This is to certify that the
dissertation entitled

THREE TOPICS ON ELECTIONS

presented by

MAKOTO TANAKA

has been accepted towards fulfillment
of the requirements for the

Ph.D degree in ECONOMICS

John D. Wilson
Major Professor's Signature

8/20/08
Date

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE

THREE TOPICS ON ELECTIONS

By

Makoto Tanaka

A DISSERATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Economics

2008

ABSTRACT

THREE TOPICS ON ELECTIONS

By

Makoto Tanaka

This dissertation consists of three chapters. In chapter one and two, I consider information transmission problem from politicians to voters about possible outcomes of policies in a two-period-and-two-elections model (Chapter one) and in an infinite period model with political parties as overlapping-generations organizations (Chapter two). In Chapter three, I analyze the choice of policy instrument by an incumbent politician who can use an inefficient tariff and a less inefficient production subsidy to help domestic manufacturers.

Chapter One: "Learning Through Elections"

I consider choices of policy platform at elections when there is uncertainty about outcomes of policies, and one party has information advantage over another party and voters. This ability might reflect the ability of the politician in the party. Since the platform choice is based on the information the politician has, the choice at the election transmits some information to voters. Then, in one possible equilibrium, I show a case that the higher the ability of the politician, the worse the information transmission from politicians to voters.

Chapter Two: "Information transmission from overlapping political parties"

I analyze an information transmission problem, again, but in infinite period with two political parties as overlapping-generations organizations. Policy outcomes depend on the true state of the world, which changes in each period. Both parties receive information about the true state. Platform choices are made by old politicians in parties who will retire soon. So, the next election does not restrict their opportunistic behavior directly. Still, I show that the OLG party structure restrains the opportunistic choice of platform by old politicians.

Chapter Three: "The Choice of Inefficient Instruments in a Simple Retrospective Voting Model with Voter Abstention"

Tariffs are more inefficient than production subsidies as the instrument to help domestic industry. Still, tariffs have been used. In this chapter, I propose one explanation about why governments use inefficient tariffs. The basic idea is the manipulation of voter abstention. If an incumbent politician decides to help a domestic industry, he also needs to determine how to distribute the cost of the help among voters. Assuming that some voters' voting abstain partially depending on the utility they receive from the incumbent's policy, the incumbent could affect the choice of voters to vote or not. Naturally, he will try to take advantage of it. I analyze if the incumbent has incentive to choose tariff in two cases: same tax rate for all voter groups, and different tax rates for groups.

To My Mother, Yumiko Tanaka

ACKNOWLEDGEMENT

I have so many people I would like to thank for my dissertation. However, because of limitation of space, I have to minimize the number of people to thank. First of all, I would like to thank my committee members. I appreciate the help of my adviser Professor John D. Wilson. Without his guidance and suggestions, I could not have completed my dissertation. I thank Professor Anthony Creane for his caring. I really appreciate it when things were rough for me. I also appreciate the conversations I had with Professor Susan Zhu and Professor Eric Chang. Because of their help, I could improve my dissertation a lot. Regarding chapter three, I also want to thank Professor Nelson of Tulane University. The conversation I could fortunately have with him helped and encouraged me a lot while I was working on the chapter.

I thank Jennifer Carducci for her professional job and help. She deserves more chocolate. I also thank Stephanie O. Smith for her help on the format of this dissertation. I appreciate her kind help and apologize for all of my formatting mistakes. She deserves more cakes. Finally, but not leastly, I would like to thank all of my friends. They helped and encouraged me, and sometimes inspired me with their way of life. I appreciate all the people I met in Michigan.

TABLE OF CONTENTS

LIST OF FIGURES

1 Learning through Elections

1.1 Introduction

The objective of this chapter is the construction of a model of two consequent elections with uncertainty about policy outcomes. Traditionally, electoral competition models in Political Economy have assumed that there is no uncertainty about outcomes of government policies. Clearly, this is a strong assumption. Before the implementations of policies, people can have only beliefs about outcomes. Politicians, however, might have more information about outcomes than voters. Suppose that a politician has some private information about possible outcomes. Then, given the information asymmetry, what policy platform will the politician choose during an election? Since the politician chooses the platform based on the information he has, the choice of the platform can be considered as the information transmission from him to voters. Then, will the politician reveal the information honestly? In other words, will the politician lie to voters? This is the information transmission problem during the election.

From 1990s, many papers have analyzed the problem with various assumptions about uncertainty (Harrington (1993a), etc.) have. Observing platform choices by parties during an election, voters can update their beliefs about the uncertainty. However, there is no reason we can expect this transmission of information to be perfect. Obviously, the information asymmetry enables politicians to take opportunistic behavior; they can choose a bad policy and lie to voters that it is a good policy. The common structure of papers in the literature is the following; there is the uncertainty about policy outcomes, which depend on the state of the world. A politician (or politicians) has some private information about possible policy outcomes. He chooses a policy. This choice may be his platform during an election (Heidhues and Lagerof (2003)), or the policy he implements before the election (Majumdar and Mukand (2004)). Then, at the election, voters vote for the candidate they prefer. The important part is this; since the politician's choice gives voters some information about the uncertainty, politicians can manipulate preferences of voters on policies and/or politicians[1].

This situation gives voters two opportunities of learning about the uncertainty (i.e., the state of the world); one is when voters observe the choices of parties during the election, and the other is when voters observe the outcome of the implemented policy after the election. In multi-election models, the voting decisions at elections depend on these two opportunities of learning. Thus, politicians have to take into consideration the impact of the election platform at one election on the election after that. To model this, I consider a two-period and two-election model. Of these two elections, the first election is for the information transmission. The second one is for the retrospective voting based on the result of the implemented policy after the first election, which guarantees that at least some information is transmitted to voters during the first election. In addition, I assume that one party has the information advantage not only over voters but also over the other party. This information advantage can be considered as showing the high ability of the politician in the party. The symmetry between parties during an election is a common assumption in electoral competition models without an incumbent. Though Heidhues and Lagerof (2003) and Bernhardt, Duggan and Squintani (2007) also assume the asymmetry between parties, the asymmetry in these papers means that parties have different information. The asymmetry here is stronger since only one party receives the information. This is a strong and uncommon assumption. However, the asymmetry between parties in the sense of this chapter is not unnatural[2]. As a factor causing such asymmetry between parties during an election, "valence" has been incorporated into electoral competition models recently (Schofield (2005), etc.) So, the ability of the politician in this chapter can be considered as the valence factor in this chapter. Then, I show that this asymmetry has an interesting implication when the ability of the politician changes.

The remainder of this chapter is organized as follows; section 1.2 describes the model. Section 1.3 describes the behaviors of voters and two parties. Section 1.4 shows the existence of equilibrium and other results. Section 1.5 discusses about this model and its results. All proofs are in the Appendix of this chapter.

1.2 Model

I consider a society with two parties (Party 1 and Party 2) and voters in two periods. There are two possible policies, a and b. The society can implement only one policy in one period. There is no uncertainty about the outcome of the policy b. But the outcome of the policy a is uncertain. A natural interpretation of these two policies is that the policy b is the continuation of a current policy, and the policy a is a new policy. As I explained in the introduction, no one is perfectly sure about the outcome of the new policy, there is the uncertainty about its outcome. In each period, there is an election. The society decides which policy it implements in the period by the election. Candidates from two parties (one from each party) announce one of two policies as their policy platform. Voters vote for the candidate they prefer, and the elected politician implements his platform.

Candidates from two parties are new candidates. If a candidate loses the first election, the game is over for him. At the second election, the party that lost the first election put up a new candidate. The politician who won the election gets a fixed rent, which is normalized to one. Let δ be the common discount factor. The objective of candidates is the maximization of expected sum of rents from two periods. They are risk neutral. The outcome of the policy b is the utility $\beta \in (0, 1)$ to voters. The outcome of the policy a could be success or failure. The outcome depends on the state of the world, w^0 and w^1. If the true state is w^0, the outcome of the policy a is failure and the utility 0 to voters. If the state is w^1, the outcome is success and utility 1 to voters. No one in the society knows the true state. Instead, voters have belief $q \in (0,1)$ as the probability of w^1 being the true state. This q is lower than β. This structure of uncertainty is almost same to the Majumdar and Mukand (2004). But unlike the paper, one implementation of the policy a is enough to know the true state of the world. Party 2 candidate also has belief q, same as voters. Since $q < \beta$, without further information about the uncertainty, the society should choose the policy b since the expected utility of the policy a is only q. But Party 1 candidate receives a signal about the probability of the true state being w^1 before the first election. To make this model simple, I assume that he receives the probability of state w^1 directly. Denote it as p. This means, the best policy choice of Party 1 candidate for voters is choosing the policy a if $p > \beta$ and choosing the policy b if $p < \beta$. If $p = \beta$, his choice does not matter. This p is a random variable, and its distribution depends on the true state of the world. If the true state is w^1, the cumulative distribution function is $F(p)$. If w^0, the function is $G(p)$. For $p \in (0,1)$, $G(p) > F(p)$. Moreover, F and G are assumed to be differentiable, $F'(p) > 0$ and $G'(p) > 0$ on $(0,1)$, and $\lim_{p\to 0} F'(p)/G'(p) = 0$, $\lim_{p\to 1} F'(p)/G'(p) = \infty$, and $F'(p)/G'(p)$ is monotonically increasing (i.e., monotone likelihood ratio property.)[3] These cumulative distribution functions can be considered as representing the ability of the politician. These properties mean that when the candidate receives the high value of p it is more likely that the true state is w^1 than that the true state is w^0, and vice versa. These distributions are common knowledge.

So, there is an informational asymmetry between Party 1 candidate and voters/Party 2 candidate. Then, Party 1 candidate's policy platform is considered as a signal from the candidate about the probability of the true state being w^1. As the strategy of Party 1 candidate, I consider the following simple one; before the first election, the candidate chooses the value $\tau \in [0,1]$ as the threshold value such that if he receives $p > \tau$, he will announce the policy a, and if he receives $p < \tau$, he will announce the policy b. If $\tau = p$, he will or will not announce the policy a. Since p and β are the expected utilities of policies a and b, respectively, the value of τ different from β means inefficiency. If $\tau < \beta$, when $p \in [\tau, \beta)$, Party 1 candidate would propose the policy a even though he knows the policy b has higher expected utility for voters. If $\tau > \beta$, when $p \in (\beta, \tau)$, the candidate would propose b even though the policy a has higher expected utility. With this strategy, $\tau = 1$ is interpreted as the candidate never announces the policy a. $\tau = 0$ is that the candidate always announces the policy a. Since the choice of a when $p > \beta$ and the choice of b when $p < \beta$ is the best for voters, the best value of τ for voters is β.

Party 1 candidate's strategy during the second election is simple because of the assumption about the uncertainty. He announces the policy that voters have higher expected utility. In a case that he announced the policy a during the first election and won, then at the start of the second period, every member of society knows the true state. If the state is w^1, it is clear that the policy a is better than the policy b. Then, the candidate should announce the policy a as the platform during the second election. If the state is w^0, it is also clear that the policy b is better. In the case that he announced the policy b during the first election and won, he announces the policy b again during the second election. This is because his platform during the second election does not give any information to voters. This is intuitively clear. Since the candidate cares only about the office rent, not the outcome of policies, if announcing the policy a during the second election is beneficial for him with a certain value of p, then it is beneficial for the candidate with any value of p to announce the policy a during the second election because his expected utility during the second election does not depend on p. Thus, whatever the value of p, Party 1 candidate announces the policy a. Then, the platform a does not give any information to voters. There is no reason for voters to update their belief. Thus they do not update their belief. Since the belief during the second election is the updated one through Party 1 candidate's platform during the first election and it is lower than β (since Party 1 candidate announced the policy b during the first election), it is beneficial for the candidate to announce the policy b.

The strategy of Party 2 candidate is as follows; during the first election, expecting the value of τ, the candidate uses the mixed strategy. Denote as c the probability of his announcing the policy a during the first election. His strategy during the second election is same as Party 1 candidate.

Since during the second election, it is clear that which policy should be proposed, I do not need to consider about the strategy of the new candidate put up by the party that lost in the first election.

Knowing Party 1 candidate's strategy, voters (and Party 2 candidate) try to infer the value of τ rationally and update their beliefs about the state w^1 after observing Party 1 candidate's platform, by the Bayesian rule. If Party 1 candidate announces the policy a, the updated belief is higher than the initial one. If the candidate announces the policy b, the updated belief is lower.

At each election, voters vote for a party according to their period preference. The preference consists of three parts. The first part is the expected utility from the outcome of implemented policy. The second is a random noise. Thus, this model has an element of probabilistic voting model (Persson and Tabellini (2002)). All voters receive same noise during each election. This noise represents factors about parties and candidates other than the utility derived directly from the implemented policy, like the advent of charismatic candidate or the eruption of scandals. This noise ε is an i.i.d. random variable with a distribution H. This ε is voters' bias for Party 1 and independent of the announced policies by candidates. It can take negative value as well as positive value. If it is negative, it means bias against Party 1. I assume that $H(0) = 0.5$ and $H'(\varepsilon) > 0$ on its support, the support of H is connected and large enough, and H is symmetric.

The third part is the assumption of the retrospective voting behavior on the implementation of the policy a by Party 1 in the first period.

(A) : If the policy a is implemented by Party 1 after the first period and the outcome is success, voters vote for Party 1 candidate at the second election. If the outcome is failure, voters vote for Party 2 candidate.

As I wrote above, if policy a was implemented in the first period, there is no uncertainty in the second period and both parties propose same policy as their platform during the second election. So, the choice of the party to vote for does not matter for voters. Thus, voters can use the second election to reward or punish Party 1 (i.e., the retrospective voting.) It is natural to assume that the success increases the probability of winning the reelection and that the failure decreases the probability. This is the retrospective voting; voting based on past performance. This (A) is the simplified and extreme version of such reaction. This works as the incentive for Party 1 candidate to announce the policy a. Without this, as I will show later, there is no equilibrium.

This voting reaction to the policy a is applied only to Party 1 during the second election. In other cases, voters vote for Party 1 candidate if,

$$EU(Party\ 1\ candidate's\ platform) + \varepsilon > EU(Party\ 2\ candidate's\ platform)$$

where, EU means the expected utility

$$\Leftrightarrow \varepsilon > EU(Party\ 2's\ platform) - EU(Party\ 1's\ platform)$$

If the opposite inequality holds, voters vote for Party 2 candidate. If the equality holds, voters vote for Party 1 with probability 0.5. Since ε is a random variable, the probability of Party 1 candidate winning the election is

$$1 - H(EU(Party\ 2's\, platform) - EU(party\ 1's\, platform))$$

The flow of the game in the first period (see figure 1.1)

First period

Flow of Game

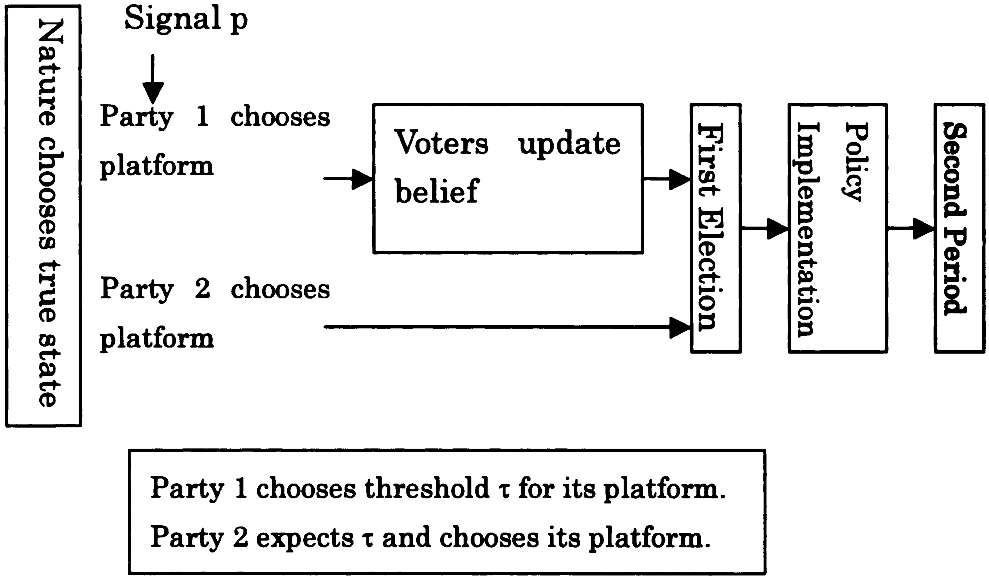

Figure 1.1:

Party 1 candidate receives the signal p.

Candidates announce platforms simultaneously.

Voters observe platforms and update their belief.

Voters receive noise ε, and vote for the party they prefer.

The Elected party implements its platform.

The outcome of the implemented policy is observed.

If the policy was a, then every member's belief about the state of world is updated.

Second period.

The equilibrium is defined as the profile of values of τ and c such that,

(1) Voters and Party 2 expect the value of τ and update their belief rationally (Bayesian),

(2) With the expectation of τ, c is the optimal value for Party 2 candidate.

(3) With the value of c and the expectation about the value of τ by voters and

Party 2 candidate, τ is the optimal for Party 1 candidate in the sense that it is the optimal for a candidate with $p > \tau$ to announce the policy a, and that it is optimal for a candidate with $p < \tau$ to announce the policy b.

(4) All expectations are correct.

1.3 Updates and Platforms

In this section, I consider behaviors of voters and party candidates in turn. For notational simplicity, I abuse notations in denoting not only the threshold value choice by Party 1 candidate but also all expectations about it as τ. I also treat τ and p interchangeably when I consider belief updates.

Voters. (see figure 1.2)

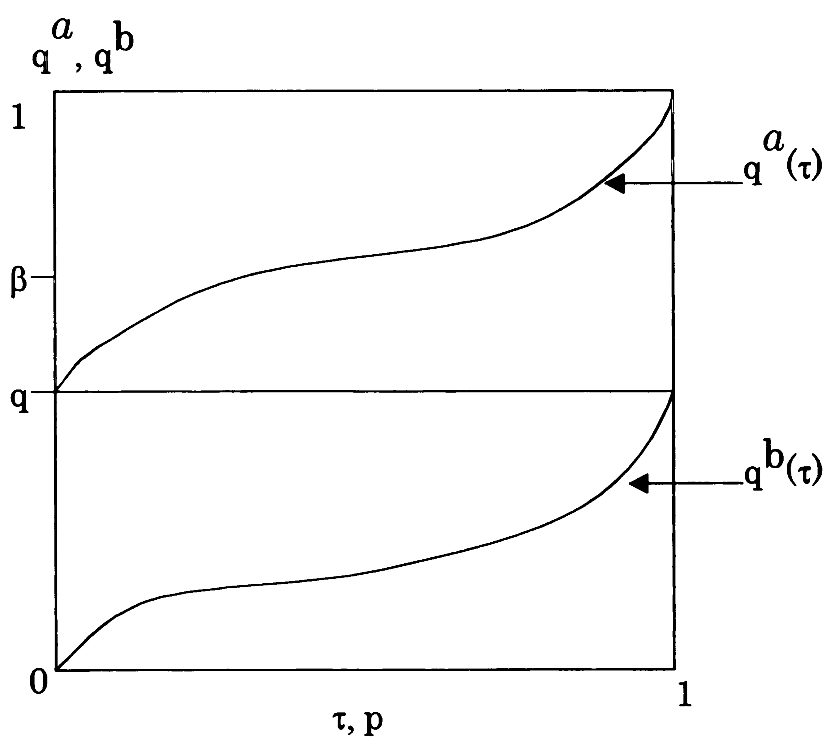

Figure 1.2

Voters update their belief about the probability of the true state being w^1 after observing the platform of Party 1 candidate. Let τ be the threshold value expected by voters.

In the case that Party 1 candidate announces the policy a, then the updated belief is,

$$(1.1)\ q^a(\tau) = \frac{q\{1-F(\tau)\}}{q\{1-F(\tau)\}+(1-q)\{1-G(\tau)\}}$$

This is well defined for $\tau \in [0,1)$.This updated belief has the following properties.

Lemma 1.1.

$$q^a(\tau) = \frac{q}{q+(1-q)\dfrac{1-G}{1-F}} > q \textit{ for } \tau \in (0,1)$$

$\lim_{\tau\to 1} q^a(\tau) = 1$

$$q^a(0) = \frac{q}{q+1-q} = q.$$

$q^{a\prime}(\tau) > 0$ *for all* $\tau \in (0,1)$

In the case that Party 1 candidate announces the policy b, the updated belief is,

$$(1.2)\ q^b(\tau) = \frac{qF(\tau)}{qF(\tau)+(1-q)G(\tau)}$$

This is well defined for $\tau \in (0,1]$. This updated belief has the following properties.

Lemma 1.2.

$$q^b(\tau) = \frac{q}{q+(1-q)\dfrac{G}{F}} < q \textit{ for } \tau \in (0,1)$$

$\lim_{\tau\to 0} q^b(\tau) = 0$

$$q^b(1) = \frac{q}{q+(1-q)} = q$$

$q^{b\prime}(\tau) \geq 0$

Between $q^a(\tau)$ and $q^b(\tau)$, there is the following relations. From Lemma 1 and 2, the following is obvious.

Lemma 1.3.

$$q^a(\tau) > q^b(\tau)$$

$$q^b(\tau) = q^a(\tau) - \frac{q^a(\tau) - q}{qF + (1-q)G} \tag{1.3}$$

Party 1 candidate. (see figure 1.3)

Party 1 candidate's choice of τ during the first election depends on three things; signal p, his expectation of voters' expectation of τ and Party 2 candidate's mixed strategy c during first election. Let q^a and q^b be Party 1 candidate's expectation of voters' updates (these come from the candidate's expectation of voters' expectation of τ.) I consider the candidate's choice of τ depending on different values of c.

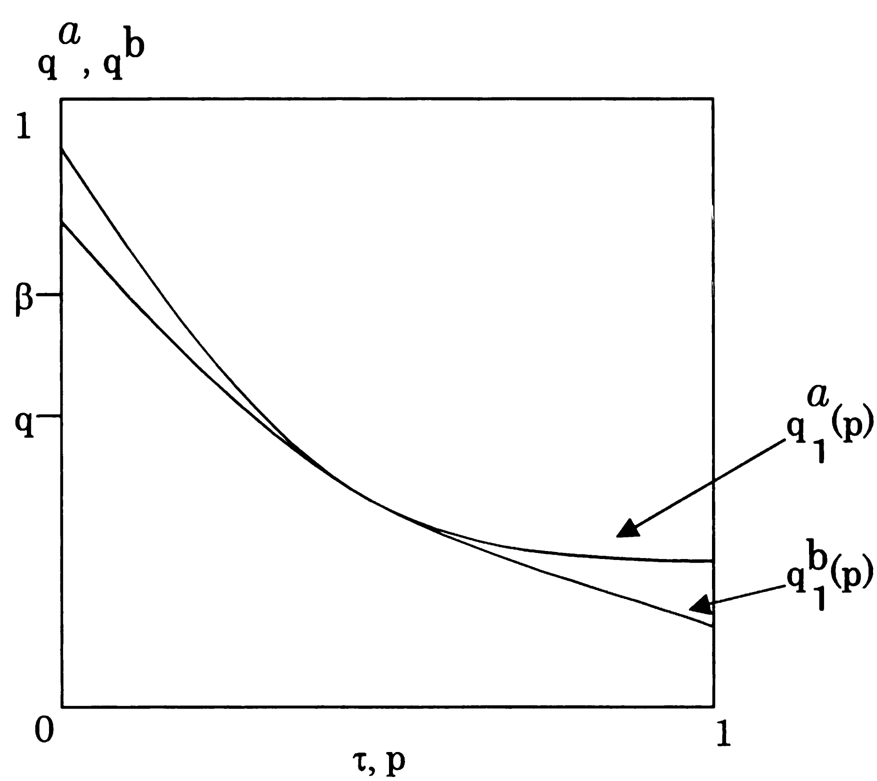

Figure 1.3

The case of $c = 0$: this is the case that Party 2 candidate announces the policy b during the first election. Then, the expected sum of rents of announcing the policy a

for Party 1 candidate with p is,

$$\{1 - H(-q^a + \beta)\}[1 + \delta p]$$

The expected sum of rents of announcing the policy b is,

$$\{1 - H(0)\}[1 + \delta\{1 - H(0)\}] = \frac{2+\delta}{4}$$

It is beneficial for the candidate to announce the policy a if,

$$\begin{aligned} &\{1 - H(-q^a + \beta)\}[1 + \delta p] \geq \frac{2+\delta}{4} \\ (1.4) \qquad &\Rightarrow 1 - \frac{2+\delta}{4(1+\delta p)} \geq H(-q^a + \beta) \end{aligned}$$

From this, the necessary minimum level of q^a for the Party 1 candidate with p to announce the policy a can be defined. Let $q_1^a(p)$ be the value of q^a defined from the above condition satisfied with equality. Then, the following result can be obtained.

Lemma 1.4.
$\dfrac{\partial q_1^a}{\partial p} < 0$
$q_1^a(0) > \beta > q_1^a(1)$

Since $q^a(0) = q < \beta$ and $q^a(1) = 1$, this Lemma implies there is a crossing point of $q_1^a()$ and $q^a()$.

The case of $c = 1$. This is the case that Party 2 candidate announces the policy a during the first election. Then, the expected sum of rents of announcing the policy a for Party 1 candidate with p is,

$$\{1 - H(0)\}[1 + \delta p] = \frac{1+\delta p}{2}$$

The expected sum of rents of announcing the policy b is,

$$\{1 - H(-\beta + q^b)\}[1 + \delta\{1 - H(0)\}] = \frac{2+\delta}{2}\{1 - H(-\beta + q^b)\}$$

It is beneficial for the candidate to announce the policy a if,

$$\frac{1+\delta p}{2} \geq \frac{2+\delta}{2}\{1 - H(-\beta + q^b)\}$$

$$\Rightarrow H(-\beta + q^b) \geq 1 - \frac{1+\delta p}{2+\delta} \tag{1.5}$$

Same as q_1^a, the necessary minimum level of q^b for a candidate with p to announce the policy a can be defined from this condition. Let $q_1^b(p)$ be the value of q^b defined from the above condition satisfied with equality. For this $q_1^b(p)$, the following results can be obtained.

Lemma 1.5.

$$\frac{\partial q_1^b}{\partial p} < 0$$

$$q_1^b(0) > \beta > q_1^b(1)$$

Unlike the case of $q^a(\tau)$ and $q_1^a(p)$, it is possible that there is no crossing point of $q^b(\tau)$ and $q_1^b(p)$. Between $q_1^a(p)$ and $q_1^b(p)$, there is a following relation.

Lemma 1.6.

$$q_1^a(p) \geq q_1^b(p)$$

The equality holds when $p = 0.5$.

The case of $c \in (0, 1)$ is the combination of above two cases.

Party 2 candidate. (see figure 1.4)

Since Party 2 candidate has the same belief q as voters, it seems at first that he always announces the policy b. But he knows that the belief will be updated after Party 1 candidate announces its platform. Following the update, the probability of Party 2 candidate winning the first election by announcing the policy b changes. If Party 2 candidate believes that it is very likely that Party 1 candidate announces the policy a and that the updated belief after that is high enough, it might be beneficial for the candidate to announce the policy a. Notice that there is a trade-off between the likelihood of Party 1 candidate announcing the policy a and the updated belief. Since Party 2 candidate does not have any private information about the probability of the true state being w^1, announcing the policy b is completely an opportunistic behavior. But, since the outcome of the policy depends on the true state not on which candidate implements the policy, voters do not punish such an opportunistic behavior in this model.

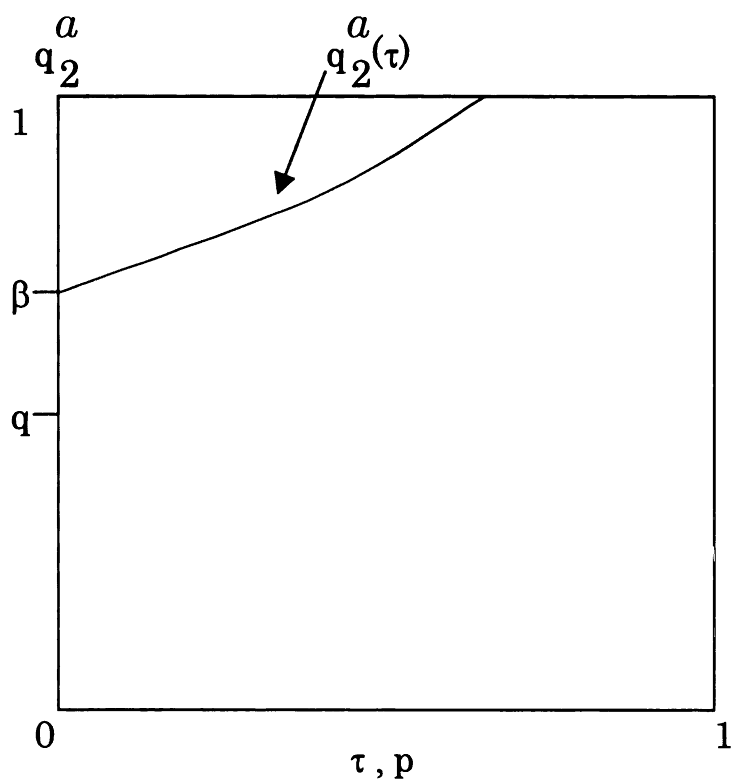

Figure 1.4

Again abusing the notation, I denote the value of τ expected by Party 2 candidate as τ. Then, Party 2 candidate thinks the probability of Party 1 candidate announcing the policy a during the first election is,

$$q\{1 - F(\tau)\} + (1 - q)\{1 - G(\tau)\} = 1 - qF(\tau) - (1 - q)G(\tau)$$

The expected sum of rents of Party 2 candidate critically depends on the assumptions about what happens after Party 2 candidate winning the first election. I restate two related assumptions. The new candidate put up by Party 1 at the second election after the defeat of the Party at the first election does not have any private information (actually it does not matter whether or not he has private information since no information can be transmitted credibly to voters during the second election.) Voters do not apply (A) to Party 2 candidate even if the candidate implements the policy a in the first period.

Then, the expected sum of rents for Party 2 candidate when the candidate announces the policy a during the first election is,

$$\begin{aligned}&\{1-qF(\tau)-(1-q)G(\tau)\}H(0)[1+\delta H(0)]\\&\quad+\{qF(\tau)+(1-q)G(\tau)\}H(-\beta+q^b(\tau))[1+\delta H(0)]\\&=[0.5+\{qF+(1-q)G\}\{H(-\beta+q^b)-0.5\}]\frac{2+\delta}{2}\end{aligned}$$

The expected sum of rents when the candidate announces the policy b is,

$$\begin{aligned}&\{1-qF(\tau)-(1-q)G(\tau)\}H(-q^a+\beta)[1+\delta H(0)]\\&\quad+\{qF(\tau)+(1-q)G(\tau)\}H(0)[1+\delta H(0)]\\&=[H(-q^a+\beta)+\{qF+(1-q)G\}\{0.5-H(q^a+\beta)\}]\frac{2+\delta}{2}\end{aligned}$$

Then, the candidate announces the policy a at the first election if,

$$\begin{aligned}&0.5+\{qF+(1-q)G\}\{H(-\beta+q^b)-0.5\}\\&\quad\geq H(-q^a+\beta)+\{qF+(1-q)G\}\{0.5-H(q^a+\beta)\}\end{aligned}$$

$$\begin{aligned}(1.6)\quad&\Rightarrow\{qF+(1-q)G\}\{H(-\beta+q^b)+H(-q^a+\beta)-1\}\\&\qquad\geq H(-q^a+\beta)-0.5\end{aligned}$$

Unlike two conditions for Party 1 candidate, this condition (1.6) has both q^a and q^b. Using (1.3), this condition is converted into the one with only q^a.

$$(1.7)\ \{qF + (1-q)G\}\{H(-\beta + q^a - \frac{q^a - q}{qF + (1-q)G}) + H(-q^a + \beta) - 1\} \geq H(-q^a + \beta) - 0.5$$

Then, the minimum necessary level of q^a for a given τ is defined as the value of q^a satisfying this condition with equality. Denote the minimum level as $q_2^a(\tau)$. Since this condition is not easy to handle, I consider also the condition with a given value of q^b. Given the fixed value of q^b, the minimum necessary level of q^a for a given τ is obtained from the condition satisfied with equality. Denote the minimum necessary level for a given q^b as $q_2^a(\tau; q^b)$. Then, the following Lemma can be obtained.

Lemma 1.7

For $\tau \in (0,1)$,

$$\frac{\partial q_2^a(\tau; q^b)}{\partial \tau} > 0$$

$$\frac{\partial q_2^a(\tau; q^b)}{\partial q^b} < 0$$

$$q_2^a(\tau) \geq q_2^a(\tau; q^b)$$

$$q_2^a(0; q^b) = q_2^a(0) = \beta$$

the value of τ *such that* $q_2^a(\tau) = 1$ *is lower than* 1.

1.4 Equilibrium

Firstly, I show there is no equilibrium without voting behavior (A) in this model.

Lemma 1.8. *Suppose that voters vote during the second election according to their period preference without* (A). *Then, there is no equilibrium.*

Thus, (A) is a necessary condition in this model. Next, I state the Proposition of the equilibrium existence.

Proposition 1.1. *With* (A), *there is an equilibrium in the model.*

There are four possible cases of the existence of equilibrium. Since these cases are rather complicated and do not give clear and intuitive characterizations of equilibrium, I concentrate on only one case that $q_2^a(\tau)$ and $q^a(\tau)$ do not cross (see figure 1.5.)

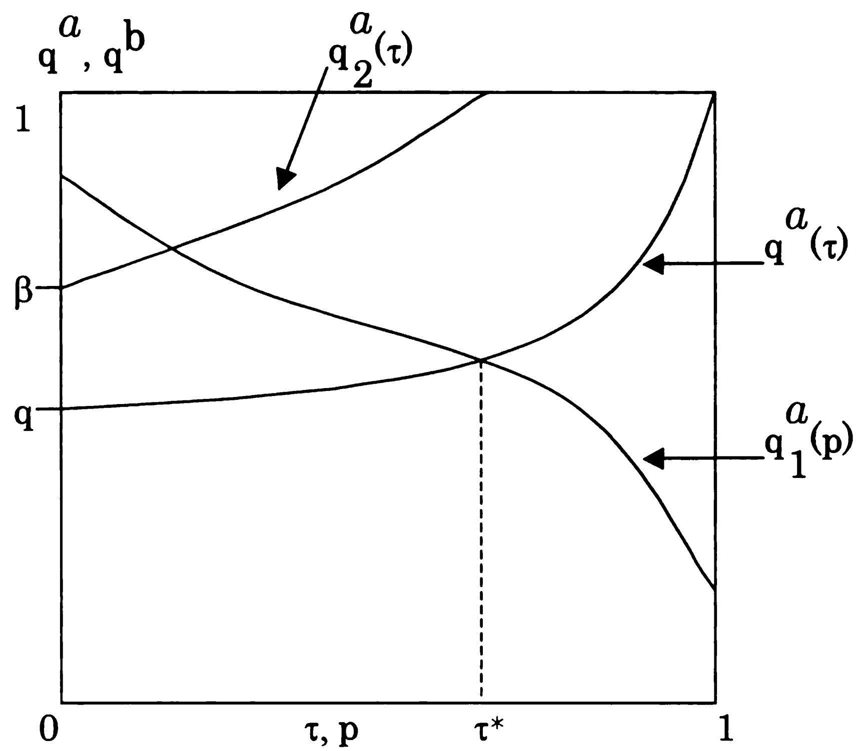

Figure 1.5

In this case, the equilibrium value of τ and c are τ^* such that $q^a(\tau^*) = q_1^a(\tau^*)$ and $c = 0$. Party 2 candidate never proposes the policy a since $q_2^a(\tau) > q^a(\tau)$; $q^a(\tau)$ is never high enough for Party 2 candidate to propose the policy. For this case, there is a simple sufficient condition. Let $H(\varepsilon)$, the distribution of ε, be the uniform distribution (its support is assumed to be large enough.) Then, (1.7) becomes $q \geq \beta$. Since $\beta > q$, $q_2^a(p)$ is always larger than $q^a(p)$ and Party 2 candidate never announces the policy a. Thus, the case like figure 1.5 is obtained. The equilibrium choice of τ is determined at the crossing of $q^a(\tau)$ and $q_1^a(p)$. Denote the equilibrium value of τ as τ^*. Then, there is no guarantee that τ^* is equal to β. τ^* could be higher or lower than β. Whichever it is, that means the inefficient choice of Party 1 candidate for voters.

In this case, a more interesting result can be obtained; the higher the ability of Party 1 candidate to discern the true state, the lower the equilibrium outcome value of τ. I give a more formal explanation. As I wrote before, the cumulative distribution functions F and G can be considered as representing the ability of Party 1 candidate. Then, a high ability candidate is the one who is more likely to receive high values of p when the true state is w^1 and less likely to receive high values of p when the true state is w^0. Let $\bar{F}\&\bar{G}$ and $F\&G$ be two different sets of distributions. Then, if $\bar{F}(p) \leq F(p)$ and $\bar{G}(p) \geq G(p)$ for any $p \in (0,1)$ and if $\bar{F}(p) < F(p)$ or $\bar{G}(p) > G(p)$ for some $p \in (0,1)$, then the candidate with $\bar{F}\&\bar{G}$ can be considered as having the higher ability than the one with $F\&G$. Since voter's updated belief is,

$$\begin{aligned} q^a(\tau) &= \frac{q\{1-F(\tau)\}}{q\{1-F(\tau)\}+(1-q)\{1-G(\tau)\}} \\ &= \frac{q}{q+(1-q)\dfrac{1-G(\tau)}{1-F(\tau)}} \end{aligned}$$

the updated belief with the higher ability candidate is never smaller than the updated belief with the lower ability candidate, and strictly higher for some τ.

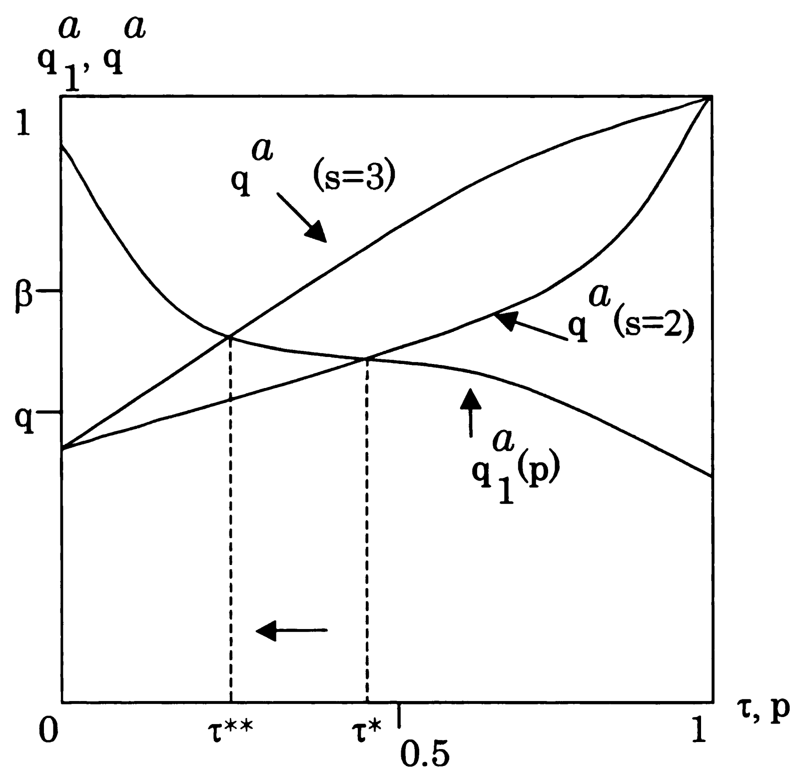

Figure 1.6

In the figure 1.6, the two different curves of voter's updated beliefs, q^a for $s = 2$ and q^a for $s = 3$ are drawn based on $F(p) = p^s$ and $G(p) = 1 - (1-p)^s$. With these function forms, higher value of s means higher ability. On $(0,1)$, the curve of q^a $(s = 3)$ is above the curve of q^a $(s = 2)$. Since the change of F and G does not affect q_1^a, the equilibrium value of τ moves to left. Since $\tau^* < \beta = 0.6$ in figure 1.6, the outcome with $s = 2$ is inefficient as the policy b can guarantee the utility β to voters. The candidate with $p \in (\tau^*, \beta)$ exploits the voters' trust on the ability of the candidate. Since $\tau^{**} < \tau^*$, with the higher ability candidate the inefficiency worsens in this case. I state this as a Proposition.

Proposition 1.2. *If H is a uniform distribution, the equilibrium exists with τ^* such that $q^a(\tau^*) = q_1^a(\tau^*)$ and $c = 0$. Moreover, in this case, the higher the ability of Party 1 candidate, the lower τ^*.*

This proposition is not saying that having a higher ability candidate is always bad. If the ability of current candidate is so low that τ^* with him is to the right of β, changing him with a higher ability candidate could help voters with τ^* moving closer to β. However, changing candidates with higher and higher ability candidates, τ^* eventually passes β. After that, higher ability candidates make things worse for voters.

1.5 Conclusion of Chapter 1

In this chapter, I considered the platform choice in a two-election-and-two-period model under uncertainty with one party having information advantage. Most papers in the literature of the policy choice with the uncertainty consider models with only one election. Since government offices usually allow multiple terms with multiple elections, candidates must consider about the future elections. Thus it is worth considering the policy choice of candidates considering future elections as well as the current election. Harrington (1993b) also considers a two-election model. However, the interest of the paper is whether politicians will keep their campaign platform after they win the election, which is different from this chapter since the interest of this chapter is on what platform a politician will choose.

This chapter assumes strong assumptions. The assumption that only one party receives the signal is clearly a strong one. In the literature of the information transmission during the election, the more common assumption is that both parties receive signals (Heidhues and Lagerof (2003) and Bernhardt, Duggan and Squintani (2007), etc.) However, it is not clear that which is more natural assumption. During elections, often one party has advantage over the other party (see the Rasmussen reports (June 21, 2008).) The assumption is a simplified description of such reality. The other strong assumption is that voters apply (A) only to Party 1 candidate. When Party 2 candidate implements the policy a and its outcome is failure, the voters does not punish the candidate. Since Party 2 candidate's choice of policy a is completely an opportunistic behavior, this seems unrealistic. It seems very likely that voters punish the candidate strongly. However, it does not change results qualitatively.

The equilibrium value of τ depends on functions and parameters. So, there is no reason to expect $\tau^* = \beta$. If $\tau^* < \beta$, the candidate with $p \in (\tau^*, \beta)$ announces the policy a while believing the policy is actually worse than the policy b. Thus, this could be considered as the inefficiency result. Similar inefficiency happens when $\tau^* > \beta$. Then the interesting result of this chapter that the higher the candidate's ability the lower the equilibrium value of τ in the case I explained in the previous section means that a higher ability candidate could make the inefficiency worse. This happens because the high ability of the candidate causes voters to update their belief higher, which gives more opportunity to the candidate to exploit. Though the result is interesting, to obtain more general result, it might be necessary to consider if the same result would happen even when two parties receive signals, which is a future research topic.

Appendix of Chapter 1

Lemma 1.1.

$$q^a(\tau) = \frac{q}{q + (1-q)\dfrac{1-G}{1-F}} > q \; for \; \tau \in (0,1)$$

$\lim_{\tau \to 1} q^a(\tau) = 1$

$$q^a(0) = \frac{q}{q+1-q} = q.$$

$q^{a\prime}(\tau) > 0 \; for \; all \; \tau \in (0,1)$

Proof. Since $G(p) > F(p)$, $\dfrac{1-G(p)}{1-F(p)} < 1$ for $p \in (0,1)$. Then, $q^a(\tau) = \dfrac{q}{q + (1-q)\dfrac{1-G}{1-F}} > q$ for $\tau \in (0,1)$, and $q^a(0) = \dfrac{q}{q+1-q} = q$. From L'Hôpital's rule and $\lim_{\tau \to 1} \dfrac{F'(\tau)}{G'(\tau)} = \infty$, $\lim_{\tau \to 1} q^a(\tau) = 1$. The derivative of $q^a(\tau)$ with respect to τ is,

$$q^{a\prime}(\tau) = \frac{q\{1-q\}}{[q\{1-F\} + (1-q)\{1-G\}]^2} F'\{1-G\}[\frac{G'}{F'}\frac{1-F}{1-G} - 1]$$

Since $\lim_{\tau \to 0} \dfrac{F'}{G'} = 0$ and $\dfrac{1-F}{1-G} > 1$ for $\tau \in (0,1)$, $q^{a\prime}(\tau)$ is positive for low value of τ. Suppose that $q^{a\prime}(t) \leq 0$ at $\tau = t \in (0,1)$. This implies $\dfrac{G'}{F'}\dfrac{1-F}{1-G} - 1 \leq 0$. Then, $\dfrac{d}{d\tau}\dfrac{1-F}{1-G} = \dfrac{F'}{1-G}[\dfrac{G'}{F'}\dfrac{1-F}{1-G} - 1] \leq 0$. With this, $\dfrac{d}{d\tau}[\dfrac{G'}{F'}\dfrac{1-F}{1-G}] = \dfrac{1-F}{1-G}\dfrac{d}{d\tau}\dfrac{G'}{F'} + \dfrac{G'}{F'}\dfrac{d}{d\tau}\dfrac{1-F}{1-G} < 0$ since $\dfrac{d}{d\tau}\dfrac{G'}{F'} < 0$. This implies, if $\dfrac{G'}{F'}\dfrac{1-F}{1-G} - 1 \leq 0$at a certain value of τ, for all $\tau' \geq \tau$, the value of $\dfrac{G'}{F'}\dfrac{1-F}{1-G} - 1 < 0$. Then, $q^a(\tau)$ is downward sloping for $\tau' \geq \tau$. Since $\lim_{\tau \to 1} q^a(\tau) = 1$, this is a contradiction. Thus $q^{a\prime}(\tau) > 0$ for all $\tau \in (0,1)$. The updated belief $q^a(\tau)$ goes up from q to 1 as increases from 0 to 1. ■

Lemma 1.2.

$$q^b(\tau) = \frac{q}{q + (1-q)\dfrac{G}{F}} < q \; for \; \tau \in (0,1)$$

$\lim_{\tau \to 0} q^b(\tau) = 0$

$$q^b(1) = \frac{q}{q + (1-q)} = q$$

$q^{b\prime}(\tau) \geq 0$

Proof. From the definition, $q^b(\tau) = \dfrac{q}{q+(1-q)\dfrac{G}{F}} < q$ for $\tau \in (0,1)$. Clearly, $q^b(1) = \dfrac{q}{q+(1-q)} = q$. From $\lim_{\tau\to 0}\dfrac{F'(\tau)}{G'(\tau)} = 0$, $\lim_{\tau\to 0} q^b(\tau) = 0$. The derivative of $q^b(\tau)$ with respect to τ is $q^{b\prime}(\tau) = \dfrac{q\{1-q\}}{[qF+(1-q)G]^2}FG'[\dfrac{F'}{G'}\dfrac{G}{F} - 1]$. Since $\lim_{\tau\to 1}\dfrac{F'}{G'} = \infty$ and $\dfrac{G}{F} > 1$ for $\tau \in (0,1)$, $q^{b\prime}(\tau)$ is positive for τ close to 1. Moreover, $\lim_{\tau\to 0}\dfrac{F'}{G'}\dfrac{G}{F} - 1 = \lim_{\tau\to 0}\dfrac{F'}{G'}[\lim_{\tau\to 0}\dfrac{G}{F} - \lim_{\tau\to 0}\dfrac{G'}{F'}] = \lim_{\tau\to 0}\dfrac{F'}{G'}[\lim_{\tau\to 0}\dfrac{G}{F} - \lim_{\tau\to 0}\dfrac{G'}{F'}] = 0$. Suppose that $q^{b\prime}(t) = 0$ at $\tau = t \in (0,1)$. This means $\dfrac{F'}{G'}\dfrac{G}{F} - 1 = 0$. Then, $\dfrac{d}{d\tau}\dfrac{G}{F} = \dfrac{G'F}{F^2}[1 - \dfrac{F'}{G'}\dfrac{G}{F}] = 0$. With this, $\dfrac{d}{d\tau}[\dfrac{F'}{G'}\dfrac{G}{F}] = \dfrac{G}{F}\dfrac{d}{d\tau}\dfrac{F'}{G'} > 0$ since $\dfrac{d}{d\tau}\dfrac{F'}{G'} > 0$. Since $q^b(0) = 0$, this guarantees $q^{b\prime}(\tau) \geq 0$. Thus, the updated belief $q^b(\tau)$ goes up from 0 to q as τ increases from 0 to 1. ■

Lemma 1.4.

$\dfrac{\partial q_1^a}{\partial p} < 0$

$q_1^a(0) > \beta > q_1^a(1)$

Proof. (1.4) $1 - \dfrac{2+\delta}{4(1+\delta p)} \geq H(-q^a + \beta)$

The left hand side of (1.4) is the increasing function of p. The right hand side of (1.4) is decreasing function of q^a as $H'(\varepsilon) > 0$. Thus, from the implicit function theorem, the derivative of $q_1^a(p)$, $\dfrac{\partial q_1^a}{\partial p} = \dfrac{\delta\{2+\delta\}}{4\{1+\delta p\}^2 H'(-q^a+\beta)} < 0$. To obtain $q_1^a(0) > \beta > q_1^a(1)$, substitute $p = 0$ into the left hand side of (1.4), $1 - \dfrac{(2+\delta)}{4} < 0.5$. Since $H(0) = 0.5$, $-q^a + \beta < 0$. Thus, $q_1^a(0) > \beta$. Substituting $p = 1$ into the right hand side of (1.4), $H(-q_1^a(1) + \beta) = 1 - \dfrac{2+\delta}{4(1+\delta)} > 0.5$. Then, $\beta > q_1^a(1)$. ■

Lemma 1.5.

$\dfrac{\partial q_1^b}{\partial p} < 0$

$q_1^b(0) > \beta > q_1^b(1)$

Proof. (1.5) $H(-\beta + q^b) \geq 1 - \dfrac{1+\delta p}{2+\delta}$

The left hand side of (1.5) is the increasing function of q^b. The right hand side of (1.5) is decreasing function of p. Thus, from the implicit function theorem, the derivative of $q_1^b(p)$, $\frac{\partial q_1^b}{\partial p} = -\frac{\delta}{(2+\delta)H'(-\beta+q^b)} < 0$. Moreover, $q_1^b(0) > \beta > q_1^b(1)$. To get this result, substitute $p = 0$ into the right hand side of (1.5), $1 - \frac{1}{2+\delta} > 0.5$. Then, $-\beta + q^b > 0 \Rightarrow q_1^a(0) > \beta$. Substituting $p = 1$ into the right hand side of (1.5), $1 - \frac{1+\delta}{2+\delta} < 0.5$. Then, $H(-\beta + q_1^b(1)) < 0.5 \Rightarrow q_1^b(1) < \beta$. ■

Lemma 1.6.

$q_1^a(p) \geq q_1^b(p)$

The equality holds when $p = 0.5$.

Proof. To show this result, solve (1.4) and (1.5) satisfied with equality for $q_1^a(p)$ and $q_1^b(p)$. Then, $q_1^a(p) = \beta - H^{-1}(1 - \frac{2+\delta}{4(1+\delta p)})$, and $q_1^b(p) = \beta + H^{-1}(1 - \frac{1+\delta p}{2+\delta})$. Thus $q_1^a(p) - q_1^b(p) = -\{H^{-1}(1 - \frac{2+\delta}{4(1+\delta p)}) + H^{-1}(1 - \frac{1+\delta p}{2+\delta})\}$. Since H is symmetric and $H^{-1}(k) \geq (<) 0 \Leftrightarrow k \geq (<) 0.5$, the sigh of the left hand side is same to the sign of, $-\{1 - \frac{2+\delta}{4(1+\delta p)} - 0.5\} - \{1 - \frac{1+\delta p}{2+\delta} - 0.5\} = \frac{\delta^2}{4(1+\delta p)(2+\delta)}(2p-1)^2$
$\Rightarrow q_1^a(p) \geq q_1^b(p)$ (the equality holds at $p = 0.5$.) ■

Lemma 1.7.

For $\tau \in (0,1)$,

$$\frac{\partial q_2^a(\tau; q^b)}{\partial \tau} > 0$$

$$\frac{\partial q_2^a(\tau; q^b)}{\partial q^b} < 0$$

$$q_2^a(\tau) \geq q_2^a(\tau; q^b)$$

$$q_2^a(0; q^b) = q_2^a(0) = \beta$$

the value of τ *such that* $q_2^a(\tau) = 1$ *is lower than* 1.

Proof. (1.6) $\{qF + (1-q)G\}\{H(-\beta + q^b) + H(-q^a + \beta) - 1\} \geq H(-q^a + \beta) - 0.5$

The derivatives of $q_2^a(\tau; q^b)$ with respect to τ and q^b for $\tau \in (0,1)$ are, $\frac{\partial q_2^a(\tau; q^b)}{\partial \tau} =$

$$\frac{\{qF' + (1-q)G'\}\{1 - H(-\beta + q^b) - H(-q^a + \beta)\}}{\{1 - qF - (1-q)G\}H'(-q^a + \beta)}$$

The sign of this is same to the sign of $1 - H(-\beta + q^b) - H(-q^a + \beta)$. Since

$\beta \geq q^b$, $q^a > \beta$ guarantees that the sign is positive. Unfortunately, it is possible that $\beta \geq q^a$. Suppose this. Since H is symmetric, the sign is same to $\beta - q^b + q^a - \beta = q^a - q^b$. The sign is, $q^a - q^b = q^a - \dfrac{q - \{1 - qF - (1-q)G\}q^a}{qF + (1-q)G} = \dfrac{q^a - q}{qF + (1-q)G} \geq 0$ Only case this is zero is $q^a - q = 0$. This implies $\tau = 0$. It is easy to obtain $\dfrac{\partial q_2^a(\tau; q^b)}{\partial q^b} = -\dfrac{\{qF + (1-q)G\}H'(-\beta + q^b)}{\{1 - qF - (1-q)G\}H'(-q^a + \beta)} < 0$. Then, $\dfrac{\partial q_2^a(\tau; q^b)}{\partial q^b} < 0$ implies $q_2^a(\tau) \geq q_2^a(\tau; q)$. If $\tau = 0$, the condition (1.6) becomes, $0.5 \geq H(-q^a + \beta) \Rightarrow q^a \geq \beta$. This does not depend on the value of q^b. So, $q_2^a(0; q^b) = q_2^a(0) = \beta$. If $\tau = 1$, the condition (1.6) becomes,

$$
\begin{aligned}
& H(-\beta + q^b) + H(-q^a + \beta) - 1 \geq H(-q^a + \beta) - 0.5 \\
& \Rightarrow H(-\beta + q^b) \geq 0.5 \\
& \Rightarrow q^b \geq \beta
\end{aligned}
$$

Since$\beta > q \geq q^b$, this condition is never satisfied. This implies that the value of τ such that $q_2^a(\tau) = 1$ is lower than 1. ∎

Lemma 1.8. *Suppose that voters vote during the second election according to their period preference without* (A). *Then there is no equilibrium.*

Proof. No (A) means that the probability of winning second election after the implementation of the policy a by Party 1 candidate is same to the one after the implementation of the policy b. Thus, the expected second period rent is same for both cases. This implies the decision-making during the first election is same to the case there is only one period. Then, there is no reason for voters to update their belief. This is the reason of no equilibrium. The following is a formal proof.

Suppose that there is an equilibrium without (A). Let τ^* and c^* be the equilibrium value of τ and c. Let q^{a*} and q^{b*} be the updated beliefs at the equilibrium.

The expected sum of rents for announcing the policy a is,

$$(1 - c^*)\{1 - H(-q^{a*} + \beta)\}[1 + \delta\{1 - H(0)\}] + c^*\{1 - H(0)\}[1 + \delta\{1 - H(0)\}]$$

The expected sum of rents for announcing the policy b is,

$$(1-c^*)\{1-H(0)\}[1+\delta\{1-H(0)\}]+c^*\{1-H(-\beta+q^{b*})\}[1+\delta\{1-H(0)\}]$$

The candidate with p announces the policy a if,

$$\begin{aligned}
&(1-c^*)\{1-H(-q^{a*}+\beta)\}+c^*\{1-H(0)\} \\
&\quad \geq (1-c^*)\{1-H(0)\}+c^*\{1-H(-\beta+q^{b*})\} \\
&\Rightarrow (1-c^*)\{H(0)-H(-q^{a*}+\beta)\}+c^*\{H(-\beta+q^{b*})-H(0)\}\geq 0 \\
&\Rightarrow H(0)-H(-q^{a*}+\beta)+c^*\{H(-\beta+q^{b*})+H(-q^{a*}+\beta)-1\}\geq 0
\end{aligned}$$

Since this condition does not depend on p, if this condition holds for some Party 1 candidate with p, it holds for any Party 1 candidate. This implies $\tau = 0$. Since $q^{a*}(0) < q < \beta$, $H(0)-H(-q^{a*}+\beta) < 0$. As shown in a previous footnote, $H(-\beta+q^{b*})+H(-q^{a*}+\beta)-1\leq 0$. Thus the condition does not hold. This implies that Party 1 candidate never announces the policy a. Then, it has to be $\tau^* = 1$. In this case, c^* has to be 0. Since $H(0)-H(-q^a(1)+\beta) > 0$, the condition satisfied, a contradiction. ■

Proposition 1.1. *With (A), there is an equilibrium in the model.*

Proof. Since different sets of functional forms and parameters produce different kind of equilibrium, I cannot obtain a simple general result. So I am going to show the existence of equilibrium in all possible cases in turn.

The first case is the simplest one. Since $q_2^a(0) = \beta$ and $q_2^a(\tau)$ reaches 1 before reaches 1, and $q^a(0) = q$ and $q^a(1) = 1$, it is possible that $q_2^a(\tau)$ and $q^a(\tau)$ do not cross.

The Case 1: $q_2^a(\tau)$ and $q^a(\tau)$ do not cross.

Since $q_1^a(0) > \beta > q_1^a(1)$ and $q^a(0) = q$ and $\lim_{\tau\to 1} q^a(\tau) = 1$, $q_1^a(p)$ and $q^a(\tau)$ always cross. Moreover, since they are monotone, the crossing point is unique. Let the point be (τ^*, q^{a*}). Then, $\tau = \tau^* \in (0,1)$ and $c = 0$ is the equilibrium. Since the assumption that $q_2^a(\tau)$ and $q^a(\tau)$ do not cross implies $q_2^a(\tau) > q^a(\tau)$, clearly it is never beneficial for Party 2 candidate to announce the policy a during the first election. Thus $c = 0$. This implies that $q_1^b(p)$ is not relevant to Party 1 candidate's choice of platform. Only $q_1^a(p)$ is relevant. Since $\frac{dq_1^a}{d\tau} < 0$, for $p < \tau^*$, the updated belief q^{a*} is lower than the necessary minimum level of the updated belief $q_1^a(p)$. Thus, the candidate with $p < \tau^*$ does not announce the policy a. On the contrary, for the candidate $p > \tau^*$, $q^{a*} > q_1^a(p)$. Thus the candidate with p announces the policy a. So, the set of $\tau = \tau^*$ and $c = 0$ is the equilibrium.

The case of that $q_2^a(\tau)$ and $q^a(\tau)$ cross is more complicated. In the case, both $q_1^a(p)$ and $q_1^b(p)$ are relevant to the choice of Party 1 candidate, but not the entire part of them. Let $\Omega \equiv \{\tau \in [0,1] : q_2^a(\tau) \geq q^a(\tau)\}$. Then, on Ω, $q_1^b(p)$ is relevant to Party 1 candidate's choice of platform because on Ω Party 2 candidate announces the policy a.[4] On $\Omega^c \equiv \{\tau \in [0,1] : q_2^a(\tau) < q^a(\tau)\}$, $q_1^a(p)$ is relevant to Party 1 candidate's choice of its platform since Party 2 candidate does not announce the policy a. Depending on functional forms and parameter values, this Ω may be a convex set or may be a collection of disjoint convex sets. Let $\Omega^* \subset \Omega$ be a convex set such that $\omega^* \equiv \min \omega \in \Omega^*$ and $\omega^{**} \equiv \max \omega \in \Omega^*$ are boundary points of Ω. Since $q_2^a(\tau)$ reaches 1 before τ reaches 1, $\omega^{**} < 1$. It may be or may not be $\Omega^* = \Omega$. To show the existence of the equilibrium, I only need to consider Ω^*. So, in the following, I simply assume $\Omega = \Omega^*$.

The Case 2 (Figure 1.7): $q_2^a(\tau)$ and q^a cross, and $q_1^a(p)$ and $q^a(\tau)$ cross on Ω^c and/or $q_1^b(p)$ and $q^b(\tau)$ cross on Ω.

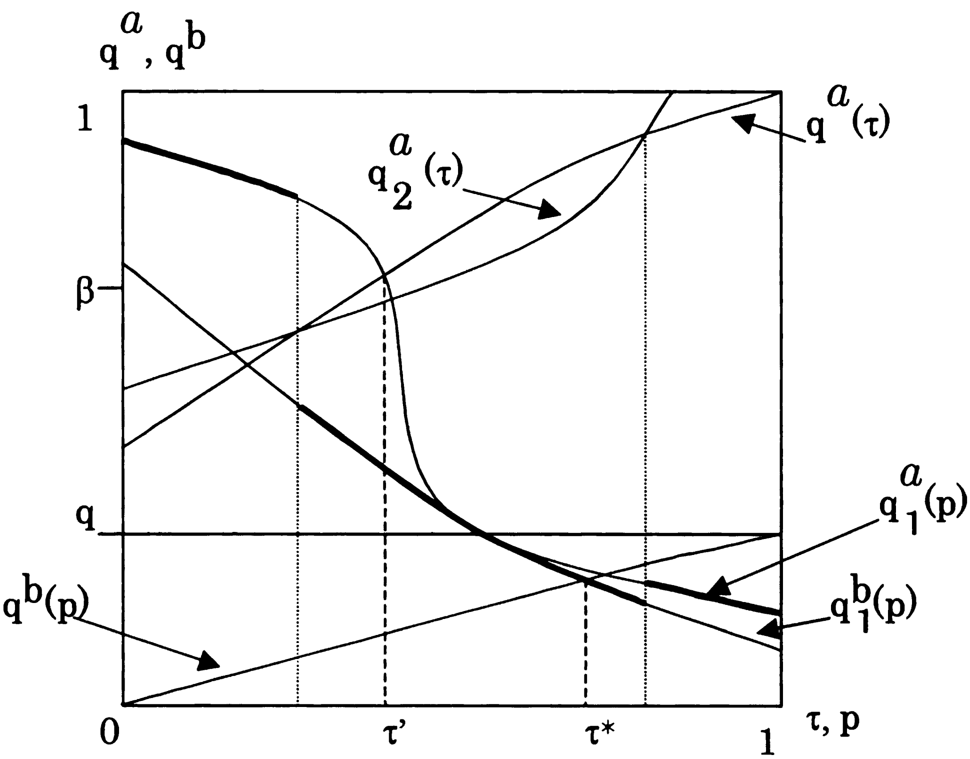

Figure 1.7

Suppose $q_1^a(p)$ and $q^a(\tau)$ cross on Ω^c. Let the crossing point be $(\tau', q^{a\prime})$. Since $\tau' \in \Omega^c$, $q_2^a(\tau') < q^{a\prime}$. This implies that Party 2 candidate does not announce the policy a at $\tau = \tau'$. Then, with the same logic in the case 1, it can be proven that $\tau = \tau'$ and $c = 0$ is the equilibrium.

Next suppose that $q_1^b(p)$ and $q^b(\tau)$ cross on Ω. Let the crossing point be (τ^*, q^{b*}). Since $\tau^* \in \Omega$, $q_2^a(\tau^*) \geq q^{a*}$. So, with $\tau = \tau^*$, Party 2 candidate announces the policy a. With Party 2 candidate announcing the policy a, among $q_1^a(p)$ and $q_1^b(p)$, only $q_1^b(p)$ is relevant to Party 1 candidate's choice of platform at $p = \tau^*$ (remember that τ is the value in $[0, 1]$. Both p and τ take their value on the same axis.) Since $\dfrac{dq_1^b}{d\tau}$ is negative, the same logic used in the case 1 shows that the set of $\tau = \tau^*$ and $c = 1$ is the equilibrium.

As implied in the Case 2 above, it is possible to have equilibria. Next I consider the case such that $q_1^a(p)$ and $q^a(p)$ do not cross on Ω^c and $q_1^b(p)$ and $q^b(p)$ do not cross on Ω. This case also is divided into two cases. Firstly, notice that since $q_1^a(p) \geq q_1^b(p)$ and $q^a(\tau) > q^b(\tau)$, that $q_1^a(p)$ and $q^a(p)$ do not cross on Ω^c implies that $q_1^a(\omega^*) > q^a(\omega^*)$ and $q^a(\omega^{**}) > q_1^a(\omega^{**})$. Then two cases are such that; the case of $q^b(\omega^*) > q_1^b(\omega^*)$ and the case that $q_1^b(\omega^{**}) > q^b(\omega^{**})$.

The Case 3 (Figure 1.8): $q_2^a(\tau)$ and q^a cross, and $q^b(\omega^*) > q_1^b(\omega^*)$, no crossing between $q_1^a(p)$ and $q^a(\tau)$ on Ω^c nor between $q_1^b(p)$ and $q^b(\tau)$ on Ω.

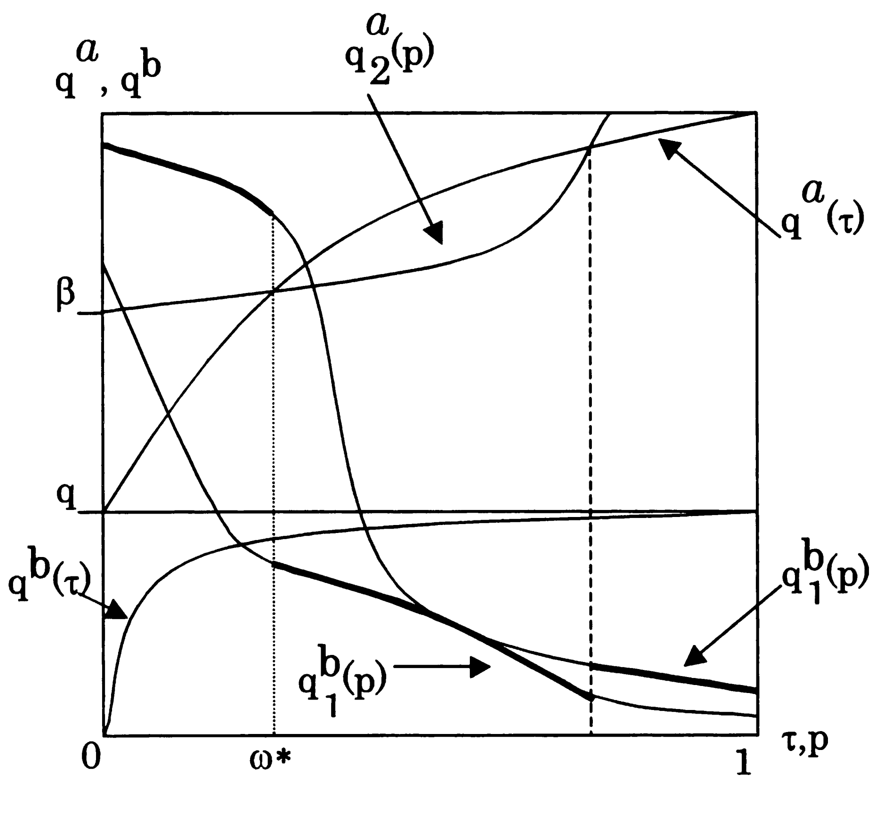

Figure 1.8

In this case, $\tau = \omega^*$ and some mixed strategy $c \in (0,1)$ is an equilibrium. Since $q_1^a(\omega^*) > q^a(\omega^*)$, if $c = 0$, announcing the policy a is not beneficial for Party 1 candidate. But $q^b(\omega^*) > q_1^b(\omega^*)$ implies that if $c = 1$ it is beneficial for Party 1 candidate to announce the policy a. Since Party 1 candidate with $p = \omega^*$ is not indifferent between announcing the policy a and announcing the policy b if $c = 0$ or $c = 1$, $\tau = \omega^*$ and $c = 0$ (or $c = 1$) is not an equilibrium. But there is a value $c^* \in (0, 1)$ such that $\tau = \omega^*$ and $c = c^*$ is the equilibrium. Let $q^{a*} \equiv q^a(\omega^*)$ and $q^{b*} = q^b(\omega^*)$. Given some mixed strategy c, the expected sum of rents of Party 1 candidate for announcing the policy a is,

$$(1-c)\{1 - H(-q^{a*} + \beta)\}\{1 + \delta p\} + c\frac{1+\delta p}{2}$$

The expected sum of rents for announcing the policy b is,

$$(1-c)\frac{2+\delta}{4} + c\frac{2+\delta}{2}\{1 - H(-\beta + q^{b*})\}$$

Party 1 candidate announces the policy a if,

$$\begin{aligned}
&(1-c)\{1 - H(-q^{a*} + \beta)\}\{1 + \delta p\} + c\frac{1+\delta p}{2} \\
&\quad \geq (1-c)\frac{2+\delta}{4} + c\frac{2+\delta}{2}\{1 - H(-\beta + q^{b*})\} \\
&\Rightarrow (1-c)[\{1 - H(-q^{a*} + \beta)\}\{1 + \delta p\} - \frac{2+\delta}{4}] \\
&\quad + c[\frac{1+\delta p}{2} - \frac{2+\delta}{2}\{1 - H(-\beta + q^{b*})\}] \geq 0
\end{aligned}$$

With $\tau = \omega^*$, the coefficient of $(1-c)$ is negative, and the coefficient of c is positive. Denote them as $(-)$ and $(+)$. Then, the condition is,

$$(1-c)(-) + c(+) = (-) + c\{(+) - (-)\} \geq 0$$

From this, the value $c^* \equiv \dfrac{-(-)}{(+) - (-)} \in (0, 1)$ is defined as the value that satisfies the condition with equality. Then the set of $\tau = \omega^*$ and $c = c^*$ is an equilibrium. Party 1 candidate with $p = \omega^*$ is indifferent between announcing the policy a and announcing the policy b. Since both $q_1^a(p)$ and $q_1^b(p)$ are decreasing functions of p, the values of $(+)$ and $(-)$ above increase with p. Thus, Party 1 candidate with $p \geq \omega^*$ announces the policy a. At same time, this means that Party 1 candidate with $p < \omega^*$ does not announce the policy a. For Party 2 candidate, since the fact that ω^* is a point of

Ω means $q_2^a(\omega^*) = q^a(\omega^*)$, Party 2 candidate is indifferent between announcing the policy a and announcing the policy b.

The Case 4 (Figure 1.9): $q_2^a(\tau)$ and q^a cross, and $q_1^b(\omega^{**}) > q^b(\omega^{**})$, no crossing between $q_1^a(p)$ and $q^a(\tau)$ on Ω^c nor between $q_1^b(p)$ and $q^b(\tau)$ on Ω.

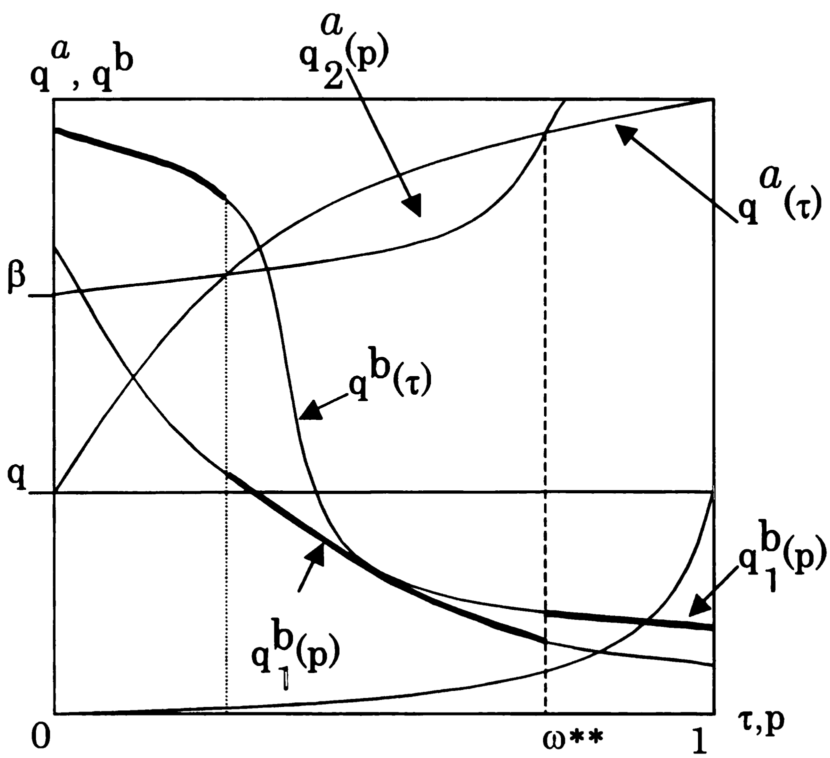

Figure 1.9

In this case, the combination of $\tau = \omega^{**}$ and some mixed strategy $c \in (0, 1)$ is an equilibrium. The proof is almost same to the Case 3. Since $q^a(\omega^{**}) > q_1^a(\omega^{**})$, if $c = 0$, announcing the policy a is beneficial for Party 1 candidate. But $q_1^b(\omega^{**}) > q^b(\omega^{**})$ implies that if $c = 1$ it is not beneficial for Party 1 candidate to announce the policy a. Since Party 1 candidate with $p = \omega^*$ is not indifferent between announcing the policy a and announcing the policy b if $c = 0$ or $c = 1$, then $\tau = \omega^{**}$ and $c = 0$ (or $c = 1$) is not an equilibrium. But there is a value $c^{**} \in (0, 1)$ such that $\tau = \omega^{**}$

and $c = c^{**}$ is the equilibrium. Let $q^{a**} \equiv q^a(\omega^{**})$ and $q^{b**} \equiv q^b(\omega^{**})$. Given some mixed strategy c, the expected sum of rents of announcing the policy a for Party 1 candidate is,

$$(1-c)\{1-H(-q^{a**}+\beta)\}\{1+\delta p\}+c\frac{1+\delta p}{2}$$

The expected sum of rents of announcing the policy b is,

$$(1-c)\frac{2+\delta}{4}+c\frac{2+\delta}{2}\{1-H(-\beta+q^{b**})\}$$

Party 1 candidate announces the policy a if,

$$\begin{aligned}
&(1-c)\{1-H(-q^{a**}+\beta)\}\{1+\delta p\}+c\frac{1+\delta p}{2} \\
&\quad \geq (1-c)\frac{2+\delta}{4}+c\frac{2+\delta}{2}\{1-H(-\beta+q^{b**})\} \\
&\Rightarrow (1-c)[\{1-H(-q^{a**}+\beta)\}\{1+\delta p\}-\frac{2+\delta}{4}] \\
&\quad +c[\frac{1+\delta p}{2}-\frac{2+\delta}{2}\{1-H(-\beta+q^{b**})\}] \geq 0
\end{aligned}$$

With $\tau = \omega^{**}$, the coefficient of $(1-c)$ is positive, and the coefficient of c is negative. Denote them as $(+)$ and $(-)$. Then, the condition is,

$$(1-c)(+)+c(-) = (+)+c\{(-)-(+)\} \geq 0$$

From this, the value $c^{**} \equiv \frac{(+)}{(+)-(-)} \in (0,1)$ is defined as the value that satisfies the condition with equality. Then the set of $\tau = \omega^{**}$ and $c = c^{**}$ is an equilibrium. Party 1 candidate with $p = \omega^{**}$ is indifferent between announcing the policy a and announcing the policy b. Since both $q_1^a(p)$ and $q_1^b(p)$ are decreasing functions of p, $(+)$ and $(-)$ above increase with p. Thus, Party 1 candidate with $p \geq \omega^*$ announces the policy a. At same time, this means that Party 1 candidate with $p < \omega^*$ does not announce the policy a. For Party 2 candidate, since $\omega^{**} = \max \omega \in \Omega$ means $q_2^a(\omega^{**}) = q^a(\omega^{**})$, Party 2 candidate is indifferent between announcing the policy a and announcing the policy b. Thus the set of $\tau = \omega^{**}$ and $c = c^{**}$ is the equilibrium.

Since above four cases cover all possible cases, the proof is completed. ∎

Endnotes of Chapter 1

1. The uncertainty is not restricted only to the policy outcomes in the literature. The uncertainty about the ability of politicians is often assumed (Majumdar and Mukand (2004), etc.) Cukierman and Tommasi (1998) and Schultz (2002) consider the uncertainty about the preference of politicians as well as the policy outcomes.

2. According to the Rasmussen reports (June 21, 2008), voters trust the Democratic party more than the Republican party on ten key issues including Economy, National Security, and the war on Iraq. Since the natural interpretation of this "trust" is competency, this result could be considered as implying that voters assume the ability gap between two parties. Though even this interpretation is actually inconsistent with the model of this chapter, at least this implies the meaning of considering the asymmetry between two parties.

3. As I wrote, this is a simplified, short cut way of modeling the signal. Following Majumdar and Mukand, this can be stated formally as the following; at first Party 1 candidate has the belief q same as voters. Then the candidate receives a random signal x, which is in $[\underline{x}, \bar{x}]$. This random signal x has a density ϕ^0 if the state is w^0 and a density ϕ^1 if the state is w^1. I assume that $\lim_{x \to \bar{x}} \dfrac{\phi^1(x)}{\phi^0(x)} = \infty$ and $\lim_{x \to \underline{x}} \dfrac{\phi^1(x)}{\phi^0(x)} = 0$. Then from this signal, the candidate deduce the new belief p of the state being w^1 using the Bayesian rule,

$$p(x) = \frac{q\phi^1(x)}{q\phi^1(x) + (1-q)\phi^0(x)} = \frac{q\dfrac{\phi^1(x)}{\phi^0(x)}}{q\dfrac{\phi^1(x)}{\phi^0(x)} + (1-q)}$$

Assuming $\dfrac{\phi^1(x)}{\phi^0(x)}$ is strictly increasing as a function of x, x has a one to one correspondence. Then, from ϕ^1 and ϕ^0, two distribution F and G can be constructed.

4. Of course, on the bound of Ω, Party 2 candidate is indifferent between announcing the policy a and announcing the policy b. So, the candidate might not announce the policy a on the bound of Ω. Actually, the condition that they are indifferent is

necessary for the existence of an equilibrium.

References

[1] Adams, James AND Samuel Merrill III (2008): "Candidate and Party Strategies in Two-Stage Elections Beginning with a Primary" *American Journal of Political Science*, 52(2), 334-359.

[2] Bernhardt, Dan, John Duggan AND Francesco Squintani (2007): "Electoral competition with privately-informed candidates" *Games and Economic Behavior*, 2007, 58, 1-29.

[3] Cukierman, Alex AND Mariano Tommasi (1998): "When does it take a Nixon to go to China?" *The American Economic Review*, 88(1), 180-197.

[4] Harrington, JR., Joseph, E. (1993a): "Economic policy, Economic performance, and elections" *The American Economic Review*, 83(1), 27-42.

[5] Harrington, JR., Joseph, E. (1993b): "The Impact of Reelection Pressures on the Fulfillment of Campaign Promises" *Games and Economic Behavior*, 5, 71-97.

[6] Heidhues, Paul AND Johan Lagerof (2003): "Hiding information in electoral competition" *Games and Economic Behavior*, 42, 48- 74.

[7] Majumdat, Sumon AND Sharun W. Mukand (2004): "Policy Gambles" *The American Economic Review*, 94(4), 1204-22.

[8] Persson, Torsten AND Guido Tabellini (2002): *Political Economics Explaining Economic Policy*, The MIT press.

[9] Rasmussen reports (2008): "Trust on Issues Democrats Trusted More on All Ten Electoral Issues"
http://www.rasmussenreports.com/public_content/politics/mood_of_america
/trust_on_issues/trust_on_issues

[10] Schultz, Christian (2002): "Policy biases with voters' uncertainty about the economy and the government" *European Economic Review*, 46, 487-506.

[11] Schofield, Norman (2005): "A valance model of political competition in Britain: 1992-1997" *Electoral Studies*, 24, 347-370

2 Information Transmission from Overlapping Generations Parties

2.1 Introduction

This chapter considers a model of electoral competition between two overlapping-generations-organization political parties with uncertainty about policy outcomes. The interest is on how to restrain opportunistic behavior of politicians who are in their last period. Suppose that there is an information asymmetry between politicians and voters of which politicians can take advantage. In multi-period models, usually the concern about future elections can restrain opportunistic behavior of politicians to some degree (Austen-Smith and Banks (1989)). This is an intuitive result. However, what would happen if they are in their last period so that they do not need to worry about future elections? It seems that in such situations opportunistic behavior of politicians, i.e., moral hazard, is inevitable. However, unless political parties consisted only of members who would retire soon, not everyone in parties would be happy to see such opportunistic behavior. This suggests that there would be conflict of interests in parties that voters could exploit to restrain politicians. Though the problem of opportunistic behavior by politicians in their last periods sounds very artificial since not all politicians in power are in their last period/term anyway, what I want to show is that the natural party structure consisting of politicians with various political life spans can give voters some leverage to restrain party bosses. The opportunistic behavior of politicians in their last period is a good example to show such possibility.

For this objective, I employ a model with OLG political parties that follows Alesina and Spear (1988) and Harrington (1992).

Traditionally, literature of Political Economy has not paid much attention to the difference between politicians and political parties. Parties in many electoral competition models (Calvert (1985), Wittman (1983), etc.) are nothing but labels put on candidates. Although there are also many papers in which parties are main players and different subjects from politicians, in those papers parties exist as independent and consistent entities (for example, Grossman and Helpman (1996).) However, parties in real world are organizations consisting of many individuals who often have different preferences[5]. Then, recognizing parties as organizations of individual politicians, not as independent entities, a natural question is what implication parties have on the behavior of politicians and election outcomes[6].

Recently, papers are appearing that consider this question. Snyder and Ting (2002) formalize the intuitive idea of political parties as brand names and shows how parties can work as signals about preferences of their candidates to voters. Caillaud and Tirole (2002) model the intra-party competition among politicians in the same party and study what kind of impacts such competitions have on the trust of voters for politicians and on elections. Levy (2004) considers the endogenous party formation through stable coalition formations and shows that the existence of parties increases the number of credible election platforms, i.e., that the existence of parties makes some outcomes feasible that would not be possible otherwise. Most of papers are finite period models. Alesina and Spear (1988) and Harrington (1992), however, consider infinite period models in which parties are overlapping-generations-organizations. They consider the problem of policy commitment.

This chapter is one of such papers, and I consider the problem of the information transmission from politicians to voters. In short, this is a problem about if politicians lie or not. I show the party organization can help voters restrain opportunistic behavior of the party boss. The situation I consider is this; there is the uncertainty about policy outcomes. No one is sure about outcomes of policies until policies are really implemented. This is a natural assumption for government policy decision-makings. The outcomes of almost any political decisions, especially big ones, are uncertain. Consider big regulation changes, or tax system changes. Maybe the decision of launching a war is the most evident example. In almost any policy decision-makings, there is the uncertainty about policy outcomes. However, it can be safely assumed that politicians have the informational advantage about possible outcomes over voters. In democracy, politicians choose policies and voters choose politicians. Thus, to achieve good policy outcomes, it is critically important to make sure that politicians honestly reveal their information to voters during elections. Given the informational asymmetry, however, this is not guaranteed. There are many papers on this problem (Harrington (1993), Majumdar and Mukand (2004) and Heidhues and Lagerlöf (2003), etc.) in finite period models. Exact results depend on the structure of models. But the general conclusion is that it is possible that politicians choose policies of which they know outcomes are likely to be bad for voters since such policies are their favorites or since choosing those policies increase the probability of their winning the next election. As an example, take Heidhues and Lagerlöf (2003). They consider a one-period model in which two parties go to an election with their election platforms. The outcome of implemented policy depends on the state of the world, and there is the uncertainty about the state of the world. Parties (or candidates) receive private signals about the state of the world before the election. Candidates have no preference on policies. Voters, however, have a belief on which state is likely. Since the objective of the candidates is winning the election, candidates have the incentive to follow voters' belief. Thus, it is possible that candidates would propose a policy that they know is not good for voters. This result is intuitive and also individually rational for candidates. But, looking at this from the viewpoints of their parties, not from the viewpoints of individual candidates, this is clearly a bad result. Suppose that an elected politician in his last political term implements his proposed platform and a bad outcome realizes. Though the politician does not need to worry about the next election, the party of the politician will have to face voters in future elections. Then, it is likely that party will be disadvantaged in the next election, i.e., the lower probability of the party winning the election (because of the retrospective voting behavior.) Thus, there is a conflict of interests in the party; the conflict between current candidates and other party members (especially candidates who will go to the next election). I consider a possibility of exploiting this conflict for voters' benefit.

Since this is an electoral competition model, not only parties but also how people vote is important. I assume voting depends on party reputations. By "party reputation," I mean the reputation on a party that the party is revealing information honestly to voters or not. This party reputation depends, in part, on the behavior of party insiders. To understand how this works, suppose an election between two parties. Then, pick a voter who does not have any particular party affiliation.

Suppose also that the voter is given an opportunity of touring the headquarters of both parties. Then, in one party's H.Q., the voter finds that all staff are shirking and complaining about their candidate. In other party's H.Q., he finds that all staff are working hard and talking about what great works their candidate can do for the country. Then, during the election, for which party will he vote? It is very likely that he will vote for the party with hard-working staff. Next, suppose that a candidate from one party chooses a policy based only on his own self-interest and does not care about what will happen to the party in the future. Then, will the staff work hard for the candidate? It seems unlikely. This suggests a possible route of a signal from the party to voters and a possibility of restraining the opportunistic behavior of current candidates through the route, especially through future candidates. In the model, I assume that young politicians are supposed to make the election campaign effort for old politicians during elections. The effort causes negative utility to young politicians. But they still make the campaign effort since that helps their party and strong party will help them in future. If old politicians make the choice that hurts the party, however, the choice will reduce the young's willingness to make the campaign effort. The campaign effort is assumed to be observable to voters. At first, I assume that voters require a certain level of effort (this will be changed later.) The required effort level is the required disutility level for young politicians during elections. Then, I consider about the reaction of young politicians to the opportunistic behavior of the party boss and if it can restrain such opportunistic behavior.

In section 2.2, I explain the model, the behavior of politicians and the definition of equilibrium. I state results in section 2.3. Section 2.4 is the conclusion with a brief literature review. All proofs and calculations are in the Appendix of this chapter.

2.2 Model and Behavior of Politicians

Before explaining the model, I state a brief description of the flow in each period. At the start of each period, the nature chooses the state of the world in the period. Then, two parties receive the same signal about the state of the world. Old politicians in both parties choose their election platforms, and young politicians choose campaign effort levels. Observing effort levels, voters update their party reputations and vote for the party during the election based on the party reputations. The winner of the election implements the policy of its election platform. The outcome of the policy is observed by voters and they update party reputations again. In the reputation update, basically, party reputations get downgraded as punishment for not making campaign efforts (suggesting suspicious behavior of old politicians) and for choosing a bad policy (i.e. retrospective voting). These two punishments are instrument and incentive for young politicians. The reputations are carried over to the next period. To not have a bad party reputation in the next period, young politicians use the campaign effort as their instrument to restrain old politicians.

The party reputation in this model means the belief by voters about if the party honestly reveals its information to voters (i.e., choosing the good policy as its platform given the information) or not during the election. Basically, old politicians want to abuse the good reputation and young ones want to keep the good reputation for their election in the future. This "reputation" is different from the one in the literature of "reputation" like Tadelis (2002), which is a paper of an OLG reputation model. In this chapter, there is no "good" politician and no "bad" politician. Party reputations are about if the party is trustable or not. Also, even though reputations are updated as voters receive new information, the update is not Bayesian. Following Harrington (1992), the reputation and its update in this model are very simple, and are given by an assumption[7]. This is because assuming the Bayesian updating in this model makes the model unnecessarily complicated; more elaborate update systems naturally lead to more strategic behavior among politicians, even assuming non-strategic behavior of voters. Since such complication is not essential to the idea of this chapter, I assume a simple "reputation" structure in the following.

2.2.1 Model

Society and its members: The society consists of voters and politicians of two parties (party 1 and party 2.) They are risk neutral. All voters are same and they do not have any particular affiliation with either party[8]. Voters care only about the current period utility; they do not try to maximize the expected utility over the infinite period. This may be because they are short-lived or myopic. In any interpretation, this assumption rules out the possibility of their employing complex strategies.

Parties and politicians: Party 1 and 2 are modeled as overlapping-generations-organizations with an old politician and a young politician in each party in each period. Following Alesina and Spear (1988), I call the old politicians at the period t as Presidential candidates (PCs) at the period t. They decide the party platforms of the period. Correspondingly, I call the young politicians at period t as Vice Presidential candidates (VPCs) at period t. VPCs make the campaign effort for their parties. The effort is observable to voters and causes the disutility to VPCs. I call politicians by the words "PC" and "VPC" in regardless of their being in an election or in office. A generation t politician of party $i \in \{1, 2\}$ enters into the party at period t as a VPC and becomes a PC of party i at period $t + 1$. Then he retires and leaves the party. All politicians in a party are same except for their generations. Then, in the election at the start of each period, old politicians in two parties choose the election platform of their parties and young politicians make the campaign effort. Assuming the policy commitment[9], the winning party will implement its platform after the election.

Policies and the state of the world: There are two possible policies: a^1 and a^2. Because of the policy commitment, the election determines which policy will be implemented in the period. However, the policy outcome depends on the state of the world, and there is the uncertainty about the true state. The number of possible states of the world is two: w^1 and w^2. The state of the world is chosen by Nature at the start of every period with probability $\frac{1}{2}$ for both states. This is common knowledge. If the policy a^i ($i \in \{1, 2\}$) is implemented in the state w^i, then voters receive utility 1. If the policy a^i is implemented in the state w^j ($j \neq i$), voters receive utility -1. Thus, voters can tell the true state in the period after a policy is implemented.

Information: Before the election in each period, parties receive a common signal about the state of the world in the period. From the signal, politicians can derive the probability of the true state being w^1. For the simplicity, I assume that parties receive the probability directly as the signal. Let r be the probability. This r is a random variable with the uniform distribution[10]. Voters do not observe the signal.

Campaign effort: during elections, only VPCs can make the campaign effort. VPCs suffer the disutility c to make the effort c. Voters can observe the campaign effort. Since the probability of the realization of each state is common knowledge and voters can perfectly observe party platforms after they are proposed, the campaign effort does nothing on informing or deceiving voters about abilities of parties or platforms. The role of the effort in this model is that the level of the effort can work as a signal from VPCs to voters about the honesty of PCs. However, if c is too high, VPCs never make the effort c. So, I set the upper bound of c at $\frac{B}{4}$.

Utilities of politicians: If maximizing the welfare of voters is the objective of politicians, parties would propose their platforms honestly, i.e., proposing policy a^1 when they receive signal $r > \frac{1}{2}$ and a^2 when $r < \frac{1}{2}$. But that is not the objective of politicians. Preferences of politicians are different from voters. Since voters do not observe the signal, politicians have incentive to lie to voters. Party $i \in \{1, 2\}$ politicians receive (dis)utilities at three occasions in their two-period political life.

Politicians's (dis)utility:

1. Winning election: when Party i wins the election at period t, the PC of party i at the period receives a fixed benefit $B > 1$.

2. Implementing policy: PC of party i receives utility 1 when he implements

the policy a^i and receive -1 when he implements a^j, $j \neq i$. If party j politician implements a policy, party i politician receives zero utility[11]. VPCs receive zero utility from the implementation of policies.

3. Making campaign efforts: If VPCs make the campaign effort c during an election, this gives utility $-c$ to them. PCs do not make the campaign effort.

Thus, PCs and VPCs do not receive same utility in same period[12]. The objective of politicians is the maximization of the sum of utilities from two periods. Since political life is just two periods, I do not use a discount factor for calculations of politicians' utility. Although this specification is for simplifying the model, this reflects reality to some degree except the assumption that PCs do not make the campaign effort. What is important for the model is, however, that they receive different utilities in same period.

In the model, Party 1 and Party 2 are mirror images of each other. PCs of Party 1 prefer the policy a^1 to the policy a^2 and PCs of Party 2 prefer a^2 to a^1. Both parties receive the same probability r of w^1 being the true state as a signal. Also each state of the world realizes with the same probability $\frac{1}{2}$. Since voters are assumed to have no affiliation with either party, they are basically indifferent between parties. However, they have preference on party strategies. If Party 1 is honest on the choice of platforms ($r \geq \frac{1}{2} \to a^1$ and $r < \frac{1}{2} \to a^2$) but Party 2 is dishonest (always proposing the policy a^2), voters prefer honest Party 1 to dishonest Party 2. So, if they thought Party 1 was more honest than Party 2, they would have a derived preference for Party 1. Since party strategies are unobservable to voters, however, voters need a something that can induce politicians to behave honestly. That "something" is the reputation on parties and the voting response based on it. Appropriate voting responses based on party reputations give incentives for VPCs to restrain the opportunistic behavior of PCs, though there are no honest politicians in the model. When new information about parties is obtained, the reputation on parties will be updated. I assume very simple strategies for parties and only two types of party reputations.

Party reputation and Voting response: Voters have only two types of reputations on parties; Honest (H) or Suspicious (S). Honest means that voters think that the party is likely to propose a^1 when $r > \frac{1}{2}$ and a^2 when $r < \frac{1}{2}$ (when $r = \frac{1}{2}$, choice does not matter.) Suspicious means that voters think that the party is likely to propose its preferred policy regardless of the signal the party receives. Notice that reputations are on parties, not on a particular politician, so this is not about politicians' type. This reputation system is very coarse compared to the set of PC's strategies, which will be explained later. Still, this coarse reputation system can improve voters' welfare.

During the election, voters prefer the party with Honest reputation to the party with Suspicious reputation. They will vote for the honest party over the suspicious party. Since all voters are assumed to be same, the honest party definitely wins. If both parties have same reputation (Honest and Honest, or Suspicious and Suspicious), following the tradition of the electoral competition literature, both parties have probability $\frac{1}{2}$ of winning the election.

Reputation update by voters: Voters cannot see signals. Instead, they use the observable campaign effort of both parties and the policy outcome as proxies of the signal and the strategies. Thus, when they observe VPCs' effort levels and when they observe the policy outcome so that they can tell the true state from the outcome, they update reputations. Zero effort chosen by VPC suggests that the PC of the zero effort party chose a policy based on his preference, not based on the signal received. After the policy outcome is observed, voters can tell the true state and which party chose the bad policy (a^i when w^j, $i \neq j$.) Since parties receive a signal about the true state, the bad policy also suggests that the PC who chose the bad policy chose it based on his preference, not on the signal. Then, voters downgrade the reputations of parties that proposed the bad policy as its platform, regardless of whether the party won the election in the period or not. The update based on VPC's effort levels can affect the election outcome in the period. The update based on the policy outcome can affect the election outcome in the next period. Though these updates are given as assumptions exogenously, these are basically simple and intuitive.

Reputation update:

(1) Before the election, if VPC of one party makes the effort level less than c, the reputation of the party is updated to Suspicious.

(2) After observing the policy outcome, the reputations of parties that proposed the bad policy are updated to Suspicious.

Above two updates are about downgrading reputations. Since parties have positive probabilities of choosing the bad policy even if they choose their policies completely honestly, the opportunities for upgrading reputations are necessary. I assume following two for upgrading.

(3) Before the election, if (1) happens to one party and the reputation of the other party was Suspicious at the start of the period, the reputation of the other party is updated to be honest.

(4) After the election, the reputation of the party that was Suspicious at the start of the period is updated to Honest.

(5) If both parties have Suspicious reputation at the end of period, voters update their reputations to Honest.

The party reputations are carried over to the next period. (3) strengthens the punishment of no effort for the case that only one party has Honest reputation. Without this, it is possible that the PC of the party with Honest reputation prefers no effort with his opportunistic choice to the effort level c with honest choice. (1) and (3) make sure that the party of which a VPC does not make the required campaign effort gets punished during the election. Combined with (2), (4) is essentially saying that the party that proposed the good policy (a^i when w^i) get rewarded for the choice. (5) is for the case that both parties obtain Suspicious reputation. Since the fact that keeping Honest reputation increases the prospect of winning election is the reason that VPCs can restrain the opportunistic behavior of PCs, upgrading both parties to Honest is beneficial for voters than keeping two Suspicious parties.

This updating system is just one example of many. Other systems, especially more complicated systems, might be able to achieve better results. However, the point is that even this simple system can improve voter welfare.

Party strategies: Politicians have two occasions to make decisions; once in each period of their political life. PCs choose party platforms. Observing PC's choice, VPCs decide the level of the campaign effort to make in the period. Then, during each election, there is an intra-party game in each party between its PC and its VPC; though it is PCs who determine party platforms, VPCs can have a say in platform choices through their campaign effort choices. The entire strategy of a politician is the profile of decisions as a VPC and as a PC. I assume that the strategy is based only on party reputations and the signal in the current period. So, all politicians in the same party have the same strategy. I consider the strategy as follows.

Take a PC of Party 1 at period t. Let (A, B) be a profile of Party 1's reputation A and Party 2's B ($A, B \in \{H, S\}$, H for Honest and S for Suspicious) at the start of the period. Let r be the signal parties receive. Then, the strategy of the Party 1's PC is,

$$
\begin{aligned}
& r \geq \frac{1}{2} \Rightarrow a^1 \\
& r < \frac{1}{2} \\
& \qquad \Rightarrow (H,H) : r \geq \frac{1}{2} - \varepsilon^0 \Rightarrow a^1 \\
& \qquad\qquad\qquad\quad r < \frac{1}{2} - \varepsilon^0 \Rightarrow a^2 \\
& \qquad\quad (H,S) : r \geq \frac{1}{2} - \varepsilon^1 \Rightarrow a^1 \\
& \qquad\qquad\qquad\quad r < \frac{1}{2} - \varepsilon^1 \Rightarrow a^2 \\
& \qquad\quad (S,H) : a^1
\end{aligned}
$$

Thus, the strategy of Party 1 PCs is the choice of thresholds ε^0 and ε^1, or probabilities of lying, for cases of (H, H) and (H, S) (see figure 2.1.)

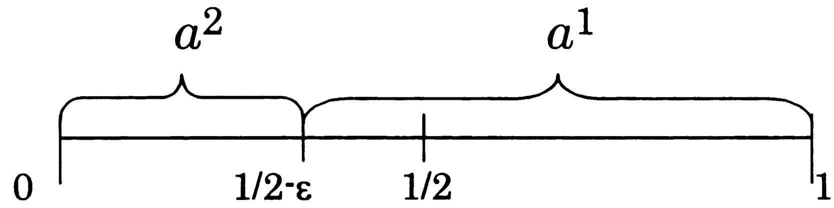

Figure 2.1

If $r \geq \frac{1}{2}$, then there is no reason for Party 1 PCs to lie. It is the happy coincidence of their preferred policy and the socially optimal choice. If $r < \frac{1}{2}$, there is an incentive for PCs to lie to voters. Since they do not need to worry about the next election, if they could choose ε^0 and ε^1 freely, they would choose $\varepsilon^0 = \varepsilon^1 = \frac{1}{2}$. Since VPCs in the current period will be PCs in the next period, higher values of ε^0 and ε^1 could benefit VPC if they would have reputation H in the next period. However, such choice would hurt the prospect of VPCs having reputation H in the next period. Thus, if PCs choose to lie for very low r, VPCs will not make the required campaign effort c. The result of this zero effort is that the smaller expected utilities of PCs. To avoid such loss, PCs might choose values of ε^0 and ε^1 that are lower than $\frac{1}{2}$. As for the case of (S, H), as long as Party 2 makes the required campaign effort, Party 1 never wins the election. Thus, actual choice does not matter for PCs. One implication of this and assumptions about the reputation system is that if (S, H) then Party 1 VPCs will not make the positive campaign effort. This is because the behavior of VPCs does not matter for the reputation in the next period. This could be changed so that VPCs in such situation makes the positive campaign effort with appropriate changes of reputation update assumptions. Since such change does not alter results qualitatively, I keep the current update system for the sake of simplicity. I do not consider the case (S, S) because this happening at the start of a period is ruled out by the reputation update. The strategies of Party 2 PC, $\hat{\varepsilon}^0$ and $\hat{\varepsilon}^1$, can be obtained by changing (H, S) to (S, H), the inequality signs, negative sings of ε^0 and ε^1 to positive signs, and exchanging a^1 and a^2(see Figure 2.2.)

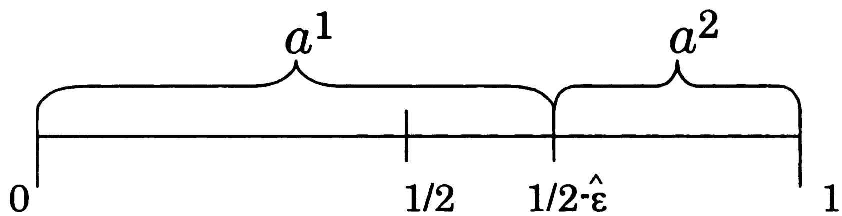

Figure 2.2

As for VPC's strategy, since VPCs move after PCs in stage games, their decision is simply about if they should accept the PC's choice and make the effort c demanded by voters or should not accept it and make zero effort (when voters demand the positive level c of the campaign effort, it is meaningless to choose an effort level between 0 and c.) So, the strategy of VPCs is simply to accept PC's choice of platform only if it gives expected utility no less than the expected utility of not accepting it. Thus, the complete set of the strategy for Party 1 politicians can be denoted as $(\varepsilon^0, \varepsilon^1)$ (the strategy for Party 2 politicians is $(\hat{\varepsilon}^0, \hat{\varepsilon}^1)$.)

The sequence of the game in period t:

The start of the period t.

If (S, S) then (H, H)

Nature chooses the state of the world for the period.

Both parties receive r.

$$\text{Intra-party game:}\left\{\begin{array}{c}\text{PC chooses platform.}\\[1em]\text{VPC decides campaign effort level.}\end{array}\right\}$$

Middle of the period.

Voters observe party platforms and campaign effort levels.

Voters update party reputations.

Election: Voters vote according to party reputations.

Suspicious reputation at the start of the period is updated to Honest.

The winner of the election implements its platform.

End of the period.

Voters receive utility from the policy outcome.

Voters update party reputations.

Next period.

2.2.2 Behavior of Politicians

In this subsection, I explain behaviors and conditions for PCs and VPCs in symmetric equilibrium. In each period, in both parties, a PC and a VPC play a stage game to decide the party strategy and the campaign effort level in the period. I derive the expected utilities of politicians in the equilibrium with VPCs making the required effort c when the party reputation is H.

Behaviors of PCs First, I calculate expected utilities of a Party 1 PC at the start of a period so he knows party reputations but has not received the signal r yet. Given an equilibrium strategy $(\varepsilon^0, \varepsilon^1)$ with VPC making the effort c when the party reputation is H, let $V_{H,H}$ be the expected utility of Party 1 PC at the start of a period with (H, H). Then, this is defined as,

$$V_{H,H} = (\frac{1}{2} - \varepsilon^0)\frac{B-1}{2} + (\frac{1}{2} + \varepsilon^0)\frac{B+1}{2} = \frac{B+2\varepsilon^0}{2}$$

$(\frac{1}{2} - \varepsilon^0)$ of the first term is the probability of $r < \frac{1}{2} - \varepsilon^0$ (r has the uniform distribution.) If $r < \frac{1}{2} - \varepsilon^0$, the PC proposes a^2. The probability of his winning election is $\frac{1}{2}$, and he will receive utility $B - 1$ if he wins. $(\frac{1}{2} + \varepsilon^0)$ of the second term is the probability of $r \geq \frac{1}{2} - \varepsilon^0$. If $r \geq \frac{1}{2} - \varepsilon^0$, he proposes a^1. The probability of his winning election is $\frac{1}{2}$, and if he wins, he will receive utility $(B + 1)$.

If (H, S),

$$V_{H,S} = (\frac{1}{2} - \varepsilon^1)(B-1) + (\frac{1}{2} + \varepsilon^1)(B+1) = B + 2\varepsilon^1$$

If (S, H), Party 1 cannot win as long as Party 2 VPC makes the positive effort. So,

$$V_{S,H} = 0$$

Expected utilities of Party 2 PC can be calculated similarly with $(\hat{\varepsilon}^0, \hat{\varepsilon}^1)$. Since PCs are in their last period, they do not need to worry about the next election. Thus, they always want to increase values of ε^0 and ε^1 to their maximum, i.e., $\frac{1}{2}$. The behavior of VPCs is more complicated since it involves calculations of expected utilities.

Behavior of VPCs Given an equilibrium strategy $(\varepsilon^0, \varepsilon^1)$ with VPC making the effort c when the party reputation is H, a Party 1 VPC's expected utilities of accepting and not accepting PC's choice are calculated as follows. Notice that when VPCs make decisions they already know the value of r in their periods. Since PCs' choice when $r > \frac{1}{2}$ is honest, I do not need to consider the case of $r > \frac{1}{2}$. So, I consider only the case of $r \leq \frac{1}{2}$.

(H, H)

$$\begin{aligned}(2.1)\ a^1 :\ & r[\frac{1}{2}V_{H,S} + \frac{1}{2}V_{H,S}] + (1-r)[\frac{1}{2}V_{S,H} + \frac{1}{2}V_{S,H}] - c \\ & = r(B + 2\varepsilon^1) - c\end{aligned}$$

This is the expected utility of accepting PC's choice of a^1 when (H, H). The true state is w^1 with probability r and w^2 with the probability $1 - r$. With (H, H), both parties have probability $\frac{1}{2}$ of winning the election. Party 1 chooses a^1 and Party 2 chooses a^2 (because $r < \frac{1}{2}$.) After the outcome of the policy implemented is observed, voters can tell which party chose the bad policy (i.e., a^i for the true state w^j.) So, the party that chose the bad policy is punished with the bad reputation.

$$\begin{aligned}(2.2)\ a^2 :\ & r[\frac{1}{2}V_{H,H} + \frac{1}{2}V_{H,H}] + (1-r)[\frac{1}{2}V_{H,H} + \frac{1}{2}V_{H,H}] - c \\ & = \frac{B + 2\varepsilon^0}{2} - c\end{aligned}$$

This is the expected utility of accepting PC's choice a^2 when (H, H). In this case, both parties choose the same policy, so, the expected utility for VPCs does not depend on r.

$$(2.3)\ 0 :\ rV_{H,H} + (1-r)V_{S,H} = \frac{r}{2}(B + 2\varepsilon^0)$$

This is the expected utility of not accepting PC's choice. Since the reputation of

Party 1 becomes Suspicious for no campaign effort, the only case that VPC can have positive expected utility at the start of the next period is when the choice of Party 2 turns out to be bad. In the case, the reputation of Party 2 would become Suspicious, too. At the start of the next period, their reputations would be updated to Honest.

(H, S)

(2.4) $a^1: \; rV_{H,S} + (1-r)V_{S,H} - c = r(B + 2\varepsilon^1) - c$

(2.5) $a^2: \; rV_{H,H} + (1-r)V_{H,H} - c = \dfrac{B + 2\varepsilon^0}{2} - c$

(2.6) $0: \; rV_{H,H} + (1-r)V_{S,H} = \dfrac{r}{2}(B + 2\varepsilon^0)$

If Party 1 VPCs do not make the effort c when (H, S), reputations become (S, H) before the election. After the implementation of policy, this will be kept same or be changed to (H, H), depending the outcome of the implemented policy.

(2.2) and (2.5) are same. Though (2.1) and (2.3) are also same to (2.4) and (2.6) for the same value of r, respectively, they should be treated as different conditions because they are for different threshold values. (2.1) and (2.3) are related to ε^0 and (2.4) and (2.6) are related to ε^1. In the case of (S, H), PCs always choose a^1 and VPCs always make zero effort. So, I do not need to consider expected values of VPCs when (S, H).

Given (H, H) and r, if (2.1) is lower than (2.3), VPCs never accept PCs' choice of a^1, since VPCs prefer making no campaign effort to accepting a^1. Same for (2.2) and (2.3). So, what is important is the difference between (2.1) and (2.3), and between (2.2) and (2.3). Same thing can be said for (H, S). Thus, I subtract (2.3) from (2.1) and (2.2), and (2.6) from (2.4) and (2.5), and denote them as (2.1), (2.2), (2.4) and

(2.5), again.

(2.1) $\frac{r}{2}(B - 2\varepsilon^0 + 4\varepsilon^1) - c$

(2.2) $\frac{1-r}{2}(B + 2\varepsilon^0) - c$

(2.4) $\frac{r}{2}(B - 2\varepsilon^0 + 4\varepsilon^1) - c$

(2.5) $\frac{1-r}{2}(B + 2\varepsilon^0) - c$

Since $B > 1$, the derivatives of (2.1) and (2.4) with respect to r is positive. So, if (2.1) or (2.4) are non-negative at the threshold $r^0 \equiv \frac{1}{2} - \varepsilon^0$ or $r^1 \equiv \frac{1}{2} - \varepsilon^1$, it is positive for r higher than the threshold values. Thus, for any r higher than the threshold values of r, PCs always propose a^1 and VPCs always accept it. Moreover, since the upper bound of c is $\frac{B}{4}$, (2.2) and (2.5) never become negative for any r and ε^0 in $[0, \frac{1}{2}]$. Because of this, I can ignore (2.2) and (2.5). This means VPCs always accept a^2 if PCs choose a^2. Thus I need to consider only (2.1) and (2.4). Since what I am looking for are threshold values of ε^0 and ε^1, I substitute the corresponding threshold values of r, $r^0 \equiv \frac{1}{2} - \varepsilon^0$ and $r^1 \equiv \frac{1}{2} - \varepsilon^1$, into above (2.1) and (2.4). To distinguish expressions with r and without r, I put "$*$" to expressions without r.

(2.1)* $\frac{1}{2}(\frac{1}{2} - \varepsilon^0)(B - 2\varepsilon^0 + 4\varepsilon^1) - c$

(2.4)* $\frac{1}{2}(\frac{1}{2} - \varepsilon^1)(B - 2\varepsilon^0 + 4\varepsilon^1) - c$

2.2.3 Definition of Equilibrium

Consistency and Deviation: Let $(\varepsilon^{0*}, \varepsilon^{1*})$ be an equilibrium strategy with VPCs making the required effort c when the party reputation is H. Then, this $(\varepsilon^{0*}, \varepsilon^{1*})$ needs to satisfy the following condition.

(a) (Consistency) $(2.1)^*$ and $(2.4)^*$ are non-negative for thresholds $r^0 = \frac{1}{2} - \varepsilon^{0*}$ for $(2.1)^*$ and $r^1 = \frac{1}{2} - \varepsilon^{1*}$ for $(2.4)^*$.

If this condition does not hold, $(\varepsilon^{0*}, \varepsilon^{1*})$ cannot be an equilibrium strategy with VPCs making the required effort c. The condition (a) is about the consistency of $(\varepsilon^{0*}, \varepsilon^{1*})$; if (a) does not hold, VPCs will not accept a^1 for thresholds r^0 and r^1. Thus, $(\varepsilon^0 \varepsilon^1)$ that does not satisfy (a) cannot be even an appropriate strategy with VPCs making the effort c, let alone being an equilibrium strategy. In such case, $V_{H,H}$ and $V_{H,S}$ cannot be calculated like above since the behavior of VPCs does not follow $(\varepsilon^{0*}, \varepsilon^{1*})$. Then, when PCs propose an inconsistent strategy, I assume VPCs never make the required campaign effort c since they cannot tell making the effort is beneficial or not. For a successful deviation, the deviated values need to satisfy (a). Suppose that $\varepsilon^{0*} < \frac{1}{2}$ or $\varepsilon^{1*} < \frac{1}{2}$ and that the current period Party 1 PC deviates from $(\varepsilon^{0*}, \varepsilon^{1*})$ when (H, H); proposing a^1 when $r < r^0 = \frac{1}{2} - \varepsilon^{0*}$ (of course $r \geq 0$.) Since the party reputation is (H, H), the PC can deviate only from ε^{0*}. Let $r = \frac{1}{2} - \bar{\varepsilon}^0$. This deviation is equal to proposing $\bar{\varepsilon}^0$ instead of ε^{0*}. If (a) holds for $(\bar{\varepsilon}^0, \varepsilon^{1*})$, the current period VPC will accept this deviation. If (a) does not hold, VPC will not accept a^1. This is obvious if $(2.1)^*$ is negative with $(\bar{\varepsilon}^0, \varepsilon^{1*})$. However, even if $(1)^*$ is not negative with $(\bar{\varepsilon}^0, \varepsilon^{1*})$, if $(2.4)^*$ is negative, the VPC will not accept a^1 since $(\bar{\varepsilon}^0, \varepsilon^{1*})$ is not a consistent strategy. He cannot calculate $V_{H,S}$, which is needed for the calculation of his expected utility. Same holds for (H, S). Thus, for $(\varepsilon^{0*}, \varepsilon^{1*})$ to be an equilibrium strategy, it has to satisfy (a). For Party 2 with $(\hat{\varepsilon}^{0*}, \hat{\varepsilon}^{1*})$, similar things can be said with thresholds $\hat{r}^0 = \frac{1}{2} + \hat{\varepsilon}^{0*}$ and $\hat{r}^1 = \frac{1}{2} + \hat{\varepsilon}^{1*}$. Now, I state the definition of equilibrium.

Equilibrium: Given the voting response of voters and the required campaign effort level c, Party 1 strategy $(\varepsilon^{0*}, \varepsilon^{1*})$ and Party 2 strategy $(\hat{\varepsilon}^{0*}, \hat{\varepsilon}^{1*})$ are a pair of equilibrium strategies if the following conditions hold.

1) Given $(\hat{\varepsilon}^{0*}, \hat{\varepsilon}^{1*})$, there is no consistent strategy $(\varepsilon^{0}, \varepsilon^{1})$ that is Pareto superior to Party 1 PCs in the situation of (H, H) and (H, S) compared to $(\varepsilon^{0*}, \varepsilon^{1*})$.

2) Given $(\varepsilon^{0*}, \varepsilon^{1*})$, there is no consistent strategy $(\hat{\varepsilon}^{0}, \hat{\varepsilon}^{1})$ that is Pareto superior to Party 2 PCs in the situation of (H, H) and (S, H) compared to $(\hat{\varepsilon}^{0*}, \hat{\varepsilon}^{1*})$.

Notice that the definition of equilibrium is not saying anything about voters, though it is them who choose the level of c. In the following, I consider only the symmetric equilibrium so that $(\varepsilon^{0*}, \varepsilon^{1*}) = (\hat{\varepsilon}^{0*}, \hat{\varepsilon}^{1*})$.

2.3 Results

First, I state about the existence of equilibrium. I consider only $r \leq \frac{1}{2}$. Given $B > 1$, since c has the upper bound $\frac{B}{4}$, choosing a consistent strategy is always beneficial than choosing any non-consistent strategy for PCs. Since (2.1) and (2.4) have the campaign effort c, and positive c can make these values negative for small enough r. This means that VPCs prefer no campaign effort to accepting a^1 for such small r. Since VPCs always accept a^2, choosing a^2 for such small values of r is more beneficial for PCs than accepting the loss of election by no campaign. Thus, PCs choose a^2 for such small r. This means that in all equilibrium strategies VPCs make the required effort c when the party reputation is H. Moreover, since $(0, 0)$ makes $(2.1)^*$ and

(2.4)* non negative for any $c \in [0, \frac{B}{4}]$, it is obvious that an equilibrium exists for any $c \in [0, \frac{B}{4}]$ and $B > 1$. I state this as a lemma.

Lemma 2.1. *For any $B > 1$ and $c \in [0, \frac{B}{4}]$, an equilibrium exists.*

Then, what kind of characteristics does the equilibrium have? I consider the easy case of $c = 0$ first. Since $B > 1$, and ε^0 and ε^1 take values in $[0, \frac{1}{2}]$, it is easy to show that (2.1)* and (2.4)* are always non negative when $c = 0$. Thus, when $c = 0$, Party 1 VPCs always accept a^1. So, Party 1 PCs always choose a^1. The same thing also holds for Party 2 politicians about a^2.

Lemma 2.2. *If $c = 0$, the only equilibrium is $\varepsilon^0 = \varepsilon^1 = \frac{1}{2}$.*

Next, I consider if positive c can achieve better result for voters, i.e., if positive c can induce PCs to lie less often. Since both (2.1)* and (2.2)* become $-c$ when $\varepsilon^0 = \varepsilon^1 = \frac{1}{2}$, the threshold values of r for positive c must be higher than zero. This means the threshold values of ε^0 and ε^1 are lower than $\frac{1}{2}$. The following proposition is about this, i.e., about the existence of equilibrium with $\varepsilon^0 + \varepsilon^1 < 1$; Party 1 PCs do not always lie when c is positive (same for Party 2 PCs when c is positive.)

Proposition 2.1. *The equilibrium with $\varepsilon^0 + \varepsilon^1 < 1$ exists with $c \in (0, \frac{B}{4})$.*

Of course, $\varepsilon^0 + \varepsilon^1 < 1$ means the improvement of the voter welfare. I calculate the long-run average voter welfare as the welfare criterion. It is $\frac{1 - 16(\varepsilon^0\varepsilon^1)^2}{2 + 4(\varepsilon^0)^2 + 4(\varepsilon^1)^2}$ (the derivation of this is in the Appendix.) When $\varepsilon^0 = \varepsilon^1 = \frac{1}{2}$, the long-run average voter welfare is zero. Since appropriate positive values of c can achieve $\varepsilon^0 + \varepsilon^1 < 1$, positive c can make the average welfare positive. So, requiring positive campaign efforts is beneficial for voters. Then, it is natural to ask if it is possible to achieve the optimal result for voters, i.e., $\varepsilon^0 = \varepsilon^1 = 0$. It is obvious that if B is so large that utility gain/loss from policy preference does not matter to politicians; they behave more honestly, and vice versa. The following Proposition is about this point.

Proposition 2.2. *If $B < 2$ and $c \in (0, \frac{B}{4})$, there is no equilibrium with $\varepsilon^0 = \varepsilon^1 = 0$. If $B \geq 2$, the equilibrium with $\varepsilon^0 = \varepsilon^1 = 0$ exists for $c = \frac{B}{4}$.*

Thus, if $B \geq 2$, voters can achieve the optimal result by requiring $c = \frac{B}{4}$. If $B \in (1, 2)$, this is not possible. This impossibility result of $\varepsilon^0 = \varepsilon^1 = 0$ does not depend on the upper bound of c. Even if c could take values higher than $\frac{B}{4}$, $\varepsilon^0 = \varepsilon^1 = 0$ is impossible for $B \in (1, 2)$. Then, when $B \in (1, 2)$, what is the characterization of the equilibrium?

Proposition 2.3. *When $B \in (1, 2)$ and $c \in (0, \frac{B}{4})$, equilibria are characterized by,*

$$\begin{aligned} \frac{1}{2} &> \varepsilon^1 \geq \varepsilon^0 \\ \varepsilon^1 &\geq \frac{2\varepsilon^0 + 2 - B}{8} \\ \varepsilon^1 &= \frac{2 - B + 2\varepsilon^0 + \sqrt{(2 - B + 2\varepsilon^0)^2 - 32c}}{8} \end{aligned}$$

All $(\varepsilon^0, \varepsilon^1) \in [0, \frac{1}{2}] \times [0, \frac{1}{2}]$ that satisfy above three conditions are equilibrium. Given this characterization, what can be said about the optimal level of c when $B \in (1, 2)$? Since the long-run average voter welfare is $\frac{1 - 16(\varepsilon^0\varepsilon^1)^2}{2 + 4(\varepsilon^0)^2 + 4(\varepsilon^1)^2}$, its indifference curve is $\varepsilon^1 = \frac{1}{2}\sqrt{\frac{1-k}{k + 4(\varepsilon^0)^2} - k}$ (k is a fixed welfare level.) The closer the curve is to the origin, the better for voters. Then, what value of c should voters require to achieve the highest welfare? The next Proposition is about this.

Proposition 2.4. *When $B \in (1, 2)$, the best long-run average voter welfare achievable is $\frac{16}{(B-2)^2+32}$ with $\varepsilon^0 = 0$ and $\varepsilon^1 = \frac{2-B}{8}$ by requiring $c = \frac{(B+2)^2}{32}$*

Though the highest achievable average voter welfare by requiring positive c when $B \in (1, 2)$ is $\frac{16}{(B-2)^2+32}$, which is lower than the maximum average voter welfare $\frac{1}{2}$, it is still better than the average voter welfare 0 with requiring $c = 0$. However, the assumption that voters require some positive level of the election campaign effort is unrealistic. Moreover, this requires that voters know the value of B. Though the calculation of $c = \frac{(B+2)^2}{32}$ is easy if B is known, there is actually an easier way to achieve the highest average voter welfare; simply, letting parties compete in the campaign effort level. Change the reputation update during the election from,

(1) Before the election, if VPC of a party makes the effort level less than c, the reputation of the party is updated to Suspicious.

to

(1)' Before the election, if VPC of a party makes the effort level less than the effort level the VPC of the other party makes, the reputation of the party is updated to Suspicious.

Then, what is the result of this change? It is clear that PCs will choose ε^0 and ε^1 for which VPCs will make positive campaign efforts. Then, what ε^0, ε^1 and effort level will PCs and VPCs choose?

Proposition 2.5. *With the reputation update (1)', if $B \geq 2$, the equilibrium is $\varepsilon^0 = \varepsilon^1 = 0$. If $B \in (1,2)$, the equilibrium is $\varepsilon^0 = 0$ and $\varepsilon^1 = \dfrac{2-B}{8}$.*

Thus, by letting parties compete in the campaign effort level, the best that can be achieved by requiring positive c can be achieved automatically for voters.

2.4 Conclusion of Chapter 2

In this chapter, I considered a model of electoral competition between OLG political parties under the uncertainty about policy outcomes. Exploiting the natural conflict of interests within parties, I showed the possibility of restraining the opportunistic behavior of politicians. Requiring young politicians of making campaign efforts, it is possible to deter the opportunistic behavior of politicians completely when $B \geq 2$. When $B \in (1,2)$, the restraint on politicians is not perfect. But it can still achieve higher voter welfare than not requiring the effort. The last Proposition shows the importance of the competition between parties, which can achieve the best possible outcome automatically for voters. However, even with the party competition, politicians might lie if $B \in (1,2)$. (H,S) and (S,H) are situations that the trust of voters is biased favorably to one party over the other. Such biased trust makes exploiting the trust too attractive for the trusted party when B is relatively small. Reputation systems and voting behaviors different from the ones in the chapter could achieve different results. But, since such possibility of abusing the trust is a part of the reward for young politicians of restraining old politicians, the similar exploitation of the voter trust is inevitable when B is small.

The idea of this chapter is based on the party reputation; young politicians make the campaign effort even though it gives them negative utility, because it helps their party and the strong party will help them in future. The "strong party" is the party with good reputation, and young politicians invest in it. Tadelis (2002) is a OLG reputation paper and considers a model in which retiring agents can sell their firm names, or reputations. The values of firm names depend on their reputations. To keep a good reputation so that they can sell firm names at high price at the end of their life, even agents in their last period make costly effort. Similarity is clear, but difference is also clear. In this chapter, it is not old politicians in their last period but young politicians with future who make costly efforts. So, this chapter shows a different route to keep a good reputation.

The reality is not the major virtue of this model. Thus, the model depends on the strong assumptions. However, at same time, explicitly modeling parties as organizations consisting of individuals is a more realistic way of modeling parties than treating the parties as consistent existences or simply treating parties only as labels on candidates. Loosening strong assumptions would change results quantitatively. But as long as there is the intertemporal conflict of interests and there is some action observable to voters, the qualitatively same result could be obtained. As I wrote in the Introduction of this chapter, there are papers of the retrospective voting models, in which re-elections in the multi-period partially restrains the incumbent politicians from practicing selfish behaviors (Austen-Smith and Banks(1989).) Though this chapter could be counted as one of them, unlike usual retrospective voting models in which the incumbent will face the election on the results of his own policies directly, in the model of this chapter the incumbent (i.e., PC in office) will not face the election. PCs face election only once to become the incumbent. Thus this chapter assumes out the possibility of the self-restraint of the incumbent in the usual retrospective voting models. Still, improving voters' welfare is possible. Though there is no direct retrospective voting on politicians, the party structure makes it possible for voters to use the retrospective voting[13]. This is different from the case of repeated elections between two parties that are consistent entities. Though I consider a particular problem of information asymmetry, the same approach in this chapter can be applied to other moral hazard problems. What important is the conflict of interests in a party and the existence of a way for insiders to send signals to voters.

Although OLG models are often used in Political Economy literature (see Persson and Tabellini (2002)) and Alesina and Spear (1988) argue that "the 'overlapping generations' model can be usefully applied to the political arena," as far as I know, Alesina and Spear (1988) and Harrington (1992) are only other papers of electoral competition model with OLG political parties. Recent papers on how party structures affect politicians and elections consider static or finite period models (though Snyder and Ting (2003) consider infinitely repeated game.) As Alesina and Spear (1988) states, the OLG structure is a good way to capture the intertemporal conflict of interests among party members.

As I said, this chapter shares the OLG political party model with Alesina and Spear (1988) and Harrington (1992). The difference between their papers and this chapter is about what problem to analyze. Their papers analyze the problem of policy commitment. So, they consider about if politicians keep their choices during elections, and I analyze about what choices they make during elections. Usually the policy commitment is assumed in electoral competition models (this chapter also assumes it.) As Alesina (1988) shows, however, the lack of the assumption could have a big implication on credible platform choices. Then, Alesina and Spear (1988) and Harrington (1992) show that the policy commitment could be achieved as the equilibrium outcome, not as the assumption. Alesina and Spear (1988) consider a transfer scheme from young politicians to old politicians for the purpose. The young politicians make the transfer as they need the old politicians to keep their promises for their own future elections. Thus, the logic is based on the conflict of interest among party members[14].

As one more related paper, I would like to discuss about Caillaud and Tirole (1999). Based on the conflict of interests among party members, Caillaud and Tirole (1999) explain why policy convergence does not happen in reality, unlike in many electoral competition models. It is a one-period model (though implicitly it assumes the second period) and has an office-oriented party leader and ideological party rank-and-files. Because of the information asymmetry between the party leader and voters, the leader might not work hard for voters. However, because of the conflict of interests between the leader and rank-and-file party members, rank-and-file members can work as monitors on the behavior of the leader for voters. Such monitoring can increase the possibility of winning election. But, since only ideological members conduct such monitoring (since effective monitoring could be costly), only ideological parties can have such monitoring.

If the results in the paper would be taken literary, they look like implying that the higher level of election campaigns might be good for voters. Clearly this is a wrong interpretation. What the results really imply is that the party structure that helps the party boss actually binds his behavior. Since he needs his party for him, he has to take care of the party, which has the longer life span than the party boss. Otherwise, party subordinates would send a signal to let voters know what is going on inside the party. There would be some obstacles to apply this result directly to actual election campaigns. Firstly, the campaign effort in this model is different from the actual campaign effort. In the model, the campaign effort is essentially the alias of VPCs' disutility. The costs of actual election campaigns are not entirely burdened by young politicians who will go to the future elections. This obscures observability. Also it would lessen the extent of the conflict of interests. For example, consider a case that VPCs can enjoy some of campaign contributions from special interests. It might not eliminate the conflict of interests entirely. But if voters are not sure about the extent to which VPCs can receive the contribution, this would add uncertainty. Secondly, there might be the information asymmetry between party leaders and subordinates. In such case, the ability of party subordinates working for voters is diminished. However, as long as the asymmetry within the party is less severe than the asymmetry between the party boss and voters, the ability is not zero.

Voters' appropriate reaction to the actions from party insiders is also very important for the results. It does not seem possible to obtain similar results without using voters' reaction but with cooperation in OLG organizations (Cremer (1986), Kandori (1992)[15].) Although the self-restraint through internal cooperation or inter-party cooperation might be important, I think, in electoral competition models, assuming a voters' appropriate reaction is a simpler and more intuitive way.

Appendix of Chapter 2

Lemma 2.1. *For any $B > 1$ and $c \in [0, \frac{B}{4}]$, an equilibrium exists.*

Proof. Let $D \equiv \{(\varepsilon^0, \varepsilon^1) \in [0, \frac{1}{2}] \times [0, \frac{1}{2}] : (2.1)^* \geq 0 \text{ and } (4)^* \geq 0 \text{ for } (\varepsilon^0, \varepsilon^1)\}$. This D is a set of consistent strategies, closed, and non-empty from the assumption $c \in [0, \frac{B}{4}]$. If $(\varepsilon^0, \varepsilon^1) \notin D$, the expected utility of PCs from this $(\varepsilon^0, \varepsilon^1)$ is zero. If $(\varepsilon^0, \varepsilon^1) \in D$, the expected utility of PCs is positive. Thus, PCs never choose $(\varepsilon^0, \varepsilon^1) \notin D$. Then, let P be the set of $(\varepsilon^0, \varepsilon^1) \in D$ which is strictly Pareto efficient. Since D is not an empty set, P is not empty. From the definition of equilibrium, this non-empty set P is the set of equilibrium strategies. ■

Proposition 2.1. *The equilibrium with $\varepsilon^0 + \varepsilon^1 < 1$ exists with $c \in (0, \frac{B}{4})$.*

Proof. $(2.1)^*$ $\frac{1}{2}(\frac{1}{2} - \varepsilon^0)(B - 2\varepsilon^0 + 4\varepsilon^1) - c$

$(2.4)^*$ $\frac{1}{2}(\frac{1}{2} - \varepsilon^1)(B - 2\varepsilon^0 + 4\varepsilon^1) - c$

If $c = 0$, $(1)^*$ and $(4)^*$ are zero for $\varepsilon^0 = \varepsilon^1 = \frac{1}{2}$. Then, slightly increase c from zero. At $\varepsilon^0 = \varepsilon^1 = \frac{1}{2}$, $(1)^*$ and $(4)^*$ become negative. However, from the previous Lemma, an equilibrium must exist. Since $(\frac{1}{2}, \frac{1}{2}) \notin D$ and all equilibrium strategies are in D, the equilibrium with $\varepsilon^0 + \varepsilon^1 < 1$ must exist. ■

The derivation of long-run average voter welfare

Since I consider symmetric equilibrium, $\varepsilon^0 = \hat{\varepsilon}^0$ and $\varepsilon^1 = \hat{\varepsilon}^1$. Then, the transition probabilities are,

$$(H, H) \Rightarrow (H, H)\text{: } \int_0^{\frac{1}{2}-\varepsilon^0} (1 - r)dr + \int_{\frac{1}{2}-\varepsilon^0}^{\frac{1}{2}+\varepsilon^0} (\frac{r}{2} + \frac{1 - r}{2})dr + \int_{\frac{1}{2}+\varepsilon^0}^{1} r dr$$
$$= \frac{3}{4} - (\varepsilon^0)^2$$

$(H,H) \Rightarrow (H,S)\&(S,H) : \frac{1}{4} + (\varepsilon^0)^2$

$(H,S)\&(S,H) \Rightarrow (H,H) :$

$$\frac{1}{2}[\int_0^{\frac{1}{2}-\varepsilon^1} rdr + \int_{\frac{1}{2}-\varepsilon^1}^{1}(1-r)dr + \int_0^{\frac{1}{2}+\varepsilon^1} rdr + \int_{\frac{1}{2}+\varepsilon^1}^{1}(1-r)dr]$$
$$= \frac{1}{4} + (\varepsilon^1)^2$$

$(H,S)\&(S,H) \Rightarrow (H,S)\&(S,H) : \frac{3}{4} - (\varepsilon^1)^2$

From these, the long-run probability of (H,H) at the start of period is derived as $\frac{1+4(\varepsilon^1)^2}{2+4(\varepsilon^0)^2+4(\varepsilon^1)^2}$, and the long-run probability of $(H,S)\&(S,H)$ is $\frac{1+4(\varepsilon^0)^2}{2+4(\varepsilon^0)^2+4(\varepsilon^1)^2}$. Next, the expected voter welfare when reputation is (H,H) is

$$\int_0^{\frac{1}{2}-\varepsilon^0} \{1-r-r\}dr + \int_{\frac{1}{2}-\varepsilon^0}^{\frac{1}{2}+\varepsilon^0}\{\frac{1}{2}(r-(1-r)) + \frac{1}{2}((1-r)-r)\}dr$$
$$+ \int_{\frac{1}{2}+\varepsilon^0}^{1}(r-(1-r))dr$$
$$= \frac{1}{2} - 2(\varepsilon^0)^2$$

The expected voter welfare when (H,S) or (S,H) is,

$$\frac{1}{2}[\int_0^{\frac{1}{2}-\varepsilon^1} (1-2r)dr + \int_{\frac{1}{2}-\varepsilon^1}^{1}(2r-1)dr + \int_0^{\frac{1}{2}+\varepsilon^1} (1-2r)dr + \int_{\frac{1}{2}+\varepsilon^1}^{1}(2r-1)dr]$$
$$= \frac{1}{2} - 2(\varepsilon^1)^2$$

Thus, the long-run average voter welfare is,

$$\frac{1+4(\varepsilon^1)^2}{2+4(\varepsilon^0)^2+4(\varepsilon^1)^2}\frac{1-4(\varepsilon^0)^2}{2} + \frac{1+4(\varepsilon^0)^2}{2+4(\varepsilon^0)^2+4(\varepsilon^1)^2}\frac{1-4(\varepsilon^1)^2}{2}$$
$$= \frac{1-16(\varepsilon^0\varepsilon^1)^2}{2+4(\varepsilon^0)^2+4(\varepsilon^1)^2}$$

Proposition 2.2. *If $B < 2$ and $c \in (0, \frac{B}{4})$, there is no equilibrium with $\varepsilon^0 = \varepsilon^1 = 0$. If $B \geq 2$, the equilibrium with $\varepsilon^0 = \varepsilon^1 = 0$ exists for $c = \frac{B}{4}$.*

Proof. (2.1)* $\frac{1}{2}(\frac{1}{2} - \varepsilon^0)(B - 2\varepsilon^0 + 4\varepsilon^1) - c$

$(2.4)^* \ \frac{1}{2}(\frac{1}{2} - \varepsilon^1)(B - 2\varepsilon^0 + 4\varepsilon^1) - c$

Let $(\varepsilon^{0*}, \varepsilon^{1*})$ be an equilibrium strategy. I stated conditions (a) above that $(\varepsilon^{0*}, \varepsilon^{1*})$ must satisfy to be an equilibrium. Actually, there are more conditions.

(a) $(2.1)^*$ and $(2.4)^*$ are non-negative for $(\varepsilon^{0*}, \varepsilon^{1*})$.

(b) At least one of $(2.1)^*$ and $(2.4)^*$ is zero at $(\varepsilon^{0*}, \varepsilon^{1*})$.

Denote this one as $(z)^*$ (this could be one of $(2.1)^*$ and $(2.4)^*$ or both.)

(c) At least one of the following two holds.

$(c.1)$ $\varepsilon^{0*} = \varepsilon^{1*} = \frac{1}{2}$.

$(c.2)$ For at least one $(z)^*$, the derivative of $(z)^*$ with respect to ε^0 and ε^1 are non-positive[16].

If (a) holds but (b) does not hold, both $(2.1)^*$ and $(2.4)^*$ would be positive. Then, PCs could deviate from $(\varepsilon^{0*}, \varepsilon^{1*})$ with holding (a). Similar thing can be said to (c); if (c) does not hold, PCs could deviate successfully. Thus, condition (b) and (c) are necessary conditions for $(\varepsilon^{0*}, \varepsilon^{1*})$ to be an equilibrium strategy. Then, substituting $\varepsilon^{0*} = \varepsilon^{1*} = 0$ into $(z)^*$, I obtain $c = \frac{B}{4}$ the maximum of c. Actually, if $c = \frac{B}{4}$, both of $(2.1)^*$ and $(2.4)^*$ are zero for $\varepsilon^{0*} = \varepsilon^{1*} = 0$ (conditions (a) and (b) hold.) This value of c, however, does not guarantee that $\varepsilon^{0*} = \varepsilon^{1*} = 0$ is an equilibrium. Since $\varepsilon^{0*} = \varepsilon^{1*} = 0$ cannot satisfy $(c.1)$, it has to satisfy $(c.2)$. Derivatives of $(2.1)^*$ and $(2.4)^*$ with respect to and ε^0 and ε^1 at $(\varepsilon^{0*}, \varepsilon^{1*})$ are,

$$\frac{\partial(2.1)^*}{\partial\varepsilon^0} = -\frac{1}{2}(B - 2\varepsilon^0 + 4\varepsilon^1) - (\frac{1}{2} - \varepsilon^0) < 0$$

$$\frac{\partial(2.1)^*}{\partial\varepsilon^1} = 1 - 2\varepsilon^0 \geq 0$$

$$\frac{\partial(2.4)^*}{\partial\varepsilon^0} = -(\frac{1}{2} - \varepsilon^1) \leq 0$$

$$\frac{\partial(2.4)^*}{\partial\varepsilon^1} = -\frac{1}{2}(B - 2\varepsilon^0 + 4\varepsilon^1) + 1 - 2\varepsilon^1$$

Suppose that $B < 2$. Then, both $\frac{\partial(2.1)^*}{\partial\varepsilon^1}$ and $\frac{\partial(2.4)^*}{\partial\varepsilon^1}$ at $\varepsilon^{0*} = \varepsilon^{1*} = 0$ is positive. Thus, $\varepsilon^{0*} = \varepsilon^{1*} = 0$ cannot be an equilibrium when $B < 2$. This result does not depend on the upper bound of c, $\frac{B}{4}$. Even if c can go higher than $\frac{B}{4}$, $\varepsilon^0 = \varepsilon^1 = 0$ could not be obtained. If $c > \frac{B}{4}$, all of (2.1), (2.2), (2.3) and (2.4) become negative at $\varepsilon^0 = \varepsilon^1 = 0$. VPCs never accept of any of ε^0 and ε^1.

Next, suppose that $B \geq 2$ and that there is an equilibrium with $(\varepsilon^{0*}, \varepsilon^{1*}) \neq (0, 0)$ for $c = \frac{B}{4}$. Because $(\varepsilon^{0*}, \varepsilon^{1*})$ is Pareto superior to $(0, 0)$, if such equilibrium exists, $\varepsilon^{0*} = 0$ and $\varepsilon^{1*} = 0$ cannot be an equilibrium for $c = \frac{B}{4}$. Since $c = \frac{B}{4}$, the equilibrium $(\varepsilon^{0*}, \varepsilon^{1*})$ cannot satisfy $(c.1)$ since $\varepsilon^{0*} = \varepsilon^{1*} = \frac{1}{2}$ and the positive c make $(2.1)^*$ and $(2.4)^*$ negative; (a) cannot hold. So, $(c.2)$ must hold. Then, looking at $(2.1)^*$ and $(2.4)^*$, it is clear that if $\varepsilon^{0*} > \varepsilon^{1*}$ then $(2.1)^*$ is zero, and if $\varepsilon^{0*} < \varepsilon^{1*}$ then $(2.4)^*$ is zero. If $\varepsilon^{0*} = \varepsilon^{1*}$ then both are zero. I consider these three cases.

Case 1. $\varepsilon^{0*} > \varepsilon^{1*}$

In this case, $(2.1)^* = 0$ and $(2.4)^* > 0$. This case is not consistent with the assumption that $(\varepsilon^{0*}, \varepsilon^{1*})$ is an equilibrium. Since $c > 0$, ε^{0*} cannot be $\frac{1}{2}$. For $\varepsilon^{0*} < \frac{1}{2}$, $\frac{\partial(2.1)^*}{\partial\varepsilon^1} = 1 - 2\varepsilon^{0*} > 0$; condition $(c.2)$ is no satisfied.

Case 2. $\varepsilon^{1*} > \varepsilon^{0*}$

In this case, $(2.4)^* = 0$ and $(2.1)^* > 0$. Since $\frac{\partial(2.4)^*}{\partial\varepsilon^0} < 0$ (the equality holds only when $\varepsilon^{1*} = \frac{1}{2}$ and this does not happen as $c > 0$) and $\frac{\partial(2.4)^*}{\partial\varepsilon^1}$ is negative for $B \geq 2$, $(c.2)$ might be satisfied.

$$(\frac{1}{2} - \varepsilon^{1*})(B - 2\varepsilon^{0*} + 4\varepsilon^{1*}) - 2c = 0$$

$\Rightarrow -4(\varepsilon^{1*})^2 + (2 - B + 2\varepsilon^{0*})\varepsilon^{1*} - \varepsilon^{0*}$ (B and c are cancelled out as $c = \frac{B}{4}$)

$\Rightarrow \varepsilon^1 = \frac{2 - B + 2\varepsilon^0 + \sqrt{(2 + B + 2\varepsilon^0)^2 - 32c}}{8}$ 17

Since $\varepsilon^{1*} > \varepsilon^{0*}$, $\sqrt{(B - 2 - 2\varepsilon^{0*})^2 - 16\varepsilon^{0*}} > (B - 2 - 2\varepsilon^{0*}) + 8\varepsilon^{0*} \geq 0$. From this,

$(B - 2 - 2\varepsilon^{0*})^2 - 16\varepsilon^{0*} > (B - 2 - 2\varepsilon^{0*})^2 + 16\varepsilon^{0*}(B - 2 - 2\varepsilon^{0*}) + 64(\varepsilon^{0*})^2$

$\Rightarrow 32(\varepsilon^{0*})^2 + 16(B - 1)\varepsilon^0 < 0$

$\Rightarrow 32\varepsilon^{0*}(\varepsilon^{0*} + \frac{B-1}{2}) < 0$

So, $\varepsilon^{0*} \in (-\frac{B-1}{2}, 0)$. However, since ε^{0*} is not negative, this is a contradiction.

Case 3. $\varepsilon^{0*} = \varepsilon^{1*}$

In this case, $(2.1)^* = (2.4)^* = 0$. Let $\varepsilon = \varepsilon^{0*} = \varepsilon^{1*}$, and substitute ε into $(2.1)^*$ or $(2.4)^*$. Then,

$\frac{1}{2}(\frac{1}{2} - \varepsilon)(B + 2\varepsilon) - c = 0$

$\Rightarrow -2\varepsilon^2 + (1 - B)\varepsilon + \frac{B}{2} - 2c = -2\varepsilon(\varepsilon - \frac{1-B}{2}) = 0$

Since $\frac{1-B}{2} < 0$, it must be $\varepsilon = 0$. Thus, $(\varepsilon^{0*}, \varepsilon^{1*}) = 0$. ■

Proposition 2.3. *When* $B \in (1, 2)$ *and* $c \in (0, \frac{B}{4})$*, equilibria are characterized by,*

$$\frac{1}{2} > \varepsilon^1 \geq \varepsilon^0$$
$$\varepsilon^1 \geq \frac{2 - B + 2\varepsilon^0}{8}$$
$$\varepsilon^1 = \frac{2 - B + 2\varepsilon^0 + \sqrt{(2 + B + 2\varepsilon^0)^2 - 32c}}{8}$$

Proof. Let $(\varepsilon^0, \varepsilon^1)$ be an equilibrium. The proof of previous Proposition already showed that the case of $\varepsilon^0 > \varepsilon^1$ at equilibrium is not possible (this result does not depend on the value of B.) Thus, always $\varepsilon^1 \geq \varepsilon^0$. From this, $(2.4)^* = \frac{1}{2}(\frac{1}{2} - \varepsilon^1)(B - 2\varepsilon^0 + 4\varepsilon^1) - c = 0$. Solving this, $\varepsilon^1 = \frac{2 - B + 2\varepsilon^0 + \sqrt{(2 + B + 2\varepsilon^0)^2 - 32c}}{8}$. Finally, since $c > 0$, ε^1 is less than $\frac{1}{2}$. Then, $(c.2)$ in the proof of previous Proposition must hold for $(2.4)^*$. So, $\frac{\partial (2.4)^*}{\partial \varepsilon^0}$ and $\frac{\partial (2.4)^*}{\partial \varepsilon^1}$ must be non-positive. $\frac{\partial (2.4)^*}{\partial \varepsilon^0} = -(\frac{1}{2} - \varepsilon^1) \leq 0$. For $\frac{\partial (2.4)^*}{\partial \varepsilon^1} = -\frac{1}{2}(B - 2\varepsilon^0 + 4\varepsilon^1) + 1 - 2\varepsilon^1$ to be non positive, $\varepsilon^1 \geq \frac{2\varepsilon^0 + 2 - B}{8}$. Thus, any equilibrium satisfies above three conditions. Then, pick $(\varepsilon^{0*}, \varepsilon^{1*}) \in [0, \frac{1}{2}] \times [0, \frac{1}{2}]$ that satisfies above three. This $(\varepsilon^{0*}, \varepsilon^{1*})$ satisfies (a), (b) and (c) in the proof of previous Proposition. I check if PCs have an incentive to deviate from $(\varepsilon^{0*}, \varepsilon^{1*})$. Since $\frac{\partial (2.4)^*}{\partial \varepsilon^1} \leq 0$ at $(\varepsilon^{0*}, \varepsilon^{1*})$, $\frac{\partial^2 (2.4)^*}{\partial \varepsilon^1 \partial \varepsilon^1} < 0$ and $\frac{\partial (2.4)^*}{\partial \varepsilon^0} \leq 0$, any point $(\varepsilon^0, \varepsilon^1) \in [0, \frac{1}{2}] \times [0, \frac{1}{2}]$ such that $\varepsilon^0 \geq \varepsilon^{0*}$, $\varepsilon^1 \geq \varepsilon^{1*}$ and $(\varepsilon^{0*}, \varepsilon^{1*}) \neq (\varepsilon^0, \varepsilon^1)$ makes $(2.4)^*$ negative. Thus, PCs will not deviate to such points. Any points satisfying above three conditions are equilibrium. So, above three conditions are the characterization of equilibria for $B \in (1, 2)$ and $c \in (0, \frac{B}{4})$. ∎

Proposition 2.4. *When $B \in (1, 2)$, the best long-run average voter welfare achievable is $\frac{16}{(B-2)^2 + 32}$ with $\varepsilon^0 = 0$ and $\varepsilon^1 = \frac{2-B}{8}$ by setting $c = \frac{(B+2)^2}{32}$.*

Proof. Let $E \equiv \{(\varepsilon^0, \varepsilon^1) \in [0, \frac{1}{2}] \times [0, \frac{1}{2}] : \varepsilon^1 \geq \frac{2\varepsilon^0 + 2 - B}{8}\}$. If $(\varepsilon^0, \varepsilon^1)$ is not in E, that $(\varepsilon^0, \varepsilon^1)$ cannot be an equilibrium. So, I consider only $(\varepsilon^0, \varepsilon^1)$ in E. The long-run average voter welfare is $\frac{1 - 16(\varepsilon^0 \varepsilon^1)^2}{2 + 4(\varepsilon^0)^2 + 4(\varepsilon^1)^2}$, and its indifference curve is $\varepsilon^1 = \frac{1}{2}\sqrt{\frac{1-k}{k + 4(\varepsilon^0)^2} - k}$ (k is a fixed welfare level.) Since the marginal rate of substitution between ε^0 and ε^1 for voters is $\frac{d\varepsilon^1}{d\varepsilon^0} = -\frac{\varepsilon^0(1-k)}{\varepsilon^1(k + 4\varepsilon^0)}$, the indifference curve is downward sloping. Because the bottom side of E is upward sloping (on the bottom, $\frac{d\varepsilon^1}{d\varepsilon^0} = \frac{1}{4}$) and $(0,0) \notin E$, it is obvious that voter welfare is maximized at the lower left corner of E. The corner is $\varepsilon^0 = 0$ and $\varepsilon^1 = \frac{2-B}{8}$. I check if this can be an equilibrium for some value of c. Substituting $\varepsilon^0 = 0$ and $\varepsilon^1 = \frac{2-B}{8}$, $\varepsilon^1 = \frac{2 - B + 2\varepsilon^0 + \sqrt{(2 + B + 2\varepsilon^0)^2 - 32c}}{8}$ holds only when $c = \frac{(B+2)^2}{32}$. This $c = \frac{(B+2)^2}{32}$ makes $(0, \frac{2-B}{8})$ an equilibrium strategy. Since $\varepsilon^0 = 0$ and $\varepsilon^1 = \frac{2-B}{8}$ maximize (2.4)* in $[0, \frac{1}{2}] \times [0, \frac{1}{2}]$ when $B \in (1,2)$, the combination of $\varepsilon^0 = 0$ and $\varepsilon^1 = \frac{2-B}{8}$ is the only equilibrium when $c = \frac{(B+2)^2}{32}$; $(0, \frac{2-B}{8})$ is the only intersection of a curve $\varepsilon^1 = \frac{2 - B + 2\varepsilon^0 + \sqrt{(2 + B + 2\varepsilon^0)^2 - 32c}}{8}$ and E. Thus, any other point in $[0, \frac{1}{2}] \times [0, \frac{1}{2}]$ cannot satisfy even condition for $c = \frac{(B+2)^2}{32}$. Finally, at $(0, \frac{2-B}{8})$, the long-run average voter welfare is $\frac{16}{(B-2)^2 + 32}$. ∎

Proposition 2.5. *With the reputation update (1)', if* $B \geq 2$, *the equilibrium is* $\varepsilon^0 = \varepsilon^1 = 0$. *If* $B \in (1,2)$, *the equilibrium is* $\varepsilon^0 = 0$ *and* $\varepsilon^1 = \frac{2-B}{8}$.

Proof. PCs of both parties want to increase the campaign effort level to the maximum. This maximum campaign effort level is same to the maximum campaign effort level that VPCs would accept with the reputation update (1). When $B \geq 2$, since $(0,0)$ is the only equilibrium with the update assumption (1) and its required level of c is $\frac{B}{4}$, $(0,0)$ is also the equilibrium result with the assumption (1)'. When $B \in (1,2)$, the combination of $\varepsilon^0 = 0$ and $\varepsilon^1 = \frac{2-B}{8}$ maximizes the value of (2.4)* in $E \equiv \{(\varepsilon^0, \varepsilon^1) \in [0, \frac{1}{2}] \times [0, \frac{1}{2}] : \varepsilon^1 \geq \frac{2\varepsilon^0 + 2 - B}{8}\}$. Thus, in both cases, the optimal with the assumption (1) is obtained with the assumption (1)'. ∎

Endnotes of Chapter 2

5. There are many empirical research papers about special interest groups based on Grossman and Helpman (1996). They found a "puzzle" that the governments of US and other countries put much higher weights on consumer welfare than on contributions from special interests. Gawande and Hoekman (2006) also obtain such a result on agricultural special interests and argue that such results are due to the assumption of "unitary" government in Grossman and Helpman (they also talk about "policy uncertainty", which is the uncertainty about legislative outcomes, not the uncertainty about the outcomes of implemented policies in this chapter.)

6. In the literature of the theoretical studies of legislative voting in Political Science, parties and politicians are different existences, and the function of parties are big subject (Aldrich & Rohde, Krehbiel & Meirowitz (2002), etc.)

7. Of course, no good/bad politician assumption and the non-Bayesian, simple reputation update are related to each other. Because of the non-Bayesian update, I do not need good/bad politicians (or some other assumption about characteristics of politicians.) Because of no good/bad politicians, a simple Bayesian update does not work in the model.

8. Alternatively, this can be considered as follows; there are three types of voters: voters who have the affiliation with party 1, voters who have the affiliation with party 2, and independent voters who have no affiliation with any party. Neither of party 1 and party 2 voter groups is a majority group. Thus, the election outcome depends on the independents. In other words, they are median voters.

9. See Alesina and Spear (1988) and Harrington (1992) for the policy commitment in OLG political parties models.

10. This is a simplified, short cut way of modeling the signal. Following Majumdar and Mukand (2004), this can be modeled as follows; at first politicians in both parties have the belief $\frac{1}{2}$ as the probability of the true state being w^1. Then politicians receive a common stochastic signal x, which is in $[\underline{x}, \bar{x}]$. This signal x would be received with the density function ϕ^1 if the state is w_1 and the density function ϕ^2 if the state is w^2. I assume that $\lim_{x\to\bar{x}} \phi^1(x)/\phi^2(x) = \infty$ and $\lim_{x\to\underline{x}} \phi^1(x)/\phi^2(x) = 0$. Then, from this signal, the candidate deduces new probability r of the state being w^1 using the Bayesian rule,

$r(x) = \frac{1}{2}\phi^1(x)/(\frac{1}{2}\phi^1(x) + \frac{1}{2}\phi^2(x)) = \phi^1(x)/\phi^2(x)/(\phi^1(x)/\phi^2(x) + 1)$ Assuming $\phi^1(x)/\phi^2(x)$ is strictly increasing as a function of x, x has a one to one correspondence. The distribution of r is derived from ϕ^1 and ϕ^2. For example, $\phi^1(x) \equiv 2(x-\underline{x})/(\bar{x}-\underline{x})^2$ and $\phi^2(x) \equiv 2(\bar{x} - x)/(\bar{x} - \underline{x})^2$ produces the uniform distribution for r.

11. Though this simplifies calculations, this assumption makes an interpretation of this "preference" difficult. If this “preference” comes from politician’s ideological preference, politicians should receive non-zero utility whoever implements those policies.

One interpretation is that politicians feel good or bad from the act of implementing policies that do or do not fit to their ideologies. Other interpretation is that this “preference” does not come from politician’s ideologies but from benefits or punishments given by special interests according to policies they implement.

12. If VPCs receive the same utility as PCs do from implementing policies, results

described later should be changed quantitatively. But the qualitative result is same.

13. New York Times columnist Thomas Friedman argued that the fact that Vice-President Cheney was not going to run for the President partially explains the problems of President Bush in his second term (New York Times, June 22, 2005).

14. However, Harrington (1992) points out that actually the scheme of Alesina and Spear (1988) is not credible. Instead, Harrington (1992) provides a different logic. He assumes that politicians have one more period after they retire. As long as their parties implement their favored policy, there could be an incentive for them to keep their campaign platform so that their party would not be disadvantaged against their opposite party at the election when they retire.

15. The model in this paper is a stochastic one. For the difficulty of cooperation by OLG organization in stochastic situation, see Messner and Polborn (2003). For the difficulty caused by limited communication, see Lagunoff and Matsui (2004).

16. $\frac{\partial(2.1)^*}{\partial\varepsilon^0}$ is always negative. $\frac{\partial(2.1)^*}{\partial\varepsilon^1}$ and $\frac{\partial(2.4)^*}{\partial\varepsilon^0}$ can be zero only when $\varepsilon^0 = \frac{1}{2}$ and $\varepsilon^1 = \frac{1}{2}$, respectively. So, I can ignore the case that they are zero. However, $\frac{\partial(2.4)^*}{\partial\varepsilon^1}$ can be possible even when neither of ε^0 and ε^1 is $\frac{1}{2}$. When $\frac{\partial(2.4)^*}{\partial\varepsilon^1} = 0$, the argument in the condition (b) is not correct in the strict sense. But, since $\frac{\partial^2(2.4)^*}{\partial\varepsilon^1\partial\varepsilon^1} = -2\varepsilon^1 - 2 < 0$, even the slightest increase of ε^1 will eventually make $(2.4)^*$ negative. Thus, I can use the condition (b) in current form.

17. From $(2.4)^* = 0$, $\varepsilon^1 = (2 - B + 2\varepsilon^0 \pm ((2 + B + 2\varepsilon^0)^2 - 32c)^{0.5})/8$ is obtained. However, $\varepsilon^1 = (2 - B + 2\varepsilon^0 - ((2 + B + 2\varepsilon^0)^2 - 32c)^{0.5})/8$ is not answer since this is negative when c is very low.

References

[1] Aldrich, John H. AND David W. Rohde: "Measuring Conditional Party Government" Presented at the annual meeting of the Midwest Political Science Association, Chicago.

[2] Alesina, Alberto (1988): "Credibility and policy convergence in a two-party system with rational voters" *American Economic Review*, 78, 796-806.

[3] Alesina, Alberto AND Stephen E. Spear (1988): "An Overlapping Generations Model of Electoral Competition" *Journal of Public Economics*, 37, 359-379.

[4] Austen-Smith, David AND Jeffrey Banks (1989): "Electoral Accountability and Incumbency" in *Models of Strategic Choice in Politics*, ed by Peter C. Ordeshook, Ann Arbor, The University of Michigan Press

[5] Caillaud, Bernard AND Jean Tirole (2002): "Parties as Political Intermediaries" *The Quarterly Journal of Economics*, 1453-1489.

[6] Calvert, Randall (1985): "Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence" *American Journal of Political Science*, 39, 69-95.

[7] Cremer, Jacques (1986): "Cooperation in Ongoing Organizations" *The Quarterly Journal of Economics*, 101, 33-49.

[8] Friedman, Thomas L. (2005): "Run, Dick, Run" *New York Times*, June 22, 19.

[9] Gawande, Kishore AND Bernard Hoekman (2006): "Lobbying and agricultural trade policy in the United States" World bank policy research working paper 3819.

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

[11] Harrington, Jr., Joseph E. (1992): "The Role of Party Reputation in the Formation of Policy" *Journal of Public Economics*, 49, 107-121.

[12] Harrington, Jr., Joseph E. (1993): "Economic Policy, Economic Performance, and Elections" *The American Economic Review*, 83(1), 27-42.

[13] Heidhues, Paul AND Johan Lagerlöf (2003): "Hiding information in electoral competition" *Games and Economic Behavior*, 42, 48-74.

[14] Kandori, Michihiro (1992): "Repeated Games Played by Overlapping Generations of Players" *Review of Economic Studies*, 59, 81-92.

[15] Krehbiel, Keith AND Adam Meirowitz (2002): "Minority Rights and Majority Power: Theoretical Consequences of the Motion to Recommit" *Legislative Studies Quarterly*, 27, 191-217.

[16] Kreps, David M (1990): "Corporate Culture and Economic Theory" in James E. Alt and Kenneth A. Shepsle, eds., *Perspectives on Positive Political Economy.* Cambridge, Ma: Cambridge University Press, 90-143.

[17] Lagunoff, Roger AND Akihiko Matsui (2004): *"Organizations and Overlapping Generations Games: Memory, Communication, and altruism"* Review of Economic Design, 8, 383-411.

[18] Levy, Gilat (2004): "A model of Political Parties." *Journal of Economic Theory*, 115, 250-277.

[19] Majumdar, Sumon AND Sharun W Mukand (2004): "Policy Gambles" *The American Economic Review*, 94(4), 1207-1222.

[20] Messner, Matthias AND Mattias K. Polborn (2003): "Cooperation in Stochastic OLG games" Journal of Economic Theory, 108, 152-168.

[21] Morrison, Alan D. AND William J. Wilhelm, Jr. (2004): "Partnership Firms, Reputation, and Human Capital" *The American Economic Review*, 94(5), 1682-1692.

[22] Persson, Torsten AND Guido Tabellini (2002): *Political Economy Explaining Economic Policy*, M.I.T. press.

[23] Snyder, Jr., James M. AND Micheal M. Ting (2002): "An Informational Rationale for Political Parties" *American Journal of Political Science*, 46(1), 90-110.

[24] Snyder, Jr., James M. AND Micheal M. Ting (2003): "Roll Calls, Party Labels, and Elections" *Political Analysis*, 11(4), 419-444.

[25] Tadelis, Steven (2002): "The Marker for Reputations as an Incentive Mechanism" *Journal of Political Economy*, 110(4), 854-882.

[26] Wittman, Donald (1983): "Candidate Motivation: A synthesis of Alternatives" *American Political Science Review*, 77, 142-157.

3 The Choice of Inefficient Instruments in a Simple Retrospective Voting Model with Voter Abstention

3.1 Introduction

The objective of this chapter is to propose one explanation about why governments use inefficient policy instruments. Particularly, I consider a case in which an incumbent politician in office uses a tariff to help an import-competing industry as a decision-making problem of the incumbent. As every introductory International Economics textbook explains, the tariff is inefficient compared to the production subsidy as a policy instrument to help domestic industry. Moreover, as Dixit and Norman (1980) and Dixit (1986) shows, it is possible to make the transition from the trade with tariffs to the trade with no tariff Pareto-improving by using appropriate lump sum transfers or commodity taxes/subsidies. If so, why at all does the incumbent use tariffs when he could make every voter better off by using taxes/subsidies instead of tariffs? Given the natural assumption that the election outcome depends on the utility levels of voters, the combination of tax and subsidy should be better than tariffs not only for voters but also for the incumbent's prospect of winning next election! Though there is the huge literature of Endogenous Tariff Formation (Mayer (1984), Grossman and Helpman (1994), Yang (1995), etc), they do not answer this question because they assume away subsidy, which Rodrik (1995) and Dixit and Roemer (2006) criticize as restrictions of possible policy instruments. To the question, this chapter tries to provide one explanation. The hinge of the explanation is the voter abstention.

The idea of the explanation is simple; if the incumbent decides to help domestic industry, other people will have to pick up the tab. Though the incumbent can expect positive feedbacks during the election from people who receive the help (as votes, campaign contributions, etc.), the help will cost the incumbent some votes from people who finance it involuntarily. However, in reality, unlike many electoral competition models of Political Economy in which all constituency vote, many people actually abstain from voting. Then, the incumbent might be able to make the loss of votes small by distributing the cost of the help to people who do not vote. The participation rates of voter groups depend on characteristics of groups (Lijphart (1997)). Assuming several different policy instruments, such characteristics must have influence on the choice of policy instruments by the incumbent. According to empirical researches on the behavior of voters, one of the important characteristics is income level (Lijphart (1997)). Generally, the higher the income level of a group, the higher the participation rate of the group. This does not mean that all political choices are for rich people, since most of actual voters are middle class or working class anyway[18]. At margin, however, different income levels of different voter groups could have an implication on political choices.

The situation I consider is as follows; an import-competing domestic industry is lobbying an incumbent to help them[19]. The incumbent can use both a tariff and a production subsidy simultaneously. If the incumbent decides to help them, he has to choose the combination of the tariff and tax to finance the help. After the decision, the incumbent will go to election. The outcome of the election depends on utilities voters received. The incumbent is office-motivated, i.e., winning the election is his only objective. To win the election in this model, the incumbent has to take two things into his consideration: voter heterogeneity and voter abstention. Voter heterogeneity is the main idea behind the endogenous tariff formation papers (ex., Mayer (1984)). Because of the heterogeneity, different voters might prefer different policies. In the model of this chapter, the heterogeneity comes from different income levels. Combined with common non-homothetic preference, it derives different consumption patterns among voter groups. Though voter abstention was incorporated in a model of Mayer (1984), the idea of abstention has not been explored further in the literature of Endogenous Tariff Formation. There are many papers about why some people abstain from voting (Harsanyi (1980), Shachar and Nalebuff (1999), Feddersen and Sandroni (2006). As surveys of abstention, Aldrich (1993) and Feddersen (2004).) However, since my attention is not on the abstention itself but on its potential effect for the choice of policy instruments, I simply assume a voter-abstention model based on the Group Rule-Utilitarian literature (Feddersen and Sandroni (2006), Coate and Conlin (2004).) With the voter abstention, the incumbent can manipulate the participation rates of different voter groups by distributing the cost and benefit of helping domestic industry among voter groups. Then, he will choose the set of the policy instruments that maximizes the probability of his winning the election. Within this situation, I consider under what condition he would use the tariff.

In the following, I describe the model in section 3.2, results in section 3.3 (two voter groups) and in section 3.4 (three voter groups.) I review related papers briefly in section 3.5. Section 3.6 is the conclusion of this chapter.

3.2 Model

The model is a small country model with two voter groups (I change this to three groups in section 3.4) that produce two goods. I explain the economy side of the model first, then the political side.

3.2.1 Economy side

The two groups in the small country are groups of capitalists and workers, called as group K and L, respectively. The population is continuum, and groups have population N^K and N^L with $N^K + N^L = 1$ and $N^K < N^L$. Each capitalist has one unit of capital, and each worker has one unit of labor. There are two goods, X and Y. Good X is the exporting good, and one unit of good X is produced by one unit of labor. Good X is the numeraire. The good Y is the import-competing good, and requires labor and capital for its production. Let $F(k,l)$ be its production function with k and l as units of capital input and labor input. This F is an increasing, differentiable and constant return to scale technology function. However, since the supply of capital is fixed, I consider $f(l) \equiv F(N^K, l)$ instead (assuming perfect competition.) This f is a decreasing return to scale function of the country wide labor input for good Y production, l. I treat f as if it is the function of total industry of good Y. The before-tax incomes of capitalists and workers, I^K and I^L, come from profits and wages. Since one unit of labor can produce one unit of good X, workers receive income $I^L = 1$ from their one unit of labor by producing good X or as wage from working in the production of good Y. Let τ and s be the specific tariff of good Y and the production subsidy per unit of good Y, respectively. The tariff revenue and the income tax, which I explain later, are used to finance the production subsidy. This assumption might cause a conceptual problem for the case that the incumbent wants to use only tariff. However, since I consider in this paper if the incumbent might have an incentive to use the tariff when he decides to help capitalists, this assumption is innocuous for my purpose[20]. Let $\bar{p}$ and $p \equiv \bar{p}+\tau$ be the fixed world price and domestic price of good Y, respectively. Then, the total profit from production of good Y is $(\bar{p}+\tau+s)f(l)-l$. Since the population size of group K is N^K, the before-tax income of a capitalist is $I^K \equiv \dfrac{(p+s)f-l}{N^K}$. I assume that $I^K > 1$, so capitalists are richer than workers.

Let x^a and y^a be the consumption of good X and good Y by one member of group $a \in \{K, L\}$. All members of two groups have same utility function, $U(x, y)$; no disutility of labor for workers. Thus, the income tax does not cause distortions, same as lump-sum tax[21]. This utility function is an increasing, strictly concave and differentiable function in both arguments. Let $y(I^a, p)$ be the common demand function of good Y of a member of group $a \in \{K, L\}$ derived from the utility maximization of $U(x, y)$ with income I^a (here abusing notation, I^a is after-tax, disposable income) and the domestic price of good Y, p. From this, the demand for good X of a member of group $a \in \{K, L\}$ is $x^a = I^a - py(p, I^a)$. Here, deviating from the usual assumption in International Economics, I assume the common utility function U is not homothetic and that the income elasticity of good Y is less than one. The total and average demand of good Y is $\bar{y} \equiv N^K y^K + N^L y^L$. Since $f(l)$ is the total domestic production of good Y, the import of good Y is defined as $m \equiv \bar{y} - f(l)$. The indirect utility function is $V(I^a, p) = U(I^a - py(I^a, p), y(I^a, p))$, $a \in \{K, L\}$.

3.2.2 Political side

I assume that voters decide to vote for the incumbent or not based on the utility they received during the incumbent's term compared to the utility they would receive under the policy of the challenger, which is assumed to be the free trade policy. So, this is the retrospective voting. The incumbent has to take into his consideration the reactions from his constituency. The objective of the incumbent politician is winning the election, i.e., the maximization of the probability of winning the election. The only decisions that incumbent can make before the election are the decision to help capitalists or not and how to finance the help. The challenger at the election proposes the free trade policy. His role in this model is just to provide an alternative to voters and makes the retrospective voting possible for them. He can be considered as a convinced free trader.

The voting rates of groups are determined by two factors: the voting decision making, which is also based on voting costs and utility differences between the policies of the incumbent and the challenger, and stochastic terms. These two determine the actual number of votes cast and the election outcome.

Voter abstention: Some people do not vote. The problem of the voter abstention, or the paradox of not voting (Feddersen (2004)), is a big issue of Political Economy, but not a topic of this chapter. So, here, I simply use a modified version of the Group Rule-utilitarian voting model (Feddersen and Sandroni (2006), Coate and Conlin (2004).) If the incumbent decides to help capitalists, no potential voter is indifferent between two candidates. Every capitalist prefers the incumbent to the challenger, and every worker prefers the challenger to the incumbent. Then, they face two obstacles before actual voting for their preferred candidate. First, voting is costly. Take a potential voter in group $a \in \{K, L\}$. Let d^a be the utility difference between his preferred candidate and the opponent. Let c be his voting cost, which varies among potential voters. Then, if the voting cost is so high that $d^\alpha - c < 0$, there is no reason for him to vote. Thus the voting cost is the first obstacle. The second obstacle is the fact that the probability of being the pivotal voter is essentially zero unless the number of actual votes is very low. Since the population is continuum in this model, such probability is zero. So, there is actually no reason for anyone to vote. Following the Group Rule-utilitarian voting literature, however, I assume that some people still vote when $d^a - c \geq 0$.[22] I call potential voters who do not vote even when $d^a - c \geq 0$ as free-riders. They do not vote since they think other people will vote for them anyway. The fraction of such free-riders is unknown and stochastic.

Let ϕ^a be the participation rate of group $a \in \{K, L\}$ among non free-ride potential voters. This ϕ depends on the voting cost function, $c(n)$. This voting cost can be considered as consisting of time, money and other factors necessary for going to a voting booth. Different potential voters have different voting costs. Then, the voting cost function $c(n)$ gives the voting cost of n-th voter in the group when every potential voters in group a are aligned from the lowest voting cost voter to the highest voting cost voter. So, n of potential voters in a group have voting cost less or equal to $c(n)$. I assume that both groups have same cost function and that $c' > 0$ and $c'' < 0$. As a simple example giving a cost function of this type, consider a voting district with one voting booth at the center of the district. Potential voters live uniformly in the district, and the voting cost of a potential voter is represented by the distance from his home to the booth. Then, voting function becomes concave like c. Further, I assume that $c(0) < 0$, so some people always want to vote, maybe due to the sense of civic duty or strong partisan feelings. Since both groups have the same voting cost function $c(n)$, the number of votes is same for both groups when $d^K = d^L = 0$. Since $N^L > N^K$, this means $\phi^K > \phi^L$ at no utility difference. I assume that all actual capitalist voters vote for the incumbent and all actual working voters vote for the challenger even when they are indifferent between two candidates. Then, given the utility difference $d^a > 0$, the participation rate of the group a, ϕ^a, is defined by,

$$d^a - c(\phi^a N^a) = 0$$

From this,

$$\frac{\partial \phi^a}{\partial d^a} = \frac{1}{N^a c'(\phi^a N^a)} > 0$$

The participation rate ϕ^a of non free-ride potential voters in group a is determined by d^a and c. But the actual number of votes is determined by ϕ^a and stochastic terms about the fraction of free-riders. This stochastic terms maybe come from many factors like the conflict of the voters' sense of the civic duty and the fact of the impossibility of being the pivotal voter, and all other uncertain factors not resolved until the election day, like weather, possible transportation troubles, abrupt schedule changes of potential voters, etc. I assume the following setup; a potential voter in group a

is not a free-rider with probability μ^a. Thus the potential voter is a free-rider with probability $1 - \mu^a$. For simplicity, all members in group $a \in \{K, L\}$ have same μ^a. This μ^a is a random variable.

The number of votes from group K is $\mu^K N^K \phi^K$, and the number of the votes from group L is $\mu^L \phi^L N^L$. Thus, the election result is determined by $\mu^K N^K \phi^K - \mu^L N^L \phi^L$. If $\mu^K N^K \phi^K - \mu^L N^L \phi^L > 0$, the incumbent wins the election[23]. Since μs are random variables, this is a random variable, too. Let $H(\mu^K, \mu^L)$ be the cumulative distribution function of μ^K and μ^L. Then, the probability of the incumbent winning the election can be calculated as $\int_{\mu^K N^K \phi^K - \mu^L N^L \phi^L > 0} dH(\mu^K, \mu^L)$. This is the objective of the incumbent. Without knowing the exact function form of H, it is not possible to know the exact value of the probability. However, to analyze the incumbent's choice, I do not need to know the exact form. Let A be the intersection of $[0,1] \times [0,1]$ and $\{(\mu^L, \mu^K) : \mu^K \geq \mu^L \frac{N^L \phi^L}{N^K \phi^K}\}$ in $\mu^L \times \mu^K$ dimension (see figure 3.1.)

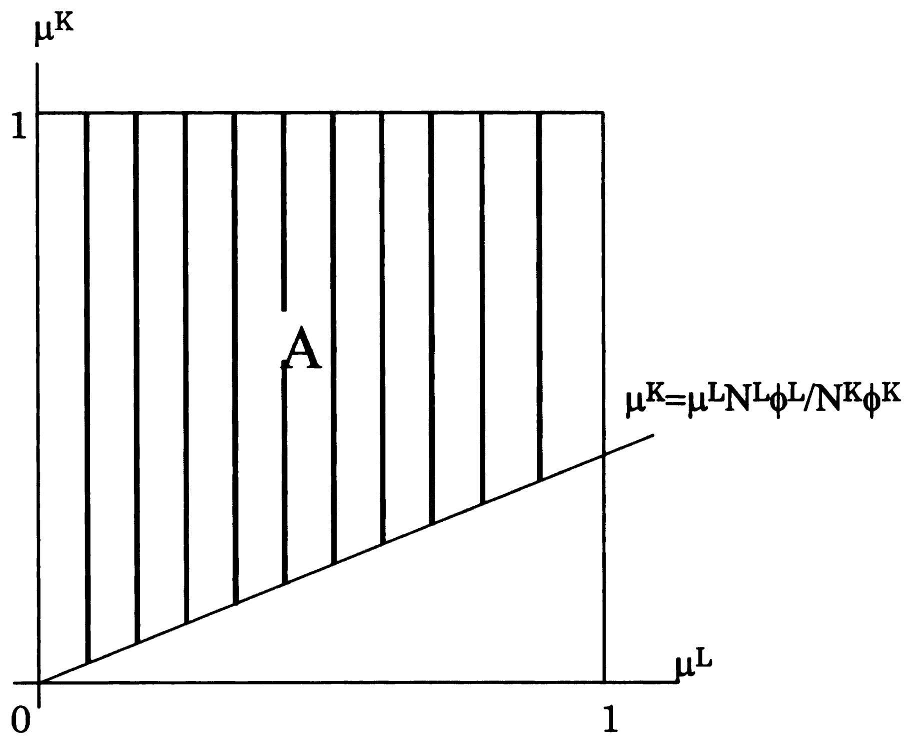

Figure 3.1

Clearly, the maximization of the probability of winning the election is equal to the maximization of the area of A. Moreover, since $\mu^K = \mu^L \frac{N^L \phi^L}{N^K \phi^K}$ always holds at $\mu^K = 0$ and $\mu^L = 0$, the maximization of the area of A is equal to the minimization of $\frac{N^L \phi^L}{N^K \phi^K}$. But this is equal to the maximization of $\frac{N^K \phi^K}{N^L \phi^L}$. Thus, the incumbent does not need to care about the distribution function H. I state this as a claim.

Claim 3.1 *The incumbent tries to maximize his probability of winning the election,* $\int_{\mu^K N^K \phi^K - \mu^L N^L \phi^L > 0} dH(\mu^K, \mu^L)$. *This is equal to the maximization of*

$$\frac{N^K \phi^K}{N^L \phi^L}$$

3.3 Results with two groups

Before stating results, let me briefly describe my strategy of obtaining results. Let the sum of the production subsidy and the tariff rate as $H \equiv s + \tau$. In the following analysis, firstly I consider if the incumbent has an incentive to use a tax to give the production subsidy s with zero tariff. Next, I consider if the incumbent has an incentive to raise the tariff rate τ for the given level of $H \equiv s + \tau$. Such increase of τ means $\frac{ds}{d\tau} = -1$. Since the objective of this chapter is to consider why the incumbent might want to use a tariff to give the help when he can give a production subsidy, this is an easy way to achieve the objective. Because of this order, the tariff does not cause additional distortions in production; it keeps the distortion in production at the level previously caused by the production subsidy, though it causes distortions in consumption. Since the tariff revenue is used to finance the production subsidy (with the tax revenue from the income tax), higher tariff means lower tax rate. Lower tax is good news for all people. But because the income elasticity of good Y is less than one, how good it is is different among people with different income levels. Regarding the tax rate, I consider two cases; the case of single uniform tax rate and the case of different tax rates for two groups.

3.3.1 Single tax rate

The incumbent will help capitalists? In this section, I consider if the incumbent would use the tax to provide the production subsidy to capitalists. I assume, for now, that the incumbent does not use the tariff. Then, the utility differences of capitalists and workers are defined as follows.

Utility difference:

$$d^K(t) = V((1-t)\frac{(p^* + s)f(l) - l}{N^K}, p^*) - V(\frac{p^* f(l^f) - l^f}{N^K}, p^*)$$

$$d^L(t) = V(1, p^*) - V(1-t, p^*)$$

Notice that they are defined as the subtraction of the utility of the non-preferred candidate's policy from the utility of the preferred candidate's policy. When the government uses only the income tax to provide the subsidy, its budget constraint is $sf(l) = t(\pi + N^L)$. Taking the derivative of this with respect to t, it is obtained that

$$\frac{ds}{dt} f(l) + sf'(l)\frac{\partial l}{\partial s}\frac{ds}{dt} = \pi + N^L \qquad \text{from } \frac{\partial \pi}{\partial l} = 0$$

The equilibrium condition for production is, $(p^* + s)f'(l) - 1 = 0$. From this,

$$\frac{dl}{ds} = -\frac{f'}{(p^* + s)f''} = -\frac{(f')^2}{f''}$$

Thus,

$$\frac{ds}{dt} = (\pi + N^L)(f - s\frac{(f')^3}{f''})^{-1} = \frac{(\pi + N^L)f''}{ff'' - s(f')^3}$$

$$d_t^K(t) = V_I^K\{-\frac{\pi}{N^K} + \frac{f(l)\frac{\partial s}{\partial t}}{N^K}\} \qquad \text{from } \frac{d\pi}{dl} = 0$$

$$d_t^L(t) = V_I^L \frac{dt}{ds}$$

Then, the numerator of the derivative of $\frac{N^K \phi^K}{N^L \phi^K}$ with respect to t is,

$$N^L \phi^L N^K \frac{1}{N^K c'(N^K \phi^K)} V_I^K \{-\frac{\pi}{N^K} + (1-t)\frac{f(l)\frac{ds}{dt}}{N^K}\}$$
$$- N^K \phi^K N^L \frac{1}{N^L c'(\phi^L N^L)} V_I^L \frac{ds}{dt}$$

When $t = 0$, it should be that $s = 0$ and $l = l^f$, so $\frac{\partial s}{\partial t}|_{t=0} = \frac{\pi^f + N^L}{f(l^f)}$, where l^f and π^f are l and π at free trade, respectively. Thus the above expression becomes,

$$N^K \frac{1}{N^K c'(\phi^K N^K)} V_I^K \frac{N^L}{N^K} N^L \phi^L - N^K \phi^K N^L \frac{1}{N^L c'(\phi^L N^L)} V_I^L$$

Moreover, when $t = 0$, $d^K = d^L = 0$. So, $\phi^K N^K = \phi^L N^L$ and $c'(N^K \phi^K) = c'(N^L \phi^L)$. Thus, the expression becomes,

$$\frac{1}{N^L N^L \phi^L c'(N^L \phi^L)} [\frac{V_I^K}{N^K} - \frac{V_I^L}{N^L}]$$

If this is positive, the incumbent would help capitalists with the production subsidy.

Lemma 3.1. *The incumbent will raise tax to help capitalists if the following condition holds at free trade*

$$\frac{V_I^K}{N^K} - \frac{V_I^L}{N^L} > 0 \tag{3.1}$$

Since $N^L > N^K$, it is likely that $\frac{V_I^K}{N^K} - \frac{V_I^L}{N^L} > 0$ unless the marginal utility of income decreases very fast. Especially if the marginal utility of income is constant, it is positive. Thus, the incumbent has an incentive to help capitalists. Though the logic of this result is a very simple and intuitive, this shows the importance of considering the voter abstention. The tax and substitution scheme I am considering here means that taking money from a larger group and giving it to a smaller group. Because of differences in group sizes, the amount of money giving to one person of the smaller group is larger than the amount of money taken from one person in the larger group. Thus, naturally, the money given to the smaller group can have bigger impact on number of votes than the money taken from the larger group. If 10 cents are taken from everyone in a group with 100 million people, it is unlikely that the 10 cents could cause the huge increase of votes from the group. However, if everyone can gain 100 dollars by going to a voting booth, that could cause a big increase of votes from a group with 100,000 people. Of course, this is same to the usual logic of special interests, so, anything but new (Persson and Tabellini (2002), pp.159).) However, if I assumed that everyone votes, this logic could not hold here since 100 million people would go to voting booths because of their loss of 10 cents! The logic would not hold anymore. Clearly, this first Lemma shows the possible impact of the voter abstention on the policy choices of politicians.

How will incumbent help capitalists? In the previous subsection, I showed that the incumbent is likely to raise the tax and provide the production subsidy. However, he can also use a tariff to help capitalists; no need for him to stick only to the production subsidy. But, if he uses both instruments, he would need to consider how they would affect voting behavior of potential voters[24]. As I said before, I consider the change of the tariff and the subsidy such that $\frac{ds}{d\tau} = -1$, i.e., $dH = 0$. Thus, for a given total level of $H \equiv s + \tau > 0$, the higher tariff means the lower tax rate. The government budget constraint is now $sf(l) = t(\pi + N^L) + \tau m$. Since $dH = 0$, higher tariff does not change the production level of good Y, the labor employment for the production of good Y and the profit. Thus, $\frac{\partial l}{\partial \tau} = \frac{\partial \pi}{\partial \tau} = 0$. Let $\bar{I} \equiv \pi + N^L$ be the average and total income. Then, $\frac{dt}{d\tau} = -\frac{f(l) + m + \tau \frac{\partial m}{\partial \tau}}{\bar{I}}$.

The effect of higher tariff on utility:

The higher tariff hurts workers, but it could benefit capitalists if the tariff is not too high. This is possible since the income elasticity of good Y is less than one.

Because of this, the consumption share of workers of good Y is higher than their income share in the total population. Thus, the shift from the tax to the tariff is beneficial for capitalists even if $dH = 0$. Given the level of H, the level of s is derived from τ. Then, from the government budget constraint, the level of t can be derived from τ. Thus, a new indirect utility function of τ and H can be defined as

$$\hat{V}^K(\tau : H) = V((1-t)I^K, p^* + \tau) = V((1-t)\frac{(p^* + s + \tau)f(l) - l}{N^K}, p^* + \tau)$$

$$V^L(\tau : H) = V(1-t, p^* + \tau)$$

The derivative of $\hat{V}^K$ with respect to τ is,

$$\hat{V}^K_\tau|_{dH=0} = \frac{V^K_I}{\bar{I}} N^L I^L I^K [\frac{y^L}{I^L} - \frac{y^K}{I^K} + \tau \frac{I^K}{N^L y^L} \frac{\partial m}{\partial \tau}]$$

The sign of this depends on the inside of the bracket. It is natural to assume that $\frac{\partial m}{\partial \tau} < 0$. However, since $I^K > I^L = 1$ and the income elasticity of good Y is less than one, $\frac{y^L}{I^L} - \frac{y^K}{I^K} > 0$[25]. Thus, the inside of the bracket could be positive. Especially when $\tau = 0$, this is definitely positive. Thus, when $H = s + \tau > 0$, capitalists prefer the positive tariffs to zero tariff.

The derivative of $\hat{V}^L$ with respect to τ is,

$$\hat{V}^L_\tau|_{dH=0} = -\frac{V^L_I}{\bar{I}} N^K I^L I^K [\frac{y^L}{I^L} - \frac{y^K}{I^K} - \frac{\tau}{N^K y^L} \frac{\partial m}{\partial \tau}]$$

The inside of the bracket is positive. Thus, this is negative. Not surprisingly, workers always prefer no tariff.

The utility differences

Then, the utility differences are defined as follows.

$$d^K(\tau : H) = \hat{V}^K(\tau : H) - \hat{V}^K(0 : 0)$$

$$d^L(\tau : H) = \hat{V}^L(0:0) - V^L(\tau : H)$$

Their derivatives with respect to τ are $d^K_\tau = \hat{V}^K_\tau$ and $d^L_\tau = -\hat{V}^L_\tau$.

The effect on $\dfrac{N^K\phi^K}{N^L\phi^L}$

The derivative of $\dfrac{N^K\phi^K}{N^L\phi^L}$ with respect to τ is,

$$\frac{1}{(N^L\phi^L)^2}[N^K\phi^K_d d^K_\tau N^L\phi^L - N^K\phi^K N^L\phi^L_d d^L_\tau]$$
$$= \frac{1}{\bar{I}(N^L\phi^L)^2}[y^L(\frac{V^K_I}{c'(K)}N^L I^L N^L\phi^L - \frac{V^L_I}{c'(L)}N^K I^L N^K\phi^K)(\frac{I^K}{I^L} - \frac{y^K}{y^L})$$
$$+\tau\frac{\partial m}{\partial\tau}(\frac{V^K_I}{c'(K)}N^L\phi^L I^K + \frac{V^L_I}{c'(L)}N^K\phi^K)]$$

where $c(\alpha) \equiv c(N^\alpha\phi^\alpha)$

At $\tau = 0$, this becomes,

$$= \frac{I^K N^K N^L}{\bar{I}(N^L\phi^L)^2}(\frac{y^L}{I^L} - \frac{y^K}{I^K})(\frac{V^K_I}{N^K c'(K)}N^L\phi^L - \frac{V^L_I}{N^L c'(L)}N^K\phi^K)$$

From the less-than-one income elasticity and the single tax rate t, $\dfrac{y^L}{I^L} - \dfrac{y^K}{I^K} > 0$. Thus, the sign of this is equal to the sign of $\dfrac{V^K_I}{N^K c'(K)}N^L\phi^L - \dfrac{V^L_I}{N^L c'(L)}N^K\phi^K$. This can be rewritten into,

$$(*)\quad N^K\phi^K N^L\phi^L[\frac{1}{N^K\phi^K}\frac{V^K_I}{N^K c'(K)} - \frac{1}{N^L\phi^L}\frac{V^L_I}{N^L c'(L)}]$$

Thus, the condition for the incumbent to raise the tariff from zero is,

$$\frac{1}{N^K\phi^K}\frac{V^K_I}{N^K c'(K)} - \frac{1}{N^L\phi^L}\frac{V^L_I}{N^L c'(L)} > 0$$
$$\Leftrightarrow \frac{N^L}{N^K}\frac{N^L\phi^L/V^L_I}{N^K\phi^K/V^K_I} > \frac{c'(K)}{c'(L)}$$

This condition can be denoted in terms of elasticities of number of voters with respect to the utility differences, ε_{nd}.

$$\frac{N^L d^L/V^L_I}{N^K d^K/V^K_I} > \frac{\varepsilon^L_{nd}}{\varepsilon^K_{nd}} \Leftrightarrow \frac{\varepsilon^K_{nd}}{N^K d^K/V^K_I} > \frac{\varepsilon^L_{nd}}{N^L d^L/V^L_I}$$

At first, this looks a bit weird. If d^L is very small and d^K is very high, then the left side of the left inequality would be very small and would make it difficult for the condition to hold, even though it sounds like a good condition of the tariff increase for the incumbent. Looking at the definition of the elasticity, however, it is clear that the sizes of ds do not matter since εs have d inside, too. Moreover, unless the marginal utility of income decreases quite fast, the left side cannot be so small. Notice that $\frac{N^a d^a}{V_I^a}$ is the group a's loss or gain in terms of income. Then, $\frac{N^L d^L / V_I^L}{N^K d^K / V_I^K}$ is the ratio of the loss of workers over the gain of capitalists. If the transfer from workers to capitalists is done efficiently, the ratio should be one. However, because of the inefficiency (i.e., dead weight loss), the ratio is likely to be more than one unless the marginal utility of income decreases very fast. Especially if the marginal utility of income is constant, the left side should be more than one from the dead weight loss (this dead weight loss comes from the production subsidy since I am considering this at $\tau = 0$.) Then, the condition says if the voting response of workers to their utility loss is not so high compared to the voting response of capitalists then the incumbent would raise the tariff. One simple way to understand the condition is this; assuming that the constant marginal utility of income, the condition above is equal to $\varepsilon_{nd}^K \frac{N^L/N^K}{d^K} > \varepsilon_{nd}^L \frac{1}{d^L}$. At $\tau = 0$, the dead weight loss from the tariff is negligible, so, the marginal transfer rate of the loss of workers to the gain of capitalists by the tariff at $\tau = 0$ is one; one dollar loss of group L means one dollar gain of group K. Suppose that every worker loses one dollar from the tariff increase. Because of the group size difference, this means $\frac{N^L}{N^K}$ dollar gain for every capitalist. Thus, $\frac{1}{d^L}$ can be considered as the ratio of the change in the worker's utility difference, and

$\dfrac{N^L/N^K}{d^K}$ as the ratio of the change in the capitalist's utility difference. Then the elasticities of votes are multiplied to them respectively. Thus, the condition simply requires that the increase of the votes from capitalists is larger than the increase of the votes from workers.

Proposition 3.1. *Suppose that the incumbent has decided to help capitalists. If*

$$\frac{\varepsilon_{nd}^K}{N^K d^K / V_I^K} > \frac{\varepsilon_{nd}^L}{N^L d^L / V_I^L} \tag{3.2}$$

at $\tau = 0$, then the incumbent will use the tariff to help them (maybe with the production subsidy.)

Though $\dfrac{\partial \dfrac{N^L \phi^L}{N^K \phi^K}}{\partial \tau} = 0$ defines the optimal value of the tariff for the incumbent, it is unlikely to be solvable analytically. So, I consider the different characterization of the optimal tariff rate. Let $\varepsilon_{d\tau} = \dfrac{d_\tau \tau}{d}$. From (*), the condition for the optimal tariff is,

$$\begin{aligned} & N^K \phi_d^K d_\tau^K N^L \phi^L - N^K \phi^K N^L \phi_d^L d_\tau^L = 0 \\ \Leftrightarrow\ & \frac{N^K \phi^K N^L \phi^L}{\tau} [\varepsilon_{nd}^K \varepsilon_{d\tau}^K - \varepsilon_{nd}^L \varepsilon_{d\tau}^L] = 0 \end{aligned}$$

Of course, this means that the impact of tariff on the voting from capitalists is same to the one on the voting from workers.

Corollary *The condition of the optimal tariff rate for the incumbent is,*

$$\varepsilon_{nd}^K \varepsilon_{d\tau}^K = \varepsilon_{nd}^L \varepsilon_{d\tau}^L$$

3.3.2 Different tax rates

In the previous subsection, I assumed that the incumbent has only one tax rate for two groups. In this section, I assume different income tax rates for different voter groups. This changes the result obtained above hugely. I call tax rates of capitalists and workers as t^K and t^L, respectively. However, to keep the previous model unchanged as much as possible, I denote those two tax rates t^K and t^L in terms of t and tax burden share $h \in [0, 1]$. So,

$$t^K \pi = ht(\pi + N^L)$$
$$t^L N^L = (1 - h)t(\pi + N^L)$$

Thus, the total income tax payment is still $t(\pi + N^L)$, but capitalists pay the share h of that total income tax payment and workers pay the share $1 - h$ of it. Then,

$$t^K = \frac{ht(\pi + N^L)}{\pi}$$
$$t^L = \frac{(1 - h)t(\pi + N^L)}{N^L}$$

I assume that the incumbent can choose the tax burden share h freely. This means that in the previous subsection there was a restriction that $t^K = t^L$, which means $h = \dfrac{\pi}{\pi + N^L}$. Of course, the change in h affects utility gains and losses of capitalists and workers. Thus, given the total income tax $t(\pi + N^L)$, the incumbent will choose the rate h that maximizes $\dfrac{N^K \phi^K}{N^L \phi^L}$. Since change in h does not change the total income tax payment, no change in s happens. So, I do not need to worry whether the change in h might change π.

In the following, firstly, I consider what rate of h the incumbent would choose for a given level of $t(\pi + N^L)$. Then, given that h, I consider if and how the incumbent would provide the protection to capitalists.

The determination of tax rates In this section, I assume $\tau = 0$, again. Thus, $H = s$. Then, given the tax rate t, the indirect utilities of capitalists and workers of h are given as follows.

$$\hat{V}^K(h:t) = V(\frac{\pi - ht(\pi + N^L)}{N^K}, p^*)$$
$$\hat{V}^L(h:t) = V(\frac{N^L - (1-h)t(\pi + N^L)}{N^L}, p^*)$$

From these, utility differences, ds, are derived. Then, the incumbent has to find the optimal level of h for him to maximize $\frac{N^K \phi^K}{N^L \phi^L}$. The derivative of $\frac{N^K \phi^K}{N^L \phi^L}$ with respect to h is,

$$\frac{1}{(N^L \phi^L)^2}[N^K \phi^K_d d^K_h N^L \phi^L - N^K \phi^K N^L \phi^L_d d^L_h]$$
$$= \frac{t(\pi + N^L)}{(N^L \phi^L)^2}[-N^L \phi^L \frac{V^K_I}{N^K c'(K)} + N^K \phi^K \frac{V^L_I}{N^L c'(L)}]$$

The incumbent chooses the value of h such that the inside of the bracket is zero or corner solutions. However, the corner solution $h = 1$ does not make sense since this means that capitalists pay all of the production subsidy they receive (in this section I assume $\tau = 0$.) Thus, the possible choice by the incumbent is $h \in [0, 1)$.

Lemma 3.2. *Given t, if the incumbent can choose the tax burden share h freely, he will set h at the level such that*

$$\frac{V^K_I}{N^K N^K \phi^K c'(K)} \geq \frac{V^L_I}{N^L N^L \phi^L c'(L)}$$

In terms of elasticity ε_{nd}, this condition is equal to

$$\frac{\varepsilon_{nd}^K}{N^K d^K / V_I^K} \geq \frac{\varepsilon_{nd}^L}{N^L d^L / V_I^L}$$

By assumption, $N^L \phi^L c'(L) = N^K \phi^K c'(K)$ when $t = 0$. Thus, the condition becomes $\frac{V_I^L}{N^L} \leq \frac{V_I^K}{N^K}$ at $t = 0$. So, unless the marginal utility of income decreases very fast, the strict inequality should hold. So, when $t = 0$, it is likely that $h = 0$ (strict inequality means that the inside of the bracket above is negative.) Moreover, it becomes $\frac{\varepsilon_{nd}^K}{N^K d^K / V_I^K} > \frac{\varepsilon_{nd}^L}{N^L d^L / V_I^L}$, which is same to (3.2) that I obtained as the condition for the choice of the positive tariff in the previous subsection.

Given the result about h above, then, will the incumbent decide to raise the tax to help capitalists? The Lemma 3.3 gives a simple condition for that.

Lemma 3.3. *If* $\frac{V_I^K}{N^K} > \frac{V_I^L}{N^L}$ *at* $t = 0$*, then the incumbent will choose the tax rate* $t > 0$ *with tax burden share* $h = 0$.

The proof is in Appendix. Since the objective of the tax is to provide the help to capitalists, $h = 0$ is a natural result. Notice that the condition in this Lemma is the same condition for the incumbent to raise tax in the case of the single tax rate. The logic for this result is also same. Moreover, I showed the same condition for the positive tariff holds. However, these same conditions imply one important difference about the tariff. Given $h = 0$, there is no incentive for the incumbent to raise the tariff rate. In the previous single tax rate case, capitalists prefer low but positive tariffs to zero tariff since the positive tariffs mean lower tax burden for them. Though the higher tariffs have the negative effect on their utility, as long as the positive effect of the lower tax outweighs it, capitalists prefer the higher tariffs. Since $h = 0$ means capitalists do not pay tax, there is no positive effect of the higher tariff for capitalists. So, the negative effect cannot be outweighed. Of course, the positive tariff rate means some transfer of the burden from workers to capitalists. So, it might be possible that workers would prefer the positive tariff to zero tariff. Given $h = 0$, the derivative of worker's indirect utility with respect to τ at $\tau = 0$ is

$$\hat{V}_\tau^L|_{dH=0} = -\frac{V_I^L}{N^L} N^K I^L I^K [\frac{y^L}{I^L} - \frac{y^K}{I^K}]$$

Because of the less-than-one income elasticity of good Y, it is $\dfrac{y^L}{(1-t)I^L} - \dfrac{y^K}{I^K} > 0$. If the tax rate is low, the inside of the bracket, $\dfrac{y^L}{I^L} - \dfrac{y^K}{I^K}$, is still positive. If the tax rate on workers so high that $\dfrac{y^L}{I^L} - \dfrac{y^K}{I^K} < 0$ (remember I^a is before-tax income), it could be positive. However, it still can be shown that positive tariffs will lower $\dfrac{N^K \phi^K}{N^L \phi^L}$ because of $\dfrac{V_I^K}{N^K N^K \phi^K c'(K)} \geq \dfrac{V_I^L}{N^L N^L \phi^L c'(L)}$. Thus, there is no reason for the incumbent to use the tariff.

Proposition 3.2. *Suppose that the incumbent can choose the tax burden share freely and that* $\dfrac{V_I^K}{N^K} > \dfrac{V_I^L}{N^L}$ *at* $t = 0$. *Then, even when the incumbent use the production subsidy to help capitalists, he will not use the tariff.*

The proof is in Appendix. In the previous single tax rate case, the condition $\dfrac{\varepsilon_{nd}^K}{N^K d^K / V_I^K} > \dfrac{\varepsilon_{nd}^L}{N^L d^L / V_I^L}$ gives the reason for the incumbent to raise the tariff since the tariff increase causes negligible dead weight loss at $\tau = 0$. By raising the tariff, not by raising the tax and subsidy, the incumbent can help capitalists with the lower dead weight loss. However, if he could change the tax burden share h, this is a better instrument for the incumbent to provide the help since the burden share change does not cause any distortion at any level. Thus, when the incumbent can choose the tax rate freely so that no tax for capitalists is the optimal for the incumbent, there is no need for the inefficient tariff. However, this result depends on the assumption that the workers are homogeneous. If they are heterogeneous, the incentive for the incumbent to raise the tariff might appear again. I consider this in the next section.

3.4 Three groups: heterogeneity in workers

In this section, I extend the previous model so that it has three groups. Three groups consist of one capitalist group and two worker groups. The difference between two worker groups is the income level. Denote two worker groups as group M and group P, with their group sizes N^M, N^P and $N^M + N^P = N^L$. I assume workers in group P has one unit of labor and workers in group M have $1 + \frac{\omega}{N^M}$, $\omega > 0$, units of labor. Thus, workers in group M are richer than workers in group P, though they are still poorer than capitalists. A natural interpretation is that group M is middle class and group P is poor workers. The total income of two worker groups is $N^L + \omega$.

The model in previous section cannot incorporate three groups smoothly. To achieve that, I assume that the group M and P are subgroups of the group L. Let $m \equiv \frac{N^M}{N^L}$. For a given level of voting cost c, I assume that among workers who have voting cost equal to or lower than c, m of them are group M workers and $1 - m$ of them are group P workers. From this assumption, the following voting cost functions for group M and P are defined.

Voting cost of group M for the $N^M\phi^M$-th voter: $c(\dfrac{N^M\phi^M}{m}) = c(N^L\phi^M)$

Voting cost of group P for the $N^P\phi^P$-th voter: $c(\dfrac{N^P\phi^P}{1-m}) = c(N^L\phi^P)$

Since the voting rate ϕ is determined in the same way, $d - c = 0$, the derivatives of ϕs of group M and P are defined as follows.

$$\frac{\partial\phi^\alpha}{\partial d^\alpha} = \frac{1}{N^L c'(\alpha)} \text{ for } \alpha = M \text{ or } P, \text{ where } c'(\alpha) = c'(N^L\phi^M)$$

Then, it is defined that $\phi^L \equiv \dfrac{N^M\phi^M + N^P\phi^P}{N^L}$. In addition, I assume that group M and P share same random variable μ^L. This means the objective of the incumbent is still $\dfrac{N^K\phi^K}{N^L\phi^L}$, which is equal to $\dfrac{N^K\phi^K}{N^M\phi^M + N^P\phi^P}$.

Though the total income of workers is now $N^L + \omega$, not N^L, with these new assumptions, the change from the previous model is minimum from the viewpoint of the incumbent about what he should do regarding capitalists. In this section, I do not consider if the incumbent decides to help capitalists or not. There should not be any qualitative difference from the results in the previous section about this. Thus, I consider only if the incumbent would use a tariff when he decides to help capitalists, and about its implication. In addition to above assumptions, I also assume that the preference is special one; $U(x, y) = x + u(y)$, no income effect on good Y and the marginal utility of income is one. I keep the assumption that the incumbent can choose tax rates freely.

3.4.1 The optimal tax rates

Before I check if the incumbent would use a tariff or not, I consider what is the optimal tax burden shares for the incumbent. Same as the subsection considering the same problem in the case of two groups, I assume $\tau = 0$ in this subsection. Let t^a be the income tax rate of the group $a \in \{K, M, P\}$. Then, since the incumbent can choose the tax rate freely, given a total tax payment from three groups, it is obvious from the previous section that he imposes $t^K = 0$ on capitalists. This is a natural result considering the objective of the tax. The implication of this is that the numerator of the derivative of $\frac{N^K \phi^K}{N^L \phi^L}$ with respect to t^K with corresponding changes in taxes on other groups so that no change in the total tax payment is $N^L \phi^L N^K \phi_d^K d_{tK}^K - N^K \phi^K N^L \phi_d^L d_{tK}^L < 0$at $t^K = 0$ (I ignore the case that the equality holds at $t^K = 0$.) Since the group L now consists of two groups, it is not clear what d^L is. However, since $N^K d_{tK}^K$is the change in the tax burden of capitalists as a group, $N^L d_{tK}^L / N^K d_{tK}^K$has a clear interpretation as the ratio of marginal transfers from capitalists to workers. Since the marginal utility of income is fixed at one and the total tax payment from three groups does not change so that no change in subsidy and no change in the production of good Y, this ratio is one: $N^L d_{tK}^L / N^K d_{tK}^K = 1$(here I am considering this with $\tau = 0$.). Thus,

$$N^L \phi^L \phi_d^K - N^K \phi^K \phi_d^L N^L d_{tK}^L / N^K d_{tK}^K > 0 \Rightarrow N^L \phi^L \phi_d^K > N^K \phi^K \phi_d^L$$

both $d_{tK}^L$and $d_{tK}^K$are negative.

Substituting partial derivatives of ϕ into this,

$$\frac{1}{N^K N^K \phi^K c'(K)} > \frac{1}{N^L N^L \phi^L} \{ \frac{N^M}{N^L c'(M)} + \frac{N^P}{N^L c'(P)} \}$$

$$\Leftrightarrow \frac{\varepsilon^K_{nd}}{N^K d^K} > \{\frac{N^M\phi^M}{N^L\phi^L}\frac{\varepsilon^M_{nd}}{N^L d^M} + \frac{N^P\phi^P}{N^L\phi^L}\frac{\varepsilon^P_{nd}}{N^L d^P}\}$$

Thus, a formula similar to the one in the previous section (i.e., $\frac{V_I^K}{N^K N^K \phi^K c'(K)} > \frac{V_I^L}{N^L N^L \phi^L c'(L)} \Leftrightarrow \frac{\varepsilon^K_{nd}}{N^K d^K / V_I^K} > \frac{\varepsilon^L_{nd}}{N^L d^L / V_I^L}$) is obtained. Then, what is the optimal tax burden share between group M and P for the incumbent? Notice that since the objective of the incumbent is the maximization of $\frac{N^K\phi^K}{N^M\phi^M + N^P\phi^P}$ and since the tax burden sharing among group M and P does not affect capitalists, the optimal share h is the one such that minimizes $N^M\phi^M + N^P\phi^P$. Here, I define t and h, again, and I use them to denote t^M and t^P. Now, t is the tax rate on the total income of worker groups $N^L + \omega$, and h is the tax burden share of group M. Then,

$$ht(N^L + \omega) = t^M(N^M + \omega)$$

$$(1-h)t(N^L + \omega) = t^P N^P$$

From these, I can obtain t^M and t^P in terms h and t.

$$t^M = ht\frac{N^L + \omega}{N^M + \omega}$$
$$t^P = \frac{(1-h)t(N^L + \omega)}{N^P}$$

By substituting these into indirect utilities and taking a derivative of $\frac{N^K\phi^K}{N^L\phi^L}$, the next Lemma can be obtained.

Lemma 3.4. *If the incumbent can choose tax rate freely, he will choose tax rates such that*

$$\frac{1}{N^K N^K \phi^K c'(K)} > \frac{1}{N^L N^L \phi^L c'(M)}$$

$$\frac{1}{c'(M)} = \frac{1}{c'(P)}$$

The proof is in Appendix.

3.4.2 Will incumbent use tariff?

Now I consider if the incumbent would use a tariff or not. Thus, I consider a situation that the incumbent is using the income tax to provide the production subsidy but not using a tariff. Then, if the derivative of $\frac{N^K \phi^K}{N^L \phi^L}$ with respect to τ at $\tau = 0$ is positive, the incumbent will use the tariff. Since the tax rate on capitalists is zero, there is no reason for capitalists to prefer positive tariffs to zero tariff. However, because of the heterogeneity in workers, workers in group M prefer positive tariffs to zero tariff. Even in the case of two groups with zero tax on capitalists, there is some possibility that workers would prefer positive tariffs to zero tariff if tax rate is very high. Here, such high tax is not required. The logic behind this is same as the one behind the case that capitalists prefer low but positive tariffs to zero tariff. Since workers in group M have higher income than workers in group P, their ratio of good Y consumption to income, $\frac{y^M}{I^M}$, is lower than the ratio of workers in group P. Thus, raising the tariff rate is equal to the transfer of the tax burden from group M to group P (as long as the tariff rate is low.) Moreover, since capitalists is not paying tax, the positive tariff is also equal to the transfer from group M to group K. Thus they have a reason to prefer low but positive tariffs to zero tariff. Of course, capitalists and group P workers do not share the reason. However, as long as workers in group M prefer positive tariffs, it might be possible for the incumbent to use tariffs for his advantage. I am going to consider if this is really the case or not. Of course, I consider the change of τ with $dH = 0$. The indirect utilities of three groups are defined as follows.

$$
\begin{aligned}
V^K &= V(\frac{\pi}{N^K}, p^* + \tau) \\
V^M &= V((1 - ht\frac{N^L + \omega}{N^M + \omega})\frac{N^M + \omega}{N^M}, p^* + \tau) \\
V^P &= V(1 - \frac{(1-h)t(N^L + \omega)}{N^P}, p^* + \tau)
\end{aligned}
$$

Derivatives of them with respect to τ at $\tau = 0$ with $dH = 0$ are,

$$
\begin{aligned}
V_\tau^K &= V_p^K = V_I^K \frac{V_p^K}{V_I^K} = -y^K \\
V_\tau^M &= \frac{h(f+m) - N^M y^M}{N^M} \\
V_\tau^P &= \frac{(1-h)(f+m) - N^P y^P}{N^P}
\end{aligned}
$$

Since $d_\tau^K = V_\tau^K$, $d_\tau^M = -V_\tau^M$ and $d_\tau^P = -V_\tau^P$, the numerator of the derivative of $\frac{N^K \phi^K}{N^L \phi^L}$ with respect to τ at $\tau = 0$ with $dH = 0$ is,

$$
\begin{aligned}
& -\frac{N^L \phi^L N^K y^K}{N^K c'(K)} + N^K \phi^K [\frac{h(f+m) - N^M y^M}{N^L c'(M)} + \frac{(1-h)(f+m) - N^P y^P}{N^L c'(L)}] \\
& < \frac{N^K \phi^K N^L \phi^L}{N^K N^K \phi^K c'(K)} [f + m - N^K y^K - N^M y^M - N^P y^P] \\
& = 0
\end{aligned}
$$

The inequality holds because of the Lemma in the previous subsection. The last equality holds since $f + m$ is the supply of good Y and $N^K y^K + N^M y^M + N^P y^P$ is the demand of good Y. Thus, even though workers in group M prefer positive tariffs, that is not enough for the incumbent to raise tariff.

Proposition 3.3. *If the incumbent can choose tax rates freely, he will not use a tariff to help capitalists.*

This is the same result as the one in the previous section. However, this result depends on the assumption that the incumbent can choose the tax rates freely. This assumption might be very problematic in three groups setting of this section. The reason is simple. As I showed, if the incumbent can choose tax rates freely, he would choose tax rates so that $\frac{1}{c'(M)} = \frac{1}{c'(P)}$. However, since $c'' < 0$, this means $d^K = d^P$. These utility differences are for each worker in group M and P. At $\tau = 0$, these utility differences are tax burdens; the tax burden of each group P worker is same to the tax burden of each group M worker. Because the income of workers in group P is lower than the income of workers in group M, the same tax burden means the higher tax rate for workers in group P than the tax rate for workers in group M. Thus, the higher tax rate is put on the lower income group. Though zero tax on group K might be politically justifiable based on some protectionist and nationalist sentiments of people for protecting domestic industries, these regressive tax rates for group M and P are quite unlikely in usual political situations. Even the zero tax rate on capitalists might not be possible because of its regressiveness, though it would be kind of ludicrous considering the purpose of the tax. So, if such regressive tax rate system is not possible, there is an incentive for the incumbent to use a tariff. The result of the single tax rate case with two groups is a clear example. With less-than-one income elasticity of the demand on good Y, the tariff has the same regressive effect. However, it is less likely that the regressiveness of tariffs become the huge political problem compared to the regressive tax system. Thus, to achieve the same regressive treatment of workers, the tariff is a politically better instrument. To show the possibility of the incumbent using a tariff when he does not have free hand on the choice of tax rates in the case of three groups, consider the following case: the the tax system is progressive so that $t^K > t^M > t^P$. Particularly, $t^P = 0$. The population of voter groups is $N^K < N^M < N^p$. Then, suppose that the incumbent is giving the help to capitalists without using a tariff. Then, the numerator of the derivative of $\dfrac{N^K \phi^K}{N^M \phi^M + N^P \phi^P}$ with respect to τ at $\tau = 0$ is,

$$\frac{d^K_\tau}{c'(K)}(N^M \phi^M + N^P \phi^P) - N^K \phi^K (\frac{d^M_\tau}{c'(M)} + \frac{d^P_\tau}{c'(P)})$$

If this is positive, the incumbent would increase the tariff from zero. Because of the logic for the single tax case in the precious section, higher tariffs reduce the burden of the help on capitalists. This increases d^K, so d^K_τ is positive. Thus, the first term is positive, i.e., $\dfrac{d^K_\tau}{c'(K)}(N^M \phi^M + N^P \phi^P) > 0$. The same logic says the burden on the group P voters rises. So, d^P_τ is positive. The effect of the tariff increase on d^M_τ is ambiguous in general. From the logic, we know a part of the burden on capitalists is transferred to group M voters. But the same logic also says a part of burden on group M also goes to group P. The sign of d^M_τ depends on which effect is stronger than the other. However, if N^M is much larger than N^P and much smaller than N^P, d^M_τ would be negative. Moreover, if $\lim_{d \to 0} c'(d) = \infty$ (like the Inada condition), $\dfrac{d^P_\tau}{c'(P)} = 0$ since $t^P = 0$. Thus, the above value is positive. The incumbent would use the tariff. The voter abstention and different group sizes give an incentive for the incumbent to redistribute resource/income. If he can use policy instruments that produce less dead weight loss, there would be no reason for him to use other instruments that produce more dead weight loss. However, if he cannot use policy instruments freely, he would use the inefficient instruments. In the setup of this chapter, if the regressive tax rate system is not possible, that might induce the incumbent to use the inefficient tariff, as shown in the the example above. Since the reason that the regressive tax system is not possible in real world is the sense of fairness in the tax system, this might be able to be said as a trade-off between efficiency and fairness. However, the main problem is the fact that the voter abstention gives the incumbent the opportunity to manipulate votes. If all potential voters vote, the incumbent would not help capitalists. Here I state the main Theorem.

Theorem 3.1 *If the incumbent can choose tax rates freely so that he can choose even regressive tax rates, he will never use the tariff to help the domestic industry. In this sense, the inefficiency of the tariff is the price of the fairness in the tax system.*

3.5 Literature Review

There are other papers that consider why tariffs are used. Generally, however, what they do is the comparison of economic welfare between the case that a tariff is used and the case that a subsidy is used; they show situations where the tariff achieves higher welfare than the subsidy. Rodrik (1986) shows that since the tariff has the character of public goods, the lobbying effort for the tariff is under-supplied, compared to the lobbying effort for the firm specific subsidy. So the distortion under the tariff could be lower than the one under the subsidy. Mayer and Riezman (1987) show that when the country is small and voters are heterogeneous in factor endowments, voters prefer the production tax cum subsidy to the tariff. So, if voters will vote on trade policy like in Mayer (1984), the tariff will not be chosen. But in cases that the country is large or that voters are heterogeneous in both of factor endowment and consumption preference, they say it is possible for all or some voters prefer tariff. Wilson (1990) shows that the efficient instrument can invite more lobbying pressures so the total distortion can be larger for the efficient instrument than for the inefficient one. Since politicians care about the level of total distortion which could affect voters' sentiment toward politicians negatively, there is an incentive for politicians to restrict the possible policy instruments to the inefficient ones. Mayer and Riezman (1990) consider several cases in which some voters prefer tariff to subsidy, like the case of heterogeneous voters, the case of risk averse voters in the uncertainty about tariff rate and tax rate, etc., though they do not show how such voter preferences will be reflected in the political process of choosing policy instruments. In this chapter, preferences on instruments are also different among different groups. The reason for such difference is similar to the one in one section of Mayer and Riezman (1990)[26]. Naito (2006) shows that given a reasonable information constraint on government, when government wants to design a tax system to redistribute income to unskilled workers, tariffs on unskilled labor intensive goods could be Pareto-improving.

Though they do not consider about tariffs directly, Coate and Morris (1995) consider why inefficient instruments are used by government in a model with uncertainties about two things; one is about the outcome of public project, and another about the type of the incumbent politician. The incumbent politician has information about the project outcome and will face re-election. Good type politician will carry on the project only when the information he has tells him that the project has high probability of producing good outcome. Bad incumbent wants to transfer money to special interests. He can do that through the direct subsidy or the public project. Voters update their reputations on the type of politician based on what the incumbent does and what its outcome is. So, to increase the probability of winning the re-election, the bad type politician chooses the project over the direct subsidy (this gives him higher utility than the project) so that he can pretend a good type believing that the project has high probability of success. Acemoglu and Robinson (2001) consider a two-period model with a special interest group whose power source is the number of voters belonging to it. In multi-periods models, policy commitment by politicians is an important problem. If it is not guaranteed by an assumption, it needs to be guaranteed by some mechanism (Harrington (1992)), or no policy commitment (Alesina (1988)). In the model of Acemoglu and Robinson (2001), the special interest prefers the inefficient policy to the efficient one since it distorts newcomers' incentive toward joining to the special interest, which guarantees the power of the special interest in the second period and secures the commitment by the government to the transferring resource to the special interest in the period.

Above, I mentioned papers in which there is no intrinsic difference in the difficulty of policy implementations. If there is difference in the difficulty of policy implementations (for example, the cheaper collection cost of tariffs and the higher cost of other taxes), however, it is actually no wonder that governments would choose the policy that is easier to implement. Gordon and Li (2005a) argues that governments of developing countries rely more on tariffs as their revenue sources than the ones of developed countries because it is more difficult to monitor economic activities of domestic private firms in developing countries than in developed countries. Gordon and

Li (2005b) claim that the model of Gordon and Li (2005a) fits actual data better than the model of Grossman and Helpman (1994). Even in current developed countries, the tax collection used to be a difficult problem in the past (maybe still is.) Gardner and Kimbrough (1992) applied to U.S. this line of explanation about tariffs[27]. Though this chapter is different in many aspects from this collection cost type explanation, there is a common element; politicians choose the policy instruments that have lower political costs. I considered the maximization problem in this chapter. But votes for the challenger at the election and the loss of votes for the incumbent can be considered as the political costs of the policy instrument choices. Then, the same problem can be considered as a political cost minimization problem, like the expenditure minimization problem[28]. Then, difference in the nature of the cost could have a big implication for welfare (maybe even a benevolent planner might decide to use tariffs when the collection costs of taxes are high. But will the planner use tariffs for the political cost reason when there is no intrinsic cost difference between tariffs and taxes?) But from the viewpoint of each politician, whatever the source of the cost of a policy is (collection costs or loss of votes for him, etc.,) they are political costs of the policy he wants to implement. So, politicians must try to minimize those political costs. In this sense, I think this chapter and the collection cost explanation are in complimentary, looking at different type of costs, not contradicting each other as the explanation about why tariffs are used.

As I wrote in previous sections, the idea of the paper is that the incumbent can take advantage of the voter abstention by redistribution costs and benefits of helping a domestic industry. From this point, this chapter is related to the literature of the Vote Buying/Turnout Buying. Nitchter (2008) analyses turnout-buyings in Argentina; political party rewarding potential voters who favor the party to its rival party for voting for the party. Even though they favor the party, they might abstain from voting without rewarding. Vote Buying is the opposite; rewarding voters who do not favor the party for voting for the party. Though the direct rewards for voting is different from what I consider in this chapter (choice of policy instruments), similar elements exist; the optimal choice of policy instruments (or the optimal distribution of rewards) could increase votes for the incumbent from people who benefit from the policy of the incumbent and decrease votes for the opponent from people who suffer from the policy.

3.6 Conclusion of Chapter 3

In this chapter, I considered why the incumbent might decide to use a tariff to help domestic industry. What important in this chapter is the voter abstention. After the incumbent makes the decision about the help, he will face the election. So, he takes the election into his consideration when he chooses policy instruments. Because of the voter abstentions, he might decide to give the help. Then, if there is no restriction on the choice of efficient policy instruments (the tax and the production subsidy here), the incumbent would not use the tariff for the help. However, its outcome does not look politically feasible since it involves the regressive tax rates. So, if the regressive tax rates are not feasible because of some restrictions on the tax system, the incumbent would use the tariff to take advantage of the voter abstention. Because of the voter abstention, there is a problem similar to the one under the externality; the undersupply of the good policy. Though no tariff is better for national welfare, this would not matter for the incumbent since he might not be rewarded fully for the good choice because of the voter abstention. So, the incumbent must try to take advantage of the voter abstention by putting more costs on people who do not vote. When there is the political restriction on the possible policy instrument choices, this could lead to the choice of tariffs. For the incumbent, it is the politically low cost distribution of the burden of his policy.

The less-than-one income elasticity is just one example of the sources for the voter heterogeneity which the incumbent can exploit. Since it means normal goods, real world examples are easy to find. Moreover, what I need in this chapter is actually the consumption pattern difference among different income groups, which is further easier to find. As an example related to trade restrictions, consider the voluntary export restriction of Japanese automobiles in 1980s. In 1980s, Japanese cars were mostly small to mid size cars, no luxury cars. Thus the rich's expenditure share of Japanese cars in 1980s should have been much lower than the middle class and working class people's share[29]. Though it is not related to trade, state lotteries also can be considered as such a case since it is an easy way for politicians to generate revenue without offending voters with higher taxes. Garrett and Coughlin (2008) estimate that the income elasticity of state lotteries is less than one.

The model in this chapter has many moving parts, which might be confusing. So I briefly explain about the relationships of some important assumptions and results here.

The challenger as Free-trader: The incumbent might not be rewarded for free-trade policy, which discourages him from pursuing the free-trade policy. He is not be rewarded not because his challenger is a convinced free trader (there would be no difference between them if the incumbent chooses the free trade policy.) Even if the challenger was a protectionist, the results in this chapter still would suggest that the incumbent might have an incentive to choose the protectionist policy. If choosing the free trade policy against the protectionist challenger were politically beneficial, he would not have chosen helping the domestic industry in this model.

Endogenous decision of abstention: This is the hinge of the idea in this chapter. Without this, the incumbent does not have the power of manipulating votes. What important is that abstention rates vary depending on the utility differences. Suppose that the some voters abstain, but the rates of election participation among groups are fixed. This is basically same to the case of no voter abstention with the different composition of population. To understand this, consider the following situation; ratios of group populations over total population are 50% for group P, 25% for group M and 25% for group K. Moreover, suppose that the participation rates are 50% for all groups. Then, ratios of actual voters over total actual voters are same to population ratios. So, this is equivalent to the case of no voter abstention with the half population size. Since group M and P always prefer challenger, the incumbent never decides to give the help. Even if participation rates are different among groups, it is still equivalent to a case of no abstention. If the participation rates are 10% for group P, 40% for group M and 100% for group K, it is same to a case of no abstention with population ratios of 12.5% for group P, 25% for group M, and 62.5% for group K. In this case, the incumbent always gives the help. What important for this chapter is that voting behaviors are influenced by utility differences, which can be manipulated by the incumbent.

Non-homothetic preference: The incumbent can manipulate voters' utility differences because of the non-homothetic preference. If preferences are homothetic, he could not do that. However, what important is not the non-homothetic preference itself, but the existence of some factors that can give the manipulating power to the incumbent.

The results in this chapter depend on strong assumptions. However, the basic logic is more general. For the logic, specific assumptions, like the less-than-one income elasticity, are not important. If there are more than one policy instruments and there is heterogeneity, similar results could be obtained. Because of the voter abstention the incumbent politician can put the cost of his policy on voter groups whose voting rates are inelastic. Similar to the argument of the Ramsey rule about minimizing the dead weight loss from commodity taxes, the incumbent can minimize the political cost of his policy (votes for his challenger) for the same level of political gain (votes for him) by putting cost on the voting-inelastic voter groups. Then it is natural to assume that the incumbent wants to use the most efficient policy instrument (e.g. the tax and production subsidy) for that purpose. However, political restrictions, like the requirement for fairness, might prevent the use of the most efficient policy instrument at the optimal level for the incumbent (e.g., regressive income tax rates.) Then, it would be beneficial for the incumbent to use less efficient instruments. In this sense, the use of inefficient tariff could be considered as the price of the fairness in the tax system.

Appendix of Chapter 3

Lemma 3.3 *If* $\frac{V_I^K}{N^K} > \frac{V_I^L}{N^L}$ *at* $t = 0$, *then the incumbent will choose the tax rate* $t > 0$ *with tax burden share* $h = 0$.

Proof. Proof: The derivative of the capitalist's indirect utility with respect to t is,

$$\hat{V}_t^K = V_I^K \frac{\frac{ds}{dt}f - h(\pi + N^L) - ht\frac{ds}{dt}f}{N^K}$$

As I showed before, $\frac{ds}{dt} = \frac{(\pi + N^L)f''}{ff'' - s(f')^3}$. Thus,

$$\hat{V}_t^K = V_I^K \frac{(\pi + N^L)}{N^K}[\frac{(1 - ht)ff''}{ff'' - s(f')^3} - h]$$

and,

$$V_t^L = -V_I^L \frac{(1 - h)(\pi + N^L)}{N^L}[1 + \frac{tff''}{ff'' - s(f')^3}]$$

Since $d_t^K = \hat{V}_t^K$ and $d_t^L = -\hat{V}_t^L$, the numerator of the derivative of $\frac{N^K\phi^K}{N^L\phi^L}$ with respect to t is,

$$(*)\ N^K\phi^K N^L\phi^L(\pi + N^L)\frac{V_I^K}{\varepsilon^K d^K N^K}[\frac{(1 - ht)ff''}{ff'' - s(f')^3} - h]$$
$$-N^K\phi^K N^L\phi^L(\pi + N^L)\frac{V_I^L}{\varepsilon_{cn}^L d^L N^L}(1 - h)[1 + \frac{tff''}{ff'' - s(f')^3}]$$

Since the incumbent chooses h so that $\frac{V_I^L}{\varepsilon_{cn}^L d^L N^L} \leq \frac{V_I^K}{\varepsilon_{cn}^K d^K N^K}$, this becomes,

$$(*) \geq N^K\phi^K N^L\phi^L(\pi + N^L)\frac{V_I^K}{\varepsilon^K d^K N^K}[\frac{(1 - ht)ff''}{ff'' - s(f')^3} - h - 1 + h - \frac{(1 - h)tff''}{ff'' - s(f')^3}]$$
$$= N^K\phi^K N^L\phi^L(\pi + N^L)\frac{V_I^K}{\varepsilon^K d^K N^K}[\frac{(1 - t)ff''}{ff'' - s(f')^3} - 1]$$

When $t = 0$, the inside of the bracket is zero. This means when the incumbent chooses the corner solution $h = 0$ so that the strict inequality holds, the derivative of

$\dfrac{N^K\phi^K}{N^L\phi^L}$ at $t=0$ is positive. When the incumbent chooses a non-corner solution so that $\dfrac{V_I^L}{N^L N^L \phi^L c'(L)} = \dfrac{V_I^K}{N^K N^K \phi^K c'(K)}$, the result means the derivative is zero at $t=0$. It is easy to prove that the only another value of t that makes the inside of the bracket zero is negative, and it is also easy to prove that $t=1$ makes the inside -1. The inside is negative for any $t \in (0,1]$. This implies that if $t>0$ and if the incumbent is choosing a a non-corner solution for h so that the equality holds, he would have an incentive to lower the tax rate. However, if $\dfrac{V_I^K}{N^K N^K \phi^K c'(K)} > \dfrac{V_I^L}{N^L N^L \phi^L c'(L)}$ at $t=0$, the incumbent has an incentive to raise the tax rate t. The only consistent combination of t and h is $t>0$ and $h=0$. As I showed before, the condition becomes $\dfrac{V_I^K}{N^K} > \dfrac{V_I^L}{N^L}$ at $t=0$. ∎

Proposition 3.2: *Suppose that the incumbent can choose the tax burden share freely and it holds that* $\dfrac{V_I^K}{N^K} > \dfrac{V_I^L}{N^L}$ *at* $t=0$*. Then, even when the incumbent use production subsidy to help capitalists, he will not use tariff.*

Proof. Proof: When $h=0$, the government budget constraint is $sf(l) = tN^L + \tau m$. The derivative of tax rate t with respect to tariff τ for the fixed $H \equiv s+\tau$ at $\tau=0$ is $\dfrac{dt}{d\tau} = -\dfrac{f+m}{N^L}$.

The indirect utilities of capitalists and workers are,

$$\hat{V}^K(\tau : H) = V(I^K, p^* + \tau)$$

$$\hat{V}^L(\tau : H) = V(1-t, p^* + \tau)$$

Then derivatives of them with respect to τ at $\tau=0$ with $dH=0$ is

$$\hat{V}_\tau^K = V_p^K = -V_I^K y^K$$

$$\hat{V}_\tau^L = V_I^L \frac{f+m}{N^L} + V_p^L = \frac{V_I^L}{N^L}(f+m-N^L y^L)$$

Here I assume $V_{\tau}^{L} > 0 \Leftrightarrow f + m - N^L y^L > 0$

Then, the numerator of derivative of $\frac{N^K\phi^K}{N^L\phi^L}$ with respect to τ at $\tau = 0$ with $dH = 0$ is,

$$N^K\phi^K N^L\phi^L[\frac{-V_I^K}{N^K N^K\phi^K c'(K)} N^K y^K + \frac{V_I^L}{N^L N^L\phi^L c'(L)}(f + m - N^L y^L)]$$

From $\frac{V_I^K}{N^K N^K\phi^K c'(K)} \geq \frac{V_I^L}{N^L N^L\phi^L c'(L)}$ and $f + m - N^L y^L > 0$,

$$\frac{-V_I^K}{N^K N^K\phi^K c'(K)} N^K y^K + \frac{V_I^L}{N^L N^L\phi^L c'(L)}(f + m - N^L y^L)$$
$$\leq \frac{V_I^K}{N^K N^K\phi^K c'(K)}[f + m - N^K y^K - N^L y^L]$$
$$= 0 \quad \blacksquare$$

Lemma 3.4: *If the incumbent can choose tax rate freely, he will choose tax rates such that*

$$\frac{1}{N^K N^K\phi^K c'(K)} > \frac{1}{N^L N^L\phi^L c'(M)}$$
$$\frac{1}{c'(M)} = \frac{1}{c'(P)}$$

Proof. Proof: The indirect utilities of workers in group M and P are defined as follows.

$$V^M = V((1 - ht\frac{N^L + \omega}{N^M + \omega})\frac{N^M + \omega}{N^M}, p^*)$$
$$V^P = V(1 - \frac{(1-h)t(N^L + \omega)}{N^P}, p^*)$$

The derivatives of these with respect to h are,

$$V_h^M = -\frac{t(N^L + \omega)}{N^M}$$
$$V_h^P = t\frac{N^L + \omega}{N^P}$$

Since the help to capitalists hurt them, $d_h^{\alpha} = -V_h^{\alpha}$. Thus, the derivative of

$N^M\phi^M + N^P\phi^P$ with respect to h is,

$$N^M\phi_d^M d_h^M + N^P\phi_d^P d_h^P = \frac{N^M}{N^L c'(M)}\frac{t(N^L+\omega)}{N^M} - \frac{N^P}{N^L c'(P)}\frac{t(N^L+\omega)}{N^P}$$
$$= \frac{t(N^L+\omega)}{N^L}\left(\frac{1}{c'(M)} - \frac{1}{c'(P)}\right)$$

If $h = 0$, all tax burden falls on group P. It would be same to the free trade situation for group M but that would induce workers in group P to vote. Thus, $c'(P) < c'(M)$ from $c'' < 0$. If $h = 1$, the opposite would happen. Thus, the incumbent should choose h so that $\frac{1}{c'(M)} - \frac{1}{c'(P)} = 0$. Substituting this into the inequality I obtain,

$$\frac{1}{N^K N^K \phi^K c'(K)} > \frac{1}{N^L N^L \phi^L}\left\{\frac{N^M}{N^L c'(M)} + \frac{N^P}{N^L c'(P)}\right\} = \frac{1}{N^L N^L \phi^L c'(M)} \quad \blacksquare$$

Endnotes of Chapter 3

18. According to CNN exit poll (CNN.com (2004)), 68% of voters of 2004 U.S. presidential election made $75,000 or less, and 82% of voters made $100,000 or less.

19. Though I used the word "lobbying", this is not a lobbying model, so, no campaign contribution from them.

20. This assumption can be changed so that the tariff revenue is distributed to people, not used to finance the subsidy. Such change makes model more complicated, but does not change qualitative results.

21. Of course, the production subsidy that is financed by the tax does cause distortions.

22. Group Rule-Utilitarian voting literature consider more complicated voting decision making (Feddersen and Sandroni (2006), and Coate and Conlin (2004)) than the one I employ here. Their decision-making is, as its name suggests, more group-based.

23. If $\mu^K N^K \phi^K - \mu^L N^L \phi^L < 0$, the challenger wins. If $\mu^K N^K \phi^K - \mu^L N^L \phi^L = 0$, then both candidates have same probability of winning the election.

24. I do not consider the impact of the tariff on the amounts of export and imports. Though the tariff increase with $dH = 0$ does not change domestic productions, it certainly changes domestic consumptions, which changes the amounts of export and import. However, I simply assume that the change is small enough so that the country is still exporting good X and importing good Y.

25. Because of the income tax, the good Y consumption - income ratio is actually

$y^a/(1-t)I^a$. Since the tax rate is same for both group, I can ignore $(1-t)$ here.

26. The reason is similar, but not same. In the paper of Mayer and Riezman (1990), the preference for tariff comes from the difference of income tax rate and the share of distribution of tariff revenue. In this chapter, the preference comes from the difference of income share and the consumption share of different consumers (in the single tax case.)

27. I thank Professor Wilson and Professor Nelson for putting my attention to those papers and this line of argument.

28. Actually, the previous version of this paper considers the minimization of votes for challenger for given level of votes for the incumbent.

29. Bordley and McDonald (1993) estimate that income elasticities of automobile in any segments (small, economy, luxury, etc.) are more than one. This is inconsistent with the assumption of this chapter. However, it is natural to suppose that income elasticities of goods change as income level changes. Bordley and McDonald (1993) also estimate that the more expensive the automobile segment is, the higher the elasticity. This is consistent with the conjecture that the elasticity of Japanese economy cars in 80's was very low for the rich.

References

[1] Aldrich, John H. (1993): "Rational Choice and Turnout" *American Journal of Political Science*, 37, 246-78.

[2] Alesina, Alberto (1988): "Credibility and Policy Convergence in a two-Party System with Rational Voters" *American Economic Review*, 78, 796-805.

[3] Bordley, Robert F. AND James B. McDonald (1993): "Estimating Aggregate Automotive Income Elasticities from the Population Income-Share Elasticity" *Journal of Business & Economic Statistics*, 11(2), 209-214.

[4] CNN.com. "Election 2004", http://www.cnn.com/ELECTION/2004/pages/results/states/US/P/00/epolls.0 .html

[5] Coate, Stephen AND Stephen Morris (1995): "On the Form of Transfers to Special Interests" *Journal of Political Economy* 103, 1210-1235.

[6] Coate, Stephen AND Michael Conlin (2004): "A Group Rule–Utilitarian Approach to Voter Turnout: Theory and Evidence", *American Economic Review,* 94, 5, 1476-1504.

[7] Dixit, Avinash (1986): "Gains from Trade without Lump Sum Compensation" *Journal of International Economics,* 21, 111-122.

[8] Dixit, Avinash AND Victor Norman (1980): *Theory of International Trade*, Cambridge: Cambridge University Press.

[9] Dixit, Avinash AND Thomas Roemer (2006): "Political Explanations of Inefficient Economic Policies – An Overview of Some Theoretical and Empirical literature" Presentation at IIPF conference *"Public Finance: Fifty Years of the Second Best -and Beyond"* in Paphos, Cyprus, August 2006.

[10] Feddersen, Timothy J (2004): "Rational Choice Theory and the Paradox of Not Voting" *Journal of Economic Perspective*, 18, 99-112.

[11] Feddersen, Timothy J. AND Alvaro Sandroni (2006): "A Theory of Participation in Election" *American Economic Review*, 96, 1271-1282.

[12] Gardner, Grant W. AND Kent P. Kimbrough (1992): "Tax Regimes, Tariff Revenues and Government Spending" *Economica*, February 1992, 59, 75-92.

[13] Garrett, Thomas A. AND Cletus C. Coughlin (2008): "Inter-temporal Differences in the Income Elasticity of Demand for Lottery Tickets" Federal Reserve Bank of St. Louis working paper series 2007-042b (revised version).

[14] Gordon, Roger AND Wei Li (2005a): "Tax Structure in Developing Countries: Many Puzzles and a Possible Explanation" NBER Working paper No.11267.

[15] Gordon, Roger AND Wei Li (2005b): "Puzzling Tax Structures in Developing Countries: A Comparison of two Alternative Explanations" NBER Working paper No.11661.

[16] Grossman, Gene M. AND Elhanan Helpman (1994): "Protection for Sale" *American Economic Review*, 84, 833-50.

[17] Harrington, Jr., Joseph E. (1992): "The role of party reputation in the formation of policy" *Journal of Public Economics*, 49, 107-21.

[18] Harsanyi, John C. (1980): "Rule Utilitarianism, Rights, Obligations and the Theory of Rational Behavior" *Theory and Decision*, 12, 115-33.

[19] Hinich, Melvin J., John Ledyard, AND Peter C. Ordeshook (1972): "Nonvoting and the existence of Equilibrium under Majority Rule" *Journal of Economic Theory*, 4, pp.144-153.

[20] Lijphart, Arend (1997): "Unequal Participation: Democracy's Unsolved Dilemma" *American Political Science Review*, 91, 1-14.

[21] Mayer, Wolfgang (1984): "Endogenous Tariff Formation" *American Economic Review*, 970-85.

[22] Mayer, Wolfgang AND Raymond Riezman (1987): "Endogenous choice of trade policy instruments" *Journal of International Economics*, 23, 377-81.

[23] Mayer, Wolfgang AND Raymond Riezman (1990): "Voter preferences for trade policy instruments" *Economics and Politics*, 2, 259-73.

[24] Nichter, Simon (2008): "Vote Buying or Turnout Buying? Machine Politics and the Secret Ballot" *American Political Science Review*, 102, 19-31.

[25] Naito, Hisahiro (2006): "Redistribution, production inefficiency and decentralized efficiency" *International Tax and Public Finance*, 13, 625-640.

[26] Persson, Torsten AND Guido Tabellini (2002): *Political Economics Explaining Economic Policy*, MIT Press

[27] Rodrik, Dani (1986): "Tariffs, subsidies, and welfare with endogenous policy" *Journal of International Economics*, 21, 285-99.

[28] Rodrik, Dani (1995): "Political Economy of Trade Policy" in Gene M. Grossman and Kenneth Rogoff, ed., *Handbook of International Economics Vol.3*, Elsevier Science B.V., Amsterdam, The Netherlands, 1457-1494.

[29] Shachar, Ron AND Barry Nalebuff (1999): "Follow the Leader: Theory and Evidence on Political Participation" *American Economic Review*, 89, 525-47.

[30] Wilson, John D. (1990): "Are Efficiency Improvements in Government Transfer Policies Self-Defeating in Political Equilibrium?" *Economics and Politics*, 2, 241-58.

[31] Yang, C. C. (1995): "Endogenous Tariff Formation under Representative Democracy: A Probabilistic Voting Model" *American Economic Review*, 85, 956-63.