INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

ABSTRACT

The Political Economy of Ideology

Luis Madrazo

Yale University

2003

This dissertation is composed of three essays. The common methodological thread is game theory and the main theme is the role of ideology in political economy. The first essay is a contribution to imperfect monitoring in an infinitely repeated game. The second and main essay uses that result to build a political economy model that explains why some efficient policies might not be carried out by a government, what type of government may overcome such difficulties, and why that same type of government may be prone to inertia. The third essay presents an electoral model in which informed political parties compete, through shifts in their platform, for the vote of an uninformed electorate. In equilibrium parties will always adhere to a single (differentiated) platform and thus become identified with a set of beliefs, a key assumption in the second essay.

The first essay shows that the ability of a principal to elicit repeated actions from an agent is severely handicapped by introducing the possibility of an alternate state of the world where such actions are not desirable if punishment is not costless to the principal. Only a principal with extreme priors will prevail in the ex-ante desirable policy. This gives rise to incentives to delegate enforcement and presents a warning as to the possible consequences of such delegation.

For the second article, I show that a government's commitment to an efficient punishment policy is compromised by introducing the possibility that the policy was designed under flawed premises if the government internalizes at least some of the costs of punishment. The only type of government that is able to implement such a policy is subjectively certain about the true state of the world. I call such a government an *ideological extremist* and observe that this ability helps explain why society may elect such governments even if

it doesn't share its beliefs. Once elected, if the policy is flawed, the *ideological extremist* is unable to interpret bad outcomes as evidence that reform is required. Inertia stemming from ideology ensues.

The final essay presents a game between a pivotal voter and two informed political parties where there is uncertainty in the effects of policy and preferences. There is a unique equilibrium, with platform divergence, that is uninformative and "dogmatic", i.e. each party proposes the same (differentiated) policy regardless of the state.

The Political Economy of Ideology

A Dissertation

Presented to the Faculty of the Graduate School

of

Yale University

in Candidacy for the Degree of

Doctor of Philosophy

by

Luis Madrazo

Dissertation Director: Stephen Morris

December 2003

UMI Number: 3109429

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Acknowledgments

My deepest gratitude to Stephen Morris for his advice and support throughout the course of my research and studies.

I would also like to acknowledge all the members of the faculty, student body and staff at Yale who made my stay in New Haven and this dissertation not only possible but pleasurable.

I'm also particularly indebted to Luis Rayo, John E. Roemer and Christopher Timmins for their direct involvement in the completion of the dissertation.

I wish to thank Bruce Ackerman, Pedro Aspe, Attila Ambrus, Santiago Bazdresch, Juan Esteban Carranza, Stephen Coate, Geoffrey Garrett, Dino Gerardi, Pepe Toño Gonzalez, Borja Gracia, Tim Guinnane, Hanming Fang, George Hall, Mara Hernandez, Tomas Lajous, Calixto Lopez, Smaranda Luca, Alejandro Madrazo, George Mailath, Alexis Milo, William Nordhaus, David Pearce, Sharun W. Mukand, Pamela O´Donnell, Marco Oviedo, Robert J. Shiller, Mariano Tommasi and Alejandro Werner.

Dedication

Para Malú, Alejandro, Catalina y _____

CONTENTS

LIST OF FIGURES

Introduction

This dissertation contains three essays in three chapters. The first one introduces a game theory result in repeated games with imperfect monitoring. The second chapter develops that result in a political economy model. It proposes a unified model that explains three distinct phenomena in politics: non-adoption of some ex-ante desirable policies, the election of ideolgues by non-ideologues, and inertia or the persistence of bad outcomes. In the third chapter I present an electoral model that shows why strategic considerations may prevent parties from converging to the middle of the political spectrum. It also sheds some light as to why ideology may simultaneously be viewed as a coherent set of beliefs about how the world works and a dogmatic and rigid approach to policy that does not yield to evidence.

Incomplete Information, Costly Punishment and Imperfect Monitoring

In the following chapter I present a model of imperfect monitoring in an infinitely repeated game where a principal would like an agent to exert costly effort and has the ability to punish him. In a standard setting, with imperfect monitoring, a scheme where the principal will "punish the innocent" may arise. That is, in equilibrium the principal will know the agent is exerting effort; nevertheless whenever he observes a deviation, although he attributes it to noise in the monitoring mechanism, he will punish the agent in order to prevent him from having an incentive to deviate. I make two substantial assumptions: first, that punishment is not costless to the principal, utility is not transferable, either because the principal cares about the agent's welfare or punishment is actually costly to carry out; the second modification is to introduce some incomplete information: the principal is not completely certain that the state of the world is one in which the punishment scheme is worth its cost. Under these two conditions the principal's ability to implement the scheme unravels.

One way of thinking about it is that in a finite setting the principal will not punish in the last period and therefore no effort will be exerted. In the second to last period, since the principal knows there are no gains to be had in the future, he cannot credibly threaten to punish and therefore the agent doesn't provide effort. By backward induction we know there

will be no effort or punishments. The infinite game with no incomplete information allows the principal to threaten punishment for bad behaviour as equilibrium consequences, the abandonment of the effort and punishment scheme in the future, will serve as a sufficiently strong incentive to carry out costly punishments in the short run. The possibility of a bad state is similar to having a finite horizon, at least under some of the paths of play. It is then possible to make a backward induction argument from those particular paths of play back to the initial node.

The argument is that there exists a bad-luck-path of play in which our monitoring mechanism will always indicate a deviation and, therefore, a need to punish. This punishment is costly and, more importantly, it deteriorates the posterior probability the principal places on being in the good state of the world, one in which we would like to continue enforcing punishment. Down the bad-luck-path of play an end period of enforcement is always reached, the principal will eventually be uncertain enough about the true state of the world so that she is not willing to pay the cost of punishment. From this end period of enforcement I make a backward induction argument and show that there are never enough incentives to carry out punishments.

I also show an exception to this result, there is a unique type of principal that may actually implement this type of punishment, one who has a prior with zero weight on the bad state. Of course, this type of principal has a disadvantage, if the bad state is realized he will be unable to interpret bad outcomes as anything other than bad luck and thus will continue to carry out fruitless punishments.

The model presented in chapter 1 is the basis on which the political economy model of chapter 2 is developed. It presents a more general point about implementation with imperfect monitoring and gives a more complete and robust treatment regarding the assumptions. However, it does not contain any additional insight, in addition to those contained in chapter 2, as to the political economy of ideology and may be skipped by those who are uninterested in the technical underpinnings of the main model.

An Application to Political Economy

The setup presented above, a principal that cannot punish without cost to herself, or

cares about the effects of punishment, and is at least slightly uncertain about the way a punishment strategy will play out in the world, is readily applicable to political economy. Indeed, the two assumptions that generate the main result found in the first essay are natural in this context. A policy maker or enforcer can't be, in general, absolutely certain how policies will affect outcomes. It is also the case that a government, at least a democratic government, cares about the welfare of its citizens, which immediately introduces a cost of punishment.

Chapter 2 presents a political economy model in which everybody understands that a credible threat of punishment by the government will make citizens behave in a desirable way. However, the government, for the reasons presented above, may be unable to commit to such a policy. Only a government that is ideological, subjectively certain that it understands how the world works, can commit to the ex-ante optimal policy. This may lead the electorate to choose ideologues, even when they don't share their beliefs. This ex-ante optimal delegation has its drawbacks: if the bad scenario is realized, one where the punishment scheme is not optimal, then the ideological government will persist in implementing an erroneous policy.

This model is an example of a more general point, that commitment to a policy is important for its implementation. This model in particular explains why some high powered incentive schemes that are suggested by mechanism design are not readily applicable in a political context. The basic structure of the political (democratic) game may be incompatible with such schemes. The model explains why we may sometimes choose people with extreme beliefs: because they are uniquely endowed with a capacity to commit. They are well suited to overcome the difficulties inherent to the political system. Not surprisingly, electing this type of government might lead to inertia, i.e. the persistence of bad outcomes. Moreover the inertia that is presented stems solely from the inability of government to interpret bad outcomes as signaling the need to change course, and does not rely on external commitment devices. However, it gives a good reason why we may have incentives to elect precisely these types of governments, and why a politician may go to great lenghths to show how firmly grounded in a certain ideology she is.

An Electoral Model with Strategic Platform Divergence

In the final chapter, I present a model that sheds light on two related puzzles that have persistently plagued political economy in general and electoral models in particular.

One is platform divergence. Many of the most commonly used models in political economy rely on a median voter argument and conclude that party platforms, particularly in a two party system, will tend to converge on the prefered outcome of the median voter. I think this puzzle is related to the way we think about parties, we are still a long way from to understanding political parties and their preferences or ideologies and this model will also point out some promising results in that direction.

The second puzzle is the definition and role of ideology. Some people who study the social sciences may agree that it is important, but will still have a hard time agreeing on its definition. In this essay I point out two particularly contradictory views on ideology. The first conception is ideology as a consistent set of beliefs about how the world works (or ought to work). The second one is that of ideology as a rigid set of ideas about politics or social issues that is not malleable to evidence or discussion.

This essay presents a model in which two informed parties must choose their platforms before an election. The electorate is less informed about the true state of the world. Parties have information that is valuable to the electorate and it seems plausible that competition amongst them might lead them to reveal it in exchange for votes. In this admitedly simple model, I find that parties, out of strategic considerations, will find it in their best interest not to reveal information. Each party will present a unique (differentiated) policy regardless of the state of the world. The basic argument is that if the electorate relies on just one party to make inference about the true state of the world it is unlikely that a party will reveal an adverse (to its electoral fortunes) state of the world. If the citizen must rely on both parties to pin down the true state, at least one will have an incentive to deviate.

The model pins down a reason why parties have a strategic incentive to adhere to a differentiated platform that is associated with certain prior beliefs about how the world works. It also sheds light on the ambiguity regarding the interpretation of ideology. To the uninformed electorate, parties will appear, in their choice of platforms, as consistent with a

coherent set of beliefs about the world. To the informed observer parties will sustain their particular platform regardless of the information they possess concerning its benefits; as a result, they will rightfully be seen as demagogues. A reason for adherence by a party to a certain (more or less rigid) set of beliefs is thereby suggested. The spirit of this result is consistent with the assumption used in chapter 2, where parties are identified with, and commited to a particular set of beliefs.

CHAPTER 1

Who Will Punish the Innocent? An Incomplete Information Extension of a Game of Imperfect Monitoring

1.1 Introduction

The model introduces an incomplete information extension to a game of imperfect monitoring where a long lived player attempts to exact a particular behavior from an agent by threatening him with punishments, that are costly to the principal, for bad outcomes. The model can be interpreted as one of repeated interaction where a principal would like to elicit costly effort from an unwilling agent. The principal is imperfectly informed about the agent's actions and this results in costly punishments being imposed on "innocent" agents on the equilibrium path. I introduce an additional level of uncertainty that takes the form of an alternate state of the world where the exertion of effort is not desirable from the principal's perspective. With this, his ability to elicit effort unravels almost completely. Only a principal with extreme beliefs is able to implement the ex-ante desirable policy.

The model shows, first, that an equilibrium where strategies that "punish the innocent" are used might not be credible when there is even a small chance that the specified model is based on the "wrong" (optimistic) assumptions about the environment. The second point of the model is to show that there is a unique exception; the ideologue. The unique type of principal that will be able to implement the ex-ante desirable mechanism will also make the outcome prone to inertia, understood as the persistence of bad outcomes. For better of for worse, only an "ideologue" will "punish the innocent".

I also outline the possibilities that arise when the enforcement task can be delegated to an agent whose priors we can pick or impose. Potential applications for this strategic delegation aspect of the model may be found in the I.O., behavioral economics and political economy.

The paper is in the tradition of the literature that began with the seminal, "Gang of Four", paper by Kreps et. al. (1982) where introducing small amounts of incomplete information can greatly alter the strategic interaction between players. Cripps, Mailath and Samulson's (forthcoming) work on impermanent reputations has a similar flavour as it shows that a long-lived player cannot sustain a commitment reputation for behaviour that is not credible for its type in the long-run.

The rest of the chapter is organized as follows: the next section introduces the model and presents the main results and the corresponding proofs, the fourth and final section concludes.

1.2 Model and Assumptions

Consider a principal that wants an agent to exert effort and can threaten him with ex-post punishments for bad behavior. If monitoring is imperfect, costly on-the-equilibrium-path punishments will be inflicted upon agents he believes to be "innocent". If the principal internalizes the costs of punishing the agent, then, introducing the possibility that punishments may be "too costly" will call into question the credibility of such a scheme. A "punish the innocent" mechanism is fragile since bad outcomes will be interpreted as evi-

dence that the bad state has been realized. If priors deteriorate sufficiently, eventually, the principal will abandon the scheme. The possibility of an end period will unravel the whole punishment scheme since short term sacrifices (ex-post punishments) are not credible once the possibility of long term benefits is discarded. I will now develop the simplest model that will allow me to demonstrate the main claims in the essay.

1.2.1 Setup

Consider an incomplete information game with the following elements: There are two types of players: A principal called player 1 and agents called 2's. Player 1 is a single infinitely lived player. 2's are a series of short lived players, that live for one period, and are indexed by time subscripts.

The game:

$$g : \theta \times a_1 \times a_2 \rightarrow \Re^2$$

2's action set consists of two elements, zero and one, for every period; they can be thought of as a choice between shirking and exercising costly effort:

$$\begin{aligned} a_2^t &\in \{0,1\}\,\forall t \\ \text{where } a_2 &= 1 \text{ stands for effort exerted} \end{aligned}$$

1's action set at each period consists of two elements, 0 and -X, the latter can be thought of as a punishment imposed on 2's:

$$a_1^t \in \{0,-X\}\,\forall t$$

Recall that player 1's action is only relevant if the bad signal is observed. Payoffs are given by punishments and a production technology that stochastically maps effort into output, Q, where the random parameter, Λ, is drawn independently every period from a distribution that is state dependent. There is also an externality parameter, $\beta > 1$, explained below:

$$\begin{aligned} Q_t &= \beta \cdot \Lambda_t \cdot a_2^t \\ \Pr(\Lambda &= 1) = \lambda \\ \Pr(\Lambda &= 0) = 1 - \lambda \end{aligned}$$

There are two possible states of the world, good and bad, from which nature picks one in period zero with unknown probability. The good state of the world is one in which λ, the probability that the production parameter Λ is realized as one, is higher that in the bad state of the world.

State of the world:

$$\begin{aligned} W &\in \{G, B\} \\ G &\Rightarrow \lambda = \overline{\lambda} \\ B &\Rightarrow \lambda = \underline{\lambda} \\ 0 &< \underline{\lambda} < \overline{\lambda} < 1 \end{aligned}$$

Player 2's prior belief that the probability that the state of the world is G, is given by ω_0. Player 1 has a prior belief that places a probability on the state of the world being G given by θ_0. Priors are common knowledge, they may be equal, as in the more traditional set-up, but this is not imposed. Players update ω_0, θ_0 into posterior beliefs ω_t, θ_t from their knowledge of equilibrium strategies and available information using Bayes's rule where possible.

Player 1 has a discount factor $\delta < 1$ ($r = \frac{1}{\delta} - 1$), and his per period payoffs depend on the externality parameter, period specific luck, costly effort, and punishments. They are given by the following equation:

$$U_1(a_1, a_2, \Lambda) = \left\{ \begin{array}{ll} \beta\Lambda a_2 - ca_2 - X & \text{if } \Lambda = 0 \text{ and } a_1 = 1 \\ \beta\Lambda a_2 - ca_2 & \text{if } \Lambda = 1 \text{ or if } a_1 = 0 \end{array} \right\}$$

Per period payoffs to 2's differ from Player 1's only in that they do not incorporate the

externality. They are given by equation[1]:

$$U_2(a_1, a_2, \Lambda) = \left\{ \begin{array}{ll} \Lambda a_2 - ca_2 - X & \text{if } \Lambda = 0 \text{ and } a_1 = 1 \\ \Lambda a_2 - ca_2 & \text{if } \Lambda = 1 \text{ or if } a_1 = 0 \end{array} \right\}$$

Notice that expected net production if 2's exercise effort is given by $\lambda\beta - c$. To make things interesting I assume the following:

$$\text{A1} \ : \ \overline{\lambda}\beta > c$$

$$\text{A2} \ : \ \underline{\lambda}\beta < c$$

$$\text{A3} \ : \ \overline{\lambda} < c$$

Assumption 1 implies that the Player 1 would prefer Player 2 to exert effort in the good state. Assumption 2 says that Player 1 is better off if 2's don't exercise effort if the bad state is realized. Assumption 3 means that before punishments, 2's will not want to exercise effort even if the state is good.

I also assume the size of the punishment is large enough to induce effort when punishment is credibly threatened, even in the bad state.

$$\text{A4:} \underline{\lambda}(1 + X) > c$$

Monitoring technology has the following structure. After production takes place, a public signal is observed, I_t. When production is positive the signal is 1. If Player 1 receives this signal he rests for the period. If he doesn't receive the signal, denoted by a zero, this implies production was nil. Player 1's decision node is reached only when the signal is bad.

$$I_t(Q_t) = \begin{pmatrix} 1 \text{ if } Q_t = \beta \\ 0 \text{ otherwise} \end{pmatrix}$$

[1] We are implicitly assuming that agents can be punished even if production is realized at zero. We may think of them as being born with a small endowment or understand punishment as a non-monetary harm that is inflicted upon them, like prision terms.

History has two elements for each period, the public signal and the action Player 1 took in that period. We denote history as the record of events up to period t as:

$$h^t \equiv [I,\, a_1]_0^{t-1}$$
$$h^t \equiv [\{1,0\},\{0,-X\}]_0^{t-1}$$

Summing up, as seen in figure 1.1, there is a game where nature chooses between the two possible states of the world. A short-lived player chooses whether to exercise costly effort. Production is determined by effort and a stochastic productivity parameter. Player 1 would like Player 2's to exert effort in the good state but not in the bad state. 2's are not willing to exercise effort in any state, unless credibly threatened with punishment. When production is positive Player 1 does not intervene. If production is realized at zero Player 1 may punish or forgive. History records whether production was positive and whether Player 1 actually carried out punishment. 1 and 2's update priors with the information recorded in history and their knowledge of equilibrium strategies. Player 1 and 2's have independent, most likely different, prior beliefs, that are common knowledge. The purpose of the model is to explore Player 1's ability to implement a painful punishment scheme with imperfect monitoring in the presence of a deeper level of uncertainty about the desirability of such scheme.

1.2.2 Play

I look at perfect bayesian nash equilibria in pure strategies. An equilibrium for the infinitely repeated game is a pair of strategies and two lists of beliefs for player 1 and player 2 respectively.

$$S_{2,t} : H^t \rightarrow \{1,0\}$$
$$S_{1,t} : H^t \rightarrow \{0,-X\}$$

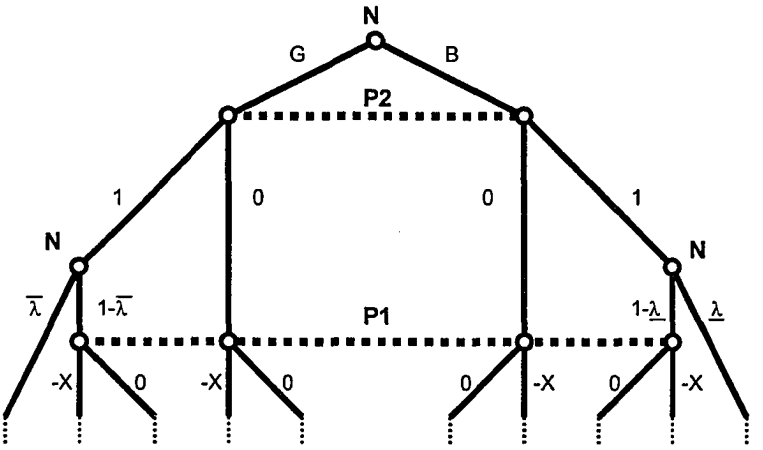

Figure 1.1:

A Perfect Bayesian Nash Equilibrium, PBNE is a set: $\{S_1^*, S_2^*, \theta, \omega\}$ s.t:

$1.E_t(U_i(S^*, \theta(S^*, H^t), \omega(S^*, H^t))) \geq E_t(U_i(S^*_{-i}, s'_i, \theta(S^*, H^t), \omega(S^*, H^t)) \qquad \forall i, s'_i, t, H^t$

$2.\theta_t \in [0,1], \qquad \omega_t \in [0,1] \qquad \forall t$

$3.\theta_t(\cdot \mid I^t) = \Pr(I^t \mid S^*, G) / \sum_W \Pr(I^t \mid S^*, W)$

$if \ \sum_W \Pr(I^t \mid S^*, W) > 0$

$4.\omega_t(\cdot \mid I^t) = \Pr(I^t \mid S^*, G) / \sum_W \Pr(I^t \mid S^*, W)$

$if \ \sum_W \Pr(I^t \mid S^*, W) > 0$

There is a stark result. A Player 1 (with non-extremist priors) will be unable to sustain a strategy that is consistent with an equilibrium in which the short run players exercise effort.

Proposition 1.1. *No implementation. If* $\theta \subset [0,1) \Rightarrow S^*_{1,2}(h^t) = 0 \ \forall t$ *on the equilibrium path. If Player 1 has non-extremist beliefs no equilibrium may arise where effort is exerted or punishments take place.*

Proof. Step 1. I define the "bad luck path" of play in an equilibrium profile as one where the signal is always bad, even when effort was exerted. I find that there exists a last period of enforcement and effort along every bad luck path of play in an equilibrium profile. This is bound to be true for two possible reasons; either strategies don't call for punishments and

effort after a sufficient number of bad outcomes or the exertion of effort accompanied by bad signals deteriorates beliefs "too much"; priors will get to the point where Player 1 is not willing to pay the immediate cost of punishment for any possible equilibrium consequences that may arise.

Define a "Bad Luck Path", BLP, as a path of play where $I_t = 0 \ \forall t$.

For every equilibrium profile, we have that along the BLP:

If $S_{2,t}(h) = 1$ *occurs a finite number of times, we label the last occurrence of* $S^*_{2,t}(h) = 1$ *as period* N'.

If $S_{2,t}(h) = 1$ *occurs infinitely often we proceed as follows:*

Note that $\left(\frac{\overline{\lambda}\beta+X}{r}\right)$ *is the maximum possible expected gain to Player 1 from equilibrium consequences at any given period. Notice this maximum gain can be obtained only in the good state for punishments can only decrease payoffs in the bad state as they are directly costly and might induce inefficient effort in a given period*

By equating the maximum possible gain, for a given set of beliefs, to the immediate cost of punishment: $\left[\left(\frac{\overline{\lambda}\beta+X}{r}\right) * \theta\right] = X$,

I find the minimal belief, $\underline{\theta}$, *that still allows Player 1 to carry out a punishment:*

$\underline{\theta} = X \div \left(\frac{\overline{\lambda}\beta+X}{r}\right)$

I now define the odds ratio for Player 1's beliefs:

$\Theta = \theta/(1-\theta)$

and note that the odds ratio is updated as follows:

$\Theta(n) = \Theta_0 \left(\frac{1-\overline{\lambda}}{1-\underline{\lambda}}\right)^n$

where n *is the number of periods along the BLP where* $S_{A,t}(h^t) = 1$

We now state the minimal belief in terms of the odds ratio:

$\underline{\Theta} = \frac{Xr}{\overline{\lambda}\beta+X-Xr}$

And define period N, *as the first period where* $\Theta(N) < \underline{\Theta} = \frac{Xr}{\overline{\lambda}\beta+X-Xr}$

Step 2. If period N, *is reached, no punishment is possible. In fact, no further punishment is possible.*

If period N *is reached,* $\Theta(N) < \underline{\Theta}$

$\Rightarrow S_{1,N}(h, I) = 0$

$\Rightarrow S_{2,N}(h) = 0$

At period $N+1$, *if* $I_{N+1} = 0$

$\Rightarrow \Theta_{N+1} \leq \Theta(N) < \underline{\Theta}$

$\Rightarrow S_{1,N+1}(h, I) = 0$

$\Rightarrow S_{2,N+1}(h) = 0$

By forward induction, $\forall \tau$:

$\Rightarrow \Theta_{N+\tau} \leq \Theta(N) < \underline{\Theta}$

$\Rightarrow S_{1,E+\tau}(h, I) = 0$

$\Rightarrow S_{2,E+\tau}(h) = 0$

Step3. If there is no credible punishment after a point in history, then there are no possible equilibrium consequences that could induce Player 1 to punish in the previous period, in the period before that one, and all previous periods.

At period $N-1$ *and* $N'-1$, *if* $I_{N-1} = 0$

Punishment has a same period cost of $-X$ *and, from above, no possible equilibrium consequences that would offset this cost*

$\Rightarrow S_{1,N-1}(h, I = 0) = 0$

$\Rightarrow S_{2,N-1}(h) = 0$

By backward induction, $\forall \tau$*:*

At period $N - \tau$ *if* $I_{N-\tau} = 0$

$\Rightarrow S_{1,N-\tau}(h, I = 0) = 0$

$\Rightarrow S_{2,N-\tau}(h) = 0$ □

I have shown that if Player 1 has prior beliefs that are not extremist, $\theta_0 \subset [0, 1)$, then no strategies where effort is exerted can constitute an equilibrium profile. If, however, Player 1 has prior, $\theta_0 = 1$, he is subjectively certain about the nature of payoffs and it can be shown that equilibria with strategies that dictate effort on the equilibrium path exist.

Proposition 1.2. *Implementation by a Player 1 with extreme beliefs.*

If $\theta_0 = 1$ *and* $\frac{\delta}{(1-\delta)} \geqslant \frac{X}{\lambda\beta - c} \Rightarrow$

$$S_2^*(h^t): \left\{ \begin{array}{ll} 0, & h^t = \{0,0\} \text{ for any } t \\ 1, & \text{otherwise} \end{array} \right\},$$

$$S_1^*(h^t): \left\{ \begin{array}{ll} 0, & h^t = \{0,0\} \text{ for any } t \\ -X, & \text{otherwise} \end{array} \right\}$$

is an equilibrium $\forall \omega_0$.

Proof. If $\theta_0 = 1$, *Bayesian updating implies* $\theta_t = 1\ \forall t$. *Player 1's beliefs are such that he understands payoffs to be associated with G.*

2's beliefs are irrelevant, as the decision to exert effort is determined by the threat of punishment irrespective of the state of the world.

On the equilibrium path:

Short lived players have no incentive to shirk, as, per S_1^* *and A3, this will simply result in a welfare decreasing punishment, for any beliefs they have about the state of the world.*

Player 1 can credibly punish since S_2^* *is backward looking and shirking entails no further effort exertion.*

Off the equilibrium path:

Players 2 can credibly shirk for strategies dictate that there are no further threats of punishment that would induce them to do so.

Finally, Player 1 will credibly abandon the punishment strategy as equilibrium strategies dictate no further effort exertion and punishment represents a pure loss at the end of the period and no gain in the future. In period 1 conditions on previous play hold vacuously and on the equilibrium path we observe effort exertion and punishments when low output is realized. □

In other words, in period 1 strategy relevant conditions on previous play hold vacuously. Nature moves first, if the world is G then effort is always exerted by type two players, output is realized with high probability; when it is not punishments are carried out. If the world is B then effort is always exerted, output is realized with low probability and punishments are carried out, more often, when the signal is zero. In expectation, the highest possible player one payoffs are observed in the first case, the worst possible equilibrium player 1 payoffs

are observed in the second case. A player one with extreme priors reduces the game to a gamble on the state of the world being good.

Notice that in the proof to proposition 1, the fact that players two act myopically is not used. Indeed substituting the short lived players for a long lived player 2 with a utility function that is the discounted sum of per period payoffs leaves our result intact. This is because the long lived player two is strictly worse off under any punishment scheme, regardless of his beliefs. In period $N-1$, the opportunity presents itself for player two to do away forever with the punishment scheme without risk of being punished. Knowing this is the case with certainty, player one will not punish in the previous period and the argument stated above falls through intact.[2]

Proposition 1.3. *Proposition 2.1 holds for a long lived player 2.*

Proof. See Proof to Proposition 1. □

Also notice that assumptions made about effort and punishments being discrete are not essential to the argument, as long as there is a lower bound on the amount of effort (or punishment) that may be exerted.

Suppose the model above is modified as follows. First, allow for a continuos action space for player two given by:

$$\begin{aligned} a_2^t &\in \{[0],[m,\infty)\}\forall t \\ &\quad \text{where } a_2 \text{ is the level of effort} \\ C^t &= c \cdot a_2^t \\ &\quad \text{is the cost of effort} \end{aligned}$$

The action space for the punishment is also allowed to be continuos and arbitrarily large:

[2]Notice that although the proof to proposition 1 does not rely on player two being short lived, the proof we have used for the existance of an equilibrium with trigger strategies in proposition 2 does require it. An equilibrium profile with long lived players can be constructed by using shorter, therefore credible, periods of punishment.

$$a_1^t \in (-\infty, 0] \quad \forall t$$

Proposition 1.4. *If there is a positive lower bound on the amount of effort that may be exerted, $m > 0$, proposition 1 holds for continuous effort and continuous punishment action spaces.*

Proof. Notice that A3, A4 imply that for the minimum effort to be exerted a sufficiently large punishment must be threatened. That is, for $a_2^N = m$ to be part of an equilibrium profile $a_1^N \geqslant \widetilde{X} > 0$ must hold. Now the proof to proposition 1 follows through if we replace the exogenously given, X with the minimal effective punishment in the extended model, $\widetilde{X}$. □

Notice that, in the absence of a lower bound on effort, an upper bound on the size of punishment suffices for the argument to hold.

1.3 Discussion and Conclusions

I've presented a dynamic model where a principal that would like to induce actions from an agent by threatening him with punishments will loose credibility if there is a possibility that the policy was designed under flawed premises. This possibility would make him abandon the policy after a sufficient number of bad outcomes. This potential end period allows the agent to shirk with impunity making the punishment lack credibility in the previous period as well, as there are no positive long run consequences to offset its immediate costs to the principal. Because of this process the whole scheme unravels. Only a principal that is subjectively certain that he understands the world will be able to implement the punishment strategy. This makes such an "ideological" player useful if we are able to pick the type of player, defined by his priors, that will actually play the game. If such a player is chosen and the world is in the bad state, a bad outcome will be persistently observed.

I point out three avenues for future research. First, an application to political economy addressed in the following chapter. Second, strategic delegation when we can pick or fix the "supervisor's" priors may also be important to the internal organization of the firm.

Finally, I would also like to explore a model of collusion where the possibility that prices are determined exogenously calls the credibility of price wars as enforcement mechanisms into question. A model of collusion would imply extending the model to a context with many long lived players that internalize the long run benefits. The fact that players internalize the benefits of collusion will increase their incentives to cooperate as the priors deteriorate. However, if priors deteriorate too rapidly or the gains from defecting increase with the number of players it is likely that sufficient conditions may be found under which players will be unable to establish a collusive regime due to the possibility of falling on the bad luck path.

1.4 Bibliography

1. Coate S. and S. Morris. (1999). "Policy Persistence." *American Economic Review* 89 (1999), 1327-1336.

2. Cripps, M., G. Mailath and L. Samuleson. (forthcoming). "Imperfect Monitoring and Impermanent Reputations". *Econometrica.*

3. Dewatripont, M and J. Tirole. (1999). "Advocates." *Journal of Political Economy* 107:1-39.

4. Green E. and R. Porter. (1984). "Noncooperative collusion under imperfect price information." *Econometrica* January 1984: 87-100.

5. Holmstrom, B. (1982). "Moral Hazard in Teams." *Bell Journal of Economics* 10: 74-91.

6. Kreps, D.M., P. Milgrom, J. Roberts, and R. Wilson (1982). "Rational Cooperation in the Finitely Repeated Prisioner's Dilemma", *Journal of Economic Theory*, 27: 245-252.

7. Mailath, G. and S. Morris (2000). "Repeated Games with Almost-Public Monitoring." *Journal of Economic Theory* 102: 189-228.

8. Morris, S. (1996). "Speculative Investor Behavior and Learning." *Quarterly Journal of Economics* 11: 1111-1133.

9. Perry, M. and P. Reny. (1994). "A Noncooperative View of Coalition Formation and the Core." *Econometrica* 62: 795-817.

CHAPTER 2

Institutional Inertia: Ideology

2.1 Introduction

Political Economy has had a hard time dealing with "ideology". At least since Downs (1957), a common assumption has been that politicians are opportunists, who maximize the probability of staying in office, the number of votes, or some form of rents from office. Other authors, as Wittman (1973), have incorporated the politician's ideology to their arguments. In their model ideology is understood as preferences over policy outcomes. It is an open issue whether politicians are mainly opportunistic, partisan or a combination of both (Roemer 2001, Chapter 8) and there are numerous examples where the differences between these two interpretations have relevant consequences (Alesina, Cohen and Roubini 1997).

Nevertheless, ideology is important. North (1981, page 47) says that: "Without an explicit theory of ideology[...]there are immense gaps in our ability to account for either current allocation of resources or historical change." I believe that ideology -and also extreme ideology- plays an important role in political decision making and that we have so far been unable to model its role satisfactorily.

As Gerrig (1997) has documented, there is a large number of different, sometimes con-

tradictory, definitions of ideology used in the social sciences. I understand ideology, not as a simple taste for one policy or another, but as something more rich and complex, basically our "world view"[1]. Ideology takes into account preferences and moral judgements, from which I abstract in this essay, as well as beliefs about how the world works. I focus on these differences of opinion over the way the world functions; in my model, they are reduced to beliefs over parameters that map policies into fundamentals.

Like Picketty (1995), I understand ideology mainly as beliefs. In this essay I model differences in ideology between competing candidates as heterogeneous priors about the true state of the world, that are common knowledge. As Roemer (2001, page 40) says: "...much political debate between parties takes place over the values of economic parameters -that is, how the economy will respond to particular policies."

As a simple example, let's look at a redistributive tax policy that expropriates all income and gives it back in equal lump-sum payments. With ideology as preferences, a conservative candidate is an individual that has a higher that average income, he objects to the scheme because it reduces his net income and he prefers to have a higher net income. With ideology as beliefs, a conservative candidate might be an individual who places a high prior on a state of the world where high taxes are undesirable because the perverse incentives they introduce to the labour market are too high. By abstracting from preferences and moral judgements, I don't mean to imply that they don't matter; only that differences in beliefs can carry us a long way towards understanding some aspects of political competition, particularly among parties.

---

[1] North, (1981), stresses three aspects about ideology:

"1. Ideology is an economizing device by which individuals come to terms with their environment and are provided with a "world view" so that the decision-making process is simplified; 2. Ideology is inextricably interwoven with moral and ethical judgements about the fairness of the world the individual perceives. This situation clearly implies a notion of possible alternatives-competing rationalizations or ideologies. A normative judgement of the "proper" distribution of income is an important part of an ideology; 3. Individuals alter their ideological perspectives when their experiences are inconsistent with their ideology."

Although I agree with these three aspects, my model uses mostly the first one which is a pretty accurate description of a belief system as understood in game theory.

In a static setting it is hard to distinguish between differences in preferences over policies and differences in beliefs about the state of the world. In a dynamic environment differences come to light, as agents with different priors gather and process information, as they may draw different conclusions and change their behavior accordingly, even if preferences over fundamentals remain identical and stable. These changes in opinion about the desirability of some policy, how steadfast is a candidate's support for a policy in the face of adversity (negative information), is a fundamental aspect about ideology that is captured by beliefs.

As I mentioned above, there are a number of often contradictory definitions of ideology. According to Gerrig (1997), the one trait that seems to be common to all definitions is *coherence*; ideology refers to a set of ideas that belong together in a non-random fashion. Gerrig adds *stability* and *contrast* as corollaries[2]. This minimal view is consistent with my understanding of ideology as beliefs about the state of the world.

By modeling ideology as heterogeneous priors, as opposed to differences in private information, we get an interesting feature: beliefs (ideology) don't necessarily converge simply because information about beliefs is shared. I think this is a desirable trait; it is not usually the case that a person with a particular ideology, say a liberal in the American sense, becomes more inclined to support a conservative position just because a conservative reveals himself as such. Persistent differences in opinion are pervasive feature of political debate. In politics it is true that we sometimes *do* agree to disagree.

This essay shows the use of modeling ideology as beliefs by applying it to a set-up in which ideology seems to play a relevant role: commitment to a policy by a policy maker in the face of adverse outcomes and information.

The model presented in this chapter is based on the one presented in the previous chapter, modified to include more than one agent in every period and extended to formalize the endogenous selection of the principal that will take part. These changes will allow for richer interpretation of the model in a political economy context.

In this essay I show how extreme ideology, construed as degenerate prior beliefs, works

[2]S*tability,* coherence throughout time, is the first corollary to the minimal definition that Gerrig comes up with. The other is *contrast*, coherence vis-a-vis a competing ideology.

as a particularly strong commitment device. I introduce a game in which citizens would like to coordinate on a cooperative equilibrium. However, they are too myopic to cooperate. A benevolent government that seeks to maximize citizens' welfare is introduced. It is able to influence citizens' behavior through the use and threat of punishments. The monitoring technology is imperfect and this leads to punishments being applied in equilibrium, even if all citizens are innocent (cooperate in equilibrium). This makes the punishment threat scheme costly. I also introduce incomplete information; there is a distinct possibility that technology may be such that the costs of the policy outweigh its gains. This possibility will undermine the capacity of most governments to commit to the policy and the whole scheme will unravel. There is an exception to this result; a government that places zero prior probability on punishments being too costly will be able to implement the policy.

Naturally, this particularly strong commitment device has particularly pernicious consequences when things turn out for the worst. This model shows how unambiguously ex-post inefficient outcomes arise and persist in equilibrium, what I call *institutional inertia.* The capacity to commit motivates the selection of extreme ideological governments that are liable to fall prey to such inertia. The electorate has an incentive to elect this type of ideological government, even when it doesn't share its system of beliefs. There is a capitulation to the extremist, as in Roemer (2001, Chapter 5) but in this case not only within a party but for society at large.

The issue of inertia or persistence has been amply, if not comprehensively, addressed by many prominent economists. Alesina and Drazen (1991) have given us a model that show how a "war of attrition" to determine who will bear the indivisible costs of a useful policy can result in a sub-optimal delay. Fernandez and Rodrik (1991) explain why individual specific uncertainty can lead to a combination of non-adoption and reversal of newly introduced policies that results in status quo bias. Coate and Morris (1999) emphasize how the actual implementation of policies will lead interest groups to undertake investments that will make them more willing to pay for the policy in the future. This last model emphasizes the relationship between inertia and non-adoption, showing how a policy that tends to persist might not be introduced if it is only useful in the short run.

The model I present here fits into this line of reasoning and exemplifies instances in which we have on-the-equilibrium-path failure both because of non-adoption and because of inertia. More recently, Mujumdar and Mukand (2002) have proposed that incoming governments may take inefficient gambles with policy to signal their confidence in their abilities. These same governments may want to persist with these policies when they go wrong, stick to their guns in order to maintain their reputation as able and self-confident players. Their paper acknowledges that ideological considerations (preferences), are an alternative and hard to distinguish motivation to the observable phenomenon they seek to explain. One interesting aspect is that failure, in the models mentioned above, comes from some sort of technological failure; indivisibility of costs, individual specific uncertainty, switching cost nature of investment, inability to credibly share information, etc. The model highlights a hurdle, an incapacity to correctly incorporate all available information and update beliefs efficiently, that, because of its use as a commitment device or simply because of lack of imagination, is likely to remain a relevant feature of political decision making; it must be acknowledged and should be incorporated into our analysis to achieve the best understanding possible of the political side of the policy making process.

The following section gives a brief motivation for the essay. Section 2.3 presents the model and main results. Section 2.4 discusses some underlying issues and concludes.

2.2 Motivation

The following paragraphs illustrate the type of policy outcomes that the model attempts to explain. The historic events described are more complicated than the model suggests and are mentioned for heuristic purposes only.

After the fall of the Berlin Wall, a painful process took place in the economies of the former Soviet block; these formerly centralized economies were transformed into market economies. The prevailing prescription called for "shock therapy" that would set in place the right incentives for growth. The state sector was privatized and a price system was introduced. The immediate outcomes in most countries were deep recessions and large

scale unemployment. These two were seen, in part, as necessary steps that would induce workers and capital to be reallocated to more productive activities. The large welfare costs of large scale unemployment were seen as a necessary evil; they were part of a "no pain, no gain" prescription. There are surely many policies that the transition economies could have undertaken to attenuate the pain from unemployment and bankruptcies; however, large scale unemployment was seen as part of the incentive structure. It was hoped that this arrangement would lead to an efficient allocation of resources and growth. Critics suggesting that the profound recessions that ensued were evidence for abandoning the crash program were mostly brushed aside.

Under the prevailing system of beliefs, the persistence of unemployment and recession, evidently painful outcomes, was interpreted not only as evidence that the correct policy was being implemented, but as crucial elements of the system. They were expected, necessary outcomes and they were never perceived as a signal that the policy was flawed. A disconnect between bad outcomes and the desirability of policy ensued.

In some countries this bitter medicine worked well. Others did not fare so well but were encouraged to stay the course and took a long time to change course. Some are still sticking to the same recipe to no avail.

A similar story can be told for numerous developing economies, such as Argentina, Brazil, and Mexico, that adopted more or less rigid, fixed exchange rate regimes during the eighties and nineties. Fixed exchange rates, but more generally, tight monetary policies, were seen as a key, necessary step to bring down inflation, an integral part of the macroeconomic stabilization programs that these countries where undertaking. I argue that tight money in the face of price shocks can be seen as a mechanism that will allow economic players to coordinate on a system where stable prices prevail; as opposed to one in which any shift in observed prices is accommodated by the monetary authority and an inflationary spiral ensues. In these circumstances tight money and the slowdown that it begets is the painful but necessary incentive needed too coordinate on stable prices. An ideological government committed to tight monetary policies, for whatever reasons, may be unable to realize that the true state of the world calls for flexibility if certain exogenous, and costly,

price shocks are unavoidable.

The relation of the arguments presented in this essay to the literature on commitment and strategic delegation in monetary policy will be explored after the model is formally introduced to enable a more thorough comparison.

2.3 Model and Assumptions

In this section I develop the simplest model that allows me to show the main points in the essay.

Consider a game where a long lived government is elected to office and interacts with cohorts of short lived pairs of citizens. The government is assumed to be benevolent and would like the citizens to coordinate on mutually beneficial actions from which they have individual incentives to deviate. In order to achieve this goal the government may punish citizens. The government would like to use the threat of punishments to induce cooperation; imperfect monitoring will make costly on-the-equilibrium-path punishments inevitable. I introduce the possibility of an alternate state of the world in which punishments are not useful and explore the government's ability to induce cooperative behavior in this set-up.

The main points of the model are: first, if there is a small probability it was designed under flawed premises, the credibility of a government's commitment to a policy that imposes painful punishments in hopes of dissuading future transgression is called into question. Second, ideologically extreme governments are uniquely positioned to overcome this hurdle and may thereby gain access to office. Finally, when extremist are elected to office and things turn out for the worse, the inadequate policy will tend to persist.

2.3.1 Setup

There are three types of players: odd citizens (1), even citizens (2), and candidates. Citizens are short lived; each cohort lives only one period. The pairs of citizens are identical except for their label. There is also a continuum of candidates of length one, their location on the continuum characterizes them on a unidimensional belief space. Candidates are infinitely

lived; they may be thought of as parties.

Citizens' action sets consist of two elements for every period; exercise costly effort or not (shirk):

$$\begin{aligned} e_{i,t} &\in \{0,1\}\,\forall(i \in \{1,2\}), t \\ \text{where } e_{i,t} &= 1 \text{ stands for effort exerted} \\ c_{i,t} &= c \text{ iff } e_{it} = 1 \\ c_{i,t} &= 0 \text{ otherwise} \end{aligned}$$

There is one "election", to be described below, the candidate who is elected becomes the government. The government may choose between two actions in every period in which he is called upon to act, punish or forgive:

$$\begin{aligned} p_t &\in \{0,-X\}\,\forall t \\ \text{where } p_t &= -X \text{ stands for punish} \end{aligned}$$

There are two possible states of the world, good and bad, from which nature picks one in period zero with unknown probability. The good state of the world is one in which Λ, a productivity parameter, is realized positive, rather than zero, with a higher independent probability each period, $\overline{\lambda}$ vs.$\underline{\lambda}$:

$$\begin{aligned} &\text{State of the world} \quad : \\ &\qquad W \in \{G,B\} \\ &\qquad G \Rightarrow \lambda = \overline{\lambda} \\ &\qquad B \Rightarrow \lambda = \underline{\lambda} \\ &\text{where } 0 < \underline{\lambda} < \overline{\lambda} < 1 \end{aligned}$$

Citizens are identical and share a common prior about the probability, ω, that the state

of the world is G. Candidates are heterogeneous and place a prior probability on the state of the world being G that is given by their location on the continuum and labeled θ_j.

In order to introduce payoffs, I now sketch how the game is played.

In the first period a representative citizen, for they are all identical, will choose a candidate to govern. Once a government is elected players engage in an infinite game. Each cohort is born, engages in joint production, receives payoffs, and dies. During production, a player may either exert effort or shirk. Production depends on the effort of both citizens and luck. The stochastic element is given by the productivity parameter λ. Production by each pair is as follows.

$$\begin{aligned} Q_t &= \Lambda\beta(e_{1.t} + e_{2,t}) \\ \text{where } \Lambda &= 1, \text{ is realized with probability } \lambda \text{ each period} \\ \Lambda &= 0, \text{ is realized with probability } 1-\lambda \text{ each period} \\ & \beta \text{ is a strictly positive parameter} \end{aligned}$$

In the absence of government (or punishments) payoffs to citizens are given by the effort exerted by each member of the pair, each player's private cost of effort, and the stochastic productivity parameter. Utility is given directly by payoffs.

$$C_{i,t} = (Q_t)/2 - (e_{i,t} * c)$$

Each pair faces the typical problem encountered by team production. Not all benefits of effort are accrued by those who incur its cost. Notice that expected net production if both players exercise effort is given by $2\lambda\beta - 2c$. The following assumptions are made:

$$\begin{aligned} \text{A1} &: \overline{\lambda}\beta > c \\ \text{A2} &: \underline{\lambda}\beta < c \\ \text{A3} &: (\overline{\lambda}\beta)/2 < c \end{aligned}$$

Assumption 1 says that it is socially efficient for a member of society to exert effort in the good state of the world. Assumption 2 states that is it is socially inefficient for citizens to exert effort if the state of the world is bad. Assumption 3 says that it is individually inefficient for the short lived citizens to exert effort, even if the state of the world is good. Individuals are rational and self interested and therefore they will never exert effort in a one-shot setup.

The above outcome suggests that there is a role to be played by a government. I now introduce this possibility. After the election takes place one of the citizens becomes the government. He has only blunt instruments at his disposal, he can punish citizens by confiscating (and destroying) their wealth. This government is blunt but has a distinct advantage when trying to achieve a good result for society, he is purely benevolent. His payoffs are the discounted sum of payoffs to all citizens.

The government works in the following manner. After each period in which production takes place, the government observes a signal, I_t. When things turn out for the best, that is the productivity parameter was realized as one and both players exerted effort, the signal is 1. If he receives this signal he rests for the period. If he doesn't receive the signal, denoted by a zero, either the productivity parameter was not realized at one or at least one player didn't exert effort, and he is called upon to act. The governments decision node is reached only when the signal is bad.

$$I_t(Q_t) = \begin{pmatrix} 1 \text{ if } Q_t = 2\beta \\ 0 \text{ otherwise} \end{pmatrix}$$

Payoffs when punishment takes place are given by:

$$C_{i,t} = (Q_t - X)/2 - (e_{i,t} * c)$$

I assume the size of the punishment is large enough so that citizens will want to exert effort, as long as the other citizen exerts effort, even if they find themselves in the bad state of the world. This assumption implies that citizens will find it in their best interests to comply even if they believe the state of the world is such that punishments are not warranted.

$$\text{A4: } \underline{\lambda}(\beta + X/2) > c$$

Citizens are short lived; however, they keep a record for their descendants about the outcome of each period (we may also think of them as long-lived, albeit impatient, citizens). This information is truthfully recorded and helps each cohort to update citizen beliefs. The content of the record, history, states only whether the government was called upon to act (implying the best outcome did not take place) and what action the government took. We denote history as the record of events up to period t as:

$$\begin{aligned} h^t &\equiv [I_\tau,\, a_{1\tau}]_{\tau=0}^{t-1} \\ h^t &\equiv [\{1,0\},\{0,-X\}]_{\tau=0}^{t-1} \end{aligned}$$

Summing up, there is a game where nature chooses between the two possible states of the world. There is a continuum of long-lived candidates, of size one, that are characterized by their position in the continuum, which represents the prior belief they place on the state of the world being "good". All candidates run for office and an "election" is held. A representative citizen picks the winner based on his predictions about equilibrium outcomes. There are two types of short-lived citizens, labeled odd and even, who are otherwise identical. Each cohort of citizens participates in joint production. Production is determined by effort and a stochastic productivity parameter (the distribution of which depends on the state of nature). It is jointly optimal for all players to exert effort in the good state and shirk in the bad state. Given property rights, joint production technology makes exerting effort individually suboptimal even in the good state. When production turns out for the best, government does not intervene. If things go wrong the government may either punish both citizens or forgive them. History records whether things turned out for the best and whether governments punished. Citizens and governments update their priors with the information recorded in history using Bayes's rule where possible. Citizens and government have different prior beliefs, this is common knowledge.

2.3.2 Policy Implementation

I first explore how equilibrium would look like under the different regimes (beliefs). I concentrate on perfect bayesian nash equilibria (PBNE) in pure strategies. An equilibrium for the infinitely repeated game is a pair of strategies and two lists of prior and posterior beliefs. Strategies have the following form:

$$S_{i,t} : H^t \rightarrow \{1,0\}$$
$$S_{g,t} : H^t \rightarrow \{0,-X\}$$

Eq: $\{S^*_{1,2}, S^*_g, \theta, \omega\}$ s.t:

$1.E_t(U_i(S^*,\theta(S^*,H^t),\omega(S^*,H^t))) \geq E_t(U_i(S^*_{-i},s'_i,\theta(S^*,H^t),\omega(S^*,H^t)) \quad \forall i,s'_i,t,H^t$

$2.E_t(U_g(S^*,\theta(S^*,H^t),\omega(S^*,H^t))) \geq E_t(U_g(S^*_{-i},s'_i,\theta(S^*,H^t),\omega(S^*,H^t)) \quad \forall i,s'_i,t,H^t$

$3.\theta_t \in [0,1], \quad \omega_t \in [0,1] \quad \forall t$

$4.\theta_t(S^*,H^t,I^t) = \Pr(H^t,I^t \mid S^*,G) / \sum_W \Pr(H^t,I^t \mid S^*,W)$

$if \sum_W \Pr(H^t,I^t \mid S^*,W) > 0$

$5.\omega_t(S^*,H^t) = \Pr(H^t \mid S^*,G) / \sum_W \Pr(H^t \mid S^*,W)$

$if \sum_W \Pr(H^t \mid S^*,W) > 0$

There is, as in chapter 1, a stark but by now familiar result. No regime, except for one, will be able to implement an equilibrium where effort is exerted.

Proposition 2.1. *No implementation. If* $\theta_0 \subset (0,1) \Rightarrow S_{1,2}(h^t) = 0 \; \forall i,t$. *If the elected government has non-extremist beliefs no equilibrium strategy may arise where effort is exerted on the equilibrium path.*

Proof. Step 1. I define a "bad luck path" of play in an equilibrium profile as one where the signal is always bad, even when effort was exerted by both citizens. I find that there exists a last period of enforcement and effort along every bad luck path of play in an equilibrium profile. This is bound to be true for two reasons: either strategies don't call for punishments and effort after a sufficient number of bad outcomes or unsuccessful effort deteriorates beliefs

up to the point where the government is not willing to pay the immediate cost of punishment for any possible equilibrium consequences that may arise.

Define a "Bad Luck Path", BLP, as a path of play where $I_t = 0\ \forall t$.

For every equilibrium profile, we have that along the BLP:

If $S_{1,2,t}(h) = 1$ *occurs a finite number of times, we label the last occurrence of* $S^*_{1,2,t}(h) = 1$ *as period* N'.

If $S_{1,2,t}(h) = 1$ *occurs infinitely often we proceed as follows:*

Note that $\left(\frac{2\bar{\lambda}\beta+X}{r}\right)$ *is the maximum possible expected gain from equilibrium consequences at any given period. Notice this maximum gain can be obtained only in the good state for punishments can only decrease payoffs in the bad state as they are directly costly and might induce inefficient effort in a given period*

By equating the maximal possible gain, for a given set of beliefs, to the immediate cost of punishment: $\left[\left(\frac{2\bar{\lambda}\beta+X}{r}\right) * \theta\right] = X$,

I am able to define the minimal belief, $\underline{\theta}$, *that still allows the government to exercise a punishment:*

$\underline{\theta} = X \div \left(\frac{2\bar{\lambda}\beta+X}{r}\right)$

I now define the odds ratio for government beliefs:

$\Theta = \theta/(1-\theta)$

and note that the odds ratio is updated as follows:

$\Theta(n) = \Theta_0 \left(\frac{1-\bar{\lambda}}{1-\underline{\lambda}}\right)^n$

where n *is the number of periods along the BLP where* $S_{1,2,t}(h^t) = 1$

We now state the minimal belief in terms of the odds ratio:

$\underline{\Theta} = \frac{Xr}{2\bar{\lambda}\beta+X-Xr}$

And define period N, *as the first period where* $\Theta(N) < \underline{\Theta} = \frac{Xr}{2\lambda\beta+X-Xr}$

Step 2. If period N *is reached, no punishment is possible. In fact, no further punishment is possible.*

If period N *is reached,* $\Theta(N) < \underline{\Theta}$

$\Rightarrow S_{g,N}(h, I) = 0$

$\Rightarrow S_{1,2,N}(h) = 0$

At period $N+1$, *if* $I_{N+1}=0$

$\Rightarrow \Theta_{N+1} \leq \Theta(N) < \underline{\Theta}$

$\Rightarrow S_{g,N+1}(h,I)=0$

$\Rightarrow S_{1,2,N+1}(h)=0$

By forward induction, $\forall\tau$:

$\Rightarrow \Theta_{N+\tau} \leq \Theta(N) < \underline{\Theta}$

$\Rightarrow S_{g,E+\tau}(h,I)=0$

$\Rightarrow S_{1,2,E+\tau}(h)=0$

Step3. If there is no credible punishment after a point in history, then there are no possible equilibrium consequences that could induce a government to punish in the previous period, in the period before that one and therefore all previous periods.

At period $N-1$ *and* $N'-1$, *if* $I_{N-1}=0$

Punishment has a same period cost of $-X$ *and, from above, no possible equilibrium consequences that would offset this cost*

$\Rightarrow S_{g,N-1}(h,I=0)=0$

$\Rightarrow S_{1,2,N-1}(h)=0$

By backward induction, $\forall\tau$:

At period $N-\tau$ *if* $I_{N-\tau}=0$

$\Rightarrow S_{g,N-\tau}(h,I=0)=0$

$\Rightarrow S_{1,2,N-\tau}(h)=0$ □

I have shown that if the government has prior beliefs that are not extremist, $\theta_{g,0} \subset [0,1)$, then no strategies that call for effort to be exerted can constitute an equilibrium profile. If, however, the government has beliefs, $\theta_{g,0}=1$, it is easy to show that there are many possible equilibria that include periods of effort exertion.

The multiplicity of equilibria we encounter when the government is an ideological extremist make the ex-ante analysis of equilibrium play by the representative citizen, and therefore the electoral decision, more complicated. Given our interpretation of degenerate priors as corresponding to an ideological extremist government that is subjectively certain

about the benefits of effort exertion, a salient benchmark for comparison purposes is a government that applies a strategy that always punishes when it observes a bad signal. Since there are many possible equilibria that are consistent with this basic strategy, namely variants in which citizens are unable to coordinate on cooperation at a varying number of periods, we arbitrarily choose a benchmark scenario: one where citizens are able to coordinate on cooperation whenever there is a credible threat of punishment.

Proposition 2.2. *Implementation by an ideologue.*

If $\theta_0 = 1$ *and* $\frac{\delta}{(1-\delta)} \geqslant \frac{X}{\lambda\beta - c} \Rightarrow$

$$S^*_{1,2}(h^t) : \left\{ \begin{array}{ll} 0, & h^t = \{0,0\} \text{ for any } t \\ 1, & \text{otherwise} \end{array} \right\},$$

$$S^*_g(h^t) : \left\{ \begin{array}{ll} 0, & h^t = \{0,0\} \text{ for any } t \\ -X, & \text{otherwise} \end{array} \right\}$$

is an equilibrium $\forall \omega_0$.

Proof. If $\theta_0 = 1$, Bayesian updating implies $\theta_t = 1 \ \forall t$. Government payoffs are therefore those associated with G.

Citizens beliefs are irrelevant, as the decision to exert effort is determined by the threat of punishment, or lack thereof, irrespective of the state of the world.

On the equilibrium path:

Citizens have no incentive to shirk, as, per S^*_g and A3, this will simply result in a welfare decreasing punishment, for any state of the world.

The government can credibly punish when necessary since it understands the consequences dictated by strategies, no effort exerted ever again, of forgoing punishment.

Off the equilibrium path:

Citizens can credibly shirk since they cannot escape punishment by **individually** deviating to effort exertion.

Finally, the government must abandon any punishment strategy since punishment is now costly and has no beneficial consequences looking forward. □

2.3.3 Election

I now look at expected payoffs when an ideologue is in office. If the true state of the world turns out to be G, then the equilibrium expected payoffs per period will be: $\overline{\lambda}2\beta + ((1 - \overline{\lambda}) - X) - 2c$. If the true state of the world is B, then the expected, per period payoff will be: $\underline{\lambda}2\beta + ((1 - \underline{\lambda}) - X) - 2c$.

From proposition 1 we gather that payoffs with a non-ideologue in office will be zero.

From A1, we know that payoffs, less the expected cost of punishment on the equilibrium path, are positive for the state G and negative for state B. The relative size of punishments on the equilibrium path with regard to the gains in efficiency will determine whether there are any citizen prior beliefs for which the ideologue will be elected.

We now formalize the electoral stage of the model:

A representative citizen, with prior beliefs ω_0, chooses the government from amongst the continuum of candidates according to the expected equilibrium outcome as to maximize expected citizen payoffs.

Proposition 2.3. *If the cost of intervention, X, is large enough, the ideologue will never be elected.*

Proof. If $X > X^* = \dfrac{\overline{\lambda}2\beta - 2c}{(1 - \overline{\lambda})} \Rightarrow \left[\overline{\lambda}2\beta + ((1 - \overline{\lambda}) - X) - 2c\right] < 0 \Rightarrow$

$\omega\left[\overline{\lambda}2\beta + ((1 - \overline{\lambda}) - X) - 2c\right] + (1 - \omega)\left[\underline{\lambda}2\beta + ((1 - \underline{\lambda}) - X) - 2c\right] < 0$

Therefore the citizen can always improve expected payoffs by choosing a non-ideological candidate. □

Proposition 2.4. *If the cost of intervention, X, is small enough,* $X \leq X^* = \dfrac{\overline{\lambda}2\beta - 2c}{(1 - \overline{\lambda})}$, *the ideologue will be elected if the odds ratio on the citizens beliefs is high enough:* $\Omega^* = \dfrac{\omega}{(1 - \omega)} > \dfrac{-\left[\underline{\lambda}2\beta + ((1 - \underline{\lambda}) - X) - 2c\right]}{\left[\overline{\lambda}2\beta + ((1 - \overline{\lambda}) - X) - 2c\right]}$

Proof. Immediately follows from the decision rule. □

2.4 Discussion and Conclusions

In summary, I have presented a model that illustrates how ideology can be modeled as differences in beliefs and how extreme ideology, in particular, can serve as a strong commitment device. This ability motivates the election of governments that hold such beliefs even when things may turn out for the worst and persist in that state. The basic result is that an apparently reasonable and desirable incentive scheme, is revealed to be unfeasible under slight extensions that are easily motivated by the context. Only what we call an ideologue, one who is subjectively certain that the scheme will work, can implement it.

The model has several features that warrant further discussion and research. Although this model presents citizens, candidates, policies, elections, and the implementation of policy it is far from a complete electoral or policy model. The lack of heterogeneity amongst citizens, and the simplified electoral process call for a more complete treatment. However, this simplified framework has served to underline the main arguments of the essay: 1) The frailty of commitment to a policy that might be flawed, 2) An ideologues ability to overcome this, and 3) The inertial consequences of electing such a government to exploit this unique ability.

Other related aspects that warrant further exploration are that beliefs can't normally be observed and that there are no political parties in the model. Although beliefs are not directly observable they can be signaled through the establishment of a platform, this is where parties come in, as only politicians who have established a platform, a set of beliefs to which they are publicly associated can give a credible signal about their future behavior they may undertake. Electors have incentives to keep away from unpredictable candidates. In this sense we may think of candidates as parties who have explicitly stated their belief about the true state of the world. Of course, strategic considerations may give incentives for parties to simulate their beliefs, but the fact that they are collectives, under public scrutiny, may diminish their ability to deceive. In the following chapter I explore some of the strategic considerations that allow us to associate a party to an outwardly rigid set of beliefs.

There are several features about the model that make it ripe for rich interpretation despite the strong simplifications. It shows a setup in which citizens reap benefits from cooperation but require government intervention to achieve this. It has governments that must implicitly respect a budget constraint and must therefore depend on punishments, not prizes, to give high powered incentives. This in turn, combined with a democratic (benevolent) government complicates the government's ability to credibly commit to such incentive schemes. Although we have an undesirable outcome without government intervention, action by government is not warranted if the tools available are too blunt. Citizens must elect a government according to how they think it will react to events under changing informational environments; they find it in their interest to abide by laws, out of fear of punishment, irrespective of their assessment of such laws. If the ideological extremist takes power we have persistent differences of opinion that lead us to institutional inertia.

At this point, a comment on the literature on time inconsistency, rules and discretion, and strategic delegation in monetary theory is in order. Rogoff's (1985) paper seems to be the closest in spirit. He argues that because of temporal inconsistencies in monetary policy society might be tempted to delegate authority on a central banker that is more "conservative", places a higher weight than the social loss function, on unemployment. However, he also argues that this excess weight must not be infinite, the central banker must not be "too conservative". His model shows that that in a world where monetary surprises may increase welfare (possibly because a labour market imperfection), agents will tend to have higher than optimal expectations in equilibrium to prevent the monetary authority from attempting such surprises. The authrity in turn will accomodate such expectations. A conservative central banker, can credibly be less accomodating of these high expectations by economic agents and will in turn make agents moderate their expectations.

In his model the central banker also faces identifiable shocks to demand that he must accomodate. If society appoints a central banker who is "too conservative" there will be a huge gain with respect to moderating inflation expectations but a loss with respect to his ability to accomodate the exogenous shocks optimally, as measured by the social loss function. This leads to an interior solution on the level of commitment to fighting inflation,

as expresse by the weight the central banker places on unemployment.

The key difference with my model is that our authority cannot always distinguish between a bad outcome that must be punished and one that must be accomodated. In terms of monetary policy I interpret this as a plausible scenario where it is not always clear that an exogenous shock has occured, only that inflationary preassures have increased for some reason. We can interpret our model as one where economic agents cooperate by moderating their inflation expectations and high expectations cannot be distinguished from exogenous demand shocks. The alternate state of the world in my model could then be interpreted as one where exogenous demand shocks are more common. Under this interpretation Rogoff's "not too conservative" monetary authority is as useless as the most reckless inflation dove. Rogoff (1985, pg. 1187) states that "it can be misleading to analyze separately the stabilization and credibility problems of the central bank". I agree and believe a scenario when it is hard to identify, at least some of the time, whether a bad outcome is due to bad luck or bad strategic interaction is a salient case in monetary policy.

I should mention that he models ideology as preferences as opposed to beliefs. In a monetary policy scenario, where only the extreme conservative can implement the best available policy, ideology as beliefs implies that the more successful central banker is one who is very hard headed, subjectively certainty that the occurence of exogenous demand shocks is so rare that it is best never to accomodate them. This interpretation seems more plausible, at least from casual observance of public statements by central bankers, that a central banker who places absolutely zero weight on unemployment at any level.

2.5 Bibliography

1. Alesina R., N. Roubini, and G. Cohen. (1997). *Political Cycles and the Macroeconomy.* MIT Press.

2. Besley T. and S. Coate. (1998). "Sources of Inefficiency in a Representative Democracy: A Dynamic Analysis." *American Economic Review* 88: 139-156.

3. Coate S. and S. Morris. (1999). "Policy Persistence." *American Economic Review* 89: 1327-1336.

4. Cukierman, A.and M. Tommasi. (1998). "When Does It Take Nixon to Go to China." *American Economic Review* 88:180-97.

5. Dewatripont, M and J. Tirole. (1999). "Advocates." *Journal of Political Economy* 107: 1-39.

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

7. Feddersen T. and W. Pesendorfer. (1999). "Elections, Information Aggregation, and Strategic Voting." *Proc. Natl. Acad. Sci. USA* 96: 10572-10574.

8. Fernandez R. and D. Rodrik. (1991). "Resistance to Reform: Status Quo Bias in the Presence of Individual-Specific Uncertainty." *American Economic Review* 81: 1146-55.

9. Gerring, J. (1997). "Ideology: A Definitional Analysis." *Political Research Quarterly* Vol. 50, Issue 4, 957-994.

10. Green E. and R. Porter. (1984). "Noncooperative Collusion Under Imperfect Price Information" *Econometrica* 74: 87-100.

11. Hardin, R. (1999). *Liberalism, Constitutionalism and Democracy.* Oxford University Press.

12. Holmstrom, B. (1982). "Moral Hazard in Teams." *Bell Journal of Economics* 10: 74-91.

13. Mailath, G and S. Morris (2000). "Repeated Games with Almost-Public Monitoring." *Journal of Economic Theory* 102: 189-228.

14. Morris, S. (1996). "Speculative Investor Behaviour and Learning" *Quarterly Journal of Economics* 11: 1111-1133.

15. Mujumdar, S. and Mukand S. (2002). "Policy Gambles". Working Paper.

16. Osborne, M.J. (1995)."Spatial Models of Political Competition Under Plurality Rule: A Survey of Some Explanations of the Number of Candidates and the Positions They Take." *Canadian Journal of Economics* 28: 261-301.

17. Perry, M. and P. Reny (1994). "A Noncooperative View of Coalition Formation and the Core." *Econometrica* 62: 795-817.

18. Persson T., and Tabellini G. (2000). *Political Economics. Explaining Economic Policy.* MIT Press.

19. Roemer, J.E. (2001). *Political Competition. Theory and Applications.* Harvard University Press.

20. Rogoff, K.(1985)."The Optimal Degree of Commitment to an Intermediate Money Target." *Quarterly Journal of Economics* 100: 1169-90.

21. Wittman, D. (1973). "Parties as Utility Maximizers" *American Political Science Review* 67: 490-98.

CHAPTER 3

Ideology: True Beliefs, Dogma or Both?

3.1 Introduction

One of the most interesting aspects about ideology is that, even if people concede that it is important to political decision making, there is little agreement about what it is or where it comes from. Gerrig (1997) has documented a large number of different sometimes contradictory definitions of ideology used in the social sciences[1].

I remark the existence of two seemingly contradictory definitions: hard normative convictions and mere rhetoric. I propose that party preferences and policy uncertainty combine

[1]Gerrig says that: "To some, ideology is dogmatic, while to others it carries connotations of political sophistication[...]not only is ideology far-flung, it also encompasses a good many definitional traits which are directly at odds with one another." He cites McClosky's (1964:362) definition: "Systems of belief that are elaborate, integrated, and coherent, that justify the exercise of power, explain and judge historical events, identify political right and wrong, set forth the interconnections (causal and moral) between politics and other spheres of activity". He also gives us Sartori's (1969: 402) very different definition of ideology: "A typically dogmatic, i.e., rigid and impermeable, approach to politics"

Gerrig cites and discussed many other definitions but I bring attention to these two as they are very close to the two conceptions I wish to address.

to explain this apparent contradiction.

Wittman (1973) proposes that politicians' preferences are important, and that parties make a trade-off by abandoning their preferred policy in favor of a more moderate proposal in order to increase the probability of getting elected. If parties are informed about the true state of the world, how policies will play out, and have potential conflicts of preference with the electorate, moving their binding platforms towards the middle ground might be counterproductive; it may give a conceding signal to the electorate that actually reduces the probability that his "moderate" platform will be elected. Platforms that are uninformative and ideological, in the dogmatic sense, may be the only equilibrium in some simple electoral games.

I present a very simple electoral model to explore the main features of the potentially informative interaction that takes place, through party platforms, between parties and the electorate. The main result and interpretation of this model rely on the use of pure strategies.

3.2 The Model

Two parties vie for the vote of a single "median" or decisive voter. There is uncertainty about the true state of the world. For example, the magnitude of labour supply elasticity may be unknown to the electorate at large, and to the decisive voter in particular. At the start of play, the state of the world is realized, elasticity may be either high or low. The state is revealed to both parties. Parties then present a binding platform. The voter then makes inference about the true state of the world and picks the party platform that he prefers.

3.2.1 Setup

Throughout the essay we will use a redistributive income tax with a uniform rate as an example. However, the model is intended to apply more generally: any unidimensional policy space with the assumptions over preferences imposed below will be encompassed.

Consider the following incomplete information, three period game:

$$G : \Theta \times A_L \times A_R \times A_v \rightarrow \Re^3$$

There are three players: a left party L, a right party R and a single pivotal voter, V. At the beginning of play nature decides on the true state of the world. With probability one half the elasticity of labour supply is either high or low. The voter's prior beliefs, μ^0, reflect this true distribution:

$$\begin{aligned} \Theta &: \theta \in \{\underline{\theta}, \overline{\theta}\} \\ 0 &< \underline{\theta} < \overline{\theta} \\ pr(\theta &= \underline{\theta}) = 1/2 = \mu^0 \end{aligned}$$

In period one parties choose a binding policy platform simultaneously. The policy (action) space consists of low, moderate and high taxes and is restricted in the following way:

$$\begin{aligned} a_L^\theta &\in \left\{t^m, t^h\right\} \\ a_R^\theta &\in \left\{t^l, t^m\right\} \\ t^l &< t^m < t^h \end{aligned}$$

In period two, V chooses to vote for one of the parties:

$$a_V \in \{0, 1\} \text{ where } a_V \text{ is equal to the probability of voting for } L$$

In period three the policy dictated by the platform of the elected party is implemented. Payoffs have the following characteristics:

$$
\begin{aligned}
u_v(t;\overline{\theta}) &= u_R(t;\overline{\theta}) \\
u_v(t;\underline{\theta}) &= u_L(t;\underline{\theta}) \\
u_R(t^l;\overline{\theta}) &> u_R(t^m;\overline{\theta}) > u_R(t^h;\overline{\theta}) \\
u_L(t^h;\underline{\theta}) &> u_L(t^m;\underline{\theta}) > u_L(t^l;\underline{\theta}) \\
u_R(t^m;\underline{\theta}) &> u_R(t^l;\underline{\theta}) > u_R(t^h;\underline{\theta}) \\
u_L(t^m;\overline{\theta}) &> u_L(t^h;\overline{\theta}) > u_L(t^l;\overline{\theta}) \\
u_L\left(t^l,\underline{\theta}\right) &= u_L\left(t^l,\overline{\theta}\right) = u_R\left(t^h,\underline{\theta}\right) = u_R\left(t^h,\overline{\theta}\right) = 0
\end{aligned}
$$

The above assumptions on payoffs can be interpreted as follows. In each state the voter's preferences agree with those of one, and only one, of the parties. If the elasticity of labour supply is high, an assessment often associated with conservative political positions, then the voter prefers low taxes, moderate taxes and high taxes, in that order. In this, "conservative" state of the world, the voter's preference ordering is the same as the right-wing party's. Similarly, if the labour supply elasticity is low, then the voter prefers high income taxes, moderate income taxes, and low income taxes, in that order. The left-wing party's preference ordering is the same as the voter's in this state of the world. When the world is "conservative", elasticity is high, the left-wing party prefer moderate taxes, high taxes and low taxes in that order. Conversely, if elasticity is low, the right-wing party will prefer moderate taxes, low taxes, and high taxes in that order. Parties concede that moderate taxes are best if the state of the world is not the one they are associated with but their least preferred policy is always the same. In this sense parties are defined by the policies they would never espouse.

These assumptions capture in a very simple manner both policy uncertainty as well as party identity uncertainty. The world is as described by conservative theorists and the right-wing party represents my interests or the world works as described by progressive theorists and the left wing party represents my interests.

One way to interpret the assumptions is that the pivotal voter would have a clear preference for one of the two more radical policies if he where informed about the true state

of the world. However, the uninformed voter who places an equal prior probability on the two states might be indifferent between them and indeed prefer the moderate platform to the other two, plausibly out of risk aversion considerations.

A strategy for parties is a mapping from the state of the world into a policy platform:

$$\Sigma_L, \Sigma_R : \Theta \rightarrow T$$

A strategy for the voter is a mapping from the observed policy platforms into the probability of voting for the left party:

$$\sigma_v : T \times T \rightarrow [0,1]$$

The voter will update his priors on the state of the world, μ, using Bayes' rule, where possible:

$$\mu : T \times T \rightarrow [0,1]$$

where $\mu(t_L, t_R)$ is the updated probability that $\theta = \underline{\theta}$

A Perfect Bayesian Nash Equilibria is a triple $(\Sigma^*, \sigma^*, \mu^*)$ s.t.:

1. $E(u_i(\theta; \sigma_v^*(\mu(\Sigma^*)); \Sigma_i)) > E(u_i(\theta; \sigma_v^*(\mu(\Sigma^*)); \Sigma_i')) \qquad \forall i, \theta, \Sigma_i'$

2. $E(u_v(\sigma_v^*(\mu(\Sigma^*, T \times T)))) > E(u_v(\sigma_v'(\mu(\Sigma^*, T \times T)))) \qquad \forall \sigma_v', T \times T$

3. $\mu(\Sigma^*, T \times T) \in [0,1]$

4. $\mu(\Sigma^*, T \times T) = \Pr(T \times T \mid \Sigma^*, \underline{\theta}) / \sum_\theta \Pr(T \times T \mid \Sigma^*, \theta)$

$if \sum_\theta \Pr(T \times T \mid \Sigma^*, \theta) > 0$

3.2.2 Pure Strategy Results

I first analyze pure strategies[2]. There are 16 combinations of pure strategies that the parties may play. They are summarized as follows:

Group 1. No PBNE Profiles

	Left Wing		Right Wing	
	$\underline{\theta}$	$\overline{\theta}$	$\underline{\theta}$	$\overline{\theta}$
1.	t^h	t^m	t^m	t^l
2.	t^h	t^m	t^l	t^l
3.	t^h	t^h	t^m	t^l
4.	t^h	t^h	t^l	t^m
5.	t^m	t^h	t^l	t^l

Group 2. Informative Equilibrium Profiles

	Left Wing		Right Wing	
	$\underline{\theta}$	$\overline{\theta}$	$\underline{\theta}$	$\overline{\theta}$
6.	t^m	t^h	t^l	t^m
7.	t^h	t^m	t^m	t^m
8.	t^h	t^m	t^l	t^m
9.	t^m	t^m	t^l	t^m
10.	t^m	t^m	t^m	t^l
11.	t^m	t^h	t^m	t^l
12.	t^m	t^h	t^m	t^m

Group 3. Uninformative Equilibrium Profiles

[2]Notice that the way pure strategies are defined for the voter automatically encompasses mixed strategies for him. A salient case: pure strategies for the parties but possibly mixed strategies for the voter is included in this manner. In general, I think the relevant case for this type of electoral model is one with pure strategies for the parties. A partial treatment of mixed strategies for the parties is included in the next subsection.

	Left Wing		Right Wing	
	$\underline{\theta}$	$\overline{\theta}$	$\underline{\theta}$	$\overline{\theta}$
13.	t^m	t^m	t^m	t^m
14.	t^m	t^m	t^l	t^l
15.	t^h	t^h	t^m	t^m
16.	t^h	t^h	t^l	t^l

The 16 combinations can be grouped in three sets. The first two groups are composed of informative or separating profiles. The voter is informed about the true state of the world and chooses the party accordingly. The third group of profiles is non-informative, the state of the world is not pinned down by the information available in equilibrium, before the voter decides. Neither party differentiates its platform according to the state of the world and therefore party platforms are not informative.

The first group of profiles, 1 through 5, can never be part of an equilibrium, there is no pair of voter beliefs and strategies that sustain such profiles as equilibria.

The second group of informative profiles may form part of a PBNE, I call profiles 7-12 the "dominant party" profiles because, in equilibrium, one of the parties always has their most favored policy, out of those available to the voter, picked. Profiles 7 through 9 may form part of left-wing party dominant equilibria and profiles 10 through 12 may constitute right-wing dominant equilibria. Profile 6 is a special case, both parties propose policies that are inverted with respect to their preferences and neither party "dominates" in the sense outlined above. The third group, all uninformative, may also be part of PBNE.

Proposition 3.1. *Profiles 1-5 may never form part of PBNE.*

Proof. See appendix A. □

The basic intuition for this group is that if both parties' platforms are needed to identify the state, as in profile 1, then the information transmission process requires punishment threats, off the equilibrium path, for both parties simultaneously, an impossibility. If the voter only depends on one party to identify the state then we have a typical moral hazard

problem with hidden information, that party will have no incentive to reveal the true state when the voter's most preferred policy is the one least preferred by the informing party.

For group two we have seven strategy profiles that inform the voter what the state of the world is; at least one of the parties varies his platform according to the state of the world. Since parties use pure strategies, beliefs are not pinned down by Bayes's rule off the equilibrium path. This flexibility is what allows the voter to have punishment strategies off the equilibrium path that support these equilibria.

Profiles 7-12 can form part of a *dominant party* equilibrium. In them, the voter will always choose the one most favored by one of the parties, out of those available, irrespective of the state of the world. This happens on and off the equilibrium path.

The reason why these profiles may form part of a PBNE is that the excluded party is "cooperating" with the equilibria by offering a policy that is no better for the citizen than the one offered by the dominant party. The excluded party has no incentive to deviate from this state since he is always ignored. The flexibility we have with respect to off-the equilibrium path beliefs is what allows the voter to ignore the excluded party off (as well as on) the equilibrium path to enforce the equilibrium. These beliefs may be unreasonable, as will be shown below.

Proposition 3.2. *Profiles 6-12 may form part of a PBNE.*

Proof. See appendix B. □

Group three is made up of uninformative profiles; both parties offer the same platform for any state of the world. The voter makes an uninformed choice after observing the party platforms. There are two profiles that seem more interesting. One, number 16, I call the Downsian Profile. In it both parties converge to the moderate platform and the voter flips a coin to choose the party. We get policy convergence despite the fact that for every state of the world at least one informed party has a strict preference for a different policy that agrees with the electorate. We also have what I call a Dogmatic Equilibrium, profile 13, where parties always offer the policy that they most prefer in the state of the world where their preference ordering coincides with that of the voter.

Proposition 3.3. *Profiles 13-16 may form part of a PBNE.*

Proof. See appendix C. □

From the above, we gather that there are many different equilibria with very different political interpretations. However, most of them depend on off-the-equilibrium-path actions by the voter that might not seem credible, that is, they are supported by unreasonable beliefs. I now look at equilibrium refinements that will allow me to separate out the less reasonable equilibria.

I use a slight variation of the Intuitive Criterion[3] to discard equilibria that are supported by unreasonable beliefs.

Definition 3.1. *Cautious Intuitive Criterion. Let* $BR(a_i, a_{-i}, \mu)$ *be the set of all best responses, possibly mixed, for player* v*, given actions* a_i, a_{-i} *and belief set* $\mu(a_i, a_{-i})$*. Let* $p_i(a_i, a_{-i}, a_v, \theta)$ *be the expected payoffs for player* i*, given an action profile and a state. For an equilibrium profile* $(\Sigma^*, \sigma_v^*, \mu^*)$*, fix a vector of equilibrium expected payoffs* $p_i^*(\cdot)$*, for one of the parties,* i*. For each* a_i *let* $J(a_i, a_{-i})$ *be the set of all* θ *for a given* a_{-i} *s.t.*

$p_i^*(\cdot) \geqslant \max_{a_v \in BR(a_i, a_{-i}, \mu)} p_i(a_i, a_{-i}, a_v, \theta)$

$\exists \mu' \in \mu(a_i, a_{-i}) s.t.$

$p_i^*(\cdot) > \max_{a_v \in BR(a_i, a_{-i}, \mu',)} p_i(a_i, a_{-i}, a_v, \theta)$

If for some a_i $\exists$ *a* θ *s.t.*

$p_i^*(\cdot) < \min_{a_v \in BR(a_i, a_{-i}, \Theta \setminus J(a_i, a_{-i}))} p_i(a_i, a_{-i}, a_v, \theta)$

Then the equilibrium fails the Cautious Intuitive Criterion.

Proposition 3.4. *Only the Dogmatic Equilibria survives the Cautious Intuitive Criterion.*

Proof. See appendix D. □

[3] Adapted from Fudenberg and Tirole's (1998) presentation of the I.C. My definition differs from theirs mostly in that the signal receiver must give zero weight to all types who's utility is non increasing for a given deviation and decreases strictly for at least one set of beliefs. Under their definition of the IC a zero weight is given only to types that have strictly decreasing utility from the deviation.

3.2.3 Mixed Strategy Results

Mixed strategies for the parties are problematic in the context of this model, that parties purposefully randomize on their platform choices seems unlikely. If mixed strategies are motivated by an involuntary "tremble" or error in perception, which seems more reasonable, we must look for results in full-support mixed strategies.

A full support mixed strategy profile for both parties also has the desirable feature that beliefs are pinned down for every possible outcome. This implies that there are no arbitrary off-the-equilibrium path beliefs that can be used to support non-credible equilibria. There is a surprising result.

Proposition 3.5. *There are no full support mixed strategy (for the parties) PBNE .*

Theorem 3.1. *Proof.* See appendix E. □

The CIC, a key element in our analysis, is harder to motivate once we allow for mixed strategies. The threat of a deviation from an equilibrium strategy when strategies are possibly mixed is hard to interpret beyond its formal statement.

I do not have a full characterization of equilibria in non-full support mixed strategies. However, it is the case that such equilibria do exist. I present an example that shows that there are many non-full support mixed strategy PBNE that allow parties to deviate from their radical platforms and do not violate the CIC.

A slight variation on the unique equilibria for pure strategies will illustrate this point. This is the case where the right wing party plays the profile from our main result, always supporting low taxes regardless of the state. In this case, if the voter always chooses the right wing party when faced with the radical profiles, the left wing party may mix (only a little) when elasticity is high.

This profile (and its symmetric counterpart) is very similar to the main result. The left party and the right party almost always support high and low taxes respectively. The voter always chooses the right wing party when faced with the radical profiles (which is allowed for in the pure strategy equilibrium profile) and chooses the radical profiles if one party deviates to the moderate policy. The left wing party, when elasticity is high, sometimes

offers the moderate platform. As the left wing party is marginalized, its actions become irrelevant and this allows it to mix.

However, there are limits to these deviations, if they occurred with a high probability then a high tax platform would become too informative and would prevent the strategy played by the voter from being optimal.

3.3 Discussion and Conclusions

In summary, I present a model where two informed parties offer binding platforms on a unidimensional space. There is uncertainty about the mapping of policy into outcomes and about the preferences that drive the behavior of the parties. A pivotal voter observes the platforms, makes inference about the state of the world and decides the outcome of the election accordingly. The world is either one where the preference ordering of the pivotal voter coincides with that of the left-wing party or it coincides with that of the right-wing party. Platforms that are consistent with the intuitive criterion have each party keeping its distance form the moderate policy in both states of the world; the left-wing party always proposes high taxes and the right-wing party always proposes low taxes. These conclusions rely on restricting the parties to pure strategies.

On the equilibrium path the voter does not receive any new information about the state of the world; his beliefs are not updated. On the equilibrium path he may mix over the parties. He votes for the extreme platform if one party deviates to a moderate platform. The beliefs that support this strategy have him maintain his priors on-the-equilibrium-path, if he observes the "dogmatic" platform profile and to believe the state of the world is "conservative" if the left-wing party deviates to a moderate platform and vice versa. He may have any beliefs if he observes a uniformly moderate policy profile.

I draw two main conclusions from the outcome outlined above. Parties will not "converge to the middle" when there is policy uncertainty and parties are informed because one of the parties will be able to univocally signal that the voter's and that party's most preferred policy is not the moderate policy by deviating to a more radical policy.

Secondly, I take the above results to explain why parties become identified with a belief system or ideology. A party will always present platforms that are consistent with beliefs about the state of the world where the party's preference ordering is consistent with those of the pivotal members of the electorate, even when the parties are informed that the state of the world is different. Seen from the outside, a party will become identified with a belief system, an ideology, and a corresponding policy. For an informed voter that shares common knowledge with the parties about the state of the world the parties will appear as dogmatic, doling out the same policy even when it is obviously inadequate. For the uninformed voter parties are true believers that are identified with a particular belief system and the consequent policy proposal. They must choose between the parties using their priors and exogenous information beyond platforms. Campaigns, in general, will not be very informative due to strategic considerations.

The next step in this research agenda will involve relaxing restrictions on the policy action space available to parties as coarseness plays an essential role in our results. A first step could be to introduce a richer, possibly continuous, policy space. Eventually, a dynamic model is desirable. This would allow us to better integrate the conclusions obtained in this model to those presented in chapters 2 and 3. In a dynamic context, with reputations to uphold and more information being revealed from policy implementation, parties might have stronger incentives to reveal the true state of the world.

Finally, an implication of the model with both strategic and normative consequences is that under a two party regime, irrespective of welfare considerations, parties have incentives to lie, or at least not reveal all the information available to them. This might not be an entirely hopeless situation. As in Dewatripont and Tirole (1999) parties, due to strategic considerations, might not be the best agents for complete information revelation. Parties, on the other hand, will have incentives to elaborate the best case for their point of view and work out the details of a program that is consistent. Meaningful information may also be delivered to the electorate by other sources and the dogmatic profile has the advantage of giving a real choice to the electorate.

3.4 Appendix A

Proof. Suppose Profile 1 is part of a PBNE:

$(a_L = t^h, a_R = t^m) \longrightarrow \mu = 1 \longrightarrow \sigma_v^*(t^h, t^m) = 1 \longrightarrow \sigma_v^*(t^h, t^l) = 1$

and

$(a_L = t^m, a_R = t^l) \longrightarrow \mu = 0 \longrightarrow \sigma_v^*(t^m, t^l) = 0 \longrightarrow \sigma_v^*(t^h, t^l) = 0,$

a contradiction.

Suppose Profile 2 is part of a PBNE:

$(a_L = t^h, a_R = t^l) \longrightarrow \mu = 1 \longrightarrow \sigma_v^*(t^h, t^l) = 1$

and

$(a_L = t^m, a_R = t^l) \longrightarrow \mu = 0 \longrightarrow \sigma_v^*(t^m, t^l) = 0 \longrightarrow \sigma_v^*(t^h, t^l) = 0$

a contradiction.

Profile 3, by symmetry to profile 2, cannot form part of a PBNE.

Suppose Profile 4 is part of a PBNE:

$(a_L = t^h, a_R = t^l) \longrightarrow \mu = 1 \longrightarrow \sigma_v^*(t^h, t^l) = 1 \longrightarrow \sigma_v^*(t^h, t^m) = 1$

and

$(a_L = t^h, a_R = t^m) \longrightarrow \mu = 0 \longrightarrow \sigma_v^*(t^h, t^m) = 0$

a contradiction.

Profile 5, by symmetry to profile 4, cannot form part of a PBNE. □

3.5 Appendix B

Proof. Profile 6. $(\Sigma^*, \sigma_v^*, \mu^*)$ is a PBNE where:

$\Sigma_L : a_L(\underline{\theta}) = t^m, a_L(\overline{\theta}) = t^h,$

$\Sigma_R : a_R(\underline{\theta}) = t^l, a_R(\overline{\theta}) = t^m$

$\mu(t^m, t^l) = 1, \mu(t^h, t^m) = 0, \mu(t^h, t^l) = 1/2, \mu(t^m, t^m) = 1/2$

$\sigma_v(t^m, t^l) = 1, \sigma_v(t^h, t^m) = 0,$

$\sigma_v(t^h, t^l) = 1/2, \sigma_v(t^m, t^m) = 1/2$

iff

$\frac{1}{2}u_L(t^h; \underline{\theta}) + \frac{1}{2}u_L(t^l; \underline{\theta}) \leqslant u_L(t^m; \underline{\theta})$

$\frac{1}{2}u_R(t^h;\overline{\theta}) + \frac{1}{2}u_R(t^l;\overline{\theta}) \leqslant u_R(t^m;\overline{\theta})$

$\frac{1}{2}u_v(t^h;\overline{\theta}) + \frac{1}{2}u_v(t^h;\underline{\theta}) = \frac{1}{2}u_v(t^l;\overline{\theta}) + \frac{1}{2}u_v(t^l;\underline{\theta})$

follows directly from the definition of a PBNE.

Profile 7. $(\Sigma^*, \sigma_v^*, \mu^*)$ is a PBNE where:

$\Sigma_L : a_L(\underline{\theta}) = t^h, a_L(\overline{\theta}) = t^m,$

$\Sigma_R : a_R(\underline{\theta}) = t^m, a_R(\overline{\theta}) = t^m$

$\mu(t^h, t^m) = 1, \mu(t^m, t^m) = 0, \mu(t^h, t^l) = 1/2, \mu(t^m, t^m) = 1/2$

$\sigma_v(t^h, t^m) = 1, \sigma_v(t^m, t^m) = 0,$

$\sigma_v(t^h, t^l) = 1/2, \sigma_v(t^m, t^m) = 1/2$

iff

$\frac{1}{2}u_L(t^h;\underline{\theta}) + \frac{1}{2}u_L(t^l;\underline{\theta}) \leqslant u_L(t^m;\underline{\theta})$

$\frac{1}{2}u_R(t^h;\overline{\theta}) + \frac{1}{2}u_R(t^l;\overline{\theta}) \leqslant u_R(t^m;\overline{\theta})$

$\frac{1}{2}u_v(t^h;\overline{\theta}) + \frac{1}{2}u_v(t^l;\overline{\theta}) = u_v(t^h;\underline{\theta}) + \frac{1}{2}u_v(t^l;\underline{\theta})$

follows directly from the definition of a PBNE.

Profile 8. $(\Sigma^*, \sigma^*, \mu^*)$ is a PBNE where:

$\Sigma_L : a_L(\underline{\theta}) = t^h, a_L(\overline{\theta}) = t^m,$

$\Sigma_R : a_R(\underline{\theta}) = t^l, a_R(\overline{\theta}) = t^m$

$\mu(t^h, t^l) = 1, \mu(t^m, t^m) = 0, \mu(t^m, t^l) = 1/2, \mu(t^h, t^m) = 1/2$

$\sigma_v(t^h, t^l) = 1, \sigma_v(t^m, t^m) = 0,$

$\sigma_v(t^m, t^l) = 1, \sigma_v(t^h, t^m) = 0$

iff

$\frac{1}{2}u_v(t^m;\overline{\theta}) + \frac{1}{2}u_v(t^m;\underline{\theta}) \geqslant \frac{1}{2}u_v(t^l;\overline{\theta}) + \frac{1}{2}u_v(t^l;\underline{\theta})$

$\frac{1}{2}u_v(t^m;\overline{\theta}) + \frac{1}{2}u_v(t^m;\underline{\theta}) \geqslant \frac{1}{2}u_v(t^h;\overline{\theta}) + \frac{1}{2}u_v(t^h;\underline{\theta})$

follows directly from the definition of a PBNE.

Profile 9. $(\Sigma^*, \sigma_v^*, \mu^*)$ is a PBNE where:

$\Sigma_L : a_L(\underline{\theta}) = t^m, a_L(\overline{\theta}) = t^m,$

$\Sigma_R : a_R(\underline{\theta}) = t^l, a_R(\overline{\theta}) = t^m$

$\mu(t^m, t^l) = 1, \mu(t^m, t^m) = 0, \mu(t^h, t^l) = 1/2, \mu(t^m, t^l) = 1/2$

$\sigma_v(t^m, t^l) = 1, \sigma_v(t^m, t^m) = 0,$

$\sigma_v(t^h, t^l) = 0, \sigma_v(t^h, t^m) = 0$

iff

$\frac{1}{2}u_v(t^l;\overline{\theta}) + \frac{1}{2}u_v(t^l;\underline{\theta}) \geqslant \frac{1}{2}u_v(t^h;\overline{\theta}) + \frac{1}{2}u_v(t^h;\underline{\theta})$

$\frac{1}{2}u_v(t^m;\overline{\theta}) + \frac{1}{2}u_v(t^m;\underline{\theta}) \geqslant \frac{1}{2}u_v(t^l;\overline{\theta}) + \frac{1}{2}u_v(t^l;\underline{\theta})$

follows directly from the definition of a PBNE.

Profiles 10, 11 and 12, by symmetry, to Profiles 7, 8 and 9 respectively, may also constitute PBNE. □

3.6 Appendix C

Proof. Profile 13. $(\Sigma^*, \sigma_v^*, \mu^*)$ is a PBNE where:

$\Sigma_L : a_L(\underline{\theta}) = t^m, a_L(\overline{\theta}) = t^m,$

$\Sigma_R : a_R(\underline{\theta}) = t^m, a_R(\overline{\theta}) = t^m$

$\mu(t^m, t^l) = 1/2, \mu(t^h, t^m) = 1/2, \mu(t^h, t^l) = 1/2, \mu(t^m, t^m) = 1/2$

$\sigma_v(t^m, t^l) = 1, \sigma_v(t^h, t^m) = 0,$

$\sigma_v(t^h, t^l) = 1/2, \sigma_v(t^m, t^m) = 1/2$

iff

$\frac{1}{2}u_v(t^m, \underline{\theta}) + \frac{1}{2}u_v(t^m, \overline{\theta}) \geqslant \frac{1}{2}u_v(t^h, \overline{\theta}) + \frac{1}{2}u_v(t^h, \underline{\theta})$

$\frac{1}{2}u_v(t^m, \underline{\theta}) + \frac{1}{2}u_v(t^m, \overline{\theta}) \geqslant \frac{1}{2}u_v(t^l, \overline{\theta}) + \frac{1}{2}u_v(t^l, \underline{\theta})$

$\frac{1}{2}u_v(t^h, \overline{\theta}) + \frac{1}{2}u_v(t^h, \underline{\theta}) = \frac{1}{2}u_v(t^l, \overline{\theta}) + \frac{1}{2}u_v(t^l, \underline{\theta})$

follows directly from the definition of a PBNE.

Profile 14. $(\Sigma^*, \sigma_v^*, \mu^*)$ is a PBNE where:

$\Sigma_L : a_L(\underline{\theta}) = t^m, a_L(\overline{\theta}) = t^m,$

$\Sigma_R : a_R(\underline{\theta}) = t^l, a_R(\overline{\theta}) = t^l$

$\mu(t^m, t^l) = 1/2, \mu(t^h, t^m) = 1/2, \mu(t^h, t^l) = 1/2, \mu(t^m, t^m) = 1/2$

$\sigma_v(t^m, t^l) = 1, \sigma_v(t^h, t^m) = 0,$

$\sigma_v(t^h, t^l) = 0, \sigma_v(t^m, t^m) = 1/2$

iff

$\frac{1}{2}u_v(t^l, \underline{\theta}) + \frac{1}{2}u_v(t^l, \overline{\theta}) \geqslant \frac{1}{2}u_v(t^h, \overline{\theta}) + \frac{1}{2}u_v(t^h, \underline{\theta})$

$\frac{1}{2}u_v(t^m,\underline{\theta}) + \frac{1}{2}u_v(t^m,\overline{\theta}) \geqslant \frac{1}{2}u_v(t^l,\overline{\theta}) + \frac{1}{2}u_v(t^l,\underline{\theta})$

follows directly from the definition of a PBNE.

Profile 15, by symmetry from profile 14.

Profile 16. $(\Sigma^*, \sigma_v^*, \mu^*)$ is a PBNE where:

$\Sigma_L : a_L(\underline{\theta}) = t^h, a_L(\overline{\theta}) = t^h,$

$\Sigma_R : a_R(\underline{\theta}) = t^l, a_R(\overline{\theta}) = t^l$

$\mu(t^m,t^l) = 1/2, \mu(t^h,t^m) = 1/2, \mu(t^h,t^l) = 1/2, \mu(t^m,t^m) = 1/2$

$\sigma_v(t^m,t^l) = 0, \sigma_v(t^h,t^m) = 1,$

$\sigma_v(t^h,t^l) = 1/2, \sigma_v(t^m,t^m) = 1/2$

iff

$\frac{1}{2}u_v(t^h;\overline{\theta}) + \frac{1}{2}u_v(t^h;\underline{\theta}) \geqslant \frac{1}{2}u_v(t^m;\overline{\theta}) + \frac{1}{2}u_v(t^m;\underline{\theta})$

$\frac{1}{2}u_v(t^l;\overline{\theta}) + \frac{1}{2}u_v(t^l;\underline{\theta}) \geqslant \frac{1}{2}u_v(t^m;\overline{\theta}) + \frac{1}{2}u_v(t^m;\underline{\theta})$

follows directly from the definition of a PBNE. □

3.7 Appendix D

Proof. Suppose Profile 6 is part of a PBNE $(\Sigma^*, \sigma_v^*, \mu^*)$:

$\Sigma_L : a_L(\underline{\theta}) = t^m, a_L(\overline{\theta}) = t^h,$

$\Sigma_R : a_R(\underline{\theta}) = t^l, a_R(\overline{\theta}) = t^m$

$\mu(t^m,t^l) = 1, \mu(t^h,t^m) = 0$

$\sigma_v(t^m,t^l) = 1, \sigma_v(t^h,t^m) = 0,$

but

$p_L^*(t^m,t^l,\sigma_v(t^m,t^l),\overline{\theta}) > p_L(t^h,t^l,\sigma_v,\overline{\theta})$, for any σ_v

$\overline{\theta} \in J(t^h,t^l) \longrightarrow BR(t^h,t^l,1) = 1$

$p_L^*(t^m,t^l,\sigma_v(t^m,t^l),\underline{\theta}) < p_L(t^h,t^l,1,\underline{\theta})$

which violates the C.I.C.

Suppose Profile 7. is part of a PBNE $(\Sigma^*, \sigma_v^*, \mu^*)$:

$\Sigma_L : a_L(\underline{\theta}) = t^h, a_L(\overline{\theta}) = t^m,$

$\Sigma_R : a_R(\underline{\theta}) = t^m, a_R(\overline{\theta}) = t^m$

$\mu(t^h, t^m) = 1, \mu(t^m, t^m) = 0$

$\sigma_v(t^h, t^m) = 1, \sigma_v(t^m, t^m) = 0,$

but

$p_R^*(t^m, t^m, \sigma_v(t^m, t^m), \underline{\theta}) \geqslant p_R(t^l, t^m, \sigma_v, \underline{\theta})$ for any σ_v

$p_R^*(t^m, t^m, \sigma_v(t^m, t^m), \underline{\theta}) > p_R(t^l, t^m, 0, \underline{\theta})$ where $(\sigma_v = 0) \in BR(t^l, t^m, 0)$

$\underline{\theta} \in J(t^l, t^m) \longrightarrow BR(t^l, t^m, 0) = 0$

$p_R^*(t^m, t^m, \sigma_v(t^m, t^m), \overline{\theta}) < p_R(t^l, t^m, 0, \overline{\theta})$

which violates the C.I.C.

Suppose Profile 8 is part of a PBNE $(\Sigma^*, \sigma_v^*, \mu^*)$:

$\Sigma_L : a_L(\underline{\theta}) = t^h, a_L(\overline{\theta}) = t^m,$

$\Sigma_R : a_R(\underline{\theta}) = t^l, a_R(\overline{\theta}) = t^m$

$\mu(t^h, t^l) = 1, \mu(t^m, t^m) = 0$

$\sigma_v(t^h, t^l) = 1, \sigma_v(t^m, t^m) = 0,$

but

$p_R^*(t^m, t^m, \sigma_v(t^m, t^m), \underline{\theta}) \geqslant p_R(t^l, t^m, \sigma_v, \underline{\theta})$ for any σ_v

$p_R^*(t^m, t^m, \sigma_v(t^m, t^m), \underline{\theta}) > p_R(t^l, t^m, 0, \underline{\theta})$ where $(\sigma_v = 0) \in BR(t^l, t^m, 0)$

$\underline{\theta} \in J(t^l, t^m) \longrightarrow BR(t^l, t^m, 0) = 0$

$p_R^*(t^m, t^m, \sigma_v(t^m, t^m), \overline{\theta}) < p_R(t^l, t^m, 0, \overline{\theta})$

which violates the C.I.C.

Suppose Profile 9 is part of a PBNE $(\Sigma^*, \sigma_v^*, \mu^*)$:

$\Sigma_L : a_L(\underline{\theta}) = t^m, a_L(\overline{\theta}) = t^m,$

$\Sigma_R : a_R(\underline{\theta}) = t^l, a_R(\overline{\theta}) = t^m$

$\mu(t^m, t^l) = 1, \mu(t^m, t^m) = 0$

$\sigma_v(t^m, t^l) = 1, \sigma_v(t^m, t^m) = 0,$

but

$p_L^*(t^m, t^l, \sigma_v(t^m, t^l), \overline{\theta}) > p_L(t^h, t^l, \sigma_v, \overline{\theta})$ for any σ_v

$\overline{\theta} \in J(t^h, t^l) \longrightarrow BR(t^h, t^l, 1) = 1$

$p_L^*(t^m, t^l, \sigma_v(t^m, t^l), \underline{\theta}) < p_L(t^h, t^l, 1, \underline{\theta})$

which violates the C.I.C.

Profiles 10,11 and 12, violate the C.I.C. by symmetry to profiles 7, 8 and 9 respectively.

Suppose Profile 13 is part of a PBNE $(\Sigma^*, \sigma^*, \mu^*)$:

$\Sigma_L : a_L(\underline{\theta}) = t^m, a_L(\overline{\theta}) = t^m,$

$\Sigma_R : a_R(\underline{\theta}) = t^m, a_R(\overline{\theta}) = t^m$

but

$p_L^*(t^m, t^m, \sigma_v(t^m, t^m), \overline{\theta}) \geqslant p_L(t^h, t^m, \sigma_v, \overline{\theta})$

$p_L^*(t^m, t^m, \sigma_v(t^m, t^m), \overline{\theta}) > p_L(t^h, t^m, 1, \overline{\theta})$ where $(\sigma_v = 1) \in BR(t^h, t^m, 1)$

$\overline{\theta} \in J(t^h, t^m) \longrightarrow BR(t^h, t^m, 1) = 1$

$p_L^*(t^m, t^m, \sigma_v(t^m, t^m), \underline{\theta}) < p_L(t^h, t^m, 1, \underline{\theta})$

which violates the C.I.C.

Suppose Profile 14 is part of a PBNE $(\Sigma^*, \sigma^*, \mu^*)$:

$\Sigma_L : a_L(\underline{\theta}) = t^m, a_L(\overline{\theta}) = t^m,$

$\Sigma_R : a_R(\underline{\theta}) = t^l, a_R(\overline{\theta}) = t^l$

$p_L^*(t^m, t^l, \sigma_v(t^m, t^l), \overline{\theta}) \geqslant p_L(t^h, t^l, \sigma_v, \overline{\theta})$ for any σ_v

$p_L^*(t^m, t^l, \sigma_v(t^m, t^l), \overline{\theta}) > p_L(t^h, t^l, 1, \overline{\theta})$ where $(\sigma_v = 1) \in BR(t^h, t^l, 1)$

$\overline{\theta} \in J(t^h, t^l) \longrightarrow BR(t^h, t^l, 1) = 1$

$p_L^*(t^m, t^l, \sigma_v(t^m, t^l), \underline{\theta}) < p_L(t^h, t^l, 1, \underline{\theta})$

which violates the C.I.C.

Profile 15, by symmetry to profile 14 violates CIC.

To verify that profile 16 can form part of an equilibrium we propose a candidate equilibrium profile and verify that the CIC holds for all deviations.

$(\Sigma^*, \sigma^*, \mu^*)$ is a PBNE where:

$\Sigma_L : a_L(\underline{\theta}) = t^h, a_L(\overline{\theta}) = t^h,$

$\Sigma_R : a_R(\underline{\theta}) = t^l, a_R(\overline{\theta}) = t^l$

$\mu(t^m, t^l) = 1/2, \mu(t^h, t^m) = 1/2, \mu(t^h, t^l) = 1/2, \mu(t^m, t^m) = 1/2$

$\sigma_v(t^m, t^l) = 0, \sigma_v(t^h, t^m) = 1,$

$\sigma_v(t^h, t^l) = 1/2, \sigma_v(t^m, t^m) = 1/2$

$\rceil(p_L^*(t^h, t^l, \sigma_v(t^h, t^l), \overline{\theta}) \geqslant p_L(t^m, t^l, \sigma_v, \overline{\theta}))$

$p_L^*(t^h, t^l, \sigma_v(t^h, t^l), \overline{\theta}) \geqslant p_L(t^m, t^l, \sigma_v, \underline{\theta})$ for any σ_v

iff $u_L(t^m, \underline{\theta}) < \frac{1}{2}u_L(t^h, \underline{\theta}) + \frac{1}{2}u_L(t^l, \underline{\theta})$

$p_L^*(t^h, t^l, \sigma_v(t^h, t^l), \overline{\theta}) > p_L(t^m, t^l, 0, \underline{\theta})$ where $(\sigma_v = 0) \in BR(t^m, t^l, 0)$

$\underline{\theta} \in J(t^m, t^l) \longrightarrow BR(t^m, t^l, 0) = 0$

but

$p_L^*(t^h, t^l, \sigma_v(t^h, t^l), \underline{\theta}) \geqslant p_L(t^m, t^l, 0, \underline{\theta})$

by symmetry CIC holds for deviations by player R. □

3.8 Appendix E

Proof. By contradiction. Let $\widetilde{\Sigma}_L \in \Delta(\{t^m, t^h\}; \theta)$, $\widetilde{\Sigma}_R \in \Delta(\{t^m, t^l\}; \theta)$, $\theta = \underline{\theta}, \overline{\theta}$ denote mixed strategies for the parties, where $\widetilde{\Sigma}_L = 1$ denotes $a_L = t^h$, $\widetilde{\Sigma}_R = 1$ denotes $a_R = t^l$ and

$$\begin{aligned} \alpha &= \widetilde{\Sigma}_L^*(\underline{\theta}) \\ \alpha' &= \widetilde{\Sigma}_L^*(\overline{\theta}) \\ \beta &= \widetilde{\Sigma}_R^*(\underline{\theta}) \\ \beta' &= \widetilde{\Sigma}_R^*(\overline{\theta}) \end{aligned}$$

Strategies for the voter are:

$$\begin{aligned} p_1 &= \sigma_v^*(t_L^h, t_R^l) \\ p_2 &= \sigma_v^*(t_L^m, t_R^l) \\ p_3 &= \sigma_v^*(t_L^h, t_R^m) \\ p_4 &= \sigma_v^*(t_L^m, t_R^m) \end{aligned}$$

Then, full-support mixed strategies for the parties imply α, α', β, $\beta' \in (0,1)$.

For a player to mix in a given state his expected utility given the other player's strategy and the voter's strategy must be equal for every action with non-zero weight. The four conditions, one for each player in each state are as follows:

$\beta(p_1(u_L(t^h, \underline{\theta})) + (1-p_1)(u_L(t^l, \underline{\theta}))) + (1-\beta)(p_3(u_L(t^h, \underline{\theta})) + (1-p_3)(u_L(t^m, \underline{\theta})))$

$= \beta(p_2(u_L(t^m, \underline{\theta})) + (1-p_2)(u_L(t^l, \underline{\theta}))) + (1-\beta)(p_4(u_L(t^m, \underline{\theta})) + (1-p_4)(u_L(t^m, \underline{\theta})))$

(3.1)

$$\beta'(p_1(u_L(t^h,\overline{\theta}))+(1-p_1)(u_L(t^l,\overline{\theta})))+(1-\beta')(p_3(u_L(t^h,\overline{\theta}))+(1-p_3)(u_L(t^m,\overline{\theta})))$$
$$=\beta'(p_2(u_L(t^m,\overline{\theta}))+(1-p_2)(u_L(t^l,\overline{\theta})))+(1-\beta')(p_4(u_L(t^m,\overline{\theta}))+(1-p_4)(u_L(t^m,\overline{\theta})))$$

(3.2)

$$\alpha(p_1(u_R(t^h,\underline{\theta}))+(1-p_1)(u_R(t^l,\underline{\theta})))+(1-\alpha)(p_3(u_R(t^m,\underline{\theta}))+(1-p_3)(u_R(t^l,\underline{\theta})))$$
$$=\alpha(p_2(u_R(t^h,\underline{\theta}))+(1-p_2)(u_R(t^m,\underline{\theta})))+(1-\alpha)(p_4(u_R(t^m,\underline{\theta}))+(1-p_4)(u_R(\underline{\theta},t^m)))$$

(3.3)

$$\alpha'(p_1(u_R(t^h,\overline{\theta}))+(1-p_1)(u_R(t^l,\overline{\theta})))+(1-\alpha')(p_3(u_R(t^m,\overline{\theta}))+(1-p_3)(u_R(t^l,\overline{\theta})))$$
$$=\alpha'(p_2(u_R(t^l,\overline{\theta}))+(1-p_2)(u_R(t^m,\overline{\theta})))+(1-\alpha')(p_4(u_R(t^m,\overline{\theta}))+(1-p_4)(u_R(t^m,\overline{\theta})))$$

(3.4)

If p_2, p_1 are given s.t.

$p_1(u_L(t^h,\underline{\theta}))+(1-p_1)(u_L(t^l,\underline{\theta}))>p_2(u_L(t^m,\underline{\theta}))+(1-p_2)(u_L(t^l,\underline{\theta}))$

then, for (3.1) to hold

$p_3(u_L(t^h,\underline{\theta}))+(1-p_3)(u_L(t^m,\underline{\theta}))<p_4(u_L(t^m,\underline{\theta}))+(1-p_4)(u_L(t^m,\underline{\theta}))$

an impossibility for any p_3, p_4.

If p_2, p_1 are given s.t.

$p_1(u_L(t^h,\underline{\theta}))+(1-p_1)(u_L(t^l,\underline{\theta}))<p_2(u_L(t^m,\underline{\theta}))+(1-p_2)(u_L(t^l,\underline{\theta}))$

$\longrightarrow p_1(u_L(t^h,\overline{\theta}))+(1-p_1)(u_L(t^l,\overline{\theta}))<p_2(u_L(t^m,\overline{\theta}))+(1-p_2)(u_L(t^l,\overline{\theta}))$

then, for (3.2) to hold

$(p_3(u_L(t^h,\overline{\theta}))+(1-p_3)(u_L(t^m,\overline{\theta}))>p_4(u_L(t^m,\overline{\theta}))+(1-p_4)(u_L(t^m,\overline{\theta}))$

an impossibility for any p_3, p_4.

Therefore

$p_1(u_L(t^h,\underline{\theta}))+(1-p_1)(u_L(t^l,\underline{\theta}))=p_2(u_L(t^m,\underline{\theta}))+(1-p_2)(u_L(t^l,\underline{\theta}))$

$\longrightarrow p_2>p_1$

Now, if p_2, p_1 are given s.t.

$p_1(u_L(t^h,\overline{\theta})) + (1-p_1)(u_L(t^l,\overline{\theta})) > p_2(u_L(t^m,\overline{\theta})) + (1-p_2)(u_L(t^l,\overline{\theta}))$

$\longrightarrow p_1(u_L(t^h,\underline{\theta})) + (1-p_1)(u_L(t^l,\underline{\theta})) > p_2(u_L(t^m,\underline{\theta})) + (1-p_2)(u_L(t^l,\underline{\theta}))$

again, for (3.1) to hold

$(p_3(u_L(t^h,\underline{\theta})) + (1-p_3)(u_L(t^m,\underline{\theta})) < p_4(u_L(t^m,\underline{\theta})) + (1-p_4)(u_L(t^m,\underline{\theta}))$

an impossibility for any p_3, p_4.

if p_2, p_1 are given s.t.

$p_1(u_L(t^h,\overline{\theta})) + (1-p_1)(u_L(t^l,\overline{\theta})) < p_2(u_L(t^m,\overline{\theta})) + (1-p_2)(u_L(t^l,\overline{\theta}))$

$\longrightarrow p_3(u_L(t^h,\overline{\theta})) + (1-p_3)(u_L(t^m,\overline{\theta})) > p_4(u_L(t^m,\overline{\theta})) + (1-p_4)(u_L(t^m,\overline{\theta}))$

an impossibility for any p_3, p_4.

Therefore

$p_1(u_L(\overline{\theta},t^h)) + (1-p_1)(u_L(\overline{\theta},t^l)) = p_2(u_L(\overline{\theta},t^m)) + (1-p_2)(u_L(\overline{\theta},t^l))$

$\longrightarrow p_2 < p_1$

a contradiction. □

3.9 Bibliography

1. Alesina R., N. Roubini and G. Cohen (1997). *Political Cycles and the Macroeconomy.* MIT Press.

2. Aragones E. (1997). "Negativity Effect and the Emergence of Ideologies." Journal of Theoretical Politics 9(2):198-210.

3. Besley T. and S. Coate. (1998). "Sources of Inefficiency in a Representative Democracy: A Dynamic Analysis." *American Economic Review* 88: 139-156.

4. Coate S. and S. Morris. (1999). "Policy Persistence." *American Economic Review* 89: 1327-1336.

5. Cukierman, A.and M. Tommasi. (1998). "When Does It Take Nixon to Go to China." *American Economic Review* 88: 180-97.

6. Dewatripont, M and J. Tirole. (1999). "Advocates." *Journal of Political Economy* 107:1-39.

7. Dewatripont, M and J. Tirole. (1999). "Advocates." *Journal of Political Economy* 107:1-39.

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

9. Feddersen T. and W. Pesendorfer. (1999). "Elections, information aggregation, and strategic voting." *Proc. Natl. Acad. Sci. USA* 96: 10572-10574.

10. Fernandez R. and D. Rodrik. (1991). "Resistance to Reform: Status Quo Bias in the Presence of Individual-Specific Uncertainty." *American Economic Review* 81: 1146-55.

11. Gerring, J. (1997). "Ideology: A Definitional Analysis." *Political Research Quarterly* Vol. 50, Issue 4, 957-994.

12. Green E. and R. Porter. (1984). "Noncooperative collusion under imperfect price information" *Econometrica* 52: 87-100.

13. Hardin, R. (1999). *Liberalism, Constitutionalism and Democracy.* Oxford University Press.

14. Holmstrom, B. (1982). "Moral Hazard in Teams." *Bell Journal of Economics* 10: 74-91.

15. Mailath, G and S. Morris (2000). "Repeated Games with Almost-Public Monitoring." *Journal of Economic Theory* 102: 189-228.

16. Mujumdar, S. and S. Mukand. (2002). "Policy Gambles". Working Paper.

17. Osborne, M.J. (1995)."Spatial model of political competition under plurality rule: A survey of some explanations of the number of candidates and the positions they take." *Canadian Journal of Economics* 28: 261-301.

18. Persson T., and G. Tabellini. (2000). *Political Economics. Explaining Economic Policy.* MIT Press.

19. Roemer, J.E. (2001). *Political Competition. Theory and Applications.* Harvard University Press.

20. Wittman, D. (1973). "Parties as utility maximizers" *American Political Science Review* 67: 490-98.