-\mathbb{E}_t \sum_{s=t}^{\infty} \beta^s M \left(h^{t+s} \right) \sum_{r=1}^{2} \xi^r \left(h^{t+s} \right) \sum_{j=1}^{N} \lambda_j^r \left(h^{t+s} \right) \times$$
$$\left[\mathbb{E}_s \sum_{s'=s}^{\infty} \beta^{s'-s} u_j \left(c_j^r \left(h^{s'} \right), l_j^r \left(h^{s'} \right) \right) - u_j \left(\sum_{i=1}^{N} l_i^r \left(h^{t+s} \right), l_j^r \left(h^{t+s} \right) \right) \right]$$

subject (10) and (11). In this problem, $r = 1$ and 2 correspond to possible randomizations over two values given any history, so, for example, $c_j^r(h^t)$ and $l_j^r(h^t)$ are the consumption and labor supply levels for group j following history h^t into two possible events $r = 1$ and $r = 2$. Here the $\xi^r(h^t)$ correspond to the probabilities of these two possible events, (naturally with $\xi^1(h^t) + \xi^2(h^t) = 1$). Consequently, the first-order conditions in the text apply to $c_j^r(h^t)$ and $l_j^r(h^t)$ for $r = 1$ and $r = 2$, and all the necessary conditions and the resulting downward distortions apply for each case separately.

7.3 Proof of Theorem 3

Part (a): Let $\boldsymbol{\alpha}^*$ be a vector of Pareto weights that correspond to the first best allocations that solve for all i

$$\max_{l_i} u_i(l_i, l_i).$$

Denote these allocations $\{c_i^*, l_i^*\}_i$. Let $V_j^W(j;\beta)$ be the value of the worst equilibrium for party j as a function of β and let κ be the smallest element in $m(j|j')$. By Assumption 2, $\kappa > 0$. Then

$$\begin{aligned} V_j^W(j;\beta) &\leq \frac{1}{1-\beta}\left[(1-\kappa)\, u_j(c_j^*, l_j^*) + \kappa u_j(0,0)\right] \\ &= \frac{1-\kappa}{1-\beta} u_j(c_j^*, l_j^*) \end{aligned}$$

Then

$$\begin{aligned} & u_j(c_j^*, l_j^*)/(1-\beta) - V_j^W(j;\beta) \\ \geq\; & \frac{\kappa}{1-\beta} u_j(c_j^*, l_j^*) \end{aligned}$$

and the right hand side goes to infinity as β approaches one. Therefore there exists $\bar{\beta}$ s.t. (17) holds.

Part (b): Recall that $s \equiv (\alpha, j)$, where $\alpha \in \Delta^{N-1}$ are ex ante weights and $j \in \mathcal{N} \equiv \{1, ..., N\}$. Let $S = \Delta^{N-1} \times \mathcal{N}$ and $\mathcal{S}$ be the Borel σ-algebra over S.

Our model implies that each s is deterministically mapped into $\alpha' = h(s)$. Thus the stochastic process for s' is determined uniquely by the Markov process $m(j' \mid j)$. Therefore we can define

$$p\left((\alpha, j), (\alpha', j')\right) = \begin{cases} 0 & \text{if } \alpha' \neq h(\alpha, j) \\ m(j' \mid j) & \text{if } \alpha' = h(\alpha, j) \end{cases}$$

Note that $p((\alpha, j), (\alpha', j'))$ is uniformly bounded above by $\max_{j,j' \in \mathcal{N}} m(j' \mid j) < 1$ (the latter inequality by Assumption 2). Therefore, for all $s \in S$ and $A \in \mathcal{S}$, we have

$$P((\alpha, j), A) = \int_A p\left((\alpha, j), (\alpha, j)'\right) \mu\left(d\left(\alpha', j'\right)\right).$$

Stokey, Lucas and Prescott, Exercise 11.4f (which is straightforward to see) shows that if there exists a bounded above function p and and a finite measure μ such that

$$P(s, A) = \int_A p\left(s, s'\right) \mu\left(ds'\right)$$

for all $s \in S$ and $A \in \mathcal{S}$, then Doeblin's condition is satisfied. Almost sure convergence to an invariant limiting distribution then follows.

Since $s\left(h^t\right)$ converges to an invariant distribution, so do $\left\{c_j\left(h^t\right), l_j\left(h^t\right)\right\}_{j=1}^N$. It remains to show that invariant distribution over $\left\{c_j\left(h^t\right), l_j\left(h^t\right)\right\}_{j=1}^N$ is non-degenerate. Suppose, to obtain a contradiction, that it is degenerate, say given by $\left\{c_j^*, l_j^*\right\}_{j=1}^N$. Since first best is not sustainable, there must exists some party j for which sustainability constraint (12) is binding when it comes to power. Therefore, by Theorem 2,

$$\frac{\partial u_j\left(c_j^*, l_j^*\right)}{\partial c} > -\frac{\partial u_j\left(c_j^*, l_j^*\right)}{\partial l}.$$

where $\left\{c_j^*, l_j^*\right\}_{j=1}^N$ is the allocations in the invariant distribution. Choose any $\varepsilon > 0$ and K as in Theorem 2. Since there is a positive probability that party j remains in power for K periods, Theorem 2 implies that

$$\frac{\partial u_j\left(c_j^*, l_j^*\right)}{\partial c} + \frac{\partial u_j\left(c_j^*, l_j^*\right)}{\partial l} < 0,$$

which yields a contradiction. Therefore $\left\{c_j\left(h^t\right), l_j\left(h^t\right)\right\}_{j=1}^N$ converges to nondegenerate distribution.

7.4 Proof of Lemma 5

Without loss of generality assume that $j = 1$. If constraint (12) does not bind, the result follows immediately. Suppose therefore that (12) binds. Then the Lagrange multiplier $\lambda_1\left(h^t\right) > 0$ and from (15) $\alpha_1\left(h^t\right) = \left(\alpha_1\left(h^{t-1}\right) + \lambda_1\left(h^t\right)\right) / \left(1 + \lambda_1\left(h^t\right)\right)$. Suppose that $\alpha_1\left(h^t\right) > \alpha_1^*$. Then it must be true that

$$u_1\left(c_1\left(\alpha\left(h^{t-1}\right),\lambda_1\left(h^t\right)\right),l_1\left(\alpha\left(h^{t-1}\right),\lambda_1\left(h^t\right)\right)\right)-v_1\left(l_2\left(\alpha\left(h^{t-1}\right),\lambda_1\left(h^t\right)\right)\right)$$
$$+\beta\left(m(1|1)V_1\left[\alpha\left(h^t\right),1\right]+m(2|1)V_1\left[\alpha\left(h^t\right),2\right]\right)=\beta V_1^W(1), \tag{27}$$

where $\{c_i(\alpha,\lambda),l_i(\alpha,\lambda)\}_{i=1}^2$ is a solution to the maximization problem

$$\max_{\{c_i,l_i\}}(\alpha_1+\lambda)u_1(c_1,l_1)+\alpha_2u_2(c_2,l_2)-\lambda v_1(l_2)$$

subject to

$$c_1+c_2\leq l_1+l_2.$$

Lemma 3 establishes that $V_1\left[\alpha\left(h^t\right),j\right]\geq V_1[\alpha^*,j]$ for all j. If we have that for any λ such that $(\alpha_1+\lambda)/(1+\lambda)>\alpha_1^*$, the following is true

$$u_1(c_1(\alpha,\lambda),l_1(\alpha,\lambda))-v_1(l_2(\alpha,\lambda))>u_1(c_1^*,l_1^*)-v_1(l_2^*),$$

then (20) would immediately imply (27) cannot hold with equality, thus leading to a contradiction and establishing the desired result. Thus to complete the proof of lemma, we will establish the following claim:

Claim 1 *For any Δ such that $(\alpha_1+\Delta)/(1+\Delta)>\alpha_1^*$, we have that*

$$u_1(c_1(\alpha,\Delta),l_1(\alpha,\Delta))-v_1(l_2(\alpha,\Delta))>u_1(c_1^*,l_1^*)-v_1(l_2^*).$$

We will first proof some intermediate results. Let us adopt the following simpler notation for the rest of the proof:

$$u_{iC}(c_i,l_i)=\frac{\partial u_i(c_i,l_i)}{\partial c_i}\text{ and }u_{iL}(c_i,l_i)=\frac{\partial u_i(c_i,l_i)}{\partial l_i}$$

and similarly denote second order derivatives. Sometimes we will also drop (c_i,l_i) and use notation u_{iC},u_{iL}, etc., whenever there is no confusion.

Claim 2 *Let $c(\Delta;\alpha)$ and $l(\Delta,\alpha)$ be the solution to the problem*

$$W(\Delta;\alpha)=\max_{c,l}(\alpha_1+\Delta)u_1(c_1,l_1)+\alpha_2u_2(c_2,l_2)-\Delta v_1(l_2)$$

subject to

$$c_1+c_2\leq l_1+l_2$$

Then

$$u_1(c_1(\Delta,\alpha), l_1(\Delta,\alpha)) - v_1(l_2(\Delta,\alpha)) \tag{28}$$

is increasing in Δ*.*

Proof. Consider two different Δ', Δ'' and denote the corresponding solutions to the above problem by $\{c_i', l_i'\}_{i=1}^2$ and $\{c_i'', l_i''\}_{i=1}^2$. By definition, this implies

$$\begin{aligned}
\left(\alpha_1 + \Delta'\right) u_1(c_1', l_1') + \alpha_2 u_2(c_2', l_2') - \Delta' v_1(l_2') &\geq \left(\alpha_1 + \Delta'\right) u_1(c_1'', l_1'') + \alpha_2 u_2(c_2'', l_2'') - \Delta' v_1(l_2'') \\
\left(\alpha_1 + \Delta''\right) u_1(c_1'', l_1'') + \alpha_2 u_2(c_2'', l_2'') - \Delta'' v_1(l_2'') &\geq \left(\alpha_1 + \Delta''\right) u_1(c_1', l_1') + \alpha_2 u_2(c_2', l_2') - \Delta'' v_1(l_2')
\end{aligned}$$

Summing these two inequalities in rearranging, we obtain

$$\begin{aligned}
\left(\Delta' - \Delta''\right)\left(u_1(c_1', l_1') - v_1(l_2')\right) &\geq \left(\Delta' - \Delta''\right)\left((c_1'', l_1'') - v_1(l_2'')\right) \\
\left(\Delta' - \Delta''\right)\left[\left(u_1(c_1', l_1') - v(l_2')\right) - \left((c_1'', l_1'') - v_1(l_2'')\right)\right] &\geq 0
\end{aligned}$$

Therefore, if $\Delta' > \Delta''$, then $(u_1(c_1', l_1') - v_1(l_2')) \geq ((c_1'', l_1'') - v_1(l_2''))$, which establishes (28).

■

Claim 3 *Suppose Assumptions 1 and 3 hold. Then for any Pareto weight* α^* *and for any* $\Delta \in [0, \alpha_1^*/\alpha_2^*]$*, define* $\alpha_{1\Delta}$ *by*

$$\frac{\alpha_{1\Delta} + \Delta}{1 + \Delta} = \alpha_1^*$$

and $\alpha_{2\Delta} = 1 - \alpha_{1\Delta}$*. Consider the maximization problem*

$$\Omega(\Delta) = \max_{c,l} \left(\alpha_{1\Delta} + \Delta\right) u_1(c_1, l_1) + \alpha_{2\Delta} u_2(c_2, l_2) - \Delta v_1(l_2) \tag{29}$$

subject to

$$c_1 + c_2 \leq l_1 + l_2.$$

Denote $\{c_i(\Delta), l_i(\Delta)\}_i$ *the solution to this problem. Then*

$$F(\Delta) \equiv u_1(c_1(\Delta), l_1(\Delta)) - v_1(l_2(\Delta)) \tag{30}$$

is increasing in Δ*. Moreover, for any* $\Delta > 0$ $F(\Delta) > F(0)$*.*

Proof. Consider the first-order conditions to the maximization problem in (29). Assumption 3 implies that $l_1(\Delta) = 1$. The other first-order conditions are

$$\left(\alpha_{1\Delta} + \Delta\right) u_{1C}\left(c_1(\Delta) - h_1(1)\right) = \alpha_{2\Delta} u_{2C}\left(c_2(\Delta) - h_2(l_2(\Delta))\right) \tag{31}$$

$$\frac{\Delta}{\alpha_{2\Delta}} \frac{v_1'(l_2(\Delta))}{u_{2C}(c_2(\Delta) - h_2(l_2(\Delta)))} = 1 - h_2'(l_2(\Delta)) \tag{32}$$

and

$$c_1(\Delta) + c_2(\Delta) = 1 + l_2(\Delta). \tag{33}$$

First, note that

$$\frac{\alpha_{1\Delta} + \Delta}{\alpha_{2\Delta}} = \frac{\alpha_1^*}{\alpha_2^*} \tag{34}$$

Substitute (34) into (31) to get

$$\alpha_1^* u_{1C}(c_1(\Delta) - h_1(1)) = \alpha_2^* u_{2C}(c_2(\Delta) - h_2(l_2(\Delta))) \tag{35}$$

Now differentiate F as defined in (30)

$$\begin{aligned} F'(\Delta) &= u_{1C} \frac{\partial c_1}{\partial \Delta} - v_1' \frac{\partial l_2}{\partial \Delta} \\ &= \frac{\partial l_2}{\partial \Delta} \left(u_{1C} \frac{\partial c_1}{\partial l_2} - v_1' \right) \end{aligned} \tag{36}$$

Substituting (33) into (35) and differentiating, we obtain

$$\alpha_1^* u_{1CC} \frac{\partial c_1(\Delta)}{\partial \Delta} = \alpha_2^* u_{2CC} \left((1 - h_2') \frac{\partial l_2}{\partial \Delta} - \frac{\partial c_1(\Delta)}{\partial \Delta} \right),$$

which implies

$$\begin{aligned} \frac{\partial c_1}{\partial l_2} &= \frac{\alpha_2^* u_{2CC}(1 - h_2')}{\alpha_1^* u_{1CC} + \alpha_2^* u_{2CC}} \\ &= \frac{\alpha_2^* u_{2CC}}{\alpha_1^* u_{1CC} + \alpha_2^* u_{2CC}} \times \frac{\Delta}{\alpha_{2\Delta}} \frac{v_1'}{u_{2C}} \\ &= \frac{\alpha_2^* u_{2CC}}{\alpha_1^* u_{1CC} + \alpha_2^* u_{2CC}} \times \frac{\Delta}{\alpha_{1\Delta} + \Delta} \frac{v_1'}{u_{1C}} \end{aligned}$$

where we used (32) in the second line and (31) in the third. Substituting this into (36), we obtain

$$\begin{aligned} F'(\Delta) &= v_1' \frac{\partial l_2}{\partial \Delta} \left(\frac{\alpha_2^* u_{2CC}}{\alpha_1^* u_{1CC} + \alpha_2^* u_{2CC}} \times \frac{\Delta}{\alpha_{1\Delta} + \Delta} - 1 \right) \\ &= v_1' \frac{\partial l_2}{\partial \Delta} \left(\frac{1}{\frac{\alpha_1^* u_{1CC}}{\alpha_2^* u_{2CC}} + 1} \times \frac{1}{1 + \frac{\alpha_{1\Delta}}{\Delta}} - 1 \right) \end{aligned}$$

The expression in the brackets is negative, therefore $F'(\Delta)$ has the opposite sign of $\frac{\partial l_2}{\partial \Delta}$. The desired result follows from $\frac{\partial l_2}{\partial \Delta} \leq 0$, which we establish next.

Consider two different Δ', Δ'' and denote the corresponding solutions to (29) by $\{c_i', l_i'\}_{i=1}^2$ and $\{c_i'', l_i''\}_{i=1}^2$. By definition, we have

$$\begin{aligned}
\left(\alpha_{1\Delta'} + \Delta'\right) u_1(c_1', l_1') + \alpha_{2\Delta'} u_2(c_2', l_2') - \Delta' v_1(l_2') &\geq \left(\alpha_{1\Delta'} + \Delta'\right) u_1(c_1'', l_1'') + \alpha_{2\Delta'} u_2(c_2'', l_2'') - \Delta' v_1(l_2'') \\
\left(\alpha_{1\Delta''} + \Delta''\right) u_1(c_1'', l_1'') + \alpha_{2\Delta''} u_2(c_2'', l_2'') - \Delta'' v_1(l_2'') &\geq \left(\alpha_{1\Delta''} + \Delta''\right) u_1(c_1', l_1') + \alpha_{2\Delta''} u_2(c_2', l_2') - \Delta'' v_1(l_2').
\end{aligned}$$

Now dividing these two inequalities by $\alpha_{2\Delta'}$ and $\alpha_{2\Delta''}$ respectively, we obtain

$$\begin{aligned}
\frac{\alpha_{1\Delta'} + \Delta'}{\alpha_{2\Delta'}} u_1(c_1', l_1') + u_2(c_2', l_2') - \frac{\Delta'}{\alpha_{2\Delta'}} v_1(l_2') &\geq \frac{\alpha_{1\Delta'} + \Delta'}{\alpha_{2\Delta'}} u_1(c_1'', l_1'') + u_2(c_2'', l_2'') - \frac{\Delta'}{\alpha_{2\Delta'}} v_1(l_2'') \\
\frac{\alpha_{1\Delta''} + \Delta''}{\alpha_{2\Delta''}} u_1(c_1'', l_1'') + u_2(c_2'', l_2'') - \frac{\Delta''}{\alpha_{2\Delta''}} v_1(l_2'') &\geq \frac{\alpha_{1\Delta''} + \Delta''}{\alpha_{2\Delta''}} u_1(c_1', l_1') + u_2(c_2', l_2') - \frac{\Delta''}{\alpha_{2\Delta''}} v_1(l_2')
\end{aligned}$$

Definition of α_Δ implies that for all Δ

$$\frac{\alpha_{1\Delta} + \Delta}{\alpha_{2\Delta}} = \frac{\alpha_1^*}{\alpha_2^*}$$

and therefore

$$\frac{\alpha_1^*}{\alpha_2^*} u_1(c_1', l_1') + u_2(c_2', l_2') - \frac{\Delta'}{\alpha_{2\Delta'}} v_1(l_2') \geq \frac{\alpha_1^*}{\alpha_2^*} u_1(c_1'', l_1'') + u_2(c_2'', l_2'') - \frac{\Delta'}{\alpha_{2\Delta'}} v_1(l_2'') \quad (37)$$

$$\frac{\alpha_1^*}{\alpha_2^*} u_1(c_1'', l_1'') + u_2(c_2'', l_2'') - \frac{\Delta''}{\alpha_{2\Delta''}} v_1(l_2'') \geq \frac{\alpha_1^*}{\alpha_2^*} u_1(c_1', l_1') + u_2(c_2', l_2') - \frac{\Delta''}{\alpha_{2\Delta''}} v_1(l_2') \quad (38)$$

These equations imply

$$\left(\frac{\Delta''}{\alpha_{2\Delta''}} - \frac{\Delta'}{\alpha_{2\Delta'}}\right) \left(v_1(l_2') - v_1(l_2'')\right) \geq 0 \quad (39)$$

Finally, from the definition of α_Δ, we have

$$\begin{aligned}
\frac{1 - \alpha_{2\Delta} + \Delta}{\alpha_{2\Delta}} &= \frac{\alpha_1^*}{\alpha_2^*} \\
\frac{1 + \Delta}{\alpha_{2\Delta}} &= 1 + \frac{\alpha_1^*}{\alpha_2^*}
\end{aligned}$$

which implies that $\alpha_{2\Delta}$ is increasing in Δ. Since $1/\alpha_{2\Delta}$ is decreasing in Δ and

$$\frac{1}{\alpha_{2\Delta}} + \frac{\Delta}{\alpha_{2\Delta}} = 1 + \frac{\alpha_1^*}{\alpha_2^*},$$

$\frac{\Delta}{\alpha_{2\Delta}}$ is increasing in Δ. Therefore from (39) $\Delta'' > \Delta'$ implies $l_2'' \leq l_2'$, completing the proof of the claim.

It remains to show that if $\Delta' = 0$ for any $\Delta'' > 0$ $F(\Delta'') > F(\Delta')$. Suppose that $F(\Delta'') = F(0)$. Previous analysis indicated that this is possible only if $l_2'' = l'_2$. But then (37) and (38) imply that

$$\alpha_1^* u_1(c_1', l_1') + \alpha_2^* u_2(c_2', l_2') = \alpha_1^* u_1(c_1'', l_1'') + \alpha_2^* u_2(c_2'', l_2'').$$

We know that $\{c_i', l_i'\}_i$ is a solution to maximizing $\alpha_1^* u_1(c_1, l_1) + \alpha_2^* u(c_2, l_2)$ subject to

$$c_1 + c_2 \leq l_1 + l_2 \tag{40}$$

Since u_i are strictly convex, this solution is unique. Therefore, any $\{c_i'', l_i''\}_i$ that satisfies (40) must have

$$\alpha_1^* u_1(c_1', l_1') + \alpha_2^* u_2(c_2', l_2') > \alpha_1^* u_1(c_1'', l_1'') + \alpha_2^* u_2(c_2'', l_2'')$$

leading to a contradiction. ■

The next claim completes the proof of lemma. We state this claim for party 1; clearly, the result is identical for party 2.

Claim 4 *Suppose Assumptions 1 and 3 hold. Let $\{c_i(\alpha, \Delta), l_i(\alpha, \Delta)\}_{i=1}^2$ be a solution to the problem*

$$\max_{\{c_i, l_i\}} (\alpha_1 + \Delta) u_1(c_1, l_1) + \alpha_2 u_2(c_2, l_2) - \Delta v_1(l_2)$$

s.t.

$$c_1 + c_2 \leq l_1 + l_2$$

for some $\Delta \geq 0$. For any Pareto weight $\alpha^* \neq \alpha$, if α and Δ are such that

$$\frac{\alpha_1 + \Delta}{1 + \Delta} > \alpha_1^*$$

then

$$u_1(c_1(\alpha, \Delta), l_1(\alpha, \Delta)) - v_1(l_2(\alpha, \Delta)) > u_1(c_1(\alpha^*, 0), l_1(\alpha^*, 0)) - v_1(l_2(\alpha^*, 0)) \tag{41}$$

Proof. Suppose $\alpha_1 < \alpha_1^*$. Let $\tilde{\Delta}$ be such that

$$\frac{\alpha_1 + \tilde{\Delta}}{1 + \tilde{\Delta}} = \alpha_1^*.$$

Since $\frac{\alpha_1 + \Delta}{1 + \Delta}$ is increasing in Δ, $0 < \tilde{\Delta} < \Delta$. From Claim 3,

$$u_1(c_1\left(\alpha, \tilde{\Delta}\right), l_1\left(\alpha, \tilde{\Delta}\right)) - v_1(l_2\left(\alpha, \tilde{\Delta}\right)) > u_1(c_1(\alpha^*, 0), l_1(\alpha^*, 0)) - v_1(l_2(\alpha^*, 0)) \tag{42}$$

and from Claim 2,

$$u_1(c_1(\alpha,\Delta), l_1(\alpha,\Delta)) - v_1(l_2(\alpha,\Delta)) \geq u_1(c_1\left(\alpha,\tilde{\Delta}\right), l_1\left(\alpha,\tilde{\Delta}\right)) - v_1(l_2\left(\alpha,\tilde{\Delta}\right))$$

establishing (41).

If $\alpha_1 > \alpha_1^*$, set $\tilde{\Delta} = 0$ and (42) follows from proof of Lemma 4. ■

8 References

Abreu, Dilip, David Pearce and Enio Stacchetti (1990) "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring" *Econometrica*, 58, 1041-1063.

Acemoglu, Daron, Georgy Egorov and Konstantin Sonin (2008) "Coalition Formation in Nondemocracies" *Review of Economic Studies*, 75(4), 987-1009, December.

Acemoglu, Daron, Michael Golosov and Aleh Tsyvinski (2006) "Markets Versus Governments: Political Economy of Mechanisms", *National Bureau of Economic Research Working Paper 1224*.

Acemoglu, Daron, Michael Golosov and Aleh Tsyvinski (2008a) "Political Economy of Mechanisms", *Econometrica,* 76(3), (2008): 619–641.

Acemoglu, Daron, Michael Golosov and Aleh Tsyvinski (2008b) "Markets Versus Governments", Journal of Monetary Economics, 55(1), 159-189

Acemoglu, Daron, Michael Golosov and Aleh Tsyvinski (2009a) "Dynamic Mirrlees Taxation under Political Economy Constraints", *Review of Economic Studies*, forthcoming.

Acemoglu, Daron, Michael Golosov and Aleh Tsyvinski (2009b) "Political Economy of Ramsey Taxation", *National Bureau of Economic Research Working paper 15302.*

Acemoglu, Daron and James A. Robinson (2000) "Why did the West Extend the Franchise? Democracy, Inequality and Growth in Historical Perspective," *Quarterly Journal of Economics*, 115, 1167-1199.

Acemoglu, Daron and James A. Robinson (2006a) *Economic Origins of Dictatorship and Democracy*, Cambridge University Press.

Acemoglu, Daron and James Robinson (2006b) "Economic Backwardness in Political Perspective" *American Political Science Review.*

Aguiar, Mark and Manuel Amador (2009) "Growth in the Shadow of Expropriation" Stanford mimeo.

Amador, Manuel (2003a) "Savings under Political Compromise", Stanford mimeo.

Amador, Manuel (2003b) "A Political Model of Sovereign Debt Repayment", Stanford mimeo.

Battaglini, Marco and Stephen Coate (2008) "A Dynamic Theory of Public Spending, Taxation and Debt", *American Economic Review*, v. 98, iss. 1, pp. 201-36.

Baron, David (1998) "Comparative Dynamics of Parliamentary Governments" *American Political Science Review*, 92, 593-609.

Barro, Robert (1973) "The Control of Politicians: an Economic Model" *Public Choice*, 14, 19-42.

Besley, Tim and Seteve Coate (1997) "An Economic Model of Representative Democracy," *Quarterly Journal of Economics*, 112, 85-114.

Besley, Timothy, and Stephen Coate (1998) "Sources of Inefficiency in a Representative Democracy: a Dynamic Analysis" *American Economic Review,* 88 139-156.

Bueno de Mesquita, Bruce D., James D. Morrow, Randolph M. Siverson and Alastair Smith (2003) *The Logic of Political Survival*, Cambridge: MIT Press.

Dixit, Avinash, Gene M. Grossman, and Faruk Gul (2000) "The Dynamics of Political Compromise." *Journal of Political Economy* 108, 3:531-568.

Dixit, Avinash and John Londregan (1995) "Redistributive Politics and Economic Efficiency." *American Political Science Review* 89, 4:856-866.

Ferejohn, John (1986) "Incumbent Performance and Electoral Control," *Public Choice* 50: 5-25.

Grossman, Gene M., and Elhanan Helpman (1994) "Protection for Sale." *American Economic Review* 84, 833-850

Grossman, Gene M., and Elhanan Helpman (1996) "Electoral Competition and Special Interest Politics." *Review of Economic Studies* 63, 265-286.

Hassler Jon, Sevi Mora, Kjandetil Storlesseten and Fabrizio Zilibotti (2003) "Survival of the Welfare State," *American Economic Review*, 93, 87-112.

Kocherlakota, Narayana (1996) "Implications of Efficient Risk Sharing without Commitment," *Review of Economic Studies.*

Krusell, Per and José-Victor Ríos-Rull (1996) "Vested Interests in a Theory of Growth and Stagnation," *Review of Economic Studies*, 63, 301-329.

Lagunoff, Roger (2005) "Markov Equilibrium in Models of Dynamic Endogenous Political Institutions" Georgetown mimeo.

Lagunoff, Roger (2006) "Dynamic Stability and reform of Political Institutions," Georgetown mimeo.

Marcet, Albert and Ramon Marimon (1998) "Recursive Contracts". Mimeo. University of Pompeu Fabra.

McGuire, Martin and Mancur Olson (1996) "The Economics of Autocracy and Majority Rule" *Journal of Economic Literature*, 34, 72-96.

Olson, Mancur C. (1993) "Dictatorship, Democracy and Development," *American Political Science Review*, 87, 567-575.

Persson, Torsten and Guido Tabellini (2000) *Political Economics: Explaining Economic Policy*, The MIT Press, Cambridge MA.

Persson, Torsten, Gerard Roland, and Guido Tabellini (1997) "Separation of Powers and Political Accountability." *Quaterly Journal of Economics* 112, 1163-1202.

Persson, Torsten, Gerard Roland and Guido Tabellini (2000) "Comparative Politics and Public Finance," *Journal of Political Economy*, 108, 1121-1161.

Prescott, Edward C. and Robert Townsend (1984a) "General Competitive Analysis in an Economy with Private Information", *International Economic Review*, 25, 1-20

Prescott, Edward C. and Robert Townsend (1984b) "Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard", *Econometrica* 52, 21-45.

Roberts, Kevin (1999) "Dynamic Voting in Clubs," London School of Economics mimeo.

Robinson James A. (2001) "When Is the State Predatory?" University of California, Berkeley, mimeo.

Sonin, Konstantin (2003) "Why the Rich May Favor Poor Protection of Property Rights" *Journal of Comparative Economics*, 31, 715-731.

Stokey, Nancy, Robert E. Lucas and Edward Prescott (1989) *Recursive Methods in Economic Dynamics*, Harvard University Press, Cambridge.

Thomas, Jonathan and Tim Worrall (1990) "Income Fluctuations and Asymmetric Information: An Example of Repeated Principle-Agent Problem", *Journal of Economic Theory* 51, 367-90.

Yared, Pierre (2009), "Politicians, Taxes, and Debt", *Review of Economic Studies* forthcoming.