A Theory of Political Competition over Military Policy and Income Redistribution

Naoki Yoshihara*
The Institute of Economic Research, Hitotsubashi University,†

This version June 2008

Abstract

We discuss political competition games between Left and Right parties, in which the policy space is two dimensional. One issue is the choice of proportional tax rate, and the second is the allocation of tax revenue between military policies and social welfare policies. On these political issues, the stylized fact is that left-wing parties prefer higher tax rates and lower military expenditure than do right-wing parties. We examine the kinds of political environments in which this fact can be rationalized as the equilibrium outcome of a given political game. By adopting the notion of the *party-unanimity Nash equilibrium* [Roemer (1998; 1999; 2001; 2005)], not only voters' economic motivations, but also their *ideological positions* are shown to be crucial factors in explaining stylized party behavior.

JEL Classification Numbers: C62, C72, D72, D78
Keywords: multidimensional political competition games, party-unanimity Nash equilibria, military expenditure, ideological concerns.

---

*Phone: (81)-42-580-8354, Fax: (81)-42-580-8333. e-mail: yosihara@ier.hit-u.ac.jp

†I am grateful to John E. Roemer for his detailed and insightful guidance on this research agenda. This research was originally initiated when I was visiting the Department of Political Sciences at Yale University as a Fulbright scholar, in August 2003. The generous hospitality of the department, as well as financial support of the Fulbright scholarship, are gratefully acknowledged.

1 Introduction

Military policy, which is in principle designed to provide for the national security of citizens, has been considered as one of the most basic roles of modern nation states. This is because the provision of national security is indispensable if the citizens of a country are to enjoy their individual liberties. The classical liberalists proposed the so-called *night watchman state* (or the *minimal state*) as the appropriate solution for protecting individual liberties from potential invaders. In contrast, social welfare policy, which seeks to provide minimum levels of income, services or other support for disadvantaged people, has been one of the main programs that contemporary nation states are expected to implement. A country that provides comprehensive social welfare programs is often identified as being a *welfare state*. In such a country, access to social welfare services is often considered a basic and inalienable right of those in need.

Since the end of World War II, most advanced countries have been expected to implement both military and social welfare policies. Thus, one of the most controversial political issues has involved what kind of tax system and budgetary allocation is needed to appropriately implement the 'optimal' provisions of national security and social welfare services. Facing that issue, left-wing parties usually place a higher priority on social welfare policies, which may involve increasing income tax rates or allocating more resources to those policies, whereas right-wing parties are less sympathetic to these concerns. In contrast, the left often strongly criticizes the expansion of military expenditure, whereas the right often justifies it on the grounds that national security is threatened.

However, the difference between the two parties in their social welfare policies may not have a game-theoretic reasoning when we consider the standard Downsian game [Downs (1957)] as a canonical model of political competition. This is because it is often the case that there is only one unique Downsian equilibrium, in which both parties propose the same policy. Moreover, the rational for the left usually preferring lower military expenditure than the right seems to be ambiguous because national security is a *pure public good* that every citizen collectively and simultaneously enjoys regardless of his or her income or social class.

In this paper, we discuss political competition games between left and right parties, in which the policy space is *two dimensional.* One policy issue is the choice of tax rule (the proportional tax rate), and the second is the allocation of tax revenue between military expenditure and the funding of social welfare provision. In these games, a rise in the tax rate and a fall in the share of tax revenue allocated to military expenditure indicates a strengthening of *welfare state policy*, which the left is expected to support. In contrast, if the proposed military expenditure is consistent with exceeding the optimal level of military force needed to defend a country from foreign aggression, and if a higher share of tax revenue is allocated to military policy, such a policy proposal may be termed *imperialism*,[1] to which the right is likely to be more sympathetic than the left.

Since the policy space in this political model is multidimensional, canonical political games, such as the Downsian party model and the Wittman party model [Wittman (1973)], cannot reliably produce a Nash equilibrium (in pure strategies).[2] To overcome this difficulty, we adopt the notion of *party-unanimity Nash equilibrium* (**PUNE**) introduced by Roemer (1998; 1999; 2001; 2005). The model is based on the idea that party decision makers have different interests, and their activists are divided into three *factions*: *opportunists*, *militants*, and *reformists*. The opportunist faction is concerned solely with winning office; the militant faction is concerned solely with publicizing the party's views; and the reformist faction is only concerned with the expected welfare of the party's members. The three factions within each party should bargain over policy proposals, given the policy proposal of the opposing party. If the policy proposal agreed on is Pareto efficient for the three factions, this policy constitutes a solution to the bargaining problem within the party. Then, a **PUNE** is a pair of policy proposals, each of which

[1] The term 'imperialism' simply indicates a *scheme of military policies* that enables a nation to act *hegemonically* against other nations by using the threat of a superior military force, which has been developed beyond the level needed for national security. This use of the term 'imperialism' may differ greatly from the classical Marxian notion of imperialism, which relates to the *economic relations* between countries (and within countries), rather than to more formal political and/or military relationships. For instance, Lenin (1916) argued that capitalism necessarily induced monopoly capitalism—which he also called 'imperialism'—that is required to find new markets and resources, and which represents the last and highest stage of capitalism.

[2] There is at least another possibility to address the existence issue of equilibrium in multidimensional games: the Besley-Coate-Osborne-Slivinski notion of citizen-candidate equilibrium [Besley and Coate (1997); Osborne and Slivinski (1996)]. However, in that models, the "citizen candidates" cannot commit to implement any policy but their own ideal policy, so that there are essentially no parties, which is inappropriate for our own motivation in this paper.

is the result of intra-party bargaining, given the other party's proposal.

We examine the competitive structure between the two parties under which the above-mentioned stylized party behavior can be rationalized as a **PUNE** of the political game. We start by analyzing two-dimensional political games in which citizens' utility functions are quasi-linear, which is a standard assumption in public-goods economies. In such a politico-economic environment, we show that the stylized party behavior does not constitute a **PUNE** of the political game. There are only two types of **PUNE**s in such a game: one is that both the left and the right propose the left's *ideal policy*, which combines the highest tax rate and the 'optimal' military expenditure; the other is that both parties propose their own ideal policies and the left wins the election with certainty. Therefore, the ideal policy of the left is always implemented. Moreover, the right's ideal military expenditure is less than the left's in equilibrium. The fact that these outcomes are unrealistic implies that standard public-goods politico-economic models fail to explain the stylized party behavior.

The above result may indicate that every citizen does not vote, solely according to his or her own economically motivated preference for military expenditure, and that voters may also have *non-economically motivated preferences*[3] for those kinds of public goods. We introduce a simple form of non-economically motivated preference for military expenditure, that reflect each citizen's *political ideology.* In such an extended model, each citizen has an economic preference over military expenditure and welfare services as well as a non-economic, ideological preference over the level of military forces. Then, under reasonable assumptions, we show that every **PUNE** of the political game in this extended model rationalizes stylized party behavior, whenever all citizens' *ideological concerns* about alternative military policies are sufficiently pronounced.

In what follows, in Section 2, we describe a basic model of the above two-dimensional political games, and also introduce the **PUNE** *with exogenous party formations.* In Sections 3 and 4, we discuss the existence and the characterization of **PUNE**s with exogenous party formations in non-ideological societies and in societies with ideological concerns, respectively. In Section 5,

[3]There have been recent developments in the literature on the theory of public goods provision with non-economically motivated preferences, which attempt to rationally explain why there are voluntary associations such as NPOs and/or NGOs that survive even under the threat of the free-rider problem and function as public goods providers. For instance, see Francois (2000; 2003; 2006).

we defines the **PUNE** *with endogenous party formations*, and obtain similar existence and characterization results to those obtained under the assumption of *exo*genous party formations. Finally, Section 6 provides concluding remarks.

2 The Model

Let the set of voter types be H, let the policy space be Υ, let the probability distribution of voter types in the polity be $\mathbb{F}$ on H, and let the utility function of type $h \in H$ over policies be $v(\cdot;h)$. Let $v(\cdot;h)$ be a non-negative real valued function for any $h \in H$. Let $(\tau^1, \tau^2) \in \Upsilon \times \Upsilon$ be a pair of policies. The set of voters who prefer τ^1 to τ^2 is denoted by $\Omega(\tau^1, \tau^2) \equiv \{h \in H \mid v(\tau^1;h) \geq v(\tau^2;h)\}$. We make the following assumption.

Assumption 1 (A1): *For any $\tau, \tau' \in \Upsilon$ with $\tau \neq \tau'$, the set of voters who are indifferent between τ and τ' is of $\mathbb{F}$-measure zero.*

Following Roemer (2001: Section 2.3, 2005), the fraction of voters who prefer policy τ^1 to policy τ^2 is $\mathbb{F}(\Omega(\tau^1, \tau^2))$. However, we assume that there is some *aggregate uncertainty* regarding how people will vote, so that the probability of victory depends on the fraction of the vote, and on a noise parameter ε which is uniformly distributed over $[-\gamma, \gamma]$, where $\gamma \in \left(0, \frac{1}{2}\right)$. Thus, the probability that τ^1 defeats τ^2 is:

$$\pi(\tau^1, \tau^2) = \begin{cases} 0 & \text{if } \mathbb{F}(\Omega(\tau^1, \tau^2)) + \gamma \leq \frac{1}{2} \\ \frac{\mathbb{F}\left(\Omega(\tau^1, \tau^2)\right) + \gamma - \frac{1}{2}}{2\gamma} & \text{if } \frac{1}{2} \in (\mathbb{F}(\Omega(\tau^1, \tau^2)) - \gamma, \mathbb{F}(\Omega(\tau^1, \tau^2)) + \gamma) \\ 1 & \text{if } \mathbb{F}(\Omega(\tau^1, \tau^2)) - \gamma \geq \frac{1}{2} \end{cases}, \quad (1)$$

whenever $\tau^1 \neq \tau^2$, and $\pi(\tau^1, \tau^2) = \frac{1}{2}$ whenever $\tau^1 = \tau^2$. Thus, one *political environment* is generally denoted by a tuple $\langle H, \mathbb{F}, \Upsilon, v, \gamma \rangle$.

In this paper, each voter is characterized by his or her income $w \in \mathbb{R}_+$ and his or her ideological position $a \in [0, 1]$, where the ideological position indicates a person's preference on the issue of how high defense expenditure should be. Thus, the set of voters is specified by:

$$H = \{(w, a) \in W \times [0, 1] \mid W \equiv [\underline{w}, \overline{w}] \subsetneq \mathbb{R}_+\}. \quad (2)$$

This population is characterized by two cumulative *marginal* distribution functions $G(w)$ and $R(a)$.

The policy issue facing this society is choosing a pair of proportional tax rates $t \in [0,1]$ and a ratio of defense expenditure over tax revenue, $\alpha \in [0,1]$. Thus, the policy space of this society is specified by:

$$\Upsilon = \{(t,\alpha) \mid t \in [0,1] \text{ and } \alpha \in [0,1]\}. \quad (3)$$

If the society chooses (t,α), then its tax revenue is $t\mu$ per capita, where μ is the average income in this society, and its defense expenditure is $\alpha t\mu$ per capita. Then, $(1-\alpha)\,t\mu$ is the subsidy that every citizen receives through an income redistribution policy. Thus, the choice of (t,α) implies a choice regarding income redistribution and military expenditure in this society. Every voter (w,a) has the same utility function,

$$v\,(t,\alpha;w,a) = (1-\beta)\,[(1-t)\,w + (1-\alpha)\,t\mu + \sigma\,(\alpha t\mu)] - \frac{\beta}{2}\,|\alpha t - a|^2 \quad (4)$$

where $\beta \in [0,1]$, and a indicates the voter's ideological position on the issue of defense expenditure. In addition, the term $(1-t)\,w + (1-\alpha)\,t\mu$ represents the voter's after-tax income when the policy (t,α) is implemented; the term $\sigma\,(\alpha t\mu)$ represents the voter's benefit from the national security supplied by the military forces; the term $-\frac{\beta}{2}\,|\alpha t - a|^2$ represents the satisfaction of the voter's political preferences over the issue of military expenditure. Assume that the function σ is *continuously differentiable*, *strictly concave*, and *monotonic*. The parameter β represents the weight on ideological views about the level of military forces.

We impose the following additional condition on the function σ.

Assumption 2 (A2): $\lim\limits_{\lambda\to 0} \frac{\partial\sigma(\lambda\mu)}{\partial\lambda\mu} = +\infty$, *and for some* $\lambda^* \in (0,1)$, $\frac{\partial\sigma(\lambda^*\mu)}{\partial\lambda^*\mu} = 1$.

The first component of **A2** implies that having no defense expenditure is exclusively undesirable in terms of national security. The second component of **A2** implies that if national income is exhausted by defense expenditure, this would result in an excessive level of military forces. In other words, the optimal defense expenditure $\lambda^*\mu$ does not require all national income to be exhausted by military expenditure. Thus, $\lambda^*\mu$ can be interpreted as a threshold: if military expenditure exceeds $\lambda^*\mu$, this implies that the main purpose of having this level of military force is not only to protect citizens from foreign aggression, but also to act hegemonically against other nations by using the threat of military force.

There are two political parties in society, *Left* (L) and *Right* (R), in the society. In Sections 3 and 4, we suppose that the *membership of both parties is exogenously fixed*, and that L represents relatively poor citizens, for whom $w_L < \mu$, and that its ideological position a_L indicates a preference for 'peace,' or 'antimilitarism.' In contrast, R represents relatively rich citizens, for whom $w_R > \mu$, and this party's ideological position a_R indicates a preference for 'relatively strong military power.' Let a^m denote the median ideological view: $R(a^m) = \frac{1}{2}$. We assume that $a_L < a^m < a_R$. Let $v_L(t, \alpha) = v(t, \alpha; w_L, a_L)$ be L's utility function, and let $v_R(t, \alpha) = v(t, \alpha; w_R, a_R)$ be R's utility function.

We define an equilibrium notion of this political game, that is, *party-unanimity Nash equilibrium* (**PUNE**), which was introduced by Roemer (1998; 1999; 2001; Chapter 8).

Definition 1: *Given a political environment* $\langle H, \mathbb{F}, \Upsilon, v, \gamma \rangle$ *as specified above, and given the two parties* Left (L) and Right (R), *a pair of policies* $(\tau^L, \tau^R) \in \Upsilon \times \Upsilon$ *with* $\tau^i = (t_i, \alpha_i)$ $(i = L, R)$ *constitutes a* party-unanimity Nash equilibrium (**PUNE**) *if* (τ^L, τ^R) *satisfies the following*:
(a) *given* τ^R*, there is no policy* $\tau \in \Upsilon$ *such that* $\pi(\tau, \tau^R) \geq \pi(\tau^L, \tau^R)$ *and* $v_L(\tau) \geq v_L(\tau^L)$*, with at least one strict inequality*;
(b) *given* τ^L*, there is no policy* $\tau \in \Upsilon$ *such that* $\pi(\tau^L, \tau) \leq \pi(\tau^L, \tau^R)$ *and* $v_R(\tau) \geq v_R(\tau^R)$*, with at least one strict inequality.*

In Definition 1, condition (a) implies that, facing the opponent's proposal τ^R, there is no policy in Υ that can improve the payoffs of all three factions in party L; and condition (b) makes the corresponding statement for the factions of party R.[4]

[4] In this definition, there is no statement of the reformist factions' payoffs, because (2a) and (2b) describe the conditions for the opportunist factions' payoffs, $\pi(\cdot, \cdot)$ and $1 - \pi(\cdot, \cdot)$, and those for the militant factions' payoffs, $v_L(\cdot)$ and $v_R(\cdot)$, only. However, as Roemer (2001: Chapter 8; Theorem 8.1(3)) showed, the equilibrium set corresponding to this simpler definition of the **PUNE** is equivalent to that of the rigorous definition of the **PUNE** given in Roemer (2001: Chapter 8; Definition 8.1).

3 Political Equilibrium in a Non-ideological Society

We first consider a non-ideological society, $\beta = 0$. When $\beta = 0$, every voter's utility function can be reduced to the following *quasi-linear* type:

$$v(t, \alpha; w, a) = (1 - t)\, w + (1 - \alpha)\, t\mu + \sigma(\alpha t\mu). \quad (5)$$

In this case, the only relevant information regarding the voter's type is his or her income level w; and hence, the voter space is unidimensional. However, the policy space remains two dimensional. Given a pair of policies $(\tau^L, \tau^R) = ((t_L, \alpha_L), (t_R, \alpha_R))$, let us define:

$$\triangle^{(w,a)}((t_L, \alpha_L), (t_R, \alpha_R)) \equiv v(t_L, \alpha_L; w, a) - v(t_R, \alpha_R; w, a). \quad (6)$$

By definition, $\triangle^{(w,a)}((t_L, \alpha_L), (t_R, \alpha_R)) > 0$ implies that the voter (w, a) prefers L to R when L offers (t_L, α_L) and R offers (t_R, α_R).

Let $\triangle t \equiv t_R - t_L$. Then, $\triangle^{(w,a)}((t_L, \alpha_L), (t_R, \alpha_R)) > 0$ holds if and only if:

$$w < \mu - \frac{\mu \cdot (\alpha_R t_R - \alpha_L t_L) + [\sigma(\alpha_L t_L \mu) - \sigma(\alpha_R t_R \mu)]}{\triangle t} \quad \text{if } \triangle t < 0, \quad (7a)$$

$$w > \mu - \frac{\mu \cdot (\alpha_R t_R - \alpha_L t_L) + [\sigma(\alpha_L t_L \mu) - \sigma(\alpha_R t_R \mu)]}{\triangle t} \quad \text{if } \triangle t > 0. \quad (7b)$$

Note that $\Omega(\tau^L, \tau^R) = \{(w, a) \in H \mid \triangle^{(w,a)}(\tau^L, \tau^R) \geq 0\}$. Thus, the fraction of voters who prefer $\tau^L = (t_L, \alpha_L)$ to $\tau^R = (t_R, \alpha_R)$ is defined as:

$$\mathbb{F}\left(\Omega(\tau^L, \tau^R)\right) = \int_{\underline{w}}^{\mu - \Theta(\tau^L, \tau^R)} g(w)\mathrm{d}w \quad \text{if } \triangle t < 0, \quad (8a)$$

$$\mathbb{F}\left(\Omega(\tau^L, \tau^R)\right) = \int_{\mu - \Theta(\tau^L, \tau^R)}^{\overline{w}} g(w)\mathrm{d}w \quad \text{if } \triangle t > 0, \quad (8b)$$

$$\text{where } \Theta(\tau^L, \tau^R) \equiv \frac{\mu \cdot (\alpha_R t_R - \alpha_L t_L) + [\sigma(\alpha_L t_L \mu) - \sigma(\alpha_R t_R \mu)]}{\triangle t},$$

and $g(w)$ is the density of W.

Now, we identify each party's ideal policy. Let $(\bar{t}_L, \overline{\alpha}_L)$ be the ideal policy of L. Then:

Lemma 1: *Let* **A2** *hold and* $\beta = 0$*. Then,* $\left(\bar{t}_L, \overline{\alpha}_L\right) = (1, \alpha^*)$*, where* $\alpha^* = \lambda^*$*.*

Note that any voter whose income is lower than μ has the same ideal policy as L's in the non-ideological society, in which $\beta = 0$. This is because every voter receives the same level of benefit $\sigma(\alpha t \mu)$ from the military policy, and L and R only differ in their respective income positions.

Let $\left(\bar{t}_R, \overline{\alpha}_R\right)$ be the ideal policy of R.

Lemma 2: *Let* **A2** *hold and* $\beta = 0$*. Then,* $\left(\bar{t}_R, \overline{\alpha}_R\right) = (t^*, 1)$*, where* t^* *is the solution of* $\frac{\partial \sigma(t^* \mu)}{\partial t \mu} = \frac{w_R}{\mu}$*.*

Note that any voter h whose income w_h is higher than μ has the same ideal military policy as R's in the non-ideological society, in which $\beta = 0$. However, the voter's ideal tax policy t^h has the property that $\frac{\partial \sigma\left(t^h \mu\right)}{\partial t \mu} = \frac{w_h}{\mu} > 1$, which differs from R's. Note that if $w_h = \mu$, then $\frac{\partial \sigma\left(t^h \mu\right)}{\partial t \mu} = 1$, which implies that $t^h = \lambda^*$. Moreover, this voter is indifferent between $(1, \lambda^*)$ and $(\lambda^*, 1)$. Since both $(1, \lambda^*)$ and $(\lambda^*, 1)$ are the voter's own ideal policies, the voter prefers L to R. Thus, some voter h with $w_h > \mu$ may still prefer L to R. In fact, any voter whose income is above μ, but below $\mu - \Theta(\tau^L, \tau^R)\,(> \mu)$ prefers L to R.

By Lemmas 1 and 2, we know that $\alpha^* \mu > t^* \mu$; that is, the ideal level of military forces supplied by L is higher than that supplied by R. This is because $w_R > \mu$. This result might be reasonable in a non-ideological society, because in this model, L represents relatively poor citizens, and R represents relatively rich citizens. Note that the poor need a higher level of national security than do the rich, because the rich, unlike the poor, can flee their home country when its national security is threatened.

We consider a society in which more than half of the voters have incomes below the mean. This implies the following reasonable assumption:

Assumption 3 (A3): $G(\mu) > \frac{1}{2}$.

We characterize the set of **PUNE**s in the game $\langle H, \mathbb{F}, L, R, \Upsilon, v, \gamma \rangle$ with $\beta = 0$. Denote any **PUNE** in this game by $(\tau^L(0), \tau^R(0)) = \left((t_L(0), \alpha_L(0)), (t_R(0), \alpha_R(0))\right)$.

Lemma 3: *Let* **A1***,* **A2***, and* **A3** *hold. Let* γ *be sufficiently small. Then, there is a unique* ***PUNE*** *such that* $t_L(0) \leq t_R(0)$*. This* ***PUNE*** *is:* $(\tau^L(0), \tau^R(0)) = \left((1, \alpha^*), (1, \alpha^*)\right)$.

Theorem 1: *Let* **A1***,* **A2***, and* **A3** *hold. Let γ be sufficiently small. Then, in the game with $\beta = 0$, there are only two* ***PUNEs****:* $\left(\left(\bar{t}_L, \overline{\alpha}_L\right), \left(\bar{t}_R, \overline{\alpha}_R\right)\right)$ *and* $\left(\left(\bar{t}_L, \overline{\alpha}_L\right), \left(\bar{t}_L, \overline{\alpha}_L\right)\right)$*, where* $\left(\bar{t}_L, \overline{\alpha}_L\right) = (1, \alpha^*)$ *and* $\left(\bar{t}_R, \overline{\alpha}_R\right) = (t^*, 1)$*. Moreover, in the first* ***PUNE****, we have* $\pi(\left(\bar{t}_L, \overline{\alpha}_L\right), \left(\bar{t}_R, \overline{\alpha}_R\right)) = 1$.

Remark 1: In this political model in which $\beta = 0$, there is no Wittman equilibrium. This is because any Wittman equilibrium is a **PUNE**, but in the game referred to in Theorem 1, there are only two types of **PUNE**s, neither of which is a Wittman equilibrium. In contrast, we can show there is one unique Downsian equilibrium in this game, which coincides with the second **PUNE** $\left(\left(\bar{t}_L, \overline{\alpha}_L\right), \left(\bar{t}_L, \overline{\alpha}_L\right)\right)$ of Theorem 1.

Theorem 1 implies that, in regard to the political issue of assigning the budget between income redistribution and defense expenditure, L's ideal policy is always implemented when voters' preferences over military expenditure are solely economically motivated. There is no room for compromise on shifting to a more right-wing policy; R may win with a 50 percent probability in the second political equilibrium, but in this case, it should implement L's ideal policy. This is exclusively implausible because, in real political competitions, R can often win and implement a more right-wing policy. In addition, typically, there are frequent changes of office in a two-party political system. A plausible model should be able to explain these frequent changes in office that occur in actual two-party systems.

4 Political Equilibrium in a Society with Ideological Concerns

Let us consider a society that has ideological concerns: $\beta > 0$. When $\beta > 0$, every voter's utility function is represented by (4). In this case, the relevant information on the voter's type is not only his or her income level, w, but also his or her ideological position, a. Thus, both the voter space and the policy space are two dimensional.

Given a pair of policies $(\tau^L, \tau^R) = ((t_L, \alpha_L), (t_R, \alpha_R))$, $\triangle^{(w,a)}((t_L, \alpha_L), (t_R, \alpha_R))$ is defined by (6) as in the case when $\beta = 0$. By definition, $\triangle^{(w,a)}((t_L, \alpha_L), (t_R, \alpha_R)) > 0$ implies that the voter (w, a) prefers L to R when L offers (t_L, α_L) and R offers (t_R, α_R).

Let $\triangle t \equiv t_R - t_L$, $\triangle \alpha t \equiv \alpha_R t_R - \alpha_L t_L$, and let $\overline{\alpha t} \equiv \frac{1}{2}(\alpha_R t_R + \alpha_L t_L)$. Then, $\triangle^{(w,a)}((t_L, \alpha_L), (t_R, \alpha_R)) > 0$ holds if and only if:

$$a < \overline{\alpha t} + \frac{(1-\beta)[\triangle t(w-\mu) + \triangle \alpha t \cdot \{\sigma(\alpha_L t_L \mu) - \sigma(\alpha_R t_R \mu)\}]}{\beta \cdot \triangle \alpha t} \text{ if } \triangle \alpha t > 0, \quad (9a)$$

$$a > \overline{\alpha t} + \frac{(1-\beta)[\triangle t(w-\mu) + \triangle \alpha t \cdot \{\sigma(\alpha_L t_L \mu) - \sigma(\alpha_R t_R \mu)\}]}{\beta \cdot \triangle \alpha t} \text{ if } \triangle \alpha t < 0, \quad (9b)$$

$$w > \mu \text{ if } \triangle \alpha t = 0 \text{ and } \triangle t > 0, \quad (9c)$$

$$w < \mu \text{ if } \triangle \alpha t = 0 \text{ and } \triangle t < 0. \quad (9d)$$

Note that $\Omega(\tau^L, \tau^R) = \{(w,a) \in H \mid \triangle^{(w,a)}(\tau^L, \tau^R) \geq 0\}$. Thus, the fraction of voters who prefer $\tau^L = (t_L, \alpha_L)$ to $\tau^R = (t_R, \alpha_R)$ is defined as:

$$\mathbb{F}(\Omega(\tau^L, \tau^R)) = \int_W \int_0^{\overline{\alpha t} + \Phi_w(\tau^L, \tau^R; \beta)} g(w) r(a; w) \, \mathrm{d}a \mathrm{d}w \text{ if } \triangle \alpha t > 0, \quad (10a)$$

$$\mathbb{F}(\Omega(\tau^L, \tau^R)) = \int_W \int_{\overline{\alpha t} + \Phi_w(\tau^L, \tau^R; \beta)}^1 g(w) r(a; w) \, \mathrm{d}a \mathrm{d}w \text{ if } \triangle \alpha t < 0, \quad (10b)$$

$$\mathbb{F}(\Omega(\tau^L, \tau^R)) = \int_\mu^{\overline{w}} g(w) \mathrm{d}w \text{ if } \triangle \alpha t = 0 \text{ and } \triangle t > 0, \quad (10c)$$

$$\mathbb{F}(\Omega(\tau^L, \tau^R)) = \int_{\underline{w}}^{\mu} g(w) \mathrm{d}w \text{ if } \triangle \alpha t = 0 \text{ and } \triangle t < 0, \quad (10d)$$

where $\Phi_w(\tau^L, \tau^R; \beta) \equiv \frac{(1-\beta)[\triangle t(w-\mu) + \triangle \alpha t \cdot \{\sigma(\alpha_L t_L \mu) - \sigma(\alpha_R t_R \mu)\}]}{\beta \cdot \triangle \alpha t}$,

and $g(w)$ is the density of W, and $r(a; w)$ is the density of ideological positions $a \in [0, 1]$ for a population of citizens with income level w.

We are ready to analyze the **PUNE**s in the political competition game $\langle H, \mathbb{F}, L, R, \Upsilon, v, \gamma \rangle$ with $\beta > 0$. Denote any **PUNE** in the game $\langle H, \mathbb{F}, L, R, \Upsilon, v, \gamma \rangle$ with $\beta > 0$ by $(\tau^L(\beta), \tau^R(\beta)) = ((t_L(\beta), \alpha_L(\beta)), (t_R(\beta), \alpha_R(\beta)))$. Denote a pair of *ideal polices* of the Militants of both parties in $\langle H, \mathbb{F}, L, R, \Upsilon, v, \gamma \rangle$ with $\beta > 0$ by $(\overline{\tau}^L(\beta), \overline{\tau}^R(\beta)) = ((\overline{t}_L(\beta), \overline{\alpha}_L(\beta)), (\overline{t}_R(\beta), \overline{\alpha}_R(\beta)))$. Note that $(\overline{\tau}^L(\beta), \overline{\tau}^R(\beta))$ is also a **PUNE** in the game $\langle H, \mathbb{F}, L, R, \Upsilon, v, \gamma \rangle$ with $\beta > 0$. We examine the existence of the following refinement of **PUNE**:

Definition 2: *Given $\langle H, \mathbb{F}, \Upsilon, v, \gamma \rangle$ with $\beta \in [0, 1]$, and the two parties* Left (L) and Right (R), *a pair of policies $(\tau^L(\beta), \tau^R(\beta)) \in \Upsilon \times \Upsilon$ constitutes a* non-trivial and non-pure **PUNE** *if $(\tau^L(\beta), \tau^R(\beta))$ is a* **PUNE** *such that* $0 < \pi(\tau^L(\beta), \tau^R(\beta)) < 1$, $\tau^L(\beta) \neq \overline{\tau}^L(\beta)$, and $\tau^R(\beta) \neq \overline{\tau}^R(\beta)$.

First, we will consider the polar case in which $\beta = 1$.

Theorem 2: *Let* **A1** *hold. Let* $\beta = 1$. *Then, for each* $\gamma > 0$, *there is some positive number* $\epsilon(\gamma) > 0$ *such that any pair of policies* $(\tau^L, \tau^R) = ((t_L, \alpha_L), (t_R, \alpha_R))$ *with* $\frac{1}{2}(\alpha_R t_R + \alpha_L t_L) \in (a^m - \epsilon(\gamma), a^m + \epsilon(\gamma))$ *and* $a_R > \alpha_R t_R \geq \alpha_L t_L > a_L$ *constitutes a* non-trivial and non-pure **PUNE** $(\tau^L(1), \tau^R(1))$ *in the game* $\langle H, \mathbb{F}, L, R, \Upsilon, v, \gamma \rangle$ *with* $\beta = 1$. *Moreover, in this game, there is no other type of* non-trivial and non-pure **PUNE**.

Theorem 2 shows there are many non-trivial **PUNE**s in the polar case in which $\beta = 1$. The set of non-trivial **PUNE**s when $\beta = 1$ is illustrated by Figure 1 as follows:

Figure 1 around here.

The result of the game in which $\beta = 1$ seems unreasonable. This is because the set of non-trivial and non-pure **PUNE**s includes the one in which $\alpha_R t_R \geq \alpha_L t_L$ with $(1 - \alpha_R) \cdot t_R > (1 - \alpha_L) \cdot t_L$, which implies that R proposes a stronger welfare policy than L. Because actual politics is consistent with the opposite behavior, the polar case in which $\beta = 1$ is inappropriate for our subject; thus, we shift our attention to the case in which $0 < \beta < 1$.

We consider the following population:

Assumption 4 (A4): *The mean income of the cohort of voters with the median ideological view* a^m *is less than mean income,* μ, *of the population:*

$$\int_W g(w) r(a^m; w) \cdot (w - \mu) \, \mathrm{d}w < 0.$$

Combined with **A3**, **A4** seems reasonable. Then:

Theorem 3: *Let* **A1**, **A2**, *and* **A3** *hold. Then, the strategy profile* $(\tau^L(\beta), \tau^R(\beta)) = ((t_L(\beta), \alpha_L(\beta)), (t_R(\beta), \alpha_R(\beta)))$ *constitutes a* non-trivial and non-pure **PUNE** *in the game* $\langle H, \mathbb{F}, L, R, \Upsilon, v, \gamma \rangle$ *with any* $\beta > 0$ *sufficiently close to unity, if and only if* $((t_L(\beta), \alpha_L(\beta)), (t_R(\beta), \alpha_R(\beta))) = ((1, \alpha_L^*), (t_R^*, 1))$ *with* $\frac{1}{2}(t_R^* + \alpha_L^*) \in (a^m - \epsilon(\gamma), a^m + \epsilon(\gamma))$ *and* $a_R > t_R^* \geq \alpha_L^* > a_L$ *for* $\epsilon(\gamma) > 0$ *given in Theorem 2. Moreover, for any* $\beta > 0$ *sufficiently close to unity, if* **A4** *holds, then there is no other type of* non-trivial and non-pure **PUNE** *for this* $\epsilon(\gamma) > 0$.[5]

[5]Note that the existence of a non-trivial and non-pure **PUNE** is guaranteed in this model, relying on Yoshihara (2008). Thus, the main contribution of this theorem is to characterize the non-trivial and non-pure **PUNE**.

The non-emptiness and the characterization of the set of non-trivial and non-pure **PUNE**s are discussed in Theorem 3. These **PUNE**s seem to describe plausible political competition in the real world; under such a **PUNE**, L proposes a stronger social welfare policy for the poor and a weaker military force, whereas R proposes a stronger military force and a lower tax rate for the rich, and neither party can win the vote with certainty, which implies that frequent changes in office between these two parties are possible. These features are consistent with the political competition that has been typical of advanced democratic countries since World War II. Theorems 1 and 3 suggest that the existence of ideological views regarding military policies provides a game-theoretic reasoning to explain the stylized party behavior in such countries.

5 The Case of Endogenous Party Formations

In this section, we consider the case in which the members of both parties are *endogenously formed*. In this case, each of the two parties represents a coalition of voter types. Thus, there is a partition of the set of voter types:

$$H = L \cup R,\ L \cap R = \varnothing.$$

Each party represents its members, in the sense that each party's preference represents the average preferences of its members. That is, we define the parties' utility functions on Υ as:

$$\begin{aligned} v_L(\tau) &= (1-\beta)\left[(1-t)\,w_L + (1-\alpha)\,t\mu + \sigma(\alpha t\mu)\right] - \frac{\beta}{2}\left|\alpha t - a_L\right|^2; \\ \text{and } v_R(\tau) &= (1-\beta)\left[(1-t)\,w_R + (1-\alpha)\,t\mu + \sigma(\alpha t\mu)\right] - \frac{\beta}{2}\left|\alpha t - a_R\right|^2, \\ \text{where } w_L &\equiv \int_{h\in L} w_h \mathrm{d}\mathbb{F}(h),\ a_L \equiv \int_{h\in L} a_h \mathrm{d}\mathbb{F}(h), \\ w_R &\equiv \int_{h\in R} w_h \mathrm{d}\mathbb{F}(h), \text{ and } a_R \equiv \int_{h\in R} a_h \mathrm{d}\mathbb{F}(h). \end{aligned}$$

The equilibrium of this multidimensional political competition game is defined as a *party-unanimity Nash equilibrium with endogenous parties* (**PUNEEP**), which was introduced by Roemer (2001: Chapter 13; 2005).

Definition 3: *A partition of voter types L, R, and a pair of policies* $(\tau^L, \tau^R) \in \Upsilon \times \Upsilon$ *with* $\tau^i = (t_i, \alpha_i)$ $(i = L, R)$ *constitutes a* party-unanimity Nash equilibrium with endogenous parties (**PUNEEP**) *if*:
(1) $H = L \cup R$ *and* $L \cap R = \varnothing$;
(2) (τ^L, τ^R) *satisfies* Definition 1(a) and (b);
(3) *for all* $h \in L$, $v(\tau^L; h) \geq v(\tau^R; h)$ *and for all* $h \in R$, $v(\tau^L; h) < v(\tau^R; h)$.

Condition (3) of Definition 3 states that party membership is stable in the sense that every party member prefers his or her party's policy to that of the opposition party. From this condition, the coalition of those who vote for a particular party and the coalition that this party represents are identical. Such a condition was used by Baron (1993) in the context of endogenous party formation, and was treated more generally by Caplin and Nalebuff (1997).

First, consider a non-ideological society in which $\beta = 0$ with endogenous party formation. Given the party membership of R, let $t\left(\frac{w_R}{\mu}\right)$ be the tax rate satisfying $\frac{\partial \sigma\left(t\left(\frac{w_R}{\mu}\right)\mu\right)}{\partial t \mu} = \frac{w_R}{\mu}$ for $w_R = \int_{h \in R} w_h \mathrm{d}\mathbb{F}(h)$. Then:

Theorem 4: *Let* **A1**, **A2**, *and* **A3** *hold. Let* γ *be sufficiently small. Then, in the game with* $\beta = 0$, *there exists* $h^* \in H$ *with* $w_{h^*} > \mu$ *such that* $L = \{h \in H \mid w_h \leq w_{h^*}\}$ *and* $R = \{h \in H \mid w_h > w_{h^*}\}$ *associated with two* ***PUNEEPs***: $\left(\left(\bar{t}_L, \overline{\alpha}_L\right), \left(\bar{t}_R, \overline{\alpha}_R\right)\right)$ *and* $\left(\left(\bar{t}_L, \overline{\alpha}_L\right), \left(\bar{t}_L, \overline{\alpha}_L\right)\right)$, *where* $\left(\bar{t}_L, \overline{\alpha}_L\right) = (1, \alpha^*)$ *and* $\left(\bar{t}_R, \overline{\alpha}_R\right) = \left(t\left(\frac{w_R}{\mu}\right), 1\right)$. *Moreover, in the first* ***PUNEEP****, we have* $\pi(\left(\bar{t}_L, \overline{\alpha}_L\right), \left(\bar{t}_R, \overline{\alpha}_R\right)) = 1$.

Second, consider the case of a society that has ideological concerns and in which there is endogenous party formation. We examine the existence of a *non-trivial and non-pure* **PUNEEP**.

Definition 4: *A partition of voter types L, R, and a pair of policies* $(\tau^L(\beta), \tau^R(\beta)) \in \Upsilon \times \Upsilon$ *with* $\tau^i = (t_i, \alpha_i)$ $(i = L, R)$ *constitutes a* non-trivial and non-pure **PUNEEP** *in the game* $\langle H, \mathbb{F}, L, R, \Upsilon, v, \gamma \rangle$ *with* β *if it is a* **PUNEEP** *such that* $\tau^L(\beta) \neq \tau^R(\beta)$, $0 < \pi(\tau^L(\beta), \tau^R(\beta)) < 1$, $\tau^L(\beta) \neq \overline{\tau}^L(\beta)$, *and* $\tau^R(\beta) \neq \overline{\tau}^R(\beta)$.

First, consider the case where $\beta = 1$.

Theorem 5: *Let* **A1** *hold. Let* $\beta = 1$. *Then, for each* $\gamma > 0$, *there is a positive number* $\epsilon(\gamma) > 0$ *such that for any* $\overline{a} \in (a^m - \epsilon(\gamma), a^m + \epsilon(\gamma))$, *any pair of policies* $(\tau^L, \tau^R) = ((t_L, \alpha_L), (t_R, \alpha_R))$ *with* $\frac{1}{2}(\alpha_R t_R + \alpha_L t_L) = \overline{a}$ *and* $\int_{h \in L} a_h \mathrm{d}\mathbb{F}(h) < \alpha_L t_L < \alpha_R t_R < \int_{h \in R} a_h \mathrm{d}\mathbb{F}(h)$, *where* $L = \{h \in H \mid a_h \leq \overline{a}\}$ *and* $R = \{h \in H \mid a_h > \overline{a}\}$, *constitutes a* non-trivial and non-pure **PUNEEP** $(\tau^L(1), \tau^R(1))$ *in the game* $\langle H, \mathbb{F}, L, R, \Upsilon, v, \gamma \rangle$ *with* $\beta = 1$. *Moreover, in this game, there is no other type of* non-trivial and non-pure **PUNEEP**.

Second, suppose that the case where $\beta > 0$ and close to unity. Then:

Theorem 6: *Let* **A1**, **A2**, *and* **A3** *hold. Then, in the game* $\langle H, \mathbb{F}, L, R, \Upsilon, v, \gamma \rangle$ *with any* $\beta > 0$ sufficiently close to unity, *there exists a* non-trivial and non-pure **PUNEEP** $(\tau^L(\beta), \tau^R(\beta)) = ((1, \alpha_L^*), (t_R^*, 1))$ *with* $\frac{1}{2}(t_R^* + \alpha_L^*) \in (a^m - \epsilon(\gamma), a^m + \epsilon(\gamma))$ *and* $\int_{h \in R} a_h \mathrm{d}\mathbb{F}(h) > t_R^* \geq \alpha_L^* > \int_{h \in L} a_h \mathrm{d}\mathbb{F}(h)$, *where* $L = \Omega(\tau^L(\beta), \tau^R(\beta))$ *and* $R = H \backslash \Omega(\tau^L(\beta), \tau^R(\beta))$, *for* $\epsilon(\gamma) > 0$ *given in Theorem 5. Moreover, for any* $\beta > 0$ *sufficiently close to unity, if* **A4** *holds and* $\int_0^\mu \int_0^1 a g(w) r(a; w) \mathrm{d}a \mathrm{d}w \leq a^m$, *then there is no other type of* non-trivial and non-pure **PUNEEP** *for this* $\epsilon(\gamma) > 0$.

Remark 2: In **Theorem 6**, if $\int_0^\mu \int_0^1 a g(w) r(a; w) \mathrm{d}a \mathrm{d}w \leq a^m$ does not hold, then $(\tau^{L\prime}(\beta), \tau^{R\prime}(\beta)) = ((\alpha_L^*, 1), (1, t_R^*))$ constitutes another type of non-trivial and non-pure **PUNEEP**. Except for these two types, there is no other type.

6 Concluding Remarks

In this paper, we have discussed the existence and characterizations of party-unanimity Nash equilibria (**PUNEs**) in two-dimensional political games for military and social welfare policies. We have shown that the existence of non-economically motivated, ideological views on military policies is crucial for providing a game-theoretic reasoning for stylized party behavior. In fact, Theorems 3 and 6 in this paper indicate that if the median ideological view a^m is in the neighborhood of the threshold level of military expenditure λ^*, then there is a **PUNE** in which the right-wing party proposes an imperialistic military policy, $t_R > \lambda^*$, regardless of whether party formation is exogenous or endogenous. Moreover, if a^m exceeds λ^*, there is a **PUNE** in which even the left-wing party proposes an imperialistic military policy, $\alpha_L > \lambda^*$. There is historical evidence to support such behavior: for example, Germany's Social Democratic Party in the lead-up to World War I and the Japanese Social Democratic Party in the lead-up to World War II. However, the political games with purely economically motivated preferences described in Section 3 cannot explain such phenomena as equilibrium behavior.

Note that Roemer (1998) also argued that the existence of non-economically motivated preferences is a crucial factor in explaining real politics in the US. Our model seems similar to that of Roemer (1998). However, because there are significant differences, our results also have different implications.[6]

7 Appendix

7.1 Proofs for Section 3

Proof of Lemma 1: To characterize L's ideal policy, we have:

$$\frac{\partial v_L\left(\bar{t}_L, \overline{\alpha}_L\right)}{\partial \alpha} = -\bar{t}_L\mu + \frac{\partial \sigma\left(\overline{\alpha}_L\bar{t}_L\mu\right)}{\partial \alpha t\mu}\bar{t}_L\mu. \quad (11)$$

From (11), we know that $\frac{\partial v_L\left(\bar{t}_L,\overline{\alpha}_L\right)}{\partial \alpha} \lesseqqgtr 0 \Leftrightarrow \frac{\partial \sigma\left(\overline{\alpha}_L\bar{t}_L\mu\right)}{\partial \alpha t\mu} \lesseqqgtr 1$. Note that if $\frac{\partial v_L\left(\bar{t}_L,\overline{\alpha}_L\right)}{\partial \alpha} > 0$, then $\overline{\alpha}_L = 1$, and if $\frac{\partial v_L\left(\bar{t}_L,\overline{\alpha}_L\right)}{\partial \alpha} < 0$, then $\overline{\alpha}_L = 0$. In addition, we have:

$$\frac{\partial v_L\left(\bar{t}_L, \overline{\alpha}_L\right)}{\partial t} = -w_L + \left(1-\overline{\alpha}_L\right)\mu + \frac{\partial \sigma\left(\overline{\alpha}_L\bar{t}_L\mu\right)}{\partial \alpha t\mu}\overline{\alpha}_L\mu. \quad (12)$$

Suppose that $\frac{\partial v_L\left(\bar{t}_L,\overline{\alpha}_L\right)}{\partial \alpha} > 0$. Then, (12) implies $\frac{\partial v_L\left(\bar{t}_L,\overline{\alpha}_L\right)}{\partial t} = -w_L + \frac{\partial \sigma\left(\overline{\alpha}_L\bar{t}_L\mu\right)}{\partial \alpha t\mu}\mu > 0$, because $w_L < \mu$ and $\frac{\partial \sigma\left(\overline{\alpha}_L\bar{t}_L\mu\right)}{\partial \alpha t\mu} > 1$. Thus, $\left(\bar{t}_L, \overline{\alpha}_L\right) = (1,1)$ should hold. However, this implies $\frac{\partial \sigma\left(\overline{\alpha}_L\bar{t}_L\mu\right)}{\partial \alpha t\mu} = \frac{\partial \sigma(\mu)}{\partial \alpha t\mu} > 1$, which is a contradiction, because of **A2**. Thus, $\frac{\partial v_L\left(\bar{t}_L,\overline{\alpha}_L\right)}{\partial \alpha} > 0$ is impossible.

Suppose that $\frac{\partial v_L\left(\bar{t}_L,\overline{\alpha}_L\right)}{\partial \alpha} < 0$. Then, (12) implies $\frac{\partial v_L\left(\bar{t}_L,\overline{\alpha}_L\right)}{\partial t} = -w_L + \mu > 0$, so that $\left(\bar{t}_L, \overline{\alpha}_L\right) = (1,0)$ should hold. However, this implies that $\frac{\partial \sigma\left(\overline{\alpha}_L\bar{t}_L\mu\right)}{\partial \alpha t\mu} =$

[6] Roemer (1998) argued that the poor do not expropriate the rich through democratic policy-making, because a poor person may vote for the right-wing party on religious grounds. Roemer (1998) obtained this result by comparing a unidimensional space of redistribution policies with a two-dimensional space of redistribution and religious policies.

$\frac{\partial\sigma(0)}{\partial\alpha t\mu} > 1$, which is a contradiction, because of $\frac{\partial v_L(\bar{t}_L,\overline{\alpha}_L)}{\partial\alpha} < 0 \Leftrightarrow \frac{\partial\sigma(\overline{\alpha}_L\bar{t}_L\mu)}{\partial\alpha t\mu} < 1$. Thus, $\frac{\partial v_L(\bar{t}_L,\overline{\alpha}_L)}{\partial\alpha} < 0$ is impossible.

Let $\frac{\partial v_L(\bar{t}_L,\overline{\alpha}_L)}{\partial\alpha} = 0$. Then, (12) implies $\frac{\partial v_L(\bar{t}_L,\overline{\alpha}_L)}{\partial t} = -w_L + \mu > 0$, so that $\bar{t}_L = 1$. Then, $\overline{\alpha}_L$ should meet the condition $\frac{\partial\sigma(\overline{\alpha}_L\mu)}{\partial\alpha\mu} = 1$, which implies $\overline{\alpha}_L = \lambda^*$. Thus, $(\bar{t}_L, \overline{\alpha}_L) = (1, \alpha^*)$, where $\alpha^* = \lambda^*$. ■

Proof of Lemma 2: Because R's ideal policy is characterized by $\frac{\partial v_R(\bar{t}_R,\overline{\alpha}_R)}{\partial\alpha} = -\bar{t}_R\mu + \frac{\partial\sigma(\overline{\alpha}_R\bar{t}_R\mu)}{\partial\alpha t\mu}\bar{t}_R\mu$, we obtain $\frac{\partial v_R(\bar{t}_R,\overline{\alpha}_R)}{\partial\alpha} \gtreqless 0$ if and only if $\frac{\partial\sigma(\overline{\alpha}_R\bar{t}_R\mu)}{\partial\alpha t\mu} \gtreqless 1$. Let us consider $\frac{\partial\sigma(\overline{\alpha}_R\bar{t}_R\mu)}{\partial\alpha t\mu} = 1$. Then,

$$\frac{\partial v_R\left(\bar{t}_R,\overline{\alpha}_R\right)}{\partial t} = -w_R + (1-\overline{\alpha}_R)\,\mu + \frac{\partial\sigma\left(\overline{\alpha}_R\bar{t}_R\mu\right)}{\partial\alpha t\mu}\overline{\alpha}_R\mu = -w_R + \mu < 0. \quad (13)$$

Thus, $\bar{t}_R = 0$, so that $\overline{\alpha}_R\bar{t}_R\mu = 0$, which implies $\frac{\partial\sigma(\overline{\alpha}_R\bar{t}_R\mu)}{\partial\alpha t\mu} > 1$, which is a contradiction. Let us consider $\frac{\partial\sigma(\overline{\alpha}_R\bar{t}_R\mu)}{\partial\alpha t\mu} < 1$. Then, because $\frac{\partial v_R(\bar{t}_R,\overline{\alpha}_R)}{\partial\alpha} < 0$, it follows that $\overline{\alpha}_R = 0$, which implies $\frac{\partial\sigma(\overline{\alpha}_R\bar{t}_R\mu)}{\partial\alpha t\mu} > 1$, which is also a contradiction.

Let us consider $\frac{\partial\sigma(\overline{\alpha}_R\bar{t}_R\mu)}{\partial\alpha t\mu} > 1$. Then, $\overline{\alpha}_R = 1$, and $\frac{\partial v_R(\bar{t}_R,\overline{\alpha}_R)}{\partial t} = -\left(\frac{w_R}{\mu} - 1\right)\mu + \left(\frac{\partial\sigma(\overline{\alpha}_R\bar{t}_R\mu)}{\partial\alpha t\mu} - 1\right)\mu$. If $\bar{t}_R = 0$, then $\frac{\partial v_R(\bar{t}_R,\overline{\alpha}_R)}{\partial\alpha} = 0 \Leftrightarrow \frac{\partial\sigma(\overline{\alpha}_R\bar{t}_R\mu)}{\partial\alpha t\mu} = 1$, a contradiction. If $\bar{t}_R = 1$, then $\overline{\alpha}_R = 1$ and **A2** together imply that $\frac{\partial v_R(\bar{t}_R,\overline{\alpha}_R)}{\partial t} < 0$, which contradicts to the fact that $\bar{t}_R = 1$. Thus, $0 < \bar{t}_R < 1$ and $\frac{\partial v_R(\bar{t}_R,\overline{\alpha}_R)}{\partial t} = 0$ hold. Note that $\frac{\partial v_R(\bar{t}_R,\overline{\alpha}_R)}{\partial t} = 0$ if and only if $\frac{w_R}{\mu} - 1 = \frac{\partial\sigma(\overline{\alpha}_R\bar{t}_R\mu)}{\partial\alpha t\mu} - 1$. Because $\lim_{\lambda\to 0}\frac{\partial\sigma(\lambda\mu)}{\partial\lambda\mu} = +\infty$ and $\frac{\partial\sigma(\lambda^*\mu)}{\partial\lambda^*\mu} = 1$ by **A2**, there exists t^* such that $\frac{\partial\sigma(t^*\mu)}{\partial t\mu} = \frac{w_R}{\mu}$. Thus, $(\bar{t}_R, \overline{\alpha}_R) = (t^*, 1)$ constitutes the solution. ■

Proof of Lemma 3: (Case 1): Let us take a pair of policies $(\tau^L, \tau^R) = ((t_L, \alpha_L), (t_R, \alpha_R))$ such that $t_L < t_R$. Then, (7b) and (8b) apply. Note that, in this case, $-\Theta(\tau^L, \tau^R) \geq 0$ if and only if

$$\Xi\left((t_L, \alpha_L), (t_R, \alpha_R)\right) \equiv (\alpha_R t_R - \alpha_L t_L)\cdot\mu + \{\sigma(\alpha_L t_L\mu) - \sigma(\alpha_R t_R\mu)\} \leq 0.$$

(Case 1-1): Let $\Xi\left(\left(t_L,\alpha_L\right),\left(t_R,\alpha_R\right)\right) \geq 0$. Then, consider a deviation from (t_L, α_L) to (t'_L, α'_L) with $\alpha'_L t'_L = \alpha_L t_L$, $\alpha'_L < \alpha_L$, and $t_R > t'_L > t_L$. Because $t_L < t_R$, we can always find such a strategy. Then, $\Xi\left(\left(t'_L,\alpha'_L\right),\left(t_R,\alpha_R\right)\right) = \Xi\left(\left(t_L,\alpha_L\right),\left(t_R,\alpha_R\right)\right) \geq 0$ and $t_R - t'_L < t_R - t_L$, which implies $0 \leq \Theta(\tau^L,\tau^R) \leq \Theta(\tau^{L\prime},\tau^R)$, so that $\mathbb{F}\left(\Omega(\tau^{L\prime},\tau^R)\right) \geq \mathbb{F}\left(\Omega(\tau^L,\tau^R)\right)$. Moreover, it follows that $v_L(\left(t'_L,\alpha'_L\right),\left(t_R,\alpha_R\right)) - v_L(\left(t_L,\alpha_L\right),\left(t_R,\alpha_R\right)) > 0$, because

$$\begin{aligned} & v_L(\left(t'_L,\alpha'_L\right),\left(t_R,\alpha_R\right)) - v_L(\left(t_L,\alpha_L\right),\left(t_R,\alpha_R\right)) \\ = & -\left(t'_L - t_L\right) w_L + \left(t'_L - t_L\right)\mu - \left(\alpha'_L t'_L - \alpha_L t_L\right)\mu + \sigma\left(\alpha'_L t'_L \mu\right) - \sigma\left(\alpha_L t_L \mu\right) \\ = & \left(t'_L - t_L\right)\left(\mu - w_L\right) > 0, \text{ because } w_L < \mu, \quad (14) \end{aligned}$$

which implies that this deviation can improve the L-militant's payoff without lowering the L-opportunist's payoff.

(Case 1-2): Let $\Xi\left(\left(t_L,\alpha_L\right),\left(t_R,\alpha_R\right)\right) < 0$. In this case, because $-\Theta(\tau^L,\tau^R) > 0$, we have $\mathbb{F}\left(\Omega(\tau^L,\tau^R)\right) < \frac{1}{2}$ by (8b) and **A3**, which implies $\pi(\tau^L,\tau^R) = 0$ for $\gamma > 0$ that is sufficiently small. Thus, L can improve its militant's payoff by deviating to $(t'_L, \alpha'_L) = (1, \alpha^*)$ without harming its opportunist.

In summary, any strategy profile $\left(\left(t_L,\alpha_L\right),\left(t_R,\alpha_R\right)\right)$ with $t_L < t_R$ cannot be a **PUNE**.

(Case2): Consider $t_L = t_R$. Then, $\triangle^{(w,a)}\left(\left(t_L,\alpha_L\right),\left(t_R,\alpha_R\right)\right) > 0$ holds if and only if $\Xi\left(\left(t_L,\alpha_L\right),\left(t_R,\alpha_R\right)\right) > 0$. L would like to maximize $\Xi\left(\left(t_L,\alpha_L\right),\left(t_R,\alpha_R\right)\right)$, whereas R would like to minimize $\Xi\left(\left(t_L,\alpha_L\right),\left(t_R,\alpha_R\right)\right)$.

(Case2-1): Let $0 < t_L = t_R < 1$. Then, L can choose $\alpha'_L t'_L$ as $\alpha'_L t'_L = \alpha_L t_L$, $\alpha'_L < \alpha_L$, and $t'_L > t_L$, so that $v_L(\left(t'_L,\alpha'_L\right),\left(t_R,\alpha_R\right)) - v_L(\left(t_L,\alpha_L\right),\left(t_R,\alpha_R\right)) > 0$. In contrast, since $t_R - t'_L < 0$, $-\Theta(\tau'^L,\tau^R) > 0$ holds, and (8a) is applied. Then, $\mathbb{F}\left(\Omega(\tau'^L,\tau^R)\right) > \frac{1}{2}$ by (8a) and **A3**, so that $\pi(\tau'^L,\tau^R) = 1$ for sufficiently small $\gamma > 0$. Thus, the L-militant is better off from this deviation without harming its opportunist.

(Case2-2): Let $t_L = t_R = 0$. Then, $\pi\left(\left(t_L,\alpha_L\right),\left(t_R,\alpha_R\right)\right) = \frac{1}{2}$. Note that when $t_L = t_R = 0$, any choice of α is irrelevant to both parties' militants and opportunists. Then, consider $\tau^{L\prime} = (t'_L, \alpha'_L)$ as $t'_L > 0$ and $\alpha'_L = 0$. Then, because $t'_L > t_R$, $\mathbb{F}\left(\Omega(\tau^{L\prime},\tau^R)\right)$ is determined by (8a) with $-\Theta(\tau^{L\prime},\tau^R) = 0$. Thus, by **A3**, $\mathbb{F}\left(\Omega(\tau^{L\prime},\tau^R)\right) > \frac{1}{2}$, which implies $\pi(\tau^{L\prime},\tau^R) \geq \frac{1}{2}$. Moreover, by (14), the L-militant's payoff is also improved by deviation from τ^L to $\tau^{L\prime}$. Thus, this case does not constitute a **PUNE** either.

(Case2-3): Let $t_L = t_R = 1$. Note that if the tax rate is fixed at unity, the optimal ratio of defense expenditure is $\alpha^* > 0$ for every individual, where the last strict inequality implies that

$$\frac{\partial v\left(1,\alpha^{*};w,a\right)}{\partial\alpha}=-\mu+\frac{\partial\sigma\left(\alpha^{*}\mu\right)}{\partial\alpha\mu}\mu=0,$$

which in turn implies that $\frac{\partial\sigma(\alpha^{*}\mu)}{\partial\alpha\mu}=1$. Even for the R-militant, $\alpha_R=\alpha^{*}$ is the best strategy, given the tax rate of $t_L=t_R=1$. That is, given $t_L=t_R=1$, α^{*} is the unanimous single-peak of satisfaction in regard to the ratio of defense expenditure. Thus, for any $(\tau^L,\tau^R)=((1,\alpha_L),(1,\alpha_R))$, if $\alpha_R\neq\alpha^{*}\neq\alpha_L$ and $\pi(\tau^L,\tau^R)<(resp.\ \geq)\,1$, then L ($resp.$ R) can improve its militant's and opportunist's payoffs by moving to $(1,\alpha^{*})$. If $\alpha_R=\alpha^{*}\neq\alpha_L$, then $\pi(\tau^L,\tau^R)=0$. Then, L can improve its militant's and opportunist's payoffs by moving to $(1,\alpha^{*})$. If $\alpha_R\neq\alpha^{*}=\alpha_L$, then $\pi(\tau^L,\tau^R)=1$, and R can become better off by moving to $(1,\alpha^{*})$.

In summary, with the exception of $((1,\alpha^{*}),(1,\alpha^{*}))$, no pair of policies $((t_L,\alpha_L),(t_R,\alpha_R))$ can constitute a **PUNE** whenever $t_L\leq t_R$.

(Case 3): Consider $t_L=t_R=1$ and $\alpha_L=\alpha_R=\alpha^{*}$. Then, L offers its ideal policy. For R, any deviation from α_R to α'_R while retaining $t'_R=1$ makes its opportunist worse off. This is because $\Xi\left((1,\alpha^{*}),(1,\alpha'_R)\right)>0$. Now, we show that, given $\alpha_L t_L=\alpha^{*}$, any deviation from $\alpha_R t_R=\alpha^{*}$ to $\alpha'_R t'_R\neq\alpha^{*}$ leads to $\Xi\left((t_L,\alpha_L),(t'_R,\alpha'_R)\right)>0$.

If $\alpha'_R t'_R<\alpha^{*}$, then $(\alpha'_R t'_R-\alpha_L t_L)\cdot\mu<0$ and $\{\sigma\left(\alpha_L t_L\mu\right)-\sigma\left(\alpha'_R t'_R\mu\right)\}>0$. By the strict concavity of σ, it follows that $\sigma\left(\alpha_L t_L\mu\right)-\sigma\left(\alpha'_R t'_R\mu\right)>|(\alpha'_R t'_R-\alpha_L t_L)\cdot\mu|$, which implies $\Xi\left((t_L,\alpha_L),(t'_R,\alpha'_R)\right)>0$. If $\alpha'_R t'_R>\alpha^{*}$, then $(\alpha'_R t'_R-\alpha_L t_L)\cdot\mu>0$ and $\{\sigma\left(\alpha_L t_L\mu\right)-\sigma\left(\alpha'_R t'_R\mu\right)\}<0$. Because $\frac{\partial\sigma\left(\alpha'_R t'_R\mu\right)}{\partial\alpha_R t_R\mu}<\frac{\partial\sigma(\alpha^{*}\mu)}{\partial\alpha\mu}=1$ by **A2**, it follows that $(\alpha'_R t'_R-\alpha_L t_L)\cdot\mu>|\sigma\left(\alpha_L t_L\mu\right)-\sigma\left(\alpha'_R t'_R\mu\right)|$, which implies $\Xi\left((t_L,\alpha_L),(t'_R,\alpha'_R)\right)>0$.

Let $t'_R<t_R=t_L=1$. Then, $\mathbb{F}\left(\Omega(\tau^L,\tau^{R\prime})\right)$ is defined by (8a), and $\Xi\left((t_L,\alpha_L),(t'_R,\alpha'_R)\right)>0$ implies $-\Theta(\tau^L,\tau^{R\prime})>0$, so that $\mathbb{F}\left(\Omega(\tau^L,\tau^{R\prime})\right)>\frac{1}{2}$ by **A3**. Thus, if $t_L=t_R=1$ and $\alpha_L=\alpha_R=\alpha^{*}$, any deviation from $\alpha_R t_R=\alpha^{*}$ to $\alpha'_R t'_R\neq\alpha^{*}$ makes the R-opportunist worse off. Next, consider a deviation strategy $\tau^{R\prime}=(t'_R,\alpha'_R)$, where $\alpha'_R t'_R=\alpha^{*}$, $\alpha'_R>\alpha^{*}$, and $t'_R<t_R$. Then, $\Xi\left((t_L,\alpha_L),(t'_R,\alpha'_R)\right)=0$, but because $t'_R<t_L$, $\mathbb{F}\left(\Omega(\tau^L,\tau^{R\prime})\right)$ is defined by (8a), and thus becomes

$$\mathbb{F}\left(\Omega(\tau^L,\tau^{R\prime})\right)=\int_{\underline{w}}^{\mu-\Theta(\tau^L,\tau^R)}g(w)\mathrm{d}w=\int_{\underline{w}}^{\mu}g(w)\mathrm{d}w>\frac{1}{2},$$

by **A3**. Hence, $\pi(\tau^L, \tau^{R\prime}) > \frac{1}{2} = \pi(\tau^L, \tau^R)$. Therefore, such a deviation makes the R-opportunist worse off, although it makes the R-militant better off. This implies that if $t_L = t_R = 1$ and $\alpha_L = \alpha_R = \alpha^*$, (t_R, α_R) is the best response for R to (t_L, α_L). In summary, $(\tau^L(0), \tau^R(0)) = ((1, \alpha^*), (1, \alpha^*))$ constitutes a **PUNE**. ■

Proof of Theorem 1: Lemma 3 shows that $\left(\left(\bar{t}_L, \overline{\alpha}_L\right), \left(\bar{t}_L, \overline{\alpha}_L\right)\right)$ is the unique **PUNE** whenever $t_L \leq t_R$. Thus, we need only check when $t_L > t_R$. Let us take such a pair of policies $(\tau^L, \tau^R) = ((t_L, \alpha_L), (t_R, \alpha_R))$. Then, $\triangle t < 0$, so $\mathbb{F}\left(\Omega(\tau^L, \tau^R)\right)$ is defined by (8a), and $-\Theta(\tau^L, \tau^R) \geq 0$ if and only if $\Xi\left((t_L, \alpha_L), (t_R, \alpha_R)\right) \geq 0$.

(Case 1): Consider $\Xi\left((t_L, \alpha_L), (t_R, \alpha_R)\right) > 0$. In this case, because $-\Theta(\tau^L, \tau^R) > 0$, we have $\mathbb{F}\left(\Omega(\tau^L, \tau^R)\right) > \frac{1}{2}$ by **A3**, which implies $\pi(\tau^L, \tau^R) = 1$ for a sufficiently small $\gamma > 0$. Then, if $\tau^R \neq \left(\bar{t}_R, \overline{\alpha}_R\right)$, R can improve its militant's payoff by moving to $\left(\bar{t}_R, \overline{\alpha}_R\right)$ without harming its opportunist. If $\tau^R = \left(\bar{t}_R, \overline{\alpha}_R\right)$ and $\tau^L \neq \left(\bar{t}_L, \overline{\alpha}_L\right)$, then L can improve its militant's payoff by moving to $\left(\bar{t}_L, \overline{\alpha}_L\right)$ without harming its opportunist. This is because $\pi(\left(\bar{t}_L, \overline{\alpha}_L\right), \left(\bar{t}_R, \overline{\alpha}_R\right)) = 1$ as we explain below.

(Case 2): Consider $\Xi\left((t_L, \alpha_L), (t_R, \alpha_R)\right) \leq 0$. Note that if $\alpha_R t_R \neq \alpha^* = \alpha_L t_L$, then $\Xi\left((t_L, \alpha_L), (t_R, \alpha_R)\right) > 0$.

(Case 2-1): If $\alpha_R t_R = \alpha^* \neq \alpha_L t_L$, then $\Xi\left((t_L, \alpha_L), (t_R, \alpha_R)\right) < 0$. In this case, L can improve its militant's and opportunist's payoffs by moving to $\tau^{L\prime} = \left(\bar{t}_L, \overline{\alpha}_L\right)$. This is because $\Xi\left(\left(\bar{t}_L, \overline{\alpha}_L\right), (t_R, \alpha_R)\right) = 0$, which implies $\mathbb{F}\left(\Omega(\tau^{L\prime}, \tau^R)\right) > \frac{1}{2}$ by **A3**, and thus $\pi(\tau^{L\prime}, \tau^R) = 1$ for a sufficiently small $\gamma > 0$. Moreover, $\left(\bar{t}_L, \overline{\alpha}_L\right)$ is the ideal policy of L.

(Case 2-2): If $\alpha_R t_R \neq \alpha^* \neq \alpha_L t_L$ with $\Xi\left((t_L, \alpha_L), (t_R, \alpha_R)\right) \leq 0$, then L can improve its militant's and opportunist's payoffs by moving to $\tau^{L\prime} = \left(\bar{t}_L, \overline{\alpha}_L\right)$. This is because $\Xi\left(\left(\bar{t}_L, \overline{\alpha}_L\right), (t_R, \alpha_R)\right) > 0$, and thus the scenario described in Case 2-1 applies.

(Case 2-3): If $\alpha_R t_R = \alpha^* = \alpha_L t_L$, then $\Xi\left((t_L, \alpha_L), (t_R, \alpha_R)\right) = 0$, which implies $\mathbb{F}\left(\Omega(\tau^L, \tau^R)\right) > \frac{1}{2}$ by **A3**, and thus $\pi(\tau^L, \tau^R) = 1$ for a sufficiently small $\gamma > 0$. Then, if $(t_L, \alpha_L) \neq \left(\bar{t}_L, \overline{\alpha}_L\right)$, then L can improve its militant's payoff by moving to $\left(\bar{t}_L, \overline{\alpha}_L\right)$ without harming its opportunist. If $(t_L, \alpha_L) = \left(\bar{t}_L, \overline{\alpha}_L\right)$, and $(t_R, \alpha_R) \neq \left(\bar{t}_R, \overline{\alpha}_R\right)$, then R can improve its militant's payoff by moving to $\left(\bar{t}_R, \overline{\alpha}_R\right)$ without harming its opportunist.

In summary, there exists no **PUNE** with $t_L > t_R$, except $(\tau^L, \tau^R) = \left(\left(\bar{t}_L, \overline{\alpha}_L\right), \left(\bar{t}_R, \overline{\alpha}_R\right)\right)$.

(Case 3): Note that $\left(\left(\bar{t}_L, \overline{\alpha}_L\right), \left(\bar{t}_R, \overline{\alpha}_R\right)\right)$ is a **PUNE**, and it is the unique

PUNE such that $\tau^L \neq \tau^R$ with $t_L > t_R$ in the game with $\beta = 0$ from the above argument. In this case,

$$\begin{aligned}\Xi\left(\left(t_L, \alpha_L\right), \left(t_R, \alpha_R\right)\right) &= \Xi\left(\left(1, \alpha^*\right), \left(t^*, 1\right)\right) \\ &= \left(t^* - \alpha^*\right) \cdot \mu + \left\{\sigma\left(\alpha^* \mu\right) - \sigma\left(t^* \mu\right)\right\}.\end{aligned}$$

Note that $t^* < \alpha^*$ by $\frac{\partial \sigma(t^* \mu)}{\partial t \mu} = \frac{w_R}{\mu} > 1$. Thus, $|t^* - \alpha^*| \cdot \mu < \sigma\left(\alpha^* \mu\right) - \sigma\left(t^* \mu\right)$, and so $\Xi\left(\left(\bar{t}_L, \overline{\alpha}_L\right), \left(\bar{t}_R, \overline{\alpha}_R\right)\right) > 0$, which implies that $\pi\left(\left(\bar{t}_L, \overline{\alpha}_L\right), \left(\bar{t}_R, \overline{\alpha}_R\right)\right) = 1$ for a sufficiently small $\gamma > 0$. By combining this result with Lemma 3, we can conclude that there are only two **PUNE**s, $\left(\left(\bar{t}_L, \overline{\alpha}_L\right), \left(\bar{t}_R, \overline{\alpha}_R\right)\right)$ and $\left(\left(\bar{t}_L, \overline{\alpha}_L\right), \left(\bar{t}_L, \overline{\alpha}_L\right)\right)$, in the game with $\beta = 0$ for a sufficiently small $\gamma > 0$. ■

7.2 Proofs for Section 4

Note that when $\triangle \alpha t > 0$, for the L-opportunist, we have:

$$\begin{aligned}&\frac{\partial \mathbb{F}\left(\Omega(\tau^L, \tau^R)\right)}{\partial t_L} \\ = \; &\int_W g(w) r\left(\overline{\alpha t} + \Phi_w(\tau^L, \tau^R; \beta); w\right) \cdot \left[\frac{\alpha_L}{2} + \frac{1-\beta}{\beta \triangle \alpha t} \Lambda^{t_L}\left(\tau^L, \tau^R; w\right)\right] \mathrm{d}w \qquad (15)\end{aligned}$$

where

$$\Lambda^{t_L}\left(\tau^L, \tau^R; w\right) \equiv \left[(\mu - w) + \triangle \alpha t \frac{\partial \sigma\left(\alpha_L t_L \mu\right)}{\partial \alpha t \mu} \alpha_L \mu + \frac{\alpha_L}{\triangle \alpha t} \triangle t \left(w - \mu\right)\right],$$

and

$$\begin{aligned}&\frac{\partial \mathbb{F}\left(\Omega(\tau^L, \tau^R)\right)}{\partial \alpha_L} \\ = \; &\int_W g(w) r\left(\overline{\alpha t} + \Phi_w(\tau^L, \tau^R; \beta); w\right) \cdot \left[\frac{t_L}{2} + \frac{1-\beta}{\beta \triangle \alpha t} \Lambda^{\alpha_L}\left(\tau^L, \tau^R; w\right)\right] \mathrm{d}w \qquad (16)\end{aligned}$$

where

$$\Lambda^{\alpha_L}\left(\tau^L, \tau^R; w\right) \equiv \triangle \alpha t \frac{\partial \sigma\left(\alpha_L t_L \mu\right)}{\partial \alpha t \mu} t_L \mu + \frac{t_L}{\triangle \alpha t} \triangle t \left(w - \mu\right).$$

In addition, for the L-militant, we have:

$$\frac{\partial v_L\left(\tau^L,\tau^R\right)}{\partial t_L} = (1-\beta)\left[-w_L+(1-\alpha_L)\mu+\frac{\partial\sigma\left(\alpha_L t_L\mu\right)}{\partial\alpha t\mu}\alpha_L\mu\right]-\beta\alpha_L\left|\alpha_L t_L-a_L\right|, \quad (17)$$

$$\frac{\partial v_L\left(\tau^L,\tau^R\right)}{\partial \alpha_L} = (1-\beta)\left[-t_L\mu+\frac{\partial\sigma\left(\alpha_L t_L\mu\right)}{\partial\alpha t\mu}t_L\mu\right]-\beta t_L\left|\alpha_L t_L-a_L\right|. \quad (18)$$

In the same way, when $\triangle\alpha t>0$, for the R-opportunist, we have:

$$\frac{\partial\mathbb{F}\left(\Omega(\tau^L,\tau^R)\right)}{\partial t_R} = \int_W g(w)r\left(\overline{\alpha t}+\Phi_w(\tau^L,\tau^R;\beta);w\right)\cdot\left[\frac{\alpha_R}{2}+\frac{1-\beta}{\beta\triangle\alpha t}\Lambda^{t_R}\left(\tau^L,\tau^R;w\right)\right]\mathrm{d}w \quad (19)$$

where

$$\Lambda^{t_R}\left(\tau^L,\tau^R;w\right)\equiv[(w-\mu)-\triangle\alpha t\frac{\partial\sigma\left(\alpha_R t_R\mu\right)}{\partial\alpha t\mu}\alpha_R\mu-\frac{\alpha_R}{\triangle\alpha t}\triangle t\left(w-\mu\right)],$$

and

$$\frac{\partial\mathbb{F}\left(\Omega(\tau^L,\tau^R)\right)}{\partial \alpha_R} = \int_W g(w)r\left(\overline{\alpha t}+\Phi_w(\tau^L,\tau^R;\beta);w\right)\cdot\left[\frac{t_R}{2}+\frac{1-\beta}{\beta\triangle\alpha t}\Lambda^{\alpha_R}\left(\tau^L,\tau^R;w\right)\right]\mathrm{d}w \quad (20)$$

where

$$\Lambda^{\alpha_R}\left(\tau^L,\tau^R;w\right)\equiv\left\{-\triangle\alpha t\frac{\partial\sigma\left(\alpha_R t_R\mu\right)}{\partial\alpha t\mu}t_R\mu-\frac{t_R}{\triangle\alpha t}\triangle t\left(w-\mu\right)\right\}.$$

In addition, for the R-militant, we have:

$$\frac{\partial v_R\left(\tau^L,\tau^R\right)}{\partial t_R} = (1-\beta)\left[-w_R+(1-\alpha_R)\mu+\frac{\partial\sigma\left(\alpha_R t_R\mu\right)}{\partial\alpha t\mu}\alpha_R\mu\right]-\beta\alpha_R\left|\alpha_R t_R-a_R\right|, \quad (21)$$

$$\frac{\partial v_R\left(\tau^L,\tau^R\right)}{\partial \alpha_R} = (1-\beta)\left[-t_R\mu+\frac{\partial\sigma\left(\alpha_R t_R\mu\right)}{\partial\alpha t\mu}t_R\mu\right]-\beta t_R\left|\alpha_R t_R-a_R\right|. \quad (22)$$

Proof of Theorem 2: Let us take a pair of policies $(\tau^L, \tau^R) = ((t_L, \alpha_L), (t_R, \alpha_R))$ with $a_R > \alpha_R t_R \geq \alpha_L t_L > a_L$. In this case, because $\beta = 1$, the fraction of voters who prefer L to R is:

$$\mathbb{F}\left(\Omega(\tau^L, \tau^R)\right) = \int_0^{\overline{\alpha t}} r(a)\,\mathrm{d}a, \quad (23)$$

where $\overline{\alpha t} = \frac{1}{2}(\alpha_R t_R + \alpha_L t_L)$. Given $\gamma \in \left(0, \frac{1}{2}\right)$, we can identify the maximal number $\epsilon(\gamma) > 0$ such that the following two conditions hold:

$$\int_0^{a^m - \epsilon(\gamma)} r(a)\,\mathrm{d}a \geq \frac{1}{2} - \gamma \text{ and } \int_0^{a^m + \epsilon(\gamma)} r(a)\,\mathrm{d}a \leq \frac{1}{2} + \gamma.$$

Note that $\epsilon(\gamma)$ is an increasing function.

Suppose that $\overline{\alpha t} \in (a^m - \epsilon(\gamma), a^m + \epsilon(\gamma))$. Then, we can show that this (τ^L, τ^R) is a non-trivial and non-pure **PUNE**. From (23), we know that $\mathbb{F}\left(\Omega(\tau^L, \tau^R)\right) \in \left(\frac{1}{2} - \gamma, \frac{1}{2} + \gamma\right)$, so that $0 < \pi(\tau^L, \tau^R) < 1$. Examining (23) and (4) with $\beta = 0$ reveals that the deviation from $\alpha_L t_L$ to $\alpha_L t_L + \varepsilon$ (resp. $-\varepsilon$) for any $\varepsilon > 0$ can improve the payoff of the L-opportunist (resp. the L-militant), but makes the L-militant (resp. the L-opportunist) worse off. Because the same argument applies to R, it follows that (τ^L, τ^R) is a non-trivial and non-pure **PUNE**.

Suppose that $\overline{\alpha t} \notin (a^m - \epsilon(\gamma), a^m + \epsilon(\gamma))$. Then, $\mathbb{F}\left(\Omega(\tau^L, \tau^R)\right) < \max\left\{\frac{1}{2} - \gamma, 0\right\}$ or $\mathbb{F}\left(\Omega(\tau^L, \tau^R)\right) > \min\left\{\frac{1}{2} + \gamma, 1\right\}$, which implies that $\pi(\tau^L, \tau^R) = 0$ or $\pi(\tau^L, \tau^R) = 1$. This does not correspond to any non-trivial non-pure **PUNE**.

Suppose that $a_R \leq \alpha_R t_R$. If $a_R = \alpha_R t_R$, then this is the ideal policy of R when $\beta = 1$. Thus, this policy does not correspond to any non-trivial non-pure **PUNE**. If $a_R < \alpha_R t_R$, deviation from $\alpha_R t_R$ to $\alpha_R t_R - \varepsilon$ can make both the militant and the opportunist better off. Similarly, $\alpha_L t_L \leq a_L$ does not correspond to any non-trivial and non-pure **PUNE**.

Suppose that $\alpha_R t_R < \alpha_L t_L$. Then, at least one of the two parties can use a deviation to make its militant and opportunist better off. Thus, this case does not correspond to any non-trivial and non-pure **PUNE**. ■

Proof of Theorem 3: From Theorem 2, it follows that any pair $((1, \alpha_L^*), (t_R^*, 1))$ with $\frac{1}{2}(t_R^* + \alpha_L^*) \in (a^m - \epsilon(\gamma), a^m + \epsilon(\gamma))$ and $a_R > t_R^* \geq \alpha_L^* > a_L$ for some $\epsilon(\gamma) > 0$ is a non-trivial and non-pure **PUNE** in the game $\langle H, \mathbb{F}, L, R, \Upsilon, v, \gamma \rangle$ with $\beta = 1$. Note that $\epsilon(\gamma)$ is determined as in Theorem 2. We show that this pair would be a non-trivial and non-pure **PUNE** in the game

$\langle H, \mathbb{F}, L, R, \Upsilon, v, \gamma \rangle$ with $\beta > 0$ close to unity. Note that for $\beta > 0$ close to unity, the optimal policy for the L-militant is $\left(\bar{t}_L(\beta), \overline{\alpha}_L(\beta)\right)$ with $\bar{t}_L(\beta) = 1$ and $\overline{\alpha}_L(\beta)$ being close to a_L, whereas the optimal policy for the R-militant is $\left(\bar{t}_R(\beta), \overline{\alpha}_R(\beta)\right)$ with $\bar{t}_R(\beta)$ being close to a_R and $\overline{\alpha}_R(\beta) = 1$. Thus, w. l. o. g, we can retain the condition $\bar{t}_R(\beta) > t_R^* \geq \alpha_L^* > \overline{\alpha}_L(\beta)$.

(1) $(1, \alpha_L^*)$ **is a best response for Left to** $(t_R^*, 1)$ **in the game** $\langle H, \mathbb{F}, L, R, \Upsilon, v, \gamma \rangle$ **with** $\beta > 0$ **close to unity.**

Suppose that there is $\beta > 0$, that is sufficiently close to unity, such that $(1, \alpha_L^*)$ is not a best response for L to $(t_R^*, 1)$ in the game $\langle H, \mathbb{F}, L, R, \Upsilon, v, \gamma \rangle$ - with this β. Then, there exists a vector $(-1, \delta(\beta))$ such that $\nabla^\beta v_L\left((1, \alpha_L^*), (t_R^*, 1)\right) \cdot (-1, \delta(\beta)) > 0$ and $\nabla^\beta \mathbb{F}_L\left((1, \alpha_L^*), (t_R^*, 1)\right) \cdot (-1, \delta(\beta)) \geq 0$, where

$$\nabla^\beta v_L\left((1, \alpha_L^*), (t_R^*, 1)\right) = \left(\frac{\partial v_L\left((1, \alpha_L^*), (t_R^*, 1); \beta\right)}{\partial t_L}, \frac{\partial v_L\left((1, \alpha_L^*), (t_R^*, 1); \beta\right)}{\partial \alpha_L}\right)$$

$$\text{and } \nabla^\beta \mathbb{F}_L\left((1, \alpha_L^*), (t_R^*, 1)\right) = \left(\frac{\partial \mathbb{F}\left(\Omega(\tau^L, \tau^R); \beta\right)}{\partial t_L}, \frac{\partial \mathbb{F}\left(\Omega(\tau^L, \tau^R); \beta\right)}{\partial \alpha_L}\right).$$

Because $\nabla^\beta v_L\left((1, \alpha_L^*), (t_R^*, 1)\right) \cdot (-1, \delta(\beta)) > 0$, we have to extrapolate from (17) and (18):

$$\delta(\beta) > \frac{(1-\beta)\left[-w_L + (1-\alpha_L^*)\mu + \frac{\partial \sigma\left(\alpha_L^* \mu\right)}{\partial \alpha t \mu} \alpha_L^* \mu\right] - \beta \alpha_L^* \left(\alpha_L^* - a_L\right)}{(1-\beta)\left[-\mu + \frac{\partial \sigma\left(\alpha_L^* \mu\right)}{\partial \alpha t \mu} \mu\right] - \beta\left(\alpha_L^* - a_L\right)}. \quad (24)$$

Given the choice of α_L^*, $\nabla^1 v_L\left((1, \alpha_L^*), (t_R^*, 1)\right) \cdot (-1, \delta(1)) > 0$, and thus $\delta(1) > \alpha_L^*$ holds. This implies that

$$\delta(\beta) > \alpha_L^* \text{ for } \beta \text{ sufficiently close to unity.}$$

Now, let us consider $\nabla^\beta \mathbb{F}_L\left((1, \alpha_L^*), (t_R^*, 1)\right) \cdot (-1, \alpha_L^*)$. Then,

$$\begin{aligned} &\nabla^\beta \mathbb{F}_L\left((1, \alpha_L^*), (t_R^*, 1)\right) \cdot (-1, \alpha_L^*) \\ = \; &-\frac{1-\beta}{\beta \triangle \alpha t} \int_W g(w) r\left(\overline{\alpha t} + \Phi_w(\tau^L, \tau^R; \beta); w\right) \cdot (w - \mu)\, \mathrm{d}w. \end{aligned}$$

We can rewrite $\delta(\beta) = \alpha_L^* + \varepsilon(\beta)$, where $\varepsilon(\beta) > 0$ for $\beta > 0$ which is sufficiently close to unity. Then:

$$\nabla^\beta \mathbb{F}_L\left((1, \alpha_L^*), (t_R^*, 1)\right) \cdot (-1, \delta(\beta))$$

$$
\begin{aligned}
&= \varepsilon(\beta)\frac{1}{2}\int_W g(w) r\left(\overline{\alpha t}+\Phi_w(\tau^L,\tau^R;\beta);w\right)\mathrm{d}w \\
&\quad +\varepsilon(\beta)\frac{1-\beta}{\beta\triangle\alpha t}\int_W g(w) r\left(\overline{\alpha t}+\Phi_w(\tau^L,\tau^R;\beta);w\right)\cdot\Lambda^{\alpha_L}\left(\tau^L,\tau^R;w\right)\mathrm{d}w \\
&\quad +\frac{1-\beta}{\beta\triangle\alpha t}\int_W g(w) r\left(\overline{\alpha t}+\Phi_w(\tau^L,\tau^R;\beta);w\right)\cdot(w-\mu)\,\mathrm{d}w \\
&\geq 0 \quad \text{(by supposition).} \qquad (25)
\end{aligned}
$$

Repeating the argument to derive (24) and (25) reveals that, for any $\beta' \in [\beta, 1]$, there exists a vector $(-1, \delta(\beta')) \equiv (-1, \delta(\beta))$ such that $\nabla^{\beta'} v_L\left((1,\alpha_L^*),(t_R^*,1)\right)\cdot(-1,\delta(\beta')) > 0$ and $\nabla^{\beta'}\mathbb{F}_L\left((1,\alpha_L^*),(t_R^*,1)\right)\cdot(-1,\delta(\beta')) \geq 0$. In particular, $\nabla^1 v_L\left((1,\alpha_L^*),(t_R^*,1)\right)\cdot(-1,\delta(1)) > 0$ and $\nabla^1\mathbb{F}_L\left((1,\alpha_L^*),(t_R^*,1)\right)\cdot(-1,\delta(1)) \geq 0$, which is a contradiction. This is because $\left((1,\alpha_L^*),(t_R^*,1)\right)$ is a non-trivial and non-pure **PUNE** when $\beta = 1$. Thus, $(1,\alpha_L^*)$ is a best response for L to $(t_R^*,1)$ in the game $\langle H,\mathbb{F},L,R,\Upsilon,v,\gamma\rangle$ with $\beta > 0$ sufficiently close to unity.

(2) $(t_R^*, 1)$ **is a best response for Right to** $(1,\alpha_L^*)$ **in the game** $\langle H,\mathbb{F},L,R,\Upsilon,v,\gamma\rangle$ **with** $\beta > 0$ **close to unity.**

Suppose that there is $\beta > 0$, that is sufficiently close to unity, such that $(t_R^*, 1)$ is not a best response for R to $(1,\alpha_L^*)$ in the game $\langle H,\mathbb{F},L,R,\Upsilon,v,\gamma\rangle$ - with this β. Then, there exists a vector $(\delta(\beta),-1)$ such that $\nabla^\beta v_R\left((1,\alpha_L^*),(t_R^*,1)\right)\cdot(\delta(\beta),-1) > 0$ and $\nabla^\beta\mathbb{F}_R\left((1,\alpha_L^*),(t_R^*,1)\right)\cdot(\delta(\beta),-1) \leq 0$, where

$$
\begin{aligned}
\nabla^\beta v_R\left((1,\alpha_L^*),(t_R^*,1)\right) &= \left(\frac{\partial v_R\left((1,\alpha_L^*),(t_R^*,1);\beta\right)}{\partial t_R}, \frac{\partial v_R\left((1,\alpha_L^*),(t_R^*,1);\beta\right)}{\partial \alpha_R}\right) \\
\text{and } \nabla^\beta\mathbb{F}_R\left((1,\alpha_L^*),(t_R^*,1)\right) &= \left(\frac{\partial\mathbb{F}\left(\Omega(\tau^L,\tau^R);\beta\right)}{\partial t_R}, \frac{\partial\mathbb{F}\left(\Omega(\tau^L,\tau^R);\beta\right)}{\partial \alpha_R}\right).
\end{aligned}
$$

Because $\nabla^\beta v_R\left((1,\alpha_L^*),(t_R^*,1)\right)\cdot(\delta(\beta),-1) > 0$, we have to extrapolate from (21) and (22):

$$
\delta(\beta) > \frac{(1-\beta)\left[-t_R^*\mu + \frac{\partial\sigma\left(t_R^*\mu\right)}{\partial\alpha t\mu}t_R^*\mu\right] - \beta t_R^*\left(t_R^* - a_R\right)}{(1-\beta)\left[-w_R + \frac{\partial\sigma\left(t_R^*\mu\right)}{\partial\alpha t\mu}\mu\right] - \beta\left(t_R^* - a_R\right)}.
$$

Given the choice of t_R^*, $\nabla^1 v_R\left((1,\alpha_L^*),(t_R^*,1)\right)\cdot(\delta(1),-1) > 0$, so that $\delta(1) > t_R^*$ holds. This implies that:

$$
\delta(\beta) > t_R^* \text{ for } \beta > 0 \text{ sufficiently close to unity.} \qquad (26)
$$

Let $\delta(\beta) = t_R^* + \varepsilon(\beta)$ where $\varepsilon(\beta) > 0$. Consider

$$\begin{aligned} & \nabla^{\beta}\mathbb{F}_R\left(\left(1, \alpha_L^*\right), \left(t_R^*, 1\right)\right) \cdot \left(t_R^*, -1\right) \\ = \quad & t_R^* \frac{1-\beta}{\beta \triangle \alpha t} \int_W g(w) r\left(\overline{\alpha t} + \Phi_w(\tau^L, \tau^R; \beta); w\right) \cdot (w - \mu)\, \mathrm{d}w. \end{aligned}$$

Thus, we have:

$$\begin{aligned} & \nabla^{\beta}\mathbb{F}_R\left(\left(1, \alpha_L^*\right), \left(t_R^*, 1\right)\right) \cdot \left(\delta(\beta), -1\right) \\ = \quad & \varepsilon\left(\beta\right) \frac{1}{2} \int_W g(w) r\left(\overline{\alpha t} + \Phi_w(\tau^L, \tau^R; \beta); w\right) \mathrm{d}w \\ & + \varepsilon\left(\beta\right) \frac{1-\beta}{\beta \triangle \alpha t} \int_W g(w) r\left(\overline{\alpha t} + \Phi_w(\tau^L, \tau^R; \beta); w\right) \cdot \Lambda^{t_R}\left(\tau^L, \tau^R; w\right) \mathrm{d}w \\ & + t_R^* \frac{1-\beta}{\beta \triangle \alpha t} \int_W g(w) r\left(\overline{\alpha t} + \Phi_w(\tau^L, \tau^R; \beta); w\right) \cdot (w - \mu)\, \mathrm{d}w \\ \leq \quad & 0 \text{ (by supposition).} \quad (27) \end{aligned}$$

This implies:

$$\begin{aligned} & \lim_{\beta \to 1} \nabla^{\beta}\mathbb{F}_R\left(\left(1, \alpha_L^*\right), \left(t_R^*, 1\right)\right) \cdot \left(\delta(\beta), -1\right) \\ = \quad & \varepsilon\left(1\right) \frac{\alpha_R}{2} \int_W g(w) r\left(\overline{\alpha t} + \Phi_w(\tau^L, \tau^R; \beta); w\right) \mathrm{d}w \\ \leq \quad & 0 \text{ (from (27)).} \quad (28) \end{aligned}$$

Note that (27) and (28) imply $\varepsilon\left(\beta\right) \leq 0$ for $\beta > 0$ sufficiently close to unity. This is a contradiction, because $\varepsilon\left(\beta\right) > 0$ given that $\nabla^{\beta} v_R\left(\left(1, \alpha_L^*\right), \left(t_R^*, 1\right)\right) \cdot \left(\delta(\beta), -1\right) > 0$ and (26). Thus, we conclude that $(t_R^*, 1)$ is a best response for R to $(1, \alpha_L^*)$ in the game $\langle H, \mathbb{F}, L, R, \Upsilon, v, \gamma \rangle$ with any $\beta > 0$ close to unity.

(3) Any other type of strategy profile cannot constitute a non-trivial and non-pure PUNE in the game $\langle H, \mathbb{F}, L, R, \Upsilon, v, \gamma \rangle$ with $\beta > 0$ close to unity.

First, let us consider any strategy (t_L, α_L) with $\alpha_L t_L = \alpha_L^*$ and $t_L < 1$, and any strategy (t_R, α_R) with $\alpha_R t_R = t_R^*$ and $\alpha_R < 1$. Then, L can improve both its militant's and its opportunist's payoffs by deviating to $(1, \alpha_L^*)$, whereas R can improve both its militant's and its opportunist's payoffs by deviating to $(t_R^*, 1)$. This is easily verified for the militants of both parties. Let us examine the opportunists of both parties.

For the L-opportunist, we must examine whether $\nabla^{\beta}\mathbb{F}_L\left(\left(t_L,\alpha_L\right),\left(t_R,\alpha_R\right)\right)\cdot \left(1-t_L,\alpha_L^*-\alpha_L\right)>0$ holds. Note that

$$\left(1-t_L\right)\alpha_L+\left(\alpha_L^*-\alpha_L\right)t_L=\left(1-t_L\right)\alpha_L+\left(t_L\alpha_L-\alpha_L\right)t_L=\left(1-t_L\right)\alpha_L-\left(1-t_L\right)\alpha_L t_L>0.$$

Then, by using (15) and (16), we obtain

$$\begin{aligned}&\nabla^{\beta}\mathbb{F}_L\left(\left(t_L,\alpha_L\right),\left(t_R,\alpha_R\right)\right)\cdot\left(1-t_L,\alpha_L^*-\alpha_L\right)\\ =&\int_W g(w)r\left(\overline{\alpha t}+\Phi_w(\tau^L,\tau^R;\beta);w\right)\cdot\left[\Psi_L+\frac{1-\beta}{\beta\triangle\alpha t}\left(1-t_L\right)\left(\mu-w\right)\right]\mathrm{d}w,\end{aligned}$$

where $\Psi_L=\left(\left(1-t_L\right)\alpha_L+\left(\alpha_L^*-\alpha_L\right)t_L\right)\left[\frac{1}{2}+\frac{1-\beta}{\beta\triangle\alpha t}\left\{\triangle\alpha t\frac{\partial\sigma(\alpha_L t_L\mu)\cdot\mu}{\partial\alpha t\mu}+\frac{\triangle t(w-\mu)}{\triangle\alpha t}\right\}\right]$. Note that $\Psi_L>0$, because $\left(1-t_L\right)\alpha_L+\left(\alpha_L^*-\alpha_L\right)t_L>0$ and $\alpha_L^*-\alpha_L<0$. By the construction of this strategy profile, $\overline{\alpha t}+\Phi(\tau^L,\tau^R;\beta)$ is close to a^m. Thus, by **A4**, we can see that

$$\frac{1-\beta}{\beta\triangle\alpha t}\left(1-t_L\right)\int_W g(w)r\left(\overline{\alpha t}+\Phi_w(\tau^L,\tau^R;\beta);w\right)\left(\mu-w\right)\mathrm{d}w>0.$$

Thus, $\nabla^{\beta}\mathbb{F}_L\left(\left(t_L,\alpha_L\right),\left(t_R,\alpha_R\right)\right)\cdot\left(1-t_L,\alpha_L^*-\alpha_L\right)>0$ holds.

For the R-opportunist, we must examine whether $\nabla^{\beta}\mathbb{F}_R\left(\left(t_L,\alpha_L\right),\left(t_R,\alpha_R\right)\right)\cdot \left(t_R^*-t_R,1-\alpha_R\right)>0$ holds. Note that

$$\left(t_R^*-t_R\right)\alpha_R+\left(1-\alpha_R\right)t_R=\left(t_R\alpha_R-t_R\right)\alpha_R+\left(1-\alpha_R\right)t_R=\left(1-\alpha_R\right)t_R-\left(1-\alpha_R\right)t_R\alpha_R>0.$$

Then, by using (19) and (20), we obtain

$$\begin{aligned}&\nabla^{\beta}\mathbb{F}_R\left(\left(t_L,\alpha_L\right),\left(t_R,\alpha_R\right)\right)\cdot\left(t_R^*-t_R,1-\alpha_R\right)\\ =&\int_W g(w)r\left(\overline{\alpha t}+\Phi_w(\tau^L,\tau^R;\beta);w\right)\cdot\left[\Psi_R+\frac{1-\beta}{\beta\triangle\alpha t}\left(t_R^*-t_R\right)\left(w-\mu\right)\right]\mathrm{d}w,\end{aligned}$$

where $\Psi_R=\left(\left(t_R^*-t_R\right)\alpha_R+\left(1-\alpha_R\right)t_R\right)\left[\frac{1}{2}+\frac{1-\beta}{\beta\triangle\alpha t}\left\{-\triangle\alpha t\frac{\partial\sigma(\alpha_L t_L\mu)\cdot\mu}{\partial\alpha t\mu}-\frac{\triangle t(w-\mu)}{\triangle\alpha t}\right\}\right]$. Note that $\Psi_R>0$ for β sufficiently close to unity, because $\left(\left(t_R^*-t_R\right)\alpha_R+\left(1-\alpha_R\right)t_R\right)>0$. Since $t_R^*-t_R<0$, we can see, by **A4**, that

$$\frac{1-\beta}{\beta\triangle\alpha t}\left(t_R^*-t_R\right)\int_W g(w)r\left(\overline{\alpha t}+\Phi_w(\tau^L,\tau^R;\beta);w\right)\left(w-\mu\right)\mathrm{d}w>0.$$

Thus, $\nabla^{\beta}\mathbb{F}_R\left(\left(t_L,\alpha_L\right),\left(t_R,\alpha_R\right)\right)\cdot\left(t_R^*-t_R,1-\alpha_R\right)>0$ holds.

Consider any strategy profile $\left(\tau^L, \tau^R\right) = ((t_L, \alpha_L), (t_R, \alpha_R))$ with $t_L = 1$, $\alpha_R = 1$, and $\frac{1}{2}(t_R + \alpha_L) \notin (a^m - \epsilon(\gamma), a^m + \epsilon(\gamma))$. Then, for any $\beta > 0$ close to unity, $\pi\left(\tau^L, \tau^R\right) \in \{0, 1\}$, and thus the profile does not correspond to any non-trivial and non-pure **PUNE**.

Consider any strategy profile $\left(\tau^L, \tau^R\right) = ((t_L, \alpha_L), (t_R, \alpha_R))$ with $t_L = 1$, $\alpha_R = 1$, $\frac{1}{2}(t_R + \alpha_L) \in (a^m - \epsilon(\gamma), a^m + \epsilon(\gamma))$, and $a_R < t_R$. Then, R can improve both its militant's and opportunist's payoffs by deviating from (t_R, α_R) to $(t_R - \varepsilon, \alpha_R)$ where $\varepsilon < t_R - a_R$, whenever $\beta > 0$ is sufficiently close to unity.

Consider any strategy profile $\left(\tau^L, \tau^R\right) = ((t_L, \alpha_L), (t_R, \alpha_R))$ with $t_L = 1$, $\alpha_R = 1$, $\frac{1}{2}(t_R + \alpha_L) \in (a^m - \epsilon(\gamma), a^m + \epsilon(\gamma))$, and $\alpha_L < a_L$. Then, L can improve both its militant's and opportunist's payoffs by deviating from (t_L, α_L) to $(t_L, \alpha_L + \varepsilon)$ where $\varepsilon < a_L - \alpha_L$, whenever $\beta > 0$ is sufficiently close to unity.

Consider any strategy profile $\left(\tau^L, \tau^R\right) = ((t_L, \alpha_L), (t_R, \alpha_R))$ with $t_L = 1$, $\alpha_R = 1$, $\frac{1}{2}(t_R + \alpha_L) \in (a^m - \epsilon(\gamma), a^m + \epsilon(\gamma))$, and $t_R < \alpha_L$. Then, at least one of these two parties can follow a deviation to make its militant and opportunist better off, whenever $\beta > 0$ is sufficiently close to unity.

In summary, there is no other type of non-trivial and non-pure **PUNE** besides the type $((1, \alpha_L^*), (t_R^*, 1))$ with $\frac{1}{2}(t_R^* + \alpha_L^*) \in (a^m - \epsilon(\gamma), a^m + \epsilon(\gamma))$ and $a_R \geq t_R^* \geq \alpha_L^* \geq a_L$ for some $\epsilon(\gamma) > 0$, whenever $\beta > 0$ is sufficiently close to unity. ■

7.3 Proofs for Section 5

Proof of Theorem 4: As we discussed in Section 3, every voter $h \in H$ with $w_h \leq \mu$ has the same ideal policy $\left(\bar{t}_h, \overline{\alpha}_h\right) = (1, \alpha^*)$, whereas every voter $h \in H$ with $w_h > \mu$ has the ideal policy $\left(\bar{t}_h, \overline{\alpha}_h\right) = \left(t\left(\frac{w_h}{\mu}\right), 1\right)$. Given $\tau_L = (1, \alpha^*)$, and given $w_{h'} > \mu$, let us consider

$$\max_{(t_R, \alpha_R) \in \Upsilon} \left[(1 - t_R) \int_{w_{h'}}^{\overline{w}} w_h \mathrm{d}\mathbb{F}(h) + (1 - \alpha_R) t_R \mu + \sigma(\alpha_R t_R \mu)\right].$$

Because the above objective function is strictly concave, there is a unique optimal solution, $\tau_R(w_{h'}) = (t_R(w_{h'}), \alpha_R(w_{h'})) \in \Upsilon$ for $w_{h'} > \mu$. Then, $\Omega(\tau^L, \tau_R(w_{h'}))$ is identified. We aim to find $h' \in H$ with $w_{h'} = \mu - \Theta\left(\tau^L, \tau_R(w_{h'})\right)$. When $w_{h'}$ increases from μ up to $\overline{w}$, then $-\Theta\left(\tau^L, \tau_R(w_{h'})\right) \geq$

0 decreases. Note that if $w_{h'} = \mu$, then $w_{h'} < \mu - \Theta\left(\tau^L, \tau_R\left(w_{h'}\right)\right)$. In addition, if $w_{h'} = \overline{w}$, then $\tau_R\left(w_{h'}\right) = \arg\max_{(t_R,\alpha_R)\in\Upsilon}\left(1 - t_R\right)\overline{w} + \left(1 - \alpha_R\right)t_R\mu + \sigma\left(\alpha_R t_R \mu\right)$, which implies $w_{h'} > \mu - \Theta\left(\tau^L, \tau_R\left(w_{h'}\right)\right)$. Thus, there exists an $h^* \in H$ with $w_{h^*} > \mu$ such that $w_{h^*} = \mu - \Theta\left(\tau^L, \tau_R\left(w_{h^*}\right)\right)$.

Define $L = \{h \in H \mid w_h \leq w_{h^*}\}$ and $R = \{h \in H \mid w_h > w_{h^*}\}$. By construction, $w_R = \int_{h\in R} w_h \mathrm{d}\mathbb{F}(h) > \mu$, and thus $w_L = \int_{h\in L} w_h \mathrm{d}\mathbb{F}(h) < \mu$. Because $R = H\backslash\Omega(\tau^L, \tau_R\left(w_{h^*}\right))$ and $L = \Omega(\tau^L, \tau_R\left(w_{h^*}\right))$ by construction, for all $h \in R$, $v(\tau^L; h) \leq v(\tau^R; h)$, and for all $h \in L$, $v(\tau^L; h) > v(\tau^R; h)$. Because $w_R > \mu$, following the argument used in the proof of Lemma 2 reveals that $\tau_R\left(w_{h^*}\right) = \left(t\left(\frac{w_R}{\mu}\right), 1\right)$, which is the R-militant's ideal policy. Because $w_L < \mu$, following the argument used in the proof of Lemma 1 shows that $\tau^L = (1, \alpha^*)$, which remains the L-militant's ideal policy. Then, by applying the arguments used in the proofs of Lemma 3 and Theorem 1, we obtain the desired result. ■

Proof of Theorem 5: As shown in the proof of Theorem 2, given $\gamma \in \left(0, \frac{1}{2}\right)$, we can identify the maximal number $\epsilon\left(\gamma\right) > 0$ such that

$$\int_0^{a^m - \epsilon(\gamma)} r\left(a\right)\mathrm{d}a \geq \frac{1}{2} - \gamma \text{ and } \int_0^{a^m + \epsilon(\gamma)} r\left(a\right)\mathrm{d}a \leq \frac{1}{2} + \gamma.$$

Take any number $\overline{a} \in \left(a^m - \epsilon\left(\gamma\right), a^m + \epsilon\left(\gamma\right)\right)$, and define $L \equiv \{h \in H \mid a_h \leq \overline{a}\}$ and $R \equiv \{h \in H \mid a_h > \overline{a}\}$. Consider any pair of policies $(\tau^L, \tau^R) = \left(\left(t_L, \alpha_L\right), \left(t_R, \alpha_R\right)\right)$ such that $\frac{1}{2}\left(\alpha_R t_R + \alpha_L t_L\right) = \overline{a}$ and $\alpha_L t_L < \alpha_R t_R$. Then, for any $h \in L$, $v(\tau^L; h) \geq v(\tau^R; h)$ and for all $h \in R$, $v(\tau^L; h) < v(\tau^R; h)$. Based on this partition, we can identify $a_L = \int_{h\in L} a_h \mathrm{d}\mathbb{F}(h)$ and $a_R = \int_{h\in R} a_h \mathrm{d}\mathbb{F}(h)$. By definition, $a_L < \overline{a} < a_R$. Then, we can appropriately choose a pair of policies $(\tau^L, \tau^R) = \left(\left(t_L, \alpha_L\right), \left(t_R, \alpha_R\right)\right)$ such that $\frac{1}{2}\left(\alpha_R t_R + \alpha_L t_L\right) = \overline{a}$, $\alpha_L t_L < \alpha_R t_R$, and $a_L < \alpha_L t_L < \overline{a} < \alpha_R t_R < a_R$. Then, as shown in the proof of Theorem 2, it follows that this (τ^L, τ^R) constitutes a non-trivial and non-pure **PUNEEP**, and there is no other type of non-trivial **PUNEEP**. ■

Proof of Theorem 6: Let us consider two types of population distribution.

(Case 1) $\int_0^{\mu}\int_0^1 ag\left(w\right)r\left(a; w\right)\mathrm{d}a\mathrm{d}w \leq a^m$ **holds.**

Given $\epsilon\left(\gamma\right) > 0$ obtained in Theorem 2, take any number $\overline{a} \in \left(a^m - \epsilon\left(\gamma\right), a^m + \epsilon\left(\gamma\right)\right)$, and consider any pair of policies $(\tau^L, \tau^R) = \left(\left(t_L, \alpha_L\right), \left(t_R, \alpha_R\right)\right) = \left(\left(1, \alpha^*\right), \left(t^*, 1\right)\right)$ such that $\frac{1}{2}\left(\alpha_R t_R + \alpha_L t_L\right) = \overline{a}$ and $\int_{h:a_h<\overline{a}} a_h \mathrm{d}\mathbb{F}(h) < \alpha_L t_L \leq \alpha_R t_R < \int_{h:a_h\geq\overline{a}} a_h \mathrm{d}\mathbb{F}(h)$. We can define $L \equiv \Omega(\tau^L, \tau^R)$ and $R \equiv H\backslash\Omega(\tau^L, \tau^R)$. Then,

$a_L = \int_W \int_0^{\overline{a}+\Phi_w(\tau^L,\tau^R;\beta)} ag(w)r\,(a;w)\mathrm{d}a\mathrm{d}w$ and $a_R = \int_W \int_{\overline{a}+\Phi_w(\tau^L,\tau^R;\beta)}^1 ag(w)r\,(a;w)\mathrm{d}a\mathrm{d}w$. Because $\beta > 0$ is sufficiently close to unity, $\Phi_w(\tau^L,\tau^R;\beta)$ is very small for each $w \in W$, which implies that a_L (resp. a_R) is very close to $\int_{h:a_h<\overline{a}} a_h \mathrm{d}\mathbb{F}(h)$ (resp. $\int_{h:a_h\geq\overline{a}} a_h \mathrm{d}\mathbb{F}(h)$). Thus, $a_L < \alpha_L t_L < \overline{a} < \alpha_R t_R < a_R$ still holds.

Note that $\int_0^{\mu}\int_0^1 ag\,(w)\,r\,(a;w)\mathrm{d}a\mathrm{d}w \leq a^m$ implies $\int_W \int_0^{a^m} wg(w)r\,(a;w)\mathrm{d}a\mathrm{d}w < \mu$. Because $\overline{a}$ is very close to a^m, it follows that $w_L = \int_W \int_0^{\overline{a}+\Phi_w(\tau^L,\tau^R;\beta)} wg(w)r\,(a;w)\mathrm{d}a\mathrm{d}w < \mu$. Similarly, $w_R = \int_W \int_{\overline{a}+\Phi_w(\tau^L,\tau^R;\beta)}^1 wg(w)r\,(a;w)\mathrm{d}a\mathrm{d}w > \mu$ holds. Thus, we can apply the argument used in the proof of Theorem 3 to show that this (τ^L,τ^R) constitutes a non-trivial and non-pure **PUNEEP**, and there is no other type of non-trivial **PUNEEP**.

(Case 2) $\int_0^{\mu}\int_0^1 ag\,(w)\,r\,(a;w)\mathrm{d}a\mathrm{d}w > a^m$ **holds.**

Let us show that the above strategy profile $(\tau^L,\tau^R) = ((1,\alpha^*),(t^*,1))$ can be a **PUNEEP** even in Case 2. Note that $\int_0^{\mu}\int_0^1 ag\,(w)\,r\,(a;w)\mathrm{d}a\mathrm{d}w > a^m$ implies $\int_W \int_0^{a^m} wg(w)r\,(a;w)\mathrm{d}a\mathrm{d}w > \mu$, so that $w_L > \mu$. Consider a deviation from τ^L to another strategy $\tau^{L\prime} = (\alpha^*,1)$. Then, although the level of defense expenditure does not change, the after-tax income $(1-\alpha^*)\,w_L$ of L under $\tau^{L\prime}$ exceeds $(1-\alpha^*)\,\mu$, which is the after-tax income of L under τ^L. Because $\beta < 1$, it follows that the L-militant's payoff is improved by this deviation. However, as shown by the proof of Theorem 3, $\nabla^{\beta}\mathbb{F}_L\,((1,\alpha^*),(t^*,1))\cdot(t'_L-1,\alpha'_L-\alpha^*) < 0$ by **A4** whenever $t'_L\alpha'_L = \alpha^*$ holds. Thus, the L-opportunist is worse off following this deviation. Note that for any other type of deviation, we can apply the same argument used in the proof of Theorem 3. Thus, $(1,\alpha^*)$ is still the best reply to $(t^*,1)$ for L. Similarly, $(t^*,1)$ is still the best reply to $(1,\alpha^*)$ for R.

Note that if $(\tau^L,\tau^R) = ((1,\alpha^*),(t^*,1))$ is a **PUNEEP** in Case 2, then it is easy to check that $(\tau^{L\prime},\tau^{R\prime}) = ((\alpha^*,1),(1,t^*))$ is also a **PUNEEP** in Case 2. ■

8 References

Baron, D. (1993): "Government formation and endogenous parties," *American Political Science Review* **87**, pp. 34-47.

Besley, T. and Coate, S. (1997): "An Economic Model of Represetative Democracy," *Quarterly Journal of Economics,* February, pp. 85-114.

Caplin, A. and Nalebuff, B. (1997): "Competition among institutions," *Journal of Economic Theory* **72**, pp. 306-342.

Downs, A. (1957): *An Economic Theory of Democracy*, New York: Harper Collins.

Francois, P. (2000): "Public service motivation as an argument for government provision," *Journal of Public Economics* **78**, pp. 275-299.

Francois, P. (2003): "Not-for-profit provision of public services," *Economic Journal* **113**, pp. C53-C61.

Francois, P. (2006): "Making a difference," forthcoming in *Rand Journal of Economics*.

Lenin, V. I. (1916): "*Imperialism, the Highest Stage of Capitalism*," in Lenin Collected Works, Moscow, Vol **22**, pp. 185-304.

Osborne, M. and Slivinski, A. (1996): "A Model of Political Competition with Citizen-Candidates," *Quarterly Journal of Economics,* February, pp. 65-96.

Roemer, J. E. (1998): "Why the poor do not expropriate the rich: an old argument in the new garb," *Journal of Public Economics* **70**, pp. 399-424.

Roemer, J. E. (1999): "The democratic political economy of progressive taxation," *Econometrica* **67**, pp. 1-19.

Roemer, J. E. (2001): *Political Competition*, Harvard University Press.

Roemer, J. E. (2005): "Will Democracy Engender Equality?" *Economic Theory* **25**, pp. 217-234.

Wittman, D. (1973): "Parties as utility maximizers," *American Political Science Review* **67**, pp. 490-498.

Yoshihara, N. (2008): "On the Existence of Party-Unanimity Nash Equilibria in Multi-dimensional Political Competition Games," *mimeo.*

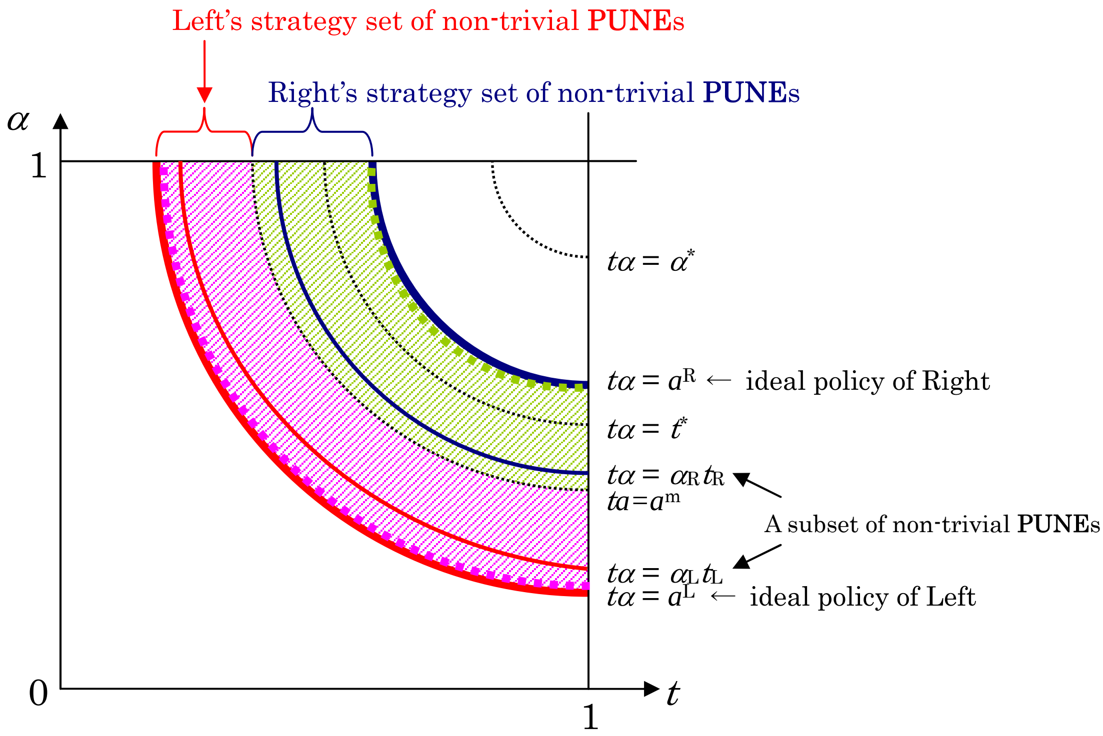

Figure 1: The set of non-trivial and non-pure **PUNE**s when β=1:

$$\{((t_L, \alpha_L),(t_R, \alpha_R)) \in \Upsilon \mid a^L < \alpha_L t_L \leq \alpha_R t_R < a^R \;\&\; a^m - \varepsilon(\gamma) \leq (1/2)\cdot(\alpha_L t_L + \alpha_R t_R) \leq a^m + \varepsilon(\gamma)\}$$