Credibility Concerns in Optimal Policy Design*

Yang K. Lu
European University Institute

First Version: November 2008
This Version: December 2009

Abstract

This paper models credibility management by a government using a simple reputation game in which government type is not directly observable by the private sector. Two non-standard features of the game produce conditions under which it is optimal for a "trustworthy" type (able to pre-commit) to separate itself from an opportunistic type (unable to pre-commit). First, policy announcement is introduced as an instrument, in addition to policy action, so that not only the opportunistic but also the trustworthy type behaves strategically. Second, time preference can differ across types. The combination of a patient trustworthy type and an impatient opportunistic type thus leads to early stages of the game marked by active policymaking (announcements and actions) on the part of the government and rapid learning on the part of the private sector, a result absent in the literature but more in line with reality.

Keywords: imperfect credibility, reputation game, optimal taxation, time inconsistency

JEL codes: E61, E62, D82

*This paper is based on the first chapter of my doctoral thesis in Boston University. I owe great debt of gratitude to Robert G. King for his constant guidance and support. The paper also benefits from discussions with Russel Cooper, Steve Davis, Piero Gottardi, Veronica Guerrieri, Peter Ireland, Bart Lipman, Lars Stole, and seminar participants at Boston University, Green Line Macro Meeting, UC-Santa Cruz, UChicago-Booth, FRB Richmond, FRB Kansas City, FRB Philidephia, Hong Kong University, Hong Kong University of Science and Technology, Chinese Univeristy of Hong Kong, City University of Hong Kong and European University Institute. All errors are my own.

1 Introduction

This paper studies the conditions under which it is optimal for a government to build and maintain its reputation. Reputation is defined as the probability of a government being trustworthy – a type that is bound to keep its promises.[1] Being trustworthy can be questionable because of a well-known *time inconsistent problem*.[2] For example, a government in power may have an incentive to announce a low tax rate plan, only to enact a high tax rate policy once it has observed the effects of the announcement on the tax base, since this yields high government revenues with little distortion of incentives to work. Similarly, it is often in the short-run best interest of a government to default on its debts, stimulate the economy with an inflation surprise, or bail out banks that are too big to fail.

Being aware of such incentives, agents in the private sector may not respond to policy announcements in the manner desired by the government. This makes the government reputation valuable. Indeed, the value of reputation is the reason why, in a typical reputation model, an opportunistic government which is subject to short-run temptation refrains from yielding to it and behaves in a trustworthy way. The conditions under which this behavior of mimicking by an opportunistic government occurs have been well studied in the literature on reputation models. This paper, however, shifts the focus to the optimal behavior of a trustworthy government. If the force of reputation is strong enough to alter the behavior of an opportunistic type, its impact on a trustworthy type should not be negligible. Therefore, what would a trustworthy type do concerning its reputation if it is endowed with the ability to set policy announcements?[3] In particular, facing optimal mimicking by an opportunistic type, will a trustworthy government set policy announcements to avoid being mimicked and thus signal its type (to build its reputation), or will policy announcements be merely accommodations to the mimicking (to maintain its reputation)?

A key condition for a trustworthy government to build its reputation is that it is sufficiently more patient than an opportunistic type. The intuition is simple. When the trustworthy type cares about the future more than an opportunistic type does, it will be more willing to invest heavily in reputation, knowing that later it will obtain better economic outcomes. This difference in reputation investment across types makes it feasible for a trustworthy government to distinguish itself and thus results in optimal reputation-building. By contrast, when an opportunistic type is as patient as a trustworthy type (which is a common assumption in the literature), the short-run temptation necessary for it to give up its reputation will also imply a too high short-run

[1] This definition of a trustworthy type is standard in the literature on reputation models. The trustworthy type is either unable to break its promises or faces severe punishment for doing so.

Examples in game theory include Kreps & Wilson (1982), Fudenberg & Levine (1989) and Cripps, Mailath & Samuelson (2004).

For applications in macroeconomic models, see for examples, Backus and Driffill [1985a,b], Barro [1986], Cukierman and Meltzer [1986] in monetary policy; Cole et. al [1995], Cole and Kehoe [1998] in sovereign debt, and Celentani and Pesendorfer [1996], Phelan [2006] in fiscal policy.

[2] Kydland and Prescott [1977].

[3] A smaller number of theoretical papers – see Mailath and Samuelson (2000) and Kim (2009) – take the approach of this paper, which is to make the trustworthy type an active player in the game. In Mailath and Samuelson (2000), the opportunistic type is an automaton. Kim (2009) studies a static model.

cost for the trustworthy type. It is therefore not optimal for the trustworthy type to set policy announcements that an opportunistic type would deviate from. As a result, the equilibrium fails to involve reputation-building by the trustworthy government.

Whether a trustworthy government builds reputation or not in equilibrium has important implications for equilibrium policies. For example, with similar setups, the optimal inflation target decreases over time in King, Lu and Pasten (2008) when the trustworthy central bank actively builds its reputation, resembling the inflation path during the period of Volker disinflation, whereas it increases over time in Cukierman and Livitan (1991), when the trustworthy central bank only seeks to maintain its reputation (what they call "accommodating imperfect credibility"). Another example is in D'Erasmo (2007), where both patient and impatient governments take reputation-building into account and the equilibrium debt-to-output ratio is significantly higher than is standard in the literature, a result that is more in line with the data.

In addition, the strategic management of reputation by a trustworthy government also implies rich dynamics of government reputation, which can be measured empirically. For examples, long-run inflation forecasts are a measure of a central bank's reputation for stabilizing inflation, and long-term interest rates are a measure of a government's reputation for repaying debts. Thus, testable implications can be drawn from theories on government reputation.

All in all, both the theoretical and empirical relevance of reputation-building by a trustworthy government call for a systematic theoretical treatment of this important matter. To the best of my knowledge, this paper is the first to point out that the combination of a patient trustworthy type and an impatient opportunistic type leads to early stages of the game marked by active policymaking on the part of the government and rapid learning on the part of the private sector, as the trustworthy type optimally builds its reputation. Moreover, this unconventional combination of a patient trustworthy type and an impatient opportunistic type can emerge in a natural way if we consider government type to be an endogenous choice based on time preference. Because a more patient government has more to gain from being able to commit, it will be more willing to pay a cost for some commitment device and become worthy of trust.

In order to have a transparent framework for analyzing government reputation management and its impact on policy design, I construct a simple fiscal model along the lines of Phelan (2006). It is a finitely-repeated reputation game, where a government that seeks to maximize the present value of its tax revenues plays against a continuum of small private-sector households who decide whether or not to engage in production with a fixed cost each period. The government in my model can choose a tax plan and announce it at the beginning of each period. Households entertain the possibility that the government is one of two types which differ with regard to their ability to commit to a policy: a 'trustworthy' government that is defined to execute announced plans for sure, an 'opportunistic' government that is given an option to deviate from its original plans. In the model, this deviation is to the highest tax rate possible, representing confiscation by the government of all households production. The government type is private information and is fixed throughout the whole game. Households update their beliefs in a Bayesian fashion after observing the government's actions at the end of each period.

This assumption that a government is able to choose a tax plan and announce it at the beginning of each period contrasts what is standard in the literature: that the trustworthy type of government is simply an automaton, following a given plan without making any choices. The endogenous tax plans enable the government, especially the trustworthy type, both to react to and to manipulate the dynamics of reputation. However, the fact that plans and announcements may convey information about government type creates a well known problem of equilibrium indeterminacy in signaling games. In this paper, I resolve the problem by adopting the refinement criterion of Mailath et al (1993).[4] In particular, I establish the uniqueness of a Markov perfect equilibrium in this reputation game that involves both intra-period signaling and inter-temporal Bayesian learning. In the unique equilibrium, an opportunistic type of government optimally imitates a trustworthy type's announcement strategy. Whether or not the opportunistic type will enact the announced plan, however, depends on its time preference and the plan itself. As a result, the announced tax plan is not informative about the government's type, but is informative about the tax action.

I show a key trade-off facing a trustworthy type when it designs optimal tax plans. On the one hand, there is an incentive to signal. That is, the trustworthy type would like to set a low planned tax rate so as to impose a high cost on the opportunistic type if it were to enact the same plan. Such a high cost makes it hard for the opportunistic type to masquerade its identity. This benefits the trustworthy type because enacting the plan then conveys more information about its identity and in turn improves its reputation, which implies a greater flow of tax revenues in the future. On the other hand, there is an incentive to accommodate. In particular, the trustworthy type would like to set a high planned tax rate so that it is easy for the opportunistic type to enact it. Such a plan is thus more credible and in turn reduces households' risk of confiscation. This benefits the trustworthy type in the short run since it secures more tax revenues for the current period.

I analyze how this trade-off is affected by the key parameters in the model, namely the patience levels of both types, the elasticity of supply, the initial reputation of the government, and the time horizon of the game. Among these, the relative patience of the trustworthy type with respect to that of the opportunistic type plays a key role in determining which side of the trade-off dominates.

When the trustworthy type is more patient than the opportunistic type, the incentive to signal dominates, so that the Markov perfect equilibrium involves initial periods of rapid learning with low taxation. As a result of this aggressive reputation-building, the trustworthy type is able to sustain robust economic outcomes in the later stages of the game despite an increasing path of planned tax rates, which converge to a constant when full credibility is achieved.

By contrast, when the trustworthy and opportunistic types share similar a time preference, the incentive to accommodate dominates, so that the Markov perfect equilibrium involves little learning with high taxation throughout the whole game. The only way households learn about the government's type is through a weak force of reputation, so that the opportunistic type chooses

[4] In my parallel work with King and Pasten (2008a,b), we adopt the same refinement to achieve uniqueness in other settings.

not to mimic in the first place, as opposed to active policymaking by the trustworthy type for the purpose of reputation building.

From a macroeconomic perspective, the time inconsistency problem of policymaking under rational expectations is at the heart of this paper. Much of the prior literature on this topic has concerned ways of "solving" this problem, including the reputation game approach initiated by Barro and Gordon (1983) and the related sustainable plan approach of Chari and Kehoe (1990).[5] Other studies focus on the modification of policymaker preferences (Rogoff, 1985) or institutions (Cukierman, 1992, Walsh, 2003) as devices for mitigating time inconsistency.

There is increasing use of Markov perfect equilibrium in macroeconomic analysis (see, for example, Klein, Krusell and Rios-Rull, 2006), as a device for studying economies with limited policy commitment. Most of this work describes optimal policy under complete information about government preferences. In that context, within my model, the government either has a perfect ability to influence households expectations on taxes or no power at all. This paper thus studies an intermediate case in which a government can partially control the expectations of households and the degree of this control is endogenous to the history of policies.

This paper is also related to the literature on career concerns initiated by Holmstrom (1999), in the sense that reputation-building is a major driving force in agent's decision-making. However, the modelling approach here is rather different. In particular, I explicitly model reputation-building as one type taking actions to avoid being mimicked by another type, and it is optimal mimicking that sometimes obscures households' learning rather than some exogenous noise.

The remainder of the paper is organized as follows. Section 2 presents the setup of the game and defines the equilibrium concept. Section 3 uses a two-period example to illustrate the main results of this paper. Section 4 extends those results to a game with an arbitrary number of periods. Section 5 establishes the uniqueness of the equilibrium. Finally, Section 6 concludes.

2 A Reputation Game

The now-classic analysis of a reputation game by Kreps and Wilson [1982] involves the interaction of two firms, an entrant and an incumbent, with the latter of unknown strength. Early in the multistage game, there are powerful forces of reputation which make the weak incumbent behave as if it were strong, which can be called a *pooling outcome,* following the terminology of the signalling game literature. There is no learning about type during this initial interval. Later, the weak incumbent follows a mixed strategy, sometimes fighting entry and sometimes not, so that there is Bayesian learning about type, which can be called a partial pooling, or *mixed outcome*. In the last period, the weak incumbent will not fight entry, so that there is full revelation of type as a *separating outcome*.

[5] In the literature on sustainable plans, the central question has been whether a Ramsey outcome can be sustained or not when there is a lack of commitment. See for example, Chari and Kehoe [1990] with an infinite horizon and complete information, and Sleet and Yeltekin [2007] with an infinite horizon and incomplete information.

In this section, I describe a fiscal policy game in which there are trustworthy and opportunistic fiscal decision-makers playing against a competitive private sector comprised of atomistic households. Unlike previous analyses of similar fiscal games, such as Phelan [2006], where the trustworthy government simply sets an exogenously fixed tax rate, I work with a case in which the trustworthy government is able to choose the optimal tax policy each period, so that we can study how the tax path optimally responds to the dynamics of imperfect credibility. The equilibrium notion in the game will be Markov perfect.

2.1 Setup

I study a finite horizon economy, indexed by $t = 1, 2, ..., T$, in which a continuum of households play against a long-lived government that collects tax on households production. Within each period, the government moves first to announce a proportional tax plan $\tau_t \in [0, 1]$ on production for this period. The households observe the tax plan and decide whether to produce or not, based on their individual production costs. After the production decisions have been made, the government then levies the taxes on their outputs at rate $\tilde{\tau}_t$.

2.1.1 Households

Each household is indexed by i with total measure 1. For convenience, we can think of i as uniformly distributed over the interval [0,1]. The action profile for households is then a measurable function $a : [0, 1] \to \{0, 1\}$. Households only live for one period, but are able to observe the whole history of the game.[6] Each household is endowed with one indivisible unit of labor that they can put into production. The production technology is linear in labor with deterministic productivity normalized to 1. The production cost c_i for household i is a random draw each period from distribution $G(\cdot)$ with support $[0, 1]$.[7] It is convenient to work with the family of distributions $G(x) = x^{\gamma}$.[8] As discussed further below, the parameter γ is interpreted as a market participation elasticity.

If household i decides to produce, $a(i) = 1$, its payoff is the after-tax production net of the production cost, $r(\tilde{\tau}, a(i) = 1) = (1 - \tilde{\tau}) - c_i$. If household i does not produce, $a(i) = 0$, it gets zero, $r(\tilde{\tau}, a(i) = 0) = 0$. We can use $\mu = \int_0^1 a(i)\, dG(c_i)$ to summarize the *aggregate participation of the households*. μ is publicly observable but the individual production decision is private information to each household. Under this assumption, no individual household's action has an impact on another's decision, nor does it affect government action.

[6] The assumption of short-lived households is not essential. As long as the rewards of production cannot be stored over periods, the maximization problem of households will be the same with a longer horizon.

[7] Macroeconomists generally study the effects of policies in settings where there is a smooth response on the part of the private sector to public interventions such as taxation and monetary policy. For example, the early 1980s analyses of monetary policy and reputation, such as Backus and Driffill [1985 a,b] and Barro [1986], worked in such settings. To this end, households in my model are heterogenous in their production costs.

[8] As will be seen later, this particular functional form of $G(\cdot)$ allows a close-form solution in this economy. At the same time, it abstracts any variations in optimal policies that are not due to credibility concerns. Hence, we can focus solely on the interaction between optimal policy designs and credibility concerns.

2.1.2 Government

The government lives for T periods. There are two types of government: trustworthy and opportunistic (TR and OP hereafter). The trustworthy government is assumed to always impose the announced tax rate τ_t at the end of each period, $\tilde{\tau}_t = \tau_t$. The opportunistic type, however, can either tax the production at rate τ_t, or confiscate all output, i.e. $\widetilde{\tau}_t = 1$, or even randomize between $\widetilde{\tau}_t = \tau_t$ and $\widetilde{\tau}_t = 1$. Let π_t denote *the probability of confiscation.* Government type is private information and stays fixed throughout the game, so that the opportunistic type is able to mimic the trustworthy type's behavior if it is in its interest to do so and thereby cover its true identity.

The government makes decisions at two distinct points within a period. At the beginning of the period, it sets the tax plan τ_t for the current period. At the end of the period, it decides whether the plan will be carried out, i.e. choosing π_t. The latter decision, however, is only relevant for the opportunistic type, since it is not an option for the trustworthy type by definition. Hence, we will refer to π_t as the action of the opportunistic type hereafter.

Both types of governments want to maximize their life-time tax revenues, discounted at rates β_{tr} and β_{op}, respectively. The objective is then defined as

$$\textstyle\sum_{t=0}^{T} \beta_{type=tr,op}^{t} \ \widetilde{\tau}_t \mu_t. \qquad (1)$$

2.2 Markov strategies and Markov perfect equilibria

Strategies are mappings from histories to actions. In this paper, I restrict attention to Markov strategies which condition actions solely on the payoff-relevant variables. When the households choose whether to produce or not, the identity of the government is the only payoff-relevant variable. That is, if an opportunistic government is in place, it may confiscate at the end of the period, leaving the household nothing but the sunk cost of production.

In this incomplete information game, the government type is not observable. So the strategies have to depend on households' best estimates of type;[9] that is, the household's belief in the government being the trustworthy type, given the history before date t. Define this belief to be the state variable of the game: ρ.

2.2.1 Credibility concepts and Markov strategies

Let us call ρ "long-term credibility," so as to distinguish it from another important credibility concept: "short-term credibility." Denoted by ψ, short-term credibility is the likelihood that the government will tax at the rate τ:

$$\psi = \rho + (1-\rho)(1-\pi).$$

[9] This extension of MPE to an incomplete information game is in the spirit of Ball (1995): "... actions depend on agents' best estimates of payoff-relevant variables".

This is the sum of two elements: the likelihood of a trustworthy type of government being present so that τ will be implemented for sure, and the likelihood of an opportunistic type being present but mimicking with probability $(1-\pi)$. Conditional on the long-term credibility level ρ, there is a one-to-one mapping between the strategy π and the short-term credibility level ψ. With $\psi = 1$, it mimics for sure: $\pi = 0$; and with $\psi = \rho$, it confiscates for sure: $\pi = 1$. Randomizing between mimicking and confiscating with $\pi \in [0,1]$ corresponds to ψ lying within $[\rho, 1]$.

Under the Markov restriction, we can write the strategies of both the households and the government as functions of long-term credibility ρ: $\{\tau(\rho), \mu(\rho), \pi(\rho)\}$. Instead of specifying each household's strategy $a_i(\rho)$, we use the participation rate $\mu(\rho)$ as a convenient aggregation of individuals' strategies. This participation rate is also the tax base for the governments, as suggested by (1) above. For the opportunistic type's strategy $\pi(\rho)$, it turns out to be more convenient and intuitive to state it in terms of ψ instead of π. Hence, we will use the strategy triple $\{\tau(\rho), \mu(\rho), \psi(\rho)\}$ to characterize the Markov perfect equilibrium (MPE) in the game, although we could have equivalently used $\{\tau(\rho), \mu(\rho), \pi(\rho)\}$.

2.2.2 Belief updating after the announcement

As the government moves first to announce the planned tax rate for the current period, the announcement itself is a signal of the government's type and can thus influence its long-term credibility directly. Hence, both types of government have to take into account the households' perception of which type has an incentive to announce the observed rate when designing their optimal signalling strategies. However, as will be clear in Section 5, always imitating a trustworthy government's announcement is the unique equilibrium strategy for an opportunistic type.[10] This result significantly simplifies the analysis of the game in two aspects. First, it makes the tax announcement uninformative about the government's identity. Therefore, there is no belief updating after households observe the announced plan. Second, it enables me to treat the trustworthy type as a leader in the game, because the tax announcement, regardless of the true identity of the policy maker, is set to maximize the trustworthy government's payoff. Hence, the stage game played between the current government and the households can be analyzed as if there were three players in the game, with a trustworthy government setting the announcement first, and the households producing accordingly, followed by the opportunistic type deciding whether to confiscate, given the tax announcement and the households' production. [11]

2.2.3 Bayesian learning and the dynamics of credibility

If the government confiscates, long-term credibility next period (ρ') will be zero, $\rho' = 0$. Even without confiscation, the strategy ψ still has an information impact on the evolution of long-term

[10] The equilibrium uniqueness is obtained after applying off-equilibrium belief refinement, as proposed by Mailath, Okuna-Fujiwara and postelwaite [1993].

[11] Even when the government in place is indeed trustworthy, it still needs to account for the reaction of the opportunistic type in terms of the confiscation probability. This is because the households, when making production decisions, are in general concerned about the possible confiscation by the opportunistic government.

credibility, which is governed by Bayes' rule:

$$\rho' = \frac{\rho}{\psi}. \tag{2}$$

That is, the marginal probability of observing $\widetilde{\tau} = \tau$ is identical to the short-term credibility of the tax policy.

2.2.4 Household expectation formation

As households make production decisions before the taxes are levied, they have to form expectations of the actual tax rate. I impose rational expectation and, in turn, symmetry across households. So the expected tax rate is the probability-weighted average between τ and 1, with the probability attached to τ being the short-term credibility ψ. The atomistic feature of each household eliminates any strategic or intertemporal concern in the production decision-making, because the decision does not individually affect the play of the government, or the future value of long-term credibility ρ. Therefore, the decision on the households' side is essentially static, with all households using a common threshold in deciding to produce or not. Only those who receive cost c_i less than the expected after-tax output $(1-\tau)\psi$ will produce, resulting in the optimized aggregate strategy:

$$\mu = G\left[(1-\tau)\psi\right]. \tag{3}$$

With the distribution $G(x) = x^{\gamma}$, the fraction of households participating is $[(1-\tau)\psi]^{\gamma}$ so that the parameter γ is the elasticity of supply with respect to the expected reward to market participation, $[(1-\tau)\psi]$.

2.2.5 Government strategies

In contrast to the individual household, the government is a big player so that it is strategic. In addition, it lives for T periods so that its optimization is intertemporal.

The opportunistic type obtains instant revenue gain by confiscating the current output. However, its true identity will then be revealed and no production will occur for future periods.[12] This loss of future tax revenues can be avoided if the opportunistic government mimics the trustworthy type's behavior by taxing at the announced rate τ. By masquerading as the trustworthy type, an opportunistic government will induce more production in the future and collect higher tax revenues, regardless of its future strategies. This trade-off between current and future tax revenues is

[12] If the government is known for sure to be the opportunistic type $\rho = 0$, the uniqe MPE in this game is 0 production. To see that: at any period, if $\pi = 1$, household optimization implies $\mu = 0$. If $\pi < 1$, government optimization requires $\tau\mu + \beta_{op}V(0) >= \mu + \beta_{op}V(0)$, which only holds for $\mu = 0$. Given $\mu = 0$, the weak type is indifferent to the choice of π. Letting $\pi = 1$, it justifies the zero production choice of the households. If no one produces in this game, the life-time tax revenue will be zero: $V(0) = 0$ for any period.

summarized by what I call the "incentive compatibility constraints," which are stated as follows:

$$\begin{array}{rrcl}\text{Pooling strategy (mimicking)} & \psi & = & 1 \text{ if } \mu < \tau\mu + \beta_{op} V'(\rho') \\ \text{Separating strategy} & \psi & = & \rho \text{ if } \mu > \tau\mu + \beta_{op} V'(\rho') \\ \text{Mixed strategy} & \psi & \in & [\rho, 1] \text{ if } \mu = \tau\mu + \beta_{op} V'(\rho'). \end{array} \tag{4}$$

If the opportunistic government confiscates, it gets all the current output as tax revenues: μ. As no production will occur after the government is revealed to be opportunistic, the future value of tax revenues is simply zero. So the value of confiscation is just μ, the left-hand side of the first inequality above. The right-hand side is the value of mimicking, i.e. collecting tax at the announced rate τ. This disciplined behavior increases the long-term credibility level to a new state: $\rho' = \rho/\psi$, generating a positive future value of tax revenues. We denote this value by V'. The prime indicates that the next-period value function differs from the current one in terms of their functional forms.

The trustworthy type, as a leader in this game, is able to use the tax announcements τ to affect both households production μ and the opportunistic type's confiscation probability π, in order to maximize its life-time tax revenues. The impact of tax announcements on the objective works through various channels. Most obviously, the announced tax rate, if actually imposed, directly affects how much revenue a government can collect each period. This payoff relevance has further impact on how likely the opportunistic type is willing to follow the announced rate. The likelihood of the opportunistic type confiscating then determines the short-term credibility, so that both current output and the evolution of credibility are affected as the households are rational and Bayesian. The trade-offs involved in such a decision making, as well as the resulting equilibrium strategy, are the focus of Sections 3 and 4.

2.2.6 Markov perfect equilibrium

Now we are ready to define a Markov perfect equilibrium of the multistage game.

Definition 1 *The sequence of Markov strategy triplets* $\{\tau_t(\rho), \mu_t(\rho), \psi_t(\rho)\}_{t=1}^{T}$ *is called Markov perfect equilibrium (MPE) if, at each stage of the game,* $\tau_t(\rho)$ *solves the trustworthy government's optimization:*

$$W_t(\rho) = \max_{\tau_t} \{\tau_t \mu_t + \beta_{tr} W_{t+1}(\rho/\psi_t)\},$$

with $\{\mu_t(\rho), \psi_t(\rho)\}$ *satisfying both the households and the government optimization conditions (3) and (4), subject to the evolution of credibility (2).*

2.2.7 Limited commitment and backward induction

This subsection discusses the approach I use to characterize a MPE solution analytically and compute it numerically. I focus on the solution to a limited commitment case in which the trustworthy type has no access to the intertemporal commitment technology, even though it can commit to the current period's tax plan.[13] In light of this limitation, when the trustworthy government sets a value τ_t in period t, it recognizes that $\{\tau_s^*\}_{s>t}^T$ will be chosen by future trustworthy governments, with similar objectives. Yet with no direct control over the future tax path, the current government can still influence future payoffs by affecting the evolution of long-term credibility, the state variable, which guides optimal future tax rates.

Hence, to obtain the solutions in this limited commitment case, I can start from the last period and apply backward induction. In each period, the trustworthy government takes the future value function as given when setting the optimal tax plan to balance its current and future payoff effects. The current tax plan affects the equilibrium reactions of both the opportunistic type and the households through the optimization conditions (3) and (4), which in turn determines the current tax revenues $\tau\mu$. This momentary payoff is weighted against the future value of revenues through the tax effect on the evolution of credibility evolution governed by Bayes' rule (2).

3 A Two-Period Model

This section solves a two-period version of the model to illustrate the basic trade-offs facing a trustworthy government with imperfect credibility. On the one hand, the tax plan tends to be higher than a benchmark rate set by a government with no credibility concerns, because implementing a plan in which households have little faith would be costly. On the other hand, a tax plan lower than the benchmark rate, once implemented, can be a powerful tool to signal the government's type and thus changes the evolution of households belief. The relative strength of these offsetting forces critically depends on the difference in time preferences of the two types β_{tr}/β_{op}, the elasticity of supply with respect to tax changes γ, and the long-term credibility level ρ. The conditions under which it is optimal for the trustworthy type to either accommodate imperfect credibility or invest in future reputation will be characterized.

3.1 Optimal tax plan with full credibility

Let us first consider a benchmark case where a trustworthy government possesses full credibility, $\rho = \psi = 1$.

With heterogeneous participation costs, only those who receive cost realizations below a threshold will participate. As the threshold is determined by the after-tax output $(1 - \tau)$, it decreases with the announced tax rate and so does the fraction of households who produce, as shown in Panel A of Figure 1. This feature replicates the celebrated Laffer curve, which traces an inverted-U relationship between taxes and revenues. Panel B in Figure 1 depicts one example of this Laffer curve.

Due to this negative effect of taxes on the production base, a trustworthy government without

[13] This restriction on the trustworthy type's committment ability is not binding in the current setting. Given the absence of shocks and real state variables, there is nothing to be gained by being able to commit to policies beyond the current period.

any credibility concern will also choose an interior optimal tax rate to maximize revenues:

$$\tau_{UC}^{*} = \arg\max_{\tau} \tau G\left(1-\tau\right).$$

I label such a rate "unconstrained optimal" because it is free of credibility concerns. If we use the specific cost distribution $G(x) = x^{\gamma}$, $\tau_{UC}^{*} = (\gamma+1)^{-1}$.

3.2 Optimal tax plan in the last period

Let us start from the last period, $t = 2$. With some inherited credibility, the opportunistic type will always confiscate, irrespective of ρ and τ. Given that behavior, the short-term credibility is fixed at the level ρ. The homogeneity feature of $G(\cdot)$ ensures that the optimal tax plan is always equal to the unconstrained optimal level for any fixed short-term credibility: $\tau_{UC}^{*} = \arg\max \tau G\left[(1-\tau)\bar{\psi}\right]$. Thus, the trustworthy type sets $\tau_2^{*} = \tau_{UC}^{*}$ to maximize current tax revenues, which yields the value functions of both types as:

$$\begin{aligned} V_2(\rho_2) &= (1-\tau_{UC}^{*})^{\gamma}\rho_2^{\gamma} \\ W_2(\rho_2) &= \tau_{UC}^{*} V_2(\rho_2). \end{aligned} \tag{5}$$

This property, brought about by the homogeneity of $G(\cdot)$, is desirable since any variation in optimal tax policies will then be attributed to the endogeneity of short-term credibility, which captures the opportunistic type's reaction to the tax plan and is the essence of what I call "credibility concerns." Consequentially, the deviation of the optimal tax rate from τ_{UC}^{*} provides a measure of credibility concerns.

3.3 Optimal tax plan with credibility concerns

Moving back to the first period, $t = 1$, the trustworthy government takes $W_2(\rho)$ and $V_2(\rho)$ as given, and chooses τ_1 by solving:

$$\max_{\tau_1} \tau_1 \mu_1 + \beta_{tr} W_2\left(\rho_1/\psi_1\right).$$

The opportunistic type, contrary to the last period, no longer has a strictly dominant strategy, since its next period payoff critically depends on its tax action in the current period. As a result, whether or not the opportunistic type will confiscate and how likely it is to confiscate varies with different levels of long-term credibility ρ_1 and tax plan τ_1. The decision of the opportunistic type in turn affects the expected tax rate, which determines how many households will produce in that period. Therefore, when the trustworthy type decides the optimal tax plan, it has to take into account the best reaction of the opportunistic type, i.e. the incentive compatibility constraint, which takes one of the following three forms:

1) If the tax plan generates a mixed outcome,

$$\tau_1\mu_1 + \beta_{op}V_2(\rho_1/\psi_1) = \mu_1$$

$$\text{with } \mu_1 = G[(1-\tau_1)\psi_1] \text{ and } \psi_1(\tau_1,\rho_1) \in [\rho_1, 1].$$

2) If the tax plan generates a pooling outcome, $\psi_1 = 1$ and the planned rate is bounded below by

$$\tau_1\mu_1 + \beta_{op}V_2(\rho_1) \geqslant \mu_1 \text{ with } \mu_1 = G[1-\tau_1].$$

3) If the tax plan generates a separating outcome, $\psi_1 = \rho_1$ and the planned rate is bounded above by

$$\tau_1\mu_1 + \beta_{op}V_2(1) \leqslant \mu_1 \text{ with } \mu_1 = G[(1-\tau_1)\rho_1].$$

Conditional on the specific nature of the outcomes induced by different planned rates, the trustworthy type maximizes in each case and picks the optimal $\tau_1(\rho_1)$ that generates the highest value.

If τ_{UC}^* satisfies the separating incentive compatibility constraint, i.e. $\beta_{op} \leqslant (1-\tau_{UC}^*)\rho_1^\gamma$, it is optimal for the trustworthy government to induce a separating outcome using τ_{UC}^*. It is because when ψ_1 is fixed at level ρ_1, by construction, τ_{UC}^* maximizes the trustworthy type's value. Similarly, if τ_{UC}^* satisfies the pooling incentive compatibility constraint, i.e. $\beta_{op}\rho_1^\gamma \geqslant (1-\tau_{UC}^*)$, τ_{UC}^* is the optimal tax plan that induces a pooling outcome.

If τ_{UC}^* fails to support either a pooling or a separating outcome, that is when

$$\rho_1 \leqslant \min\left\{\left[\beta_{op}/(1-\tau_{UC}^*)\right]^{1/\gamma}, \left[(1-\tau_{UC}^*)/\beta_{op}\right]^{1/\gamma}\right\}, \tag{6}$$

the optimal tax plan will make the incentive compatibility constraint binding, i.e. the opportunistic type being indifferent between mimicking and confiscating. The following lemma formalizes this property.

Lemma 1 *When τ_{UC}^* does not satisfy the incentive compatibility constraint of either a pooling or a separating outcome, the problem of the trustworthy type can be simplified to:*

$$\max_{\tau_1\in[0,1]} \tau_1\mu_1 + \beta_{tr}W_2(\rho_1/\psi_1) \tag{7}$$

$$s.t.\ \tau_1\mu_1 + \beta_{op}V_2(\rho_1/\psi_1) = \mu_1 \textit{ with } \mu_1 = G[(1-\tau_1)\psi_1] \textit{ and } \psi_1 \in [\rho_1, 1] \tag{8}$$

Proof. As shown above, when short-term credibility ψ_1 is fixed, the optimal solution to the unconstrained maximization (7) is τ_{UC}^*. In the case of having a pooling or a separating outcome, short-term credibility is fixed at 1 or ρ_1, respectively. Imposing the incentive compatibility constraint will alter the optimal solution if τ_{UC}^* makes the constraint binding. In this case, maximization dictates that the trustworthy type makes the least deviation from τ_{UC}^* to meet the pooling or separating incentive compatibility constraints. That is, $\hat{\tau}_1$ such that $\hat{\tau}_1\mu_1(\hat{\tau}_1,\rho_1) + \beta_{op}V_2(1) =$

$\mu_1(\hat{\tau}_1, \rho_1)$ is the optimal rate among those that generate separating outcomes, and $\tilde{\tau}_1$ such that $\tilde{\tau}_1\mu_1(\tilde{\tau}_1, 1) + \beta_{op}V_2(\rho_1) = \mu_1(\tilde{\tau}_1, 1)$ is the optimal rate among those that generate pooling outcomes. In the case of having a mixed outcome, the opportunistic type is already indifferent between mimicking and confiscating. Therefore, in any case, the indifference condition for the opportunistic type is satisfied. ■

3.3.1 The trade-offs in setting the optimal tax plan

The lemma says that in this more interesting case, where the optimal tax plan is different from the unconstrained optimal rate, the trustworthy type always sets the tax plan such that the opportunistic type is indifferent between mimicking and confiscating. The indifference condition (8) thus determines the best reaction of the opportunistic type to the tax plan τ_1 given long-term credibility ρ_1, which can be captured by the endogenous variation of short-term credibility ψ_1 as a function of both τ_1 and ρ_1 :

$$\psi_1(\tau_1, \rho_1) = \left(\frac{\beta_{op}(1-\tau_{UC}^*)^\gamma}{(1-\tau_1)^{\gamma+1}}\right)^{\frac{1}{2\gamma}} \rho_1^{\frac{1}{2}}. \tag{9}$$

It is easy to see that ψ_1 increases with τ_1:

$$\frac{\partial \psi_1}{\partial \tau_1} \geqslant 0.$$

The intuition is the following. Holding ψ_1 constant, decreasing the tax plan τ_1 increases the tax-revenue gap between following and deviating from the plan, and in turn raises the temptation to deviate. It thus increases the probability of confiscation π_1 and implies lower short-term credibility ψ_1 to maintain the indifference.

Lower short-term credibility ψ_1 in the current tax plan increases the households' risk of confiscation after production, and thus depresses the tax base $\mu_1 = (1-\tau_1)^\gamma \psi_1^\gamma$. Therefore, while a lower tax plan makes it more rewarding to produce, it is less likely to be carried out. This credibility effect offsets the conventional tax effect on the production base as described by the Laffer curve. As a result, the marginal gain of raising the planned tax rate is larger when short-term credibility responds endogenously than when it is fixed:[14]

$$\frac{\partial \tau_1 \mu_1}{\partial \tau_1} = \frac{\partial \tau_1 G\left[(1-\tau_1)\psi_1\right]}{\partial \tau_1} \geqslant \frac{\partial \tau_1 G\left[(1-\tau_1)\right]}{\partial \tau_1} \geqslant 0.$$

The tax plan maximizing current tax revenues is therefore higher than the unconstrained optimal rate τ_{UC}^*, with the discrepancy measuring the extent to which a trustworthy government accommodates imperfect credibility.

Announcing a higher tax plan may benefit current tax revenues, but it impedes the transmission of information from the government to households. Because a higher tax plan is more likely to be

[14] The derivatives are positive because a reasonable tax rate should always be on the increasing side of the Laffer curve.

carried out by the opportunistic type as well, implementing it does not convey much information about the government's type. The Bayesian learning equation (2) reflects such a learning effect, where a higher short-term credibility ψ_1 reduces future long-term credibility ρ_2. Since the value of the trustworthy type in the last period increases in long-term credibility, a higher tax plan today imposes a loss in terms of future values through a slower growth in long-term credibility:

$$\frac{\partial \beta_{tr} W_2 (\rho_1/\psi_1)}{\partial \tau_1} \leqslant 0.$$

Therefore, the credibility and learning effect of a tax plan, both of which stem from the optimal reaction of the opportunistic type, comprise the trade-offs facing the trustworthy government. In determining the tax plan, it has to weigh current revenues against future continuation values.

3.3.2 The key determinants in the trade-offs

What is then the overall effect of such trade-offs on the optimal tax plan? When is it optimal for the trustworthy type to lower the tax plan relative to the unconstrained optimal rate τ^*_{UC}, as a way of investing in future reputation? Or is it optimal to raise the tax plan relative to τ^*_{UC} so as to secure more current tax revenues? I will argue that the most important parameters that determine the overall effect are two: 1) the elasticity of supply γ relative to the number of future periods (that is 1 in the current case); and 2) the ratio of time discount factors β_{tr}/β_{op}.

To see this, it is useful to observe that the payoff to the trustworthy type differs from the opportunistic type's value of mimicking only in its weight on future continuation values:

$$W_1(\rho) = \max_{\tau_1} \left\{ \tau_1 \mu_1 + \frac{\beta_{tr}}{\beta_{op}} \tau^*_{UC} V_2 (\rho_1/\psi_1) \right\}.$$

Furthermore, because the opportunistic type is indifferent, the value of mimicking equals the current production base, as does its reaction to tax changes:

$$\frac{\partial \left[\tau_1 \mu_1 + V_2 (\rho_1/\psi_1)\right]}{\partial \tau_1} = \frac{\partial \mu_1}{\partial \tau_1}.$$

As mentioned in the previous subsection, the reaction of the production base to tax changes is a compound result of both the credibility effect, working through the endogenous short-term credibility, and the standard Laffer-curve effect. What determines the direction of the response is the level of elasticity of supply γ. When the supply is elastic, $\gamma > 1$, the production base still reacts negatively to a tax increase, $\partial\mu_1/\partial\tau_1 < 0$. However, when the supply is relatively rigid, $\gamma \leqslant 1$, the credibility effect becomes dominant and it overturns the standard reaction of the production base to a tax increase. It now increases with the planned tax rate: $\partial\mu_1/\partial\tau_1 \geqslant 0$.

The direction of μ_1's response to τ_1 captures the relative elasticity of the opportunistic type compared to households, since it is determined by how sensitively the opportunistic type's strategy ψ_1 changes with the planned tax rate τ_1. Given this relative elasticity, the extent to which the trustworthy type aims to distinguish itself through the learning effect depends critically on the relative weight it puts on future continuation values compared to the opportunistic type: $\left(\beta_{tr}/\beta_{op}\right)\tau_{UC}^{*}$. In this sense, the ratio of time preferences β_{tr}/β_{op} is a measure of the incentive to signal by the trustworthy type.

When the relative elasticity of the opportunistic type is low, i.e. $\partial\mu_1/\partial\tau_1 < 0$, a moderate level of signaling incentive, say $\beta_{tr}/\beta_{op} = \left(\tau_{UC}^{*}\right)^{-1}$, is enough to motivate the trustworthy type to invest in its future reputation. This is because decreasing the planned tax rate relative to τ_{UC}^{*} increases μ_1, and in turn increases the payoff to the trustworthy type so long as $\left(\beta_{tr}/\beta_{op}\right)\tau_{UC}^{*} \geqslant 1$. When the relative elasticity is high, however, $\partial\mu_1/\partial\tau_1 \geqslant 0$ so that decreasing the planned tax rate has an opposite effect on μ_1, as well as on the payoff to the trustworthy type if $\left(\beta_{tr}/\beta_{op}\right)\tau_{UC}^{*} = 1$. To make the lower planned tax rate desirable for the trustworthy type, a larger weight on the beneficial part, i.e. future continuation values, is necessary. That is, a higher ratio of time preference is needed for the trustworthy type to invest in its future reputation.

As the nature of the trustworthy type's optimization differentiates between low and high elasticity of the opportunistic type compared to households, I will treat the two cases separately in discussing the Markov Perfect Equilibrium.

3.3.3 MPE when $\gamma > 1$

First, take the case when $\gamma > 1$ so that $\partial\mu_1/\partial\tau_1 < 0$. In this case, there will be three ranges of the state variable ρ_1, over which a particular form of equilibrium prevails. The property of the MPE on each region depends on the ratio of time discount factors, which is summarized by the following proposition. The proof is constructive.

Proposition 1 *Under the assumption $\gamma > 1$, there exists a unique MPE at period 1 with the following properties:*

i. If $\beta_{tr}/\beta_{op} \in \left(0, \left(\gamma+1\right)/\gamma\right)$ and $\beta_{op} \geqslant 1-\tau_{UC}^{*}$, $\tau_1^{*} \geqslant \tau_{UC}^{*}$ and the "credibility regions" are:

"mixed region":	$\tau_1^{*}\left(0 \leqslant \rho_1 \leqslant l_1\right) = \bar{\tau}_1;\ \psi_1^{*} = \psi_1(\bar{\tau}_1, \rho_1)$
"constrained pooling region":	$\tau_1^{*}\left(l_1 \leqslant \rho_1 \leqslant h_1\right) = \tilde{\tau}_1\left(\rho_1\right);\ \psi_1^{*} = 1$
"unconstrained pooling region":	$\tau_1^{*}\left(h_1 \leqslant \rho_1 \leqslant 1\right) = \tau_{UC}^{*};\ \psi_1^{*} = 1$

ii. If $\beta_{tr}/\beta_{op} \in \left[\left(\gamma+1\right)/\gamma, \gamma+1\right]$, $\tau_1^{*} \leqslant \tau_{UC}^{*}$ and the "credibility regions" are:

"mixed region":	$\tau_1^{*}\left(0 \leqslant \rho_1 \leqslant l_1^{-1}\right) = \bar{\tau}_1;\ \psi_1^{*} = \psi_1\left(\bar{\tau}_1, \rho_1\right)$
"constrained separating region":	$\tau_1^{*}\left(l_1^{-1} < \rho_1 < h_1^{-1}\right) = \hat{\tau}_1\left(\rho_1\right);\ \psi_1^{*} = \rho_1$
"unconstrained separating region":	$\tau_1^{*}\left(h_1^{-1} \leqslant \rho_1 \leqslant 1\right) = \tau_{UC}^{*};\ \psi_1^{*} = \rho_1$

iii. If $\beta_{tr}/\beta_{op} \geqslant \gamma+1$, the MPE is the same as case ii, except for the absence of the mixed region.

Households production is $\mu_1^* = [(1-\tau_1^*)\,\psi_1^*]^\gamma$ *in all cases and*

$$\bar{\tau}_1 = \frac{2-\beta_{tr}/\beta_{op}}{\gamma+1-\beta_{tr}/\beta_{op}}; \tag{10}$$

$$\tilde{\tau}_1(\rho_1) = 1-\left[\beta_{op}(1-\tau_{UC}^*)^\gamma \rho_1^\gamma\right]^{\frac{1}{\gamma+1}} \text{ and } \hat{\tau}_1(\rho_1) = 1-\left[\beta_{op}(1-\tau_{UC}^*)^\gamma \rho_1^{-\gamma}\right]^{\frac{1}{\gamma+1}}; \tag{11}$$

$$h_1 = \left[\frac{1-\tau_{UC}^*}{\beta_{op}}\right]^{1/\gamma} \text{ and } l_1 = \left[\frac{(1-\bar{\tau}_1)^{\gamma+1}}{\beta_{op}(1-\tau_{UC}^*)^\gamma}\right]^{1/\gamma} \tag{12}$$

Proof. When the unconstrained optimal rate τ_{UC}^* does not support either a separating or a pooling outcome, i.e. $\rho_1 \leqslant \min\{h_1, h_1^{-1}\}$, the optimal tax rate solving the problem of the trustworthy type as described in (7) is $\bar{\tau}_1$ if we ignore the boundary condition $\psi_1 \in [\rho_1, 1]$. This rate decreases with both γ and β_{tr}/β_{op}. It is higher than τ_{UC}^* when $\beta_{tr}/\beta_{op} \leqslant (\gamma+1)/\gamma$.

Imposing the boundary condition, $\bar{\tau}_1$ is thus feasible as long as it implies short-term credibility $\psi_1(\bar{\tau}_1, \rho_1)$ within $[\rho_1, 1]$, i.e. $\rho_1 \leqslant \min\{l_1, l_1^{-1}\}$. Therefore, *the MPE is mixed with planned tax rate* $\bar{\tau}_1$ *when* $\rho_1 \leqslant \min\{l_1, l_1^{-1}\}$. l_1 decreases with β_{op} but increases with β_{tr}/β_{op}.

Whenever the boundary condition of $\psi_1(\bar{\tau}_1, \rho_1)$ binds, the MPE is either pooling or separating, depending on which side of the boundary is binding. If $\psi_1(\bar{\tau}_1, \rho_1) < \rho_1$, i.e. $\rho_1 > l_1^{-1}$, a separating outcome is the best for the trustworthy type. If $\psi_1(\bar{\tau}_1, \rho_1) > 1$, i.e. $\rho_1 > l_1$, a pooling outcome is the best. The optimal tax plans to generate separating and pooling outcomes are the least deviations from the unconstrained optimal rate, i.e. $\hat{\tau}_1$ and $\tilde{\tau}_1$, as proved in Lemma 1. $\hat{\tau}_1$ increases with ρ_1 and $\tilde{\tau}_1$ decreases with ρ_1.

Now we can combine the results established above to prove the proposition.

$l_1 \leqslant 1$ and $h_1 \leqslant 1$ if $\beta_{tr}/\beta_{op} \leqslant (\gamma+1)/\gamma$ and $\beta_{op} \geqslant 1-\tau_{UC}^*$. In this case, there exists a credibility region $\rho_1 > l_1$ such that the MPE is pooling. Within this region, when long-term credibility is high enough – $\rho_1 > h_1$, τ_{UC}^* can induce a pooling outcome and is thus the optimal rate; when $\rho_1 \in [l_1, h_1]$, τ_{UC}^* does not support a pooling outcome and thus $\tilde{\tau}_1$ is the optimal rate.

$l_1 \geqslant 1$ and $h_1 \geqslant 1$ if $\beta_{tr}/\beta_{op} \geqslant (\gamma+1)/\gamma$. In this case, β_{op} is smaller than $1-\tau_{UC}^*$ since β_{tr} cannot be greater than 1. There thus exists a credibility region $\rho_1 > l_1^{-1}$ such that the MPE is separating. Within this region, the separating outcome is constrained and induced by $\hat{\tau}_1$ as long as $\rho_1 < h_1^{-1}$. Once long-term credibility exceeds h_1^{-1}, τ_{UC}^* becomes the optimal rate.

When $\beta_{tr}/\beta_{op} \geqslant \gamma+1$, $(\beta_{tr}/\beta_{op})\,\tau_{UC}^*$ is greater than 1. The trustworthy type thus puts more weight on future continuation values than the opportunistic type. Because decreasing the tax plan always raises the value of mimicking when $\gamma > 1$, it certainly benefits the trustworthy type. In addition, $\beta_{tr}/\beta_{op} \geqslant \gamma+1$ implies $\beta_{op} < 1-\tau_{UC}^*$. So in the region where an unconstrained outcome is feasible, the MPE is also separating. Therefore, with any level of long-term credibility, having a separating outcome is the dominant choice for the trustworthy type. ■

Figure 2 plots two representative policy functions $\tau_1^*(\rho_1)$ associated with cases i and ii. The left-hand panel is the case of equally patient government types: $\beta_{tr} = \beta_{op} > 1-\tau_{UC}^*$. The right-hand panel is the case where the trustworthy type is sufficiently more patient than the opportunistic type: $\beta_{tr}/\beta_{op} > (\gamma+1)\,\gamma^{-1}$. The elasticity of supply is $\gamma = 2$ so that the unconstrained optimal tax rate τ^*_{UC} is 1/3.

The most notable difference between these two functions is their opposite monotonicity. With equally patient government types, the optimal tax plan decreases as credibility increases, until it reaches the unconstrained optimal level τ^*_{UC}. However, with the opportunistic type less patient than the trustworthy one, the optimal tax plan increases with credibility until it reaches the unconstrained optimal level τ^*_{UC}.

This contrasting feature again stems from the difference in the extent to which the trustworthy type is willing to invest in credibility capital. When it is infeasible to build up reputation rapidly enough to compensate for the current revenue loss, the trustworthy type will use the tax plan to sustain high current revenues. In turn, the optimal tax plan needs to be higher than τ^*_{UC} to prevent the opportunistic type from confiscating the outputs so that short-term credibility is large enough to induce a decent participation rate. This is the case when $\beta_{tr} = \beta_{op}$. By contrast, when $\beta_{tr} > \beta_{op}$, the main goal of the tax plan is to increase the confiscation probability of the opportunistic type so as to stimulate rapid growth in long-term credibility. As a result, the optimal tax plan is lower than τ^*_{UC} in order to impose a higher cost on the opportunistic government, if it would mimic the trustworthy type's tax action.

This divergence in the purpose of the optimal tax plans naturally results in the opposite patterns of monotonicity. In both cases, the gap between the optimal tax plan and the unconstrained optimal level τ^*_{UC} shrinks as long-term credibility grows. This reflects the diminishing credibility concerns of the trustworthy type in designing an optimal tax plan, as its credibility capital accumulates.

What remains to be discussed is the case when $\beta_{tr}/\beta_{op} < (\gamma+1)\,\gamma^{-1}$ and $\beta_{op} < 1 - \tau^*_{UC}$. By and large, the overall pattern of the policy function still follows the left-hand panel in Figure 2. That is, the optimal tax plan decreases with long-term credibility, being at the rate $\bar{\tau}_1$ that induces a mixed outcome when ρ_1 is low and at the unconstrained rate τ^*_{UC} when ρ_1 is high.

However, it is a separating outcome that is induced by τ^*_{UC}, as a result of the relative impatience of the opportunistic type compared to *Case i*. In addition, constrained pooling outcomes can only be optimal with moderate long-term credibility if $l_1 \leqslant 1$, i.e. if β_{op} is relatively high or β_{tr}/β_{op} is relatively low.[15] This result is intuitive since pooling outcomes are only desirable for the trustworthy type when the incentive to mimic is relatively strong compared to the incentive to signal. In this case, the trustworthy type's choice of a pooling or a separating outcome depends not only on the

[15] If $l_1 \leqslant 1$, we have to compare the payoff to the trustworthy type associated with the unconstrained separating outcomes to both the mixed outcomes and the constrained pooling outcomes. This is because the region where the constrained pooling outcomes may be optimal is always a subset of the region where τ^*_{UC} can induce a separating outcome. To see this, start with $\beta_{tr}/\beta_{op} = (\gamma+1)/\gamma$, the highest possible level in the current case. The constrained pooling region then coincides with the unconstrained separating region. Now, if we decrease β_{tr}/β_{op}, a lower β_{op} is needed to keep l_1 constant. Hence, with the constrained pooling region fixed at $[l_1, 1]$, decreasing β_{tr}/β_{op} expands the unconstrained separating region. It can be shown that the unconstrained separating outcomes only dominate when the long-term credibility is above a cutoff $\bar{\rho}_1$, where $\bar{\rho}_1 > l_1$.

If $l_1 > 1$, a constrained separating region exists but it will be a subset of the unconstrained separating region. Because it is always optimal to use τ^*_{UC} to induce a separating outcome whenever possible, the MPE in this case will not include the constrained separating outcomes.

ratio of time preference β_{tr}/β_{op}, but also on long-term credibility ρ_1.

3.3.4 MPE when $\gamma \leqslant 1$

When households production is more rigid, $\gamma \leqslant 1$ so that $\partial\mu_1/\partial\tau_1 \geqslant 0$, and long-term credibility plays a bigger role in determining the property of a MPE. The following proposition characterizes the MPE. The proof is constructive.

Proposition 2 *Under the assumption $\gamma \leqslant 1$, there exists a unique MPE at period 1 with the following properties:*

i. With low $\rho_1 \leqslant L_0$, the MPE is "constrained pooling" $\{\tau_1^* = \tilde{\tau}_1(\rho_1), \psi_1^* = 1\}$.

ii. With moderate $\rho_1 \in [L_0, H_1]$, the MPE is "constrained pooling" except in the following cases:

a) if $\beta_{tr}/\beta_{op} \geqslant A_1(\beta_{op}, \gamma)$, the MPE is

$$\text{"mixed"}: \quad \tau_1^*(L_0 \leqslant \rho_1 \leqslant L_1) = 0;\ \psi_1^* = \psi_1(0, \rho_1);$$
$$\text{"constrained separating"}: \quad \tau_1^*(L_1 \leqslant \rho_1 \leqslant M_1) = \hat{\tau}_1(\rho_1);\ \psi_1^* = \rho_1.$$

b) if $\beta_{tr}/\beta_{op} \geqslant A_2(\beta_{op}, \gamma) \geqslant A_1(\beta_{op}, \gamma)$, the MPE is

$$\text{"constrained separating"}: \quad \tau_1^*(M_1 \leqslant \rho_1 \leqslant H_1) = \hat{\tau}_1(\rho_1);\ \psi_1^* = \rho_1.$$

iii. With high $\rho_1 \in (H_1, 1]$, τ_{UC}^* is always the equilibrium tax plan. The opportunistic type mimics for sure if $\beta_{op} \geqslant 1 - \tau_{UC}^*$ and confiscates for sure if $\beta_{op} \leqslant 1 - \tau_{UC}^*$.

Households production is $\mu_1^* = [(1 - \tau_1^*)\psi_1^*]^\gamma$ in all cases. L_0, L_1, M_1, H_1, A_1 and A_2 are functions of $\beta_{tr}/\beta_{op}, \beta_{op}$ and γ. Their specific forms are detailed in the appendix.[16]

Proof. First of all, $\bar{\tau}_1$ is no longer an optimal choice for the trustworthy type because it becomes an minimizer in this case.[17] The trustworthy type thus chooses between pooling and separating outcomes unless the tax rate hits the feasibility constraint: $[0, 1]$. L_1 is the lowest long-term credibility with which a constrained separating outcome can be induced by a non-negative tax rate. So L_1 is obtained by solving $\hat{\tau}_1(L_1) = 0$.

Within the region $[0, L_1]$, a constrained separating outcome is not feasible, so the choice is between a constrained pooling outcome and a mixed outcome with zero tax rate. L_0 is the cutoff below which the constrained pooling outcome dominates. This is *Property i.* When $\rho_1 \geqslant L_1$, the choice is between a constrained pooling outcome and a constrained separating one. M_1 is the cutoff below which the constrained pooling outcome dominates.

[16] Notice that $A_2(\beta_{op}, \gamma) \geqslant (\gamma + 1)/\gamma$ if and only if $\beta_{op} \leqslant 1 - \tau_{UC}^*$, with the equality holds at $\beta_{op} = 1 - \tau_{UC}^*$. Also, when $\beta_{op} \geqslant 1 - \tau_{UC}^*$, $A_2(\beta_{op}, \gamma) \geqslant 1/\beta_{op} \geqslant \beta_{tr}/\beta_{op}$, which is proved in the appendix. This excludes the possibility that a constrained separating region can be next to an unconstrained pooling region.

[17] Proof is in the appendix.

Finally, when ρ_1 is high enough that τ_{UC}^* can support either a separating or a pooling outcome, the choice is between a constrained and an unconstrained outcome. If both outcomes are pooling or separating, the unconstrained outcome is always dominant whenever it is feasible, as proved in Lemma 1. In these cases, the cutoffs are h_1 and h_1^{-1}, respectively. If one outcome is pooling and the other is separating, however, $\tilde{H}_1$ is the cutoff below which a constrained pooling outcome dominates, whereas $\hat{H}_1$ is the cutoff below which a constrained separating outcome dominates. Therefore, we obtain H_1 as specified in the appendix.

Cutoffs L_0, L_1, M_1, H_1 in the credibility space are functions of β_{op}, γ and β_{tr}/β_{op}, so all these 3 factors play important roles in determining the credibility regions. Figure 3 plots $L_0, L_1, M_1, \left\{\tilde{H}_1, \hat{H}_1\right\}$ on the space of $\left(\rho_1, \beta_{tr}/\beta_{op}\right)$. The left-hand panel depicts the case where $\beta_{op} \leqslant 1 - \tau_{UC}^*$ and the right-hand panel $\beta_{op} \geqslant 1 - \tau_{UC}^*$. From the plots, we can also view L_0, M_1, H_1 as cutoffs in the space of β_{tr}/β_{op} as functions of $\left\{\rho_1, \beta_{op}, \gamma\right\}$. There are some noticeable features common to both cases:

1) $\tilde{H}_1$ is smaller than M_1 for all ρ_1 and is tangential to M_1 at $\rho_1 = h_1^{-1}$, because τ_{UC}^* dominates all the other $\tilde{\tau}_1$ that induce pooling outcomes. Similarly, $\hat{H}_1$ is larger than M_1 and is tangential to M_1 at $\rho_1 = h_1$, because τ_{UC}^* dominates all the other $\hat{\tau}_1$ that induce separating outcomes. 2) It shows in the appendix that $\lim_{\rho_1 \to 1} \tilde{H}_1 = -\infty$ and $\lim_{\rho_1 \to 1} \hat{H}_1 = +\infty$. Therefore, unconstrained outcomes are always dominant if long-term credibility is high enough. This proves *Property iii.* 3) L_0 and M_1 intersects once at $\rho_1 = L_1$ where $\hat{\tau}_1 = 0$.

Using these features together with the plots in Figure 3, we can obtain the results summarized in *Property ii.* When $\beta_{tr}/\beta_{op} \leqslant M_1\left(\rho_1 = L_1, \beta_{op}, \gamma\right)$, constrained pooling outcomes are dominant unless $\rho_1 \geqslant H_1$. $A_1\left(\beta_{op}, \gamma\right)$ is thus defined as $M_1\left(\rho_1 = L_1, \beta_{op}, \gamma\right)$. When $\beta_{tr}/\beta_{op} \geqslant \lim_{\rho_1 \to 1} M_1\left(\rho_1, \beta_{op}, \gamma\right)$, constrained separating outcomes are dominant whenever they are feasible unless $\rho_1 \geqslant H_1$. Denote the limit of M_1 as $A_2\left(\beta_{op}, \gamma\right)$. When β_{tr}/β_{op} is between A_1 and A_2, a zero planned tax rate is optimal with low ρ_1, followed by the "constrained separating region" with moderate ρ_1 and the "constrained pooling region" with high ρ_1. ■

Propositions 1 and 2 have in common the fact that:

Corollary 1 *If $\beta_{tr} = \beta_{op}$, the trustworthy type never reduces its tax plan below τ_{UC}^* in equilibrium to invest in reputation.*

Only when the trustworthy type is sufficiently more patient than the opportunistic type does a mixed outcome with $\tau_1^* < \tau_{UC}^*$ or a constrained separating outcome become optimal. In addition, the MPE's property is invariant with β_{tr}/β_{op} with extreme values of long-term credibility ρ_1. Only with moderate ρ_1 does high β_{tr}/β_{op} induce separating outcomes in equilibrium.

In contrast to most cases with $\gamma > 1$, however, the MPE with $\gamma \leqslant 1$ often has a policy function $\tau_1^*\left(\rho_1\right)$ that is discontinuous when ρ_1 moves from one credibility region to another. Such high sensitivity of optimal tax plans to credibility levels is essentially a reflection of the high sensitivity of the opportunistic type's strategy to tax changes.

4 The Effect of Time Horizon – Solutions of a T-period Model

Denote n as the number of periods left until the terminal stage of the game, so $n = 0, 1, ..., T-1$ corresponds to $t = T, T-1, ..., 1$. When T is large enough, we can always divide the game into two parts: $\gamma > n$ and $\gamma \leqslant n$. Thus, the analysis in Section 3 can be applied accordingly.[18]

When there are more future periods, i.e. n is larger, the opportunistic type's incentive to mimic is strengthened. That is $\psi_n(\rho, \tau) \geqslant \psi_m(\rho, \tau)$ if $n \geqslant m$. The strengthened incentive to mimic makes it increasingly harder for the trustworthy type to signal. On the other hand, if the trustworthy type is more patient than the opportunistic type, the difference between types in terms of their future continuation values becomes more prominent when n is larger. In this case, the trustworthy type puts an increasingly larger weight on future continuation values than the opportunistic type, and is thus more willing to signal its type.

Which force is dominant again critically depends on the ratio of time discount factors β_{tr}/β_{op}. A low ratio $\beta_{tr}/\beta_{op} < B_1$, makes the trustworthy type yield to the mimicking by the opportunistic type and exert no effort to signal. B_1 is defined as

$$B_1 = \begin{cases} (\gamma + 1)/\gamma & \text{if } \gamma \geqslant 1 \\ A_1(\beta_{op}, \gamma) & \text{if } \gamma \leqslant 1. \end{cases}$$

This is the case that will be presented in Propositions 3 and 4. A high ratio $\beta_{tr}/\beta_{op} > B_2$, on the other hand, motivates the trustworthy type to invest heavily in reputation, knowing that it will later obtain a larger tax base. B_2 is defined as a cutoff in β_{tr}/β_{op} at $\bar{n}$ ($\bar{n} \geqslant \gamma > \bar{n} - 1$ if $\gamma > 1$ and $\bar{n} = 2$ if $\gamma \leqslant 1$), above which the MPE is *not* constrained pooling with moderate long-term credibility $\rho_{\bar{n}}$:

$$B_2 > \begin{cases} (\gamma + 1)/\gamma & \text{if } \gamma \geqslant 1 \\ A_2(\beta_{op}, \gamma) & \text{if } \gamma \leqslant 1 \end{cases}.$$

Proposition 5 will present such a case. These propositions will focus on the properties of time series along the equilibrium path instead of the policy function $\tau_n^*(\rho_n)$ at each horizon, because the latter has an overall pattern similar to its counterpart elaborated in Section 3 but with the exact shape dependent on details of the model.

Proposition 3 *If $\beta_{tr}/\beta_{op} < B_1$ and $\beta_{op} \geqslant 1 - \tau_{UC}^*$,*

i) The MPE in the early part of the game ($n \geqslant \gamma$) is pooling, starting with unconstrained pooling if T is large enough.

ii) The MPE in the later part of the game ($n < \gamma$) depends on the initial long-term credibility

[18] The relative sensitivity of the weak type compared to households to tax changes increases if there are more periods left in the game. That is:

$$\frac{\partial \mu_n(\rho, \tau)}{\partial \tau} > 0 \text{ if and only if } n > \gamma.$$

ε. With $\varepsilon \in [l_{k+1}, l_k)_{k=1}^{K \in (\gamma-2, \gamma-1]}$, the MPE is pooling at $n > k$ and is mixed at $n \leqslant k$ with

$$\bar{\tau}_n = \frac{n + 1 - n\beta_{tr}/\beta_{op}}{\gamma + 1 - n\beta_{tr}/\beta_{op}}. \tag{13}$$

With $\varepsilon \geqslant l_1$, the MPE is always pooling. l_k is the boundary between the "constrained pooling" and "mixed" region at $n = k$.

The planned tax rate in equilibrium thus starts with τ^*_{UC} if T is large enough and increases over time until $n \leqslant k$. Afterwards, it decreases over time until the end of the game.

Proof. The detailed proof is left to the appendix. A feature worth noticing is that when $\gamma > 1$, as in the case of $n = 1$, each period close to the end of the game involves three regions: "unconstrained pooling," "constrained pooling," and "mixed" regions. Both the "unconstrained" and "constrained" pooling regions expand to lower long-term credibility as we move backward in the game, i.e. more periods are left for the government to care about the future. Thus, the mixed region disappears at a certain, large enough, horizon, as does the "constrained pooling region." ■

Figure 4 plots the equilibrium time series $\{\tau^*_t, \rho^*_t, \psi^*_t, \mu^*_t\}_{t=1}^T$ from a numerical example with parameters $\gamma = 5$, $\beta_{tr} = \beta_{op} \geqslant 1 - \tau^*_{UC}$, $T = 20$ and $\varepsilon = 0.01$. To understand the equilibrium paths, it is useful to recall that the gap between the equilibrium tax plan from τ^*_{UC} captures the credibility concerns that the trustworthy type has when it suffers from imperfect credibility.

In this example, T is large enough so that a tax plan τ^*_{UC} can induce the opportunistic type to mimic for sure. Therefore, the game starts with an unconstrained pooling regime and with the trustworthy type keeping the tax plan at the unconstrained optimal level τ^*_{UC}. In this regime, the economy behaves as if the trustworthy type has full credibility. As time elapses, fewer future periods remain to reward the opportunistic type for maintaining its reputation. In order to sustain a pooling outcome, the trustworthy type has to reduce the opportunistic type's temptation to confiscate. This is done by raising the planned tax rate from the unconstrained optimal level τ^*_{UC}, because not only the gap in actual tax rates between mimicking and confiscating will shrink, but the production base μ will also drop, which will further decrease the gain from confiscation. Such a credibility concern for the trustworthy type becomes severer as we move towards the end of the game, so the equilibrium tax plan departs more and more from τ^*_{UC} and the production base continues to drop.

When it becomes too expensive for the trustworthy type to support a pooling outcome, the mixed regime takes over, with the opportunistic type randomizing between mimicking and confiscating. The randomization makes long-term credibility grow over time whenever the tax plan is implemented since such an event is now informative about the government's type. The growth in long-term credibility in turn mitigates the credibility concern in the optimal policy design. As a result, the equilibrium tax rate declines over time toward the unconstrained optimal level τ^*_{UC}. Short-term credibility also declines in this case as the opportunistic type's incentive to mimic fades away and the long-term credibility grows slowly. Driven by the declining short-term credibility, the production base μ continues to drop until the end of the game.

When $\beta_{op} < 1 - \tau^*_{UC}$, the equilibrium time series are similar if initial long-term credibility is low, except that an unconstrained pooling outcome is no longer feasible. Instead, an unconstrained separating outcome becomes optimal at each horizon n if long-term credibility is high. The presence of such an "unconstrained separating region" alters the MPE with respect to the case where $\beta_{op} \geqslant 1 - \tau^*_{UC}$:

Proposition 4 *If $\beta_{tr}/\beta_{op} < B_1$ and $\beta_{op} < 1 - \tau^*_{UC}$, the equilibrium time series depends on the initial long-term credibility ε:*

i) With low ε, the MPE is constrained pooling at $n \geqslant \gamma$ and is mixed with $\bar{\tau}_n$ defined by (13) at $n < \gamma$. The equilibrium path of the planned tax rate is thus the same as the one in Proposition 3.

ii) With moderate ε, the MPE is mixed first and unconstrained separating later, with the planned tax rate decreasing over time to τ^*_{UC}.

iii) With high ε, the MPE is unconstrained separating throughout the whole game.

Proof. Details of the proof are again in the appendix. However, it is worthwhile to point out that the credibility regions in this case differ significantly from those in the case where $\beta_{op} \geqslant 1 - \tau^*_{UC}$. More specifically, at each horizon $n \geqslant \gamma$, there are three credibility regions ranked by ρ from low to high: "constrained pooling", "mixed" and "unconstrained separating" regions. At each horizon $n < \gamma$, the constrained pooling region is replaced by a mixed region with $\bar{\tau}_n$ defined by (13). As we move towards the end of the game,.this mixed region with $\bar{\tau}_n$ defined by (13) expands to higher long-term credibility. But the unconstrained separating region remains the same throughout the whole game. ■

The common message from Propositions 3 and 4 is that when $\beta_{tr}/\beta_{op} < B_1$, all planned tax rates in equilibrium, τ^*_n, are larger than τ^*_{UC}. Notice also that $B_1 > 1$ at any level of elasticity of supply γ.[19] Therefore, Corollary 1 in Section 3 is robust to any finite-horizon game:

If $\beta_{tr} = \beta_{op}$, the trustworthy type never reduces its tax plan below τ^*_{UC} in equilibrium to invest in reputation.

This result can be understood by the amount that various government types are willing to pay in terms of current tax revenues for their reputation. If the trustworthy type shares a similar time preference with the opportunistic type, the short-run cost necessary to induce confiscation by the opportunistic type does not pay off for the trustworthy type in the long run, either. If the trustworthy type is sufficiently more patient, however, it can endure more current pain to inflate the short-run temptation for the opportunistic type to confiscate, and thus gain more credibility for higher tax revenues in the future. This is the case where investment in reputation by the trustworthy type can be optimal:

Proposition 5 *If $\beta_{tr}/\beta_{op} > B_2$, the equilibrium time series depends on the initial long-term credibility ε:*

[19] $\min A_1(\beta_{op}, \gamma) = \lim_{\rho_1 \to 0} M_1(\rho_1, \beta_{op}, \gamma) = \gamma + 1$.

i) When ε is not extremely low, the MPE is mixed in the early part of the game and is separating in the later part. The planned tax rate in equilibrium starts with a zero or low rate ($\tau_n^* < \tau_{UC}^*$) and increases over time to τ_{UC}^*.

ii) When ε is extremely low, the MPE in the early part of the game ($n \geqslant \gamma$) is constrained pooling with the planned tax rate increasing over time. After $n < \gamma$, the MPE is mixed first and separating later, with the planned tax rate starting below τ_{UC}^* and increasing over time to τ_{UC}^*.

Proof. Details of the proof are again in the appendix. Recall that this is the case where the increasing incentive to signal dominates the increasing cost with more future periods left in the game, i.e. a larger n. Therefore, when we move backwards to the beginning of the game, the trustworthy type is willing to sacrifice more current tax revenues to accumulate reputation, which implies a lower planned tax rate. However, if long-term credibility is extremely low and the number of future periods in the game is large, the production base will be too small for the opportunistic type to confiscate and give up the future streams of tax revenues. A constrained pooling outcome is thus the equilibrium until the later part of the game, when the opportunistic type is less concerned about maintaining it reputation. Only by then can the trustworthy type find it both feasible and desirable to invest in reputation using a planned tax rate lower than τ_{UC}^*. ■

Using the same parameterization as the previous numerical example, except for decreasing β_{op} so that $\beta_{tr}/\beta_{op} > (\gamma + 1)/\gamma$, I plot the equilibrium time series in Figure 5 as a contrast to Figure 4 where both government types are equally patient.

With a low but not extremely low level of initial credibility, the game starts with a mixed regime where the tax rate is kept at zero by the trustworthy type to invest in its reputation. A zero tax plan generates zero tax revenue when it is enacted. This high loss gives the plan low credibility in the initial periods of the game when the government is likely to be opportunistic, since the opportunistic type will deviate from the tough plan with a very high probability. As households have little confidence in being taxed at rate zero, not many of them decide to produce, i.e, the initial production base μ remains low. However, when tough tax plans are consecutively implemented for a number of periods, the evidence that the current government is a trustworthy type quickly accumulates. With more trust in the government's behavior, the production base rises.

After long-term credibility reaches a certain level, it can be maintained with lower investments. This takes place when the trustworthy type starts to raise the planned tax rate to reduce revenue losses. As households are more convinced that the current government is trustworthy, the planned tax rate becomes more credible. Hence, despite higher tax rates, the production base continues to rise and in turn increases tax revenues substantially. Long-term credibility is still growing, but at a lower rate than before.

The mixed regime with positive taxes will last until the trustworthy government decides to distinguish itself absolutely from the opportunistic type. Such an event will occur before the final period if the long-term credibility accumulates to a high point where the revenue cost of implementing such a separation is justifiable. After gaining full credibility, the trustworthy type can tax at the unconstrained optimal level, with the production base being at its peak as well. The government then achieves the best outcomes after all the earlier effort to overcome its lack of credibility.

As mentioned at the beginning of this section, with a larger n, the trustworthy type is more willing but also find it harder to signal its type. Neither of the offsetting forces dominates the other when β_{tr}/β_{op} lies between B_1 and B_2 so the properties of the MPE are less clear in this case. The equilibrium time series largely depend on how the equilibrium path of long-term credibility evolves over time. Nonetheless, if $\gamma > 1$, low initial long-term credibility is always associated with constrained pooling outcomes early in the game, whereas high initial long-term credibility is associated with early separating outcomes. The cutoff value of the initial long-term credibility is determined by the policy function at $\bar{n}$ where $\bar{n} \geqslant \gamma > \bar{n} - 1$.

5 Optimal Imitation in Announcement

In last two sections, I derived the optimal announcement under a restriction on the signaling strategy of the opportunistic government: it has to always imitate the announcement strategy of the trustworthy type. I will revisit this restriction and implied optimal policies in this section, where I study the within-period signalling game and obtain equilibrium strategies and beliefs. As a preview of the results, I find that (1) imitation is indeed the equilibrium signaling strategy of the opportunistic government, so that the restriction is not a binding one; and (2) the optimal announcement derived under the restriction coincides with the unique equilibrium in the signaling game, when refinement by Mailath, Okuna-Fujiwara and Postelwaite [1993] is applied. I thus complete the solution to the finitely-repeated game, determining both within-period signalling and the intertemporal evolution of credibility.

5.1 The signaling game

The within-period signaling game is between the government in place (now as the sender) and the households (now as the receivers). At the beginning of one period, the households have a prior belief on the government's type ρ, which is given by the updated credibility arising from last period's government tax action. The current government, with its type unobserved by the households, then sends a public message in terms of an announced tax plan τ.[20] Contemplating this message, the households update their beliefs of the current government's type $\phi(\tau)$ from ρ in a manner that this section studies in detail.[21] The new long-term credibility then serves as the base for the interaction between households' production $\mu(\phi, \tau)$ and the government's confiscation probability $\pi(\phi, \tau)$, if the opportunistic type is in place.

[20] I focus here, as above, on the situation in which an opportunistic government, if present, has not been revealed. There is no signalling game if type has been revealed, either for trustworthy or opportunistic.

[21] Note that it will turn out that $\phi(\tau)$ will be uniformative and hence there will be no updating after the message is received. But we are at present allowing for this *potential* updating.

5.1.1 The pure sequential equilibrium

As is standard in game theory literature, a suitable equilibrium concept for such a sender-receiver game with incomplete information is the "sequential equilibrium" proposed by Kreps and Wilson [1982].[22] More specifically, let us denote the governments' signalling strategies as $m(TR)$ and $m(OP)$ for trustworthy and opportunistic types, respectively.

Definition 2 *The strategies and beliefs* $\{m(TR), m(OP), \mu(\phi,\tau), \phi(\tau)\}$ *form a pure "sequential equilibrium" within any time period if:*[23],[24]

D.3.1) $m(TR)$ *maximizes the trustworthy type's payoff:*

$$m(TR) = \arg\max_{\tau} \tau\mu(\phi,\tau) + \beta_{tr} W'(\phi/\psi)$$

D.3.2) $m(OP)$ *maximizes the opportunistic type's payoff:*

$$m(OP) = \arg\max_{\tau} [1-\pi(\phi,\tau)] \left[\tau\mu(\phi,\tau) + \beta_{op} V'(\phi/\psi)\right] + \pi(\phi,\tau)\mu(\phi,\tau)$$

D.3.3) $\mu(\phi,\tau)$ *maximizes the households' expected payoff:*

$$\mu(\phi,\tau) = G[(1-\tau)\psi]$$

where $\psi = \phi + (1-\phi)[1-\pi(\phi,\tau)]$.

D.3.4) $\phi(\tau)$ *is formed in a Bayesian fashion consistent with the strategies*

$$\phi(\tau) = \Pr(TR|\tau) = \frac{\rho \Pr(\tau|TR)}{\rho \Pr(\tau|TR) + (1-\rho)\Pr(\tau|OP)}$$

where

$$\Pr(\tau|TR) = \begin{cases} 1 & \textit{if } m(TR) = \tau \\ 0 & \textit{otherwise} \end{cases} ; \ \Pr(\tau|OP) = \begin{cases} 1 & \textit{if } m(OP) = \tau \\ 0 & \textit{otherwise} \end{cases}$$

5.1.2 The equilibrium is always pooling

The signaling strategy restriction $m(OP) = \arg\max_\tau \tau\mu(\phi,\tau) + \beta_{tr} W'(\phi/\psi)$ used in previous sections essentially imposes the belief function $\phi(\tau) = \rho$ for all τ because the opportunistic government perfectly imitates the announcement of the trustworthy type and thus makes the message

[22] Although sequential equilibrium may be a further refinement of perfect Bayesian equilibrium, the difference between these two concepts is not relevant in the current context. But following the classic papers (Krep and Wilson 1982, Grossman and Perry 1986, Mailath et al 1993) this work has been build upon, I keep using "sequential equilibrium" as the main equilibrium concept.

[23] I leave out the opportunistic type's strategy for its confiscation probability $\pi(\phi,\tau)$ in this definition because: (1) it is not part of the signaling equilibrium as it is neither the sender's signaling strategy nor the receiver's best reaction; (2) it is fully determined by the MPE which I derived in previous sections so long as ϕ replaces ρ; (3) its information is incorporated by the construction of short-term credibility, which enters in the payoff functions of both governments and households.

[24] I focus only on pure signaling strategies here. That is, the sender is not allowed to randomize over several messages.

uninformative. However, even when the imitation restriction is removed, any equilibrium of the signaling game must still be a pooled one, in which both types send the same message.

To see it, suppose $m(TR) = \bar{\tau}$ and $m(OP) = \tilde{\tau} \neq \bar{\tau}$. Then, the opportunistic type's identity is perfectly revealed if households observe message $\tilde{\tau}$: that is, $\phi(\tilde{\tau}) = 0$. In turn, the opportunistic type will confiscate for sure as the future credibility will never grow again. Anticipating this confiscation, no household will produce, which implies a zero payoff of the opportunistic type. Given this consequence of sending a message $\tilde{\tau} \neq m(TR)$, the opportunistic type can always improve its welfare by deviating to the same message as the trustworthy government. Hence, distinct messages from trustworthy and opportunistic types can never occur in the equilibrium. In other words, any equilibrium of this signalling game involves the opportunistic government imitating the trustworthy type's message-sending. Therefore, in equilibrium, $m(TR) = m(OP) = \bar{\tau}$ and $\phi(\bar{\tau}) = \rho$.

5.1.3 Multiple pooling equilibria

Thus, removing the previous restriction on opportunistic-type signaling strategy does not affect any of the equilibrium outcomes. However, one still must determine the equilibrium message and the requirement $m(TR) = m(OP)$ leaves the out-of-equilibrium beliefs unspecified. A notorious consequence of such unregulated out-of-equilibrium beliefs is a continuum of sequential equilibria that makes the optimal announcement τ^* derived previously only one of many possible outcomes.

For example, if the households somehow interpret all other messages but $\tau = a$ as indicating the current government is opportunistic, it is indeed optimal for both opportunistic and trustworthy governments to send message a in equilibrium. Therefore, any message $\tau \in [0, 1]$ can be supported in a sequential equilibrium by properly specifying the out-of-equilibrium beliefs.

5.2 Disciplining off-equilibrium beliefs

In the face of this indeterminacy problem, many equilibrium refinements have been developed to obtain sharper predictions in signaling games. It turns out that the game under study, in which it is costless for the opportunistic type to imitate the trustworthy one, is somewhat problematic for many standard refinements. However, the approach of Mailath, Okuna-Fujiwara, and Postelwhaite [1993] turns out to be an appropriate approach to this class of games and, further, selects a unique equilibrium coincident with the optimal policy prediction derived in previous sections.

In their work, Mailath, Okuno-Fujiwara, and Postelwaite elaborate the vision of Grossman and Perry [1986] of using Bayesian reasoning to discipline out-of-equilibrium beliefs. Mailath et.al. postulate three key ingredients in constructing belief restrictions. First, any out-of-equilibrium message must be one that is sent in an alternative sequential equilibrium by some set of agents. Second, the incentives that various types of agents have to send an alternative message τ are evaluated by comparison of benefits in a candidate and alternative equilibrium. Third, when such comparison induces an out-of-equilibrium probability distribution of sender types computed by Bayes' law, it is then used to generate restrictions on beliefs.

To discuss the approach of Mailath et. al., I find it useful to define "strongly coherent beliefs" as follows:[25]

Definition 3 *For an economy with a set of sender types* Ω*, a candidate equilibrium* $\hat{m}, \hat{\mu}, \hat{\phi}$ *and an alternative equilibrium* $\tilde{m}, \tilde{\mu}, \tilde{\phi}$*, the message* τ *gives rise to a "strongly coherent out-of-equilibrium belief"* φ *about a subset of types if*

D.4.1) $\forall \omega \in \Omega$, $\hat{m}(\omega) \neq \tau$*, there is a non-empty set* $K = \{\omega \in \Omega | \tilde{m}(\omega) = \tau\}$*;*

D.4.2) $\forall \omega \in K$, $R\left(\left\{\tilde{m}, \tilde{\mu}, \tilde{\phi}\right\}, \omega\right) \geqslant R(\{\hat{m}, \hat{\mu}, \hat{\phi}\}, \omega)$ *and* $\exists \omega \in K$, $R\left(\left\{\tilde{m}, \tilde{\mu}, \tilde{\phi}\right\}, \omega\right) > R(\{\hat{m}, \hat{\mu}, \hat{\phi}\}, \omega)$ *where* $R(\cdot)$ *is the payoff function of the senders;*

D.4.3) $\forall \tilde{\omega} \in K$,

$$\varphi(\tilde{\omega}|\tau) = \frac{\rho(\tilde{\omega}) \Pr(\tau|\tilde{\omega})}{\sum_{\omega \in \Omega} \rho(\omega) \Pr(\tau|\omega)}$$

where $\rho(\tilde{\omega})$ *is the prior belief of type* $\tilde{\omega}$ *and* $\Pr(\tau|\omega)$ *specifies the probability that a sender of type* ω *would issue the message* τ *so that*

$$\Pr(\tau|\omega) = \begin{cases} 1 & \text{if } \omega \in K \text{ and } R\left(\left\{\tilde{m}, \tilde{\mu}, \tilde{\phi}\right\}, \omega\right) > R(\{\hat{m}, \hat{\mu}, \hat{\phi}\}, \omega) \\ 0 & \text{if } \omega \notin K \\ [0,1] & \text{otherwise} \end{cases} \quad (14)$$

Using this definition, I can state the "Mailath, Okuno-Fujiwara, and Postelwaite Refinement" as:

Definition 4 (Mailath, Okuno-Fujiwara, and Postelwaite Refinement) *A sequential equilibrium* $\hat{m}, \hat{\mu}, \hat{\phi}$ *is "defeated" by an alternative equilibrium* $\tilde{m}, \tilde{\mu}, \tilde{\phi}$ *if there exists a type* $\tilde{\omega}$ *and a message* $\tilde{\tau} = \tilde{m}(\tilde{\omega})$ *send by* $\tilde{\omega}$ *in the alternative equilibrium, such that any strongly coherent belief on message* $\tilde{\tau}$ *is inconsistent with the supporting belief of the candidate equilibrium:*

$$\varphi(\tilde{\omega}|\tilde{\tau}) \neq \hat{\phi}(\tilde{\omega}|\tilde{\tau}) \text{ for any } \Pr(\tilde{\tau}|\tilde{\omega}) \text{ satisfying (14)}$$

Essentially, this refinement says that the candidate equilibrium is eliminated if it cannot be supported by any strongly coherent out-of-equilibrium belief, when such a belief exists.

5.3 Illustration in a static setting

To see how this equilibrium refinement of Mailath et. al. selects the unique equilibrium τ^*, I first use a static version of the current signaling game, which allows a simple graphical description of the key ideas. I will show in next subsection that the same ideas apply to the original game.

[25] This label comes from my parallel work with King and Pasten [2008a] on a related model of costless imitative signalling. There, we also study the out-of-equilibrium beliefs restriction proposed by Grossman and Perry [1986] which we call "weakly coherent belief." Use of these belief definitions allows us to state these two alternative approaches in a common manner, highlighting differences in beliefs.

In the static signaling game, the opportunistic government always confiscates and the payoffs of both types reduce to the momentary tax revenues:

$$R(\tau, \mu, TR) = \tau\mu;\; R(\tau, \mu, OP) = \mu$$

An indifference curve in (τ, μ) space is thus $\mu = w/\tau$ for the trustworthy type with payoff level w, and $\mu = v$ for the opportunistic type with payoff level v. I draw examples of these indifference curves in Panel A of Figure 6 with dashed lines. In each panel, the horizontal axis is the tax announcement τ and the vertical axis is the participation rate μ. The indifference curve of the trustworthy type $\mu = w/\tau$ is decreasing and strictly convex in τ. Given the announced tax plan τ, higher levels of the trustworthy type's payoff correspond to an upward shift of the indifference curve. By contrast, the indifference curve of the opportunistic type is independent of the announced plan τ so that it is the horizontal line in the panel. Higher participation rate μ increases the payoff level of the opportunistic type, and in turn shifts its indifference curve upwards.

To facilitate the drawing of the graph and as in the early part of section 4 above, I further specialize the elasticity of supply $\gamma = 1$ so that

$$\mu(\phi, \tau) = (1 - \tau)\,\phi(\tau)$$

specifies the participation rate depending on the belief ϕ upon receiving different message τ. This participation function then fully reflects the out-of-equilibrium beliefs, which are the central focus of this section. I use dotted lines in the Figure to draw such belief-based participation functions.

In addition, the continuum of pooling sequential equilibria can be captured by the solid line $\mu = (1 - \tau)\,\rho$ in all the panels, where ρ is the prior belief of the current government being trustworthy. Any point a on this solid line can be an equilibrium by properly specifying the out-of-equilibrium beliefs. One example is shown in Panel B, where the households believe that any $\tau \neq a$ is sent by the opportunistic government: $\phi(a) = \rho$ and $\phi(\tau \neq a) = 0$.

The solid line is also the restriction of signaling strategy, which we imposed to derive the optimal announcement τ^* in section 3 and 4 because $\phi(\tau) = \rho$ for all τ. Hence, Panel A shows the determination of τ^* in the static context where $\tau^* = \arg\max \tau\mu$ subject to $\mu = (1 - \tau)\,\rho$. In the discussion below, I will apply the refinement of Mailath et al to rule out any pooling equilibrium above and below τ^*, and further establish the "undefeated" feature of τ^* so that it is the unique signaling equilibrium surviving this refinement.

5.3.1 Ruling out high tax equilibria

As shown in Panel C in Figure 6, a candidate equilibrium with high tax announcement $h > \tau^*$ must be supported by beliefs that make τ other than h less desirable for both the trustworthy and opportunistic governments. Reflected in the graph, the belief-implied participation rates must lie below the indifference curves through h for both types. In particular, to avoid the trustworthy government strictly prefer τ^* over h, the supporting belief on τ^* must imply a participation rate lower than $\hat{\mu}$, which is the exact point on the indifference curve at message τ^*. Apparently, the requirement on $\phi(\tau^*)$ is: $(1-\tau^*)\phi(\tau^*) \leqslant \hat{\mu} < (1-\tau^*)\rho$ and, in turn,

$$\phi(\tau^*) < \rho.$$

However, such belief function is not strongly coherent with the senders' incentives in an alternative equilibrium $\tau^* < h$. For the opportunistic government, it is strictly better off by sending τ^* as its indifference curve through τ^* is above the one through h. For the trustworthy government, deviating to τ^* also strictly increases its payoff because τ^*, by construction, is the unique revenue maximizer when the belief on the equilibrium message is ρ. Therefore, the strongly coherent belief on receipt of the the out-of-equilibrium message τ^* is $\varphi(\tau^*) = \rho$ according to the Bayes' law in D.4.3). In other words, the supporting belief of the candidate equilibrium h such that $\phi(\tau^* < h) < \rho$ is not strongly coherent. Thus, $h > \tau^*$ is *defeated* by the alternative equilibrium τ^*.

5.3.2 Ruling out low tax equilibria

Panel D in Figure 6 shows a candidate equilibrium with a low tax announcement $l < \tau^*$ and its supporting out-of-equilibrium beliefs. Applying the same reasoning as above, to avoid the trustworthy type deviate from l to τ^*, the supporting belief on message τ^* has to satisfy:

$$\phi(\tau^*) < \rho.$$

Again, this belief is not strongly coherent with the senders' incentive in an alternative equilibrium τ^*. The trustworthy government is strictly better off by deviating to τ^* but the opportunistic type is not as the participation rate in the candidate equilibrium l is higher. Thus, the only strongly coherent belief on τ^* – $\varphi(\tau^*)$ – is equal to 1, which is inconsistent with the $\phi(\tau^*) < \rho$ requirement for the support of l. Thus, any candidate equilibrium with low tax announcement $l < \tau^*$ is *defeated* by the alternative equilibrium τ^*.

5.3.3 Existence

What remains to be shown is that the candidate equilibrium τ^* survives such refinement – that is, τ^* is undefeated by any alternative equilibrium with tax announcement h or l.

To see this result, first notice that the candidate equilibrium τ^* can be supported by the out-of-equilibrium beliefs

$$\begin{aligned} \phi(\tau) &= 0 \text{ for } \tau < \tau^* \\ \phi(\tau) &= \rho \text{ for } \tau \geqslant \tau^* \end{aligned}$$

Consider an alternative equilibrium $h > \tau^*$; an opportunistic government does not have incentive to deviate from τ^* to h because the alternative equilibrium involves lower participation rate and thus less tax revenues for it. A trustworthy government will not deviate to h either, since τ^* generates strictly higher revenues in the candidate equilibrium than h does in the alternative equilibrium: $\tau^*(1-\tau^*)\rho > h(1-h)\rho$. If neither government will send message h, no restrictions are placed on the out-of-equilibrium beliefs at h.

Now consider an alternative equilibrium $l < \tau^*$, which has a higher participation rate so that the opportunistic government will find it desirable to send such a message. However, the trustworthy government will not because the τ^* is already the revenue-maximizer given the equilibrium belief of ρ. Hence, the strongly coherent belief on message $l < \tau^*$ is 0, which coincides with the supporting belief function $\phi(\tau) = 0$ for $\tau < \tau^*$.

In sum, the equilibrium τ^* is the only survivor of the Mailath et. al. refinement.

5.4 Generalization to the dynamic setting

Having displayed the power of this refinement approach in the static setting, I now argue that the same idea works in the original game.

In the dynamic context, the optimal announcement τ^* derived in sections 3 and 4 is essentially the solution to the following constrained maximization problem after imposing the out-of-equilibrium belief function $\phi(\tau) = \rho$ for all τ:

$$\begin{aligned} \tau^* &= \arg\max_{\tau} \tau\mu(\rho,\tau) + \beta_{tr} W'(\rho/\psi) \qquad (15) \\ \text{subject to } \mu(\rho,\tau) &= G[(1-\tau)\psi] \end{aligned}$$

with endogenous $\psi(\rho,\tau)$ being determined by $\tau\mu(\rho,\tau) + \beta_{op} V'(\rho/\psi) = \mu(\rho,\tau)$ or its boundary value $\{\rho, 1\}$.

There are two complications added by having the governments live more than one period in the signaling game. One is that the payoff functions of governments now include the future continuation values. The other is possible mimicking of the opportunistic type so that the short-term credibility ψ is not necessarily equal to the households' beliefs about the type ϕ. However, neither changes the essence of the equilibrium selection.

5.4.1 Any candidate equilibrium $\tau \neq \tau^*$ is defeated

From the argument in the static setting, the key ingredient in ruling out equilibria above and below τ^* is the inconsistency between the supporting out-of-equilibrium belief on message τ^*: $\phi(\tau^*) < \rho$ and the strongly coherent belief on τ^*: $\varphi(\tau^*) \geqslant \rho$. In the dynamic setting, we still need $\phi(\tau^*) < \rho$ as a requirement of the supporting beliefs of any candidate equilibrium $\tau \neq \tau^*$, because $\tilde{\phi}(\tau^*) = \rho = \phi(\tau \neq \tau^*)$ generates strictly higher payoffs for the trustworthy type by the construction of τ^* so that it will induce the government deviate from τ to τ^*. Similarly, any belief $\tilde{\phi}(\tau^*) > \rho$ will also induce the deviation as the payoff of the trustworthy type always increases with the long-term credibility ϕ.

On the other hand, the strongly coherent belief on τ^*: $\varphi(\tau^*) \geqslant \rho$ is also preserved in the dynamic setting. This stems from the fact that the trustworthy government has strictly higher payoff in the alternative equilibrium τ^* than in the candidate equilibrium $\tau \neq \tau^*$ by the construction of τ^*. Then, at least the trustworthy type will have incentive to send the out-of-equilibrium message τ^*: $\Pr(\tau^*|TR) = 1$. Regardless of the incentive of the opportunistic type, Bayes' law implies $\varphi(\tau^*) \geqslant \rho$.

Therefore, any candidate equilibrium $\tau \neq \tau^*$ in the dynamic setting is defeated by the alternative equilibrium τ^* as its supporting belief cannot be strongly coherent with the governments' incentives in sending message τ^*, and hence all alternative equilibrium must be discarded from the equilibrium set.

5.4.2 Candidate equilibrium τ^* is undefeated

The undefeated nature of the candidate equilibrium τ^* in the static setting stems from two facts: (1) such an equilibrium can be supported by any belief $\phi(\tau) \leqslant \rho$ for all τ with $\phi(\tau^*) = \rho$; and (2) the strongly coherent beliefs on any message $\tau \neq \tau^*$ are consistent with this class of supporting beliefs.

Both key facts hold in the dynamic setting. For the first one, because τ^* is the unique payoff-maximizer for the constrained optimization (15) under the restriction $\phi(\tau) = \rho$ for all τ, the payoff of the trustworthy type by sending τ^* is strictly higher than the payoff by sending $\tau \neq \tau^*$. As the payoff decreases with long-term credibility ϕ, sending τ^* is certainly the best choice for the trustworthy type when $\phi(\tau) \leqslant \rho = \phi(\tau^*)$.

For the second fact, by the construction of τ^*, the trustworthy type has no incentive to send any message $\tau \neq \tau^*$ because its payoff in the candidate equilibrium τ^* is strictly higher than in any alternative equilibrium. Hence, for messages $\tilde{\tau}$ that only the opportunistic type finds it worthwhile to deviate to, the strongly coherent belief is $\varphi(\tilde{\tau}) = 0$, consistent with the supporting beliefs. For messages that even the opportunistic type does not want to send, there are no strongly coherent beliefs defined, and hence no restriction imposed on the supporting beliefs.

All in all, the candidate equilibrium τ^* is undefeated by any alternative equilibrium $\tau \neq \tau^*$ in the dynamic setting so that it is the unique signaling equilibrium surviving the refinement of Mailath, Okuna-Fujiwara and Postelwaite [1993].

6 Conclusions and Remarks

The paper has presented a simple reputation game where a trustworthy government can use policy announcement as an instrument to accommodate public doubts concerning its ability to pre-commit, or to facilitate public learning of its true identity. The unique Markov perfect equilibrium of the game reveals that it is optimal for a trustworthy government to separate itself from an opportunistic government only when the trustworthy type is sufficiently more patient. If this condition does not hold, it will be too expensive for a trustworthy type to induce a separating outcome through active policymaking, and thus reputation building will never be in equilibrium. This partially explains why optimal reputation building by a trustworthy government, which is often seen in practices, has not been adequately addressed in the literature since both types have usually been assumed to be equally patient.

In fact, assuming a trustworthy type to be more patient than an opportunistic type is more natural if we consider government type to be an endogenous choice based on time preference. Considering that it is costly to obtain a commitment device, only a more patient government will pay the cost and become trustworthy because, as the model predicts, it will have more to gain from being able to commit.[26]

Although not formally proven here, the reasoning that reputation-building by a trustworthy government cannot be a MPE when both types are equally patient should generalize to other models even when the preferences of trustworthy and opportunistic types differ along other dimensions. No matter how different payoff functions are across types, as long as types weigh current payoffs against future payoffs in a similar fashion, the short-run cost of inducing the opportunistic type to give up its reputation will not pay off for the trustworthy type in the long run, either.

References

[1] Backus, David A, and John Driffill. (1985a) "Inflation and Reputation." *American Economic Review*, 75:3, 530-38.

[2] Backus, David A, and John Driffill. (1985b). "Rational Expectations and Policy Credibility Following a Change in Regime." *The Review of Economic Studies*, 52:2, 211-221.

[3] Barro, Robert J, (1986) "Reputation in a Model of Monetary Policy with Incomplete Information", *Journal of Monetary Economics* 17, 3-20.

[4] Barro, Robert J, and David B. Gordon, (1983) "Rules, discretion and reputation in a model of monetary policy", *Journal of Monetary Economics* 12, 101-122.

[5] Celentani, Macro and Wolfgang Pesendorfer, (1996) "Reputation in dynamic games", *Journal of Eonomic Theory* 70, 109-132

[6] Chari, V. V. and Patrick J. Kehoe, (1990) "Sustainable Plan", *Journal of Political Economy* 98, No.4, 783-802.

[7] Chari, V. V. and Patrick J. Kehoe, (1989) "Sustainable Plan and Debt", working paper, Federal Reserve Bank of Minneapolis.

[8] Chari, V. V., Lawrence J. Christiano and Martin Eichenbaum, (1998) "Expectation Traps and Discretion", *Journal of Economic Theory* 81, 462-492.

[26] Note that the comparison is between one scenario where no type buys the commitment device so that the private sector produce nothing in the MPE, and the other scenario where governments with certain time preference buy the commitment device but the time preference is private information.

[9] Cole, Harold L., James Dow, and William B. English, (1995) "Default, Settlement, and Signalling: Lending Resumption in a Reputational Model of Sovereign Debt", *International Economic Review* 36, 365-385.

[10] Cole, Harold L. and Patric J. Kehoe, (1998) "Models of Sovereign Debt: Partial versus General Reputations", *International Economic Review* 39, 55-70.

[11] Cukierman, Alex, (1993) *Central Bank Strategy, Credibility, and Independence*, MIT press.

[12] Cukierman, Alex, and Nissan Liviatan, (1991) "Optimal Accommodation by Strong Policymakers under Incomplete Information." *Journal of Monetary Economics*, 27:1, 99-127.

[13] Cukierman, Alex, and Allan H. Meltzer, (1986) "A Theory of Ambiguity, Credibility, and Inflation under Asymmetric Information" *Econometrica* 54, 1099-1128.

[14] D'Erasmo, Pablo, (2007) "Government Reputation and Debt Repayment in Emerging Economics", mimeo, UTAustin.

[15] Doraszelski, Ulrich and Mark Satterthwaite, (2007) "Computable Morkov-Perfect Industry Dynamics: Existence, Purification, and Multiplicity", working paper, Harvard Univeristy.

[16] Fudenberg, Drew and David K. Levine, (1983) "Subgame-Perfect Equilibria of Finite- and Infinite-Horizon Games", *Journal of Economic Theory* 31, 251-268.

[17] Fudenberg, Drew and David K. Levine, (1989) "Reputation and Equilibrium Selection in Games with a Patient Player", *Econometrica* 57, No.4, 759-778.

[18] Grossman, Sanford J, and Motty Perry. (1986) "Perfect Sequential Equilibrium.", *Journal of Economic Theory*, 39:1, 97-119.

[19] King, Robert G., Yang K. Lu and Ernesto Pasten (2008a) "Coherent beliefs with costless imitative signalling", work in progress.

[20] King, Robert G., Yang K. Lu and Ernesto Pasten (2008b) "Managing Expectations", *Journal of Money, Credit and Banking*, forthcoming.

[21] Klein, Paul, Per Krusell and Jose-Victor Rios-Rull, (2006) "Time-Consistent Public Policy", working paper, Princeton University.

[22] Kreps, David M. and Robert Wilson, (1982a) "Sequential Equilibria", *Econometrica* 50, 863-894.

[23] Kreps, David M. and Robert Wilson, (1982b) "Reputation and Imperfet Information", *Journal of Economic Theory* 27, 253-279.

[24] Kydland, Finn E. and Edward C. Prescott, (1977) "Rules rather than discretion", *Journal of Political Economy* 85, 473-491.

[25] Mailath, George J, Masahiro Okuno-Fujiwara and Andrew Postlewaite. (1993) "Belief-Based Refinements in Singnalling Games", *Journal of Economic Theory*, 60, 241-276.

[26] Mailath, George J. and Larry Samulson, (2001) "Who Wants a Good Reputation?", *Review of Economic Studies* 68, 415-441.

[27] Maskin, Eric and Jean Tirole, (2001) "Markov Perfect Equilibrium: I. Observable Actions", *Journal of Economic Theory* 100, 191-219.

[28] Milgrom, Paul and John Roberts, (1982) "Predation, Reputation, and Entry Deterrence", *Journal of Economic Theory* 27, 280-312.

[29] Ireland, Peter N., (1999) "Expectations, Credibility, and Time-Consistent Monetary Policy", NBER working paper.

[30] Phelan, Chirstopher, (2006) "Public trust and Government Betrayal", *Journal of Economic Theory* 130, 27-43.

[31] Rogoff, Kenneth, (1985) "The Optimal Degree of Commitment to an Intermediate Monetary Target." *Quarterly Journal of Economics* 100, 1169-1190.

[32] Sleet, Christopher and Sevin Yeltekin, (2007) "Recursive monetary policy games with incomplete information", *Journal of Economic Dynamics and Control* 31, 1557-1583.

[33] Walsh, Carl E., (2003) *Montary Theory and Policy*, MIT press.

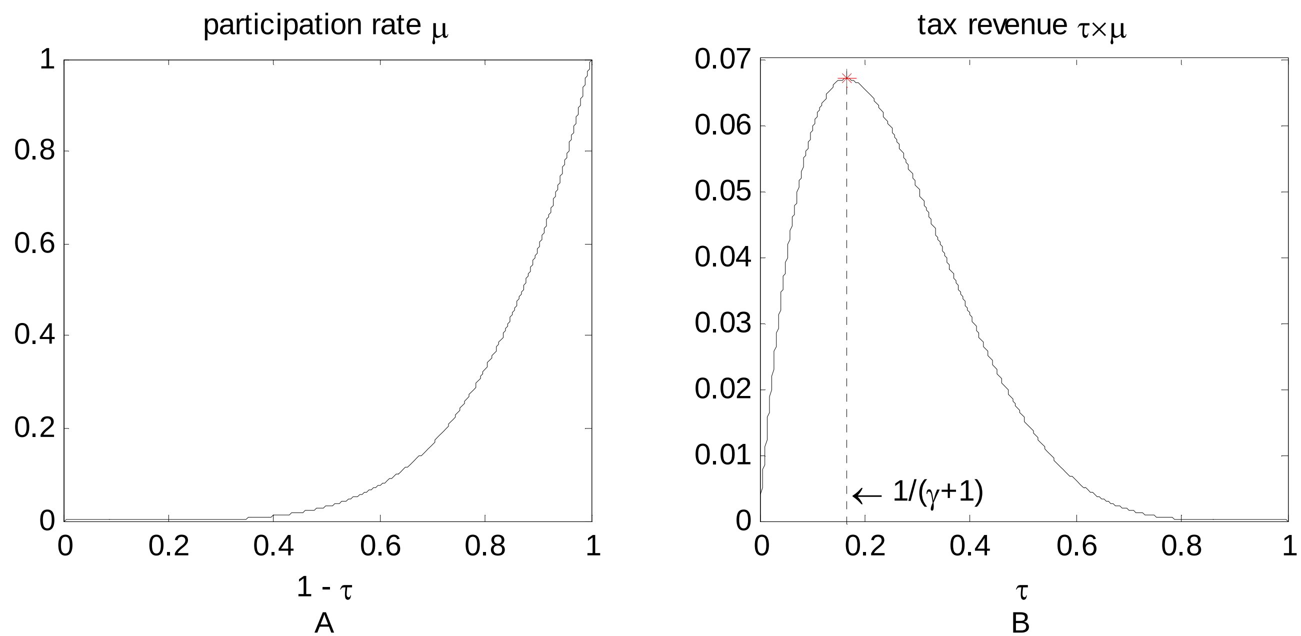

Figure 1**:** Participation rate and Laffer curve.
The parameter values are: c=0.4; ψ=1; γ=5.

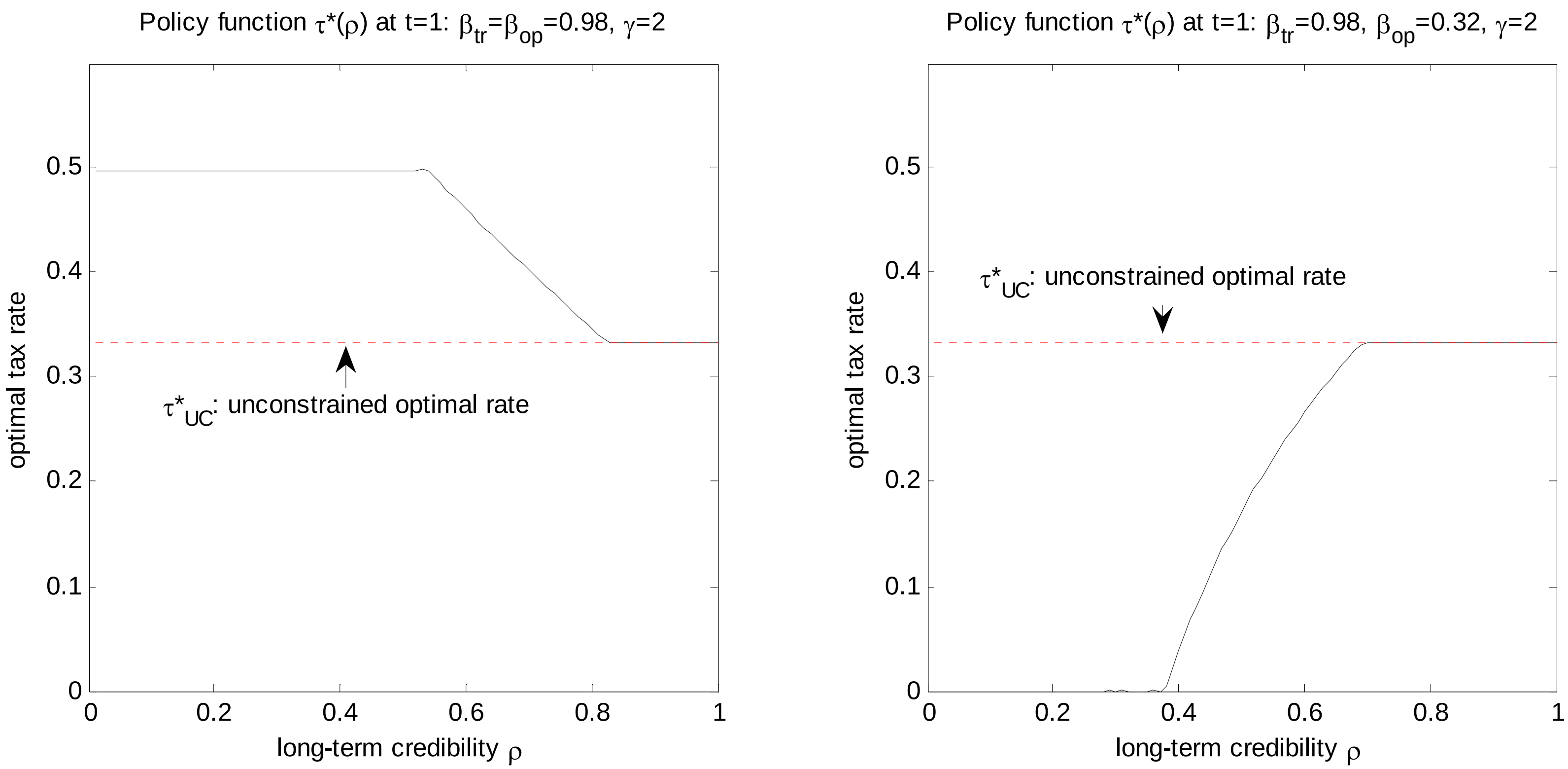

Figure 2: Policy functions at t=1 with γ=2

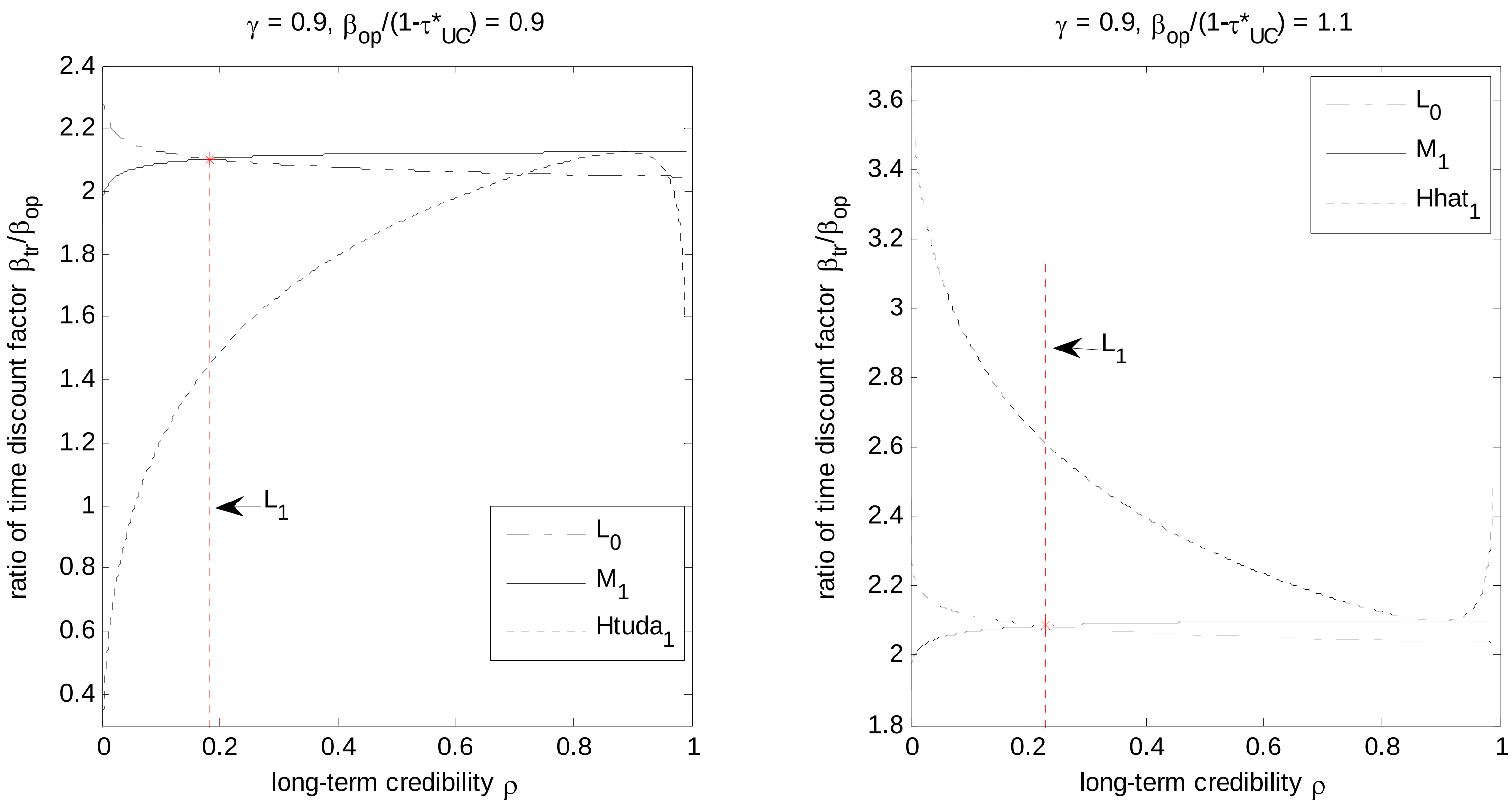

Figure 3: Cutoff values at t=1 with γ<1

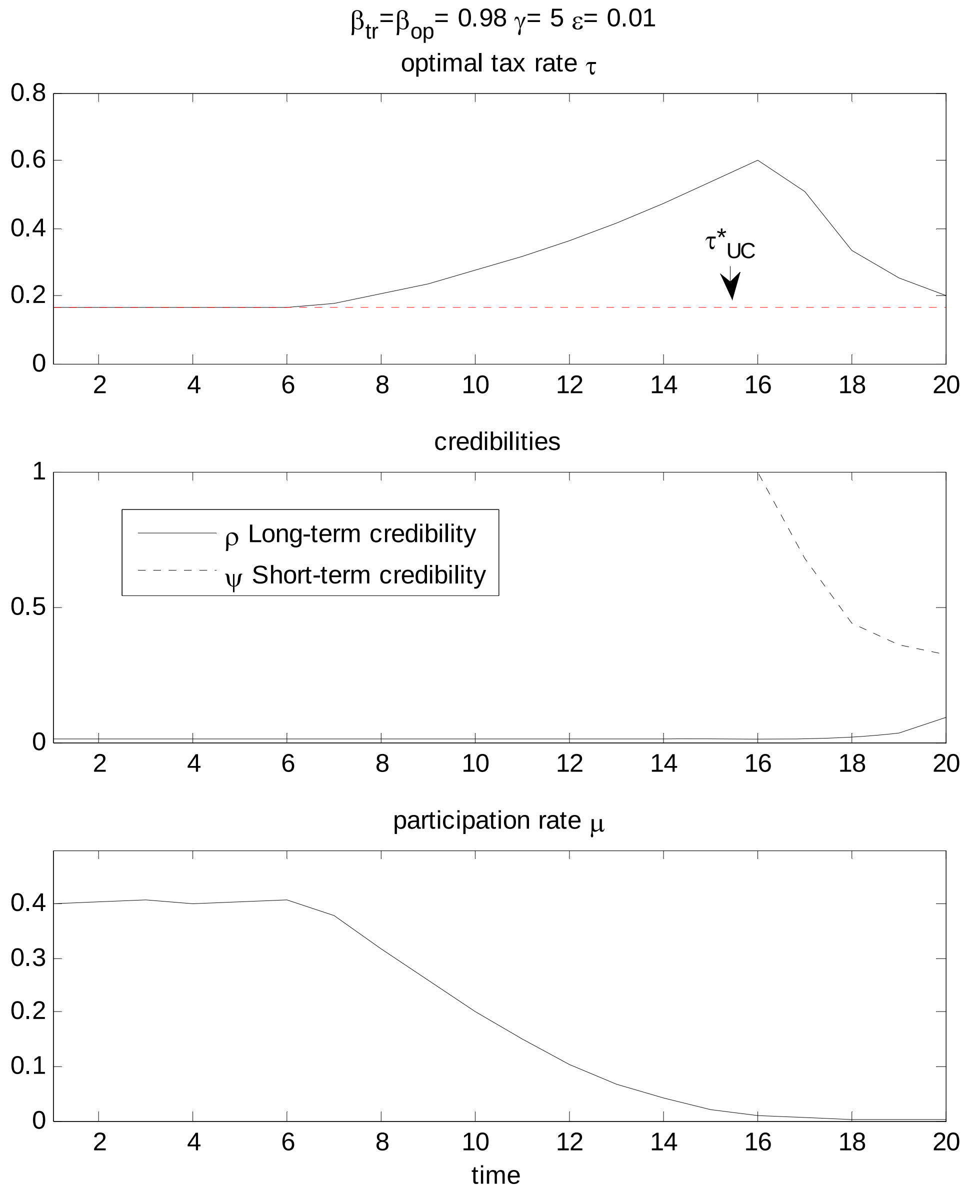

Figure 4: Time series in the case of equally patient government types

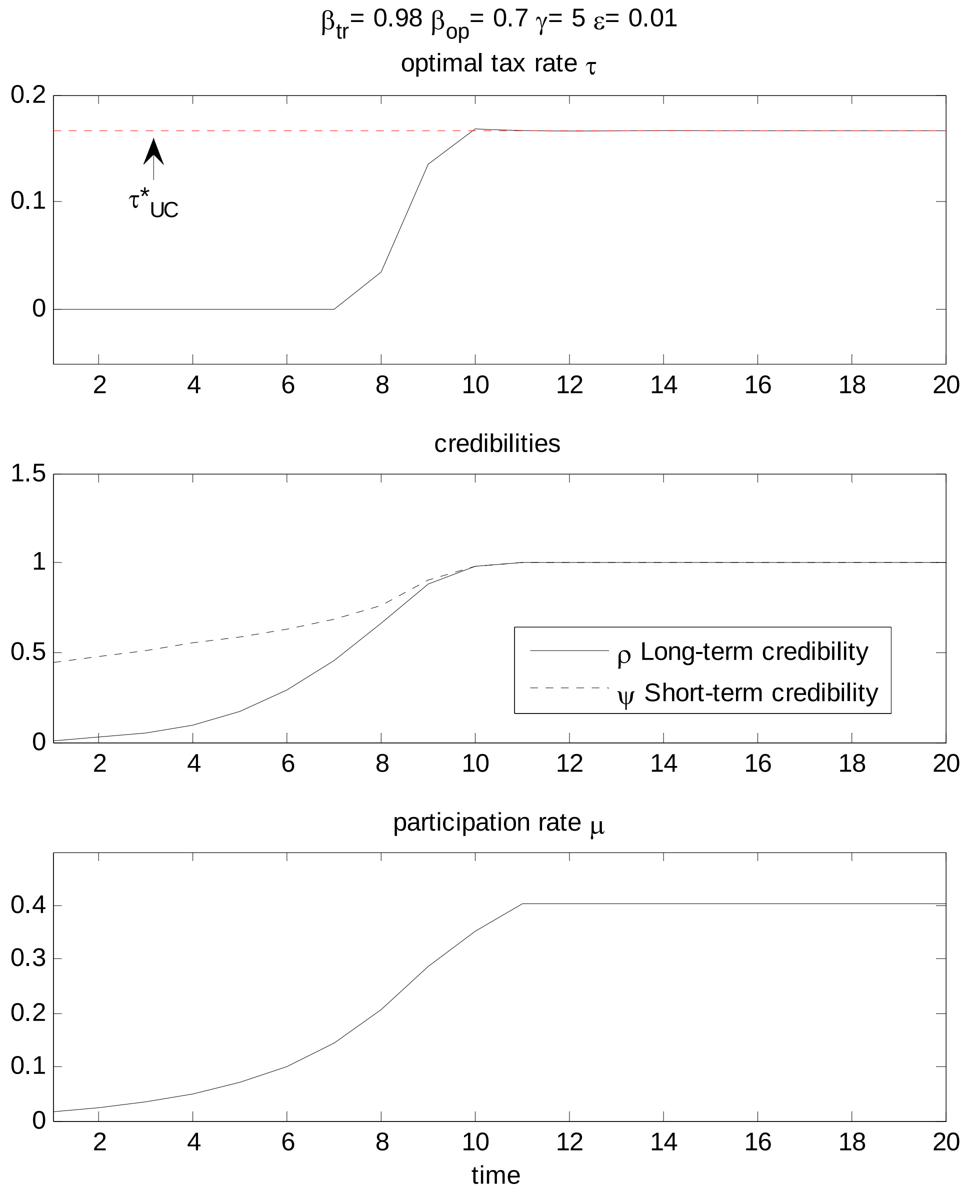

Figure 5: Time series in the case of less patient opportunistic type of government.

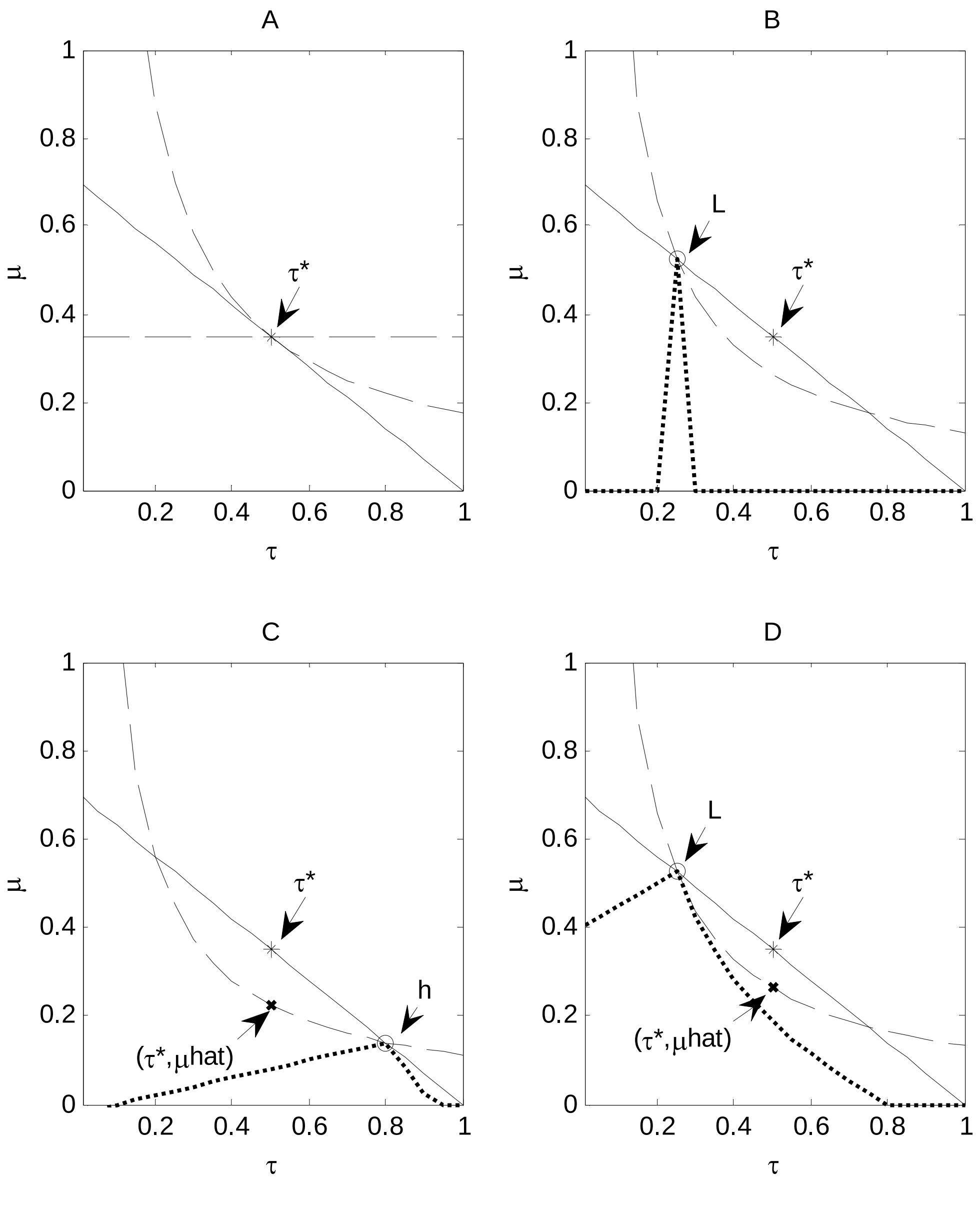

Figure 6: Signaling game in static setting.

7 Appendix

7.1 Proofs for Proposition 2

7.1.1 Parameter definitions in Proposition 2

$$L_0 \text{ solves } \frac{\beta_{tr}}{\beta_{op}}\tau^*_{UC} = \frac{\left[1-\left[\beta_{op}\left(1-\tau^*_{UC}\right)^{\gamma}L_0^{\gamma}\right]^{1/(\gamma+1)}\right]\left[\beta_{op}\left(1-\tau^*_{UC}\right)^{\gamma}L_0^{\gamma}\right]^{(\gamma-1)/(2\gamma+2)}}{1-\left[\beta_{op}\left(1-\tau^*_{UC}\right)^{\gamma}L_0^{\gamma}\right]^{1/2}}$$

$$L_1 = \beta_{op}^{1/\gamma}\left(1-\tau^*_{UC}\right)$$

$$M_1 \text{ solves } \frac{\beta_{tr}}{\beta_{op}}\tau^*_{UC} = \frac{M_1^{\gamma^2/(\gamma+1)}-M_1^{\gamma/(\gamma+1)}}{\left[\beta_{op}\left(1-\tau^*_{UC}\right)^{\gamma}\right]^{1/(\gamma+1)}\left(1-M_1^{\gamma}\right)}+1$$

$$H_1 = \begin{array}{|c|c|c|}\hline & \beta_{tr}/\beta_{op} \leqslant A_2\left(\beta_{op},\gamma\right) & \beta_{tr}/\beta_{op} \geqslant A_2\left(\beta_{op},\gamma\right) \\ \hline \beta_{op} \leqslant 1-\tau^*_{UC} & \tilde{H}_1 & h_1^{-1} \\ \hline \beta_{op} \geqslant 1-\tau^*_{UC} & h_1 & \hat{H}_1 \\ \hline \end{array}$$

$$\text{with } \tilde{H}_1 \text{ solving } \frac{\beta_{tr}}{\beta_{op}}\tau^*_{UC} = \frac{\tilde{H}_1^{\gamma}}{1-\tilde{H}_1^{\gamma}}\left[\left[1-\tilde{\tau}_1(\tilde{H}_1)\right]^{-1}-\left(1+\frac{\tau^*_{UC}}{\beta_{op}}\right)\right]$$

$$\text{and } \hat{H}_1 \text{ solving } \frac{\beta_{tr}}{\beta_{op}}\tau^*_{UC} = \frac{1}{1-\hat{H}_1^{\gamma}}\left[\left(1+\frac{\tau^*_{UC}}{\beta_{op}}\right)-\left[1-\hat{\tau}_1(\hat{H}_1)\right]^{-1}\right]$$

$$A_1\left(\beta_{op},\gamma\right) = \frac{\left[\beta_{op}\left(1-\tau^*_{UC}\right)^{\gamma}\right]^{(\gamma-1)/(\gamma+1)}-1}{1-\beta_{op}\left(1-\tau^*_{UC}\right)^{\gamma}}\left(\gamma+1\right)+\gamma+1$$

$$A_2\left(\beta_{op},\gamma\right) = \frac{1-\gamma}{\left[\beta_{op}\left(1-\tau^*_{UC}\right)^{\gamma}\right]^{1/(\gamma+1)}}+\gamma+1$$

$$\tilde{\tau}_1\left(\rho_1\right), \hat{\tau}_1\left(\rho_1\right), h_1 \text{ are defined in (11) and (12)}$$

7.1.2 $\bar{\tau}_1$ is the minimizer when $\gamma \leqslant 1$

The idea of this proof is to show that whenever the second order derivative of the trustworthy type's objective is negative, the first order derivative is also negative. Therefore, $\bar{\tau}_1$ that makes the first order derivative equal to zero can not have negative second order derivative and thus not a maximizer.

Denote U_{tr} as the objective function of the trustworthy type at $t=1$:

$$\begin{aligned} U_{tr}\left(\tau_1;\rho_1\right) &= \tau_1\mu_1\left(\tau_1;\rho_1\right)+\beta_{tr}W_2\left(\rho_1/\psi_1\right) \\ &= \tau_1\mu_1\left(\tau_1;\rho_1\right)+\frac{\beta_{tr}}{\beta_{op}}\tau_2^*\beta_{op}V_2\left(\rho_1/\psi_1\right) \\ &= \tau_1\mu_1\left(\tau_1;\rho_1\right)+\frac{\beta_{tr}}{\beta_{op}}\tau_2^*\left(1-\tau_1\right)\mu_1\left(\tau_1;\rho_1\right). \end{aligned}$$

The last equality stems from the fact that the opportunistic type is indifferent between mimicking and confiscating whenever τ^*_{UC} is not the optimal rate at $t=1$.

The first order derivative of U_{tr} is:

$$\frac{\partial U_{tr}}{\partial \tau_1} = \left(1 - \frac{\beta_{tr}}{\beta_{op}}\tau_2^*\right)\mu_1 + \left[\left(1 - \frac{\beta_{tr}}{\beta_{op}}\tau_2^*\right)\tau_1 + \frac{\beta_{tr}}{\beta_{op}}\tau_2^*\right]\frac{\partial \mu_1}{\partial \tau_1}.$$

The second order derivative of U_{tr} is:

$$\frac{\partial^2 U_{tr}}{(\partial \tau_1)^2} = 2\left(1 - \frac{\beta_{tr}}{\beta_{op}}\tau_2^*\right)\frac{\partial \mu_1}{\partial \tau_1} + \left[\left(1 - \frac{\beta_{tr}}{\beta_{op}}\tau_2^*\right)\tau_1 + \frac{\beta_{tr}}{\beta_{op}}\tau_2^*\right]\frac{\partial^2 \mu_1}{(\partial \tau_1)^2}.$$

Now let us write μ_1 and $\partial^2 \mu_1 / (\partial \tau_1)^2$ as functions of $\partial \mu_1 / \partial \tau_1$. To do so, we need the explicit form of μ_1:

$$\mu_1 = (1-\tau_1)^\gamma \psi_1^\gamma = (1-\tau_1)^\gamma \left[\frac{\beta_{op}(1-\tau_2^*)^\gamma \rho_1^\gamma}{(1-\tau_1)^{\gamma+1}}\right]^{1/2} = (1-\tau_1)^{\frac{\gamma-1}{2}}\left[\beta_{op}(1-\tau_2^*)^\gamma \rho_1^\gamma\right]^{1/2},$$

so that

$$\begin{aligned} \frac{\partial \mu_1}{\partial \tau_1} &= \frac{1-\gamma}{2}\frac{\mu_1}{1-\tau_1} = \frac{1-\gamma}{2}(1-\tau_1)^{\frac{\gamma-3}{2}}\left[\beta_{op}(1-\tau_2^*)^\gamma \rho_1^\gamma\right]^{1/2}; \\ \frac{\partial^2 \mu_1}{(\partial \tau_1)^2} &= \frac{3-\gamma}{2}\frac{1}{1-\tau_1}\frac{\partial \mu_1}{\partial \tau_1}. \end{aligned}$$

Plug them back into the derivatives:

$$\begin{aligned} \frac{\partial U_{tr}}{\partial \tau_1} &= \left[\left(1 - \frac{\beta_{tr}}{\beta_{op}}\tau_2^*\right)\frac{2(1-\tau_1)}{1-\gamma} + \left(1 - \frac{\beta_{tr}}{\beta_{op}}\tau_2^*\right)\tau_1 + \frac{\beta_{tr}}{\beta_{op}}\tau_2^*\right]\frac{\partial \mu_1}{\partial \tau_1}; \\ \frac{\partial^2 U_{tr}}{(\partial \tau_1)^2} &= \left[2\left(1 - \frac{\beta_{tr}}{\beta_{op}}\tau_2^*\right) + \left[\left(1 - \frac{\beta_{tr}}{\beta_{op}}\tau_2^*\right)\tau_1 + \frac{\beta_{tr}}{\beta_{op}}\tau_2^*\right]\frac{3-\gamma}{2(1-\tau_1)}\right]\frac{\partial \mu_1}{\partial \tau_1} \\ &= \left[2 + \frac{\tau_1}{1-\tau_1}\frac{3-\gamma}{2} + \frac{\beta_{tr}}{\beta_{op}}\tau_2^*\left(\frac{3-\gamma}{2} - 2\right)\right]\frac{\partial \mu_1}{\partial \tau_1}. \end{aligned}$$

Now, if $\partial^2 U_{tr} / (\partial \tau_1)^2 < 0$, given the fact that $\gamma \leqslant 1$ implies $\partial \mu_1 / \partial \tau_1 \geqslant 0$,

$$\begin{aligned} 2 + \frac{\tau_1}{1-\tau_1}\frac{3-\gamma}{2} + \frac{\beta_{tr}}{\beta_{op}}\tau_2^*\left(\frac{3-\gamma}{2} - 2\right) &< 0 \\ \frac{3-\gamma}{(1-\tau_1)(\gamma+1)} + 1 &< \frac{\beta_{tr}}{\beta_{op}}\tau_2^*. \end{aligned}$$

If $\partial U_{tr} / \partial \tau_1 < 0$,

$$\begin{aligned} \left(1 - \frac{\beta_{tr}}{\beta_{op}}\tau_2^*\right)\frac{2(1-\tau_1)}{1-\gamma} + \left(1 - \frac{\beta_{tr}}{\beta_{op}}\tau_2^*\right)\tau_1 + \frac{\beta_{tr}}{\beta_{op}}\tau_2^* &< 0 \\ \frac{2}{1+\gamma} + \frac{\tau_1}{1-\tau_1}\frac{1-\gamma}{1+\gamma} &< \frac{\beta_{tr}}{\beta_{op}}\tau_2^*. \end{aligned}$$

Notice that

$$\frac{3-\gamma}{(1-\tau_1)(\gamma+1)}+1>\frac{2}{1+\gamma}+\frac{\tau_1}{1-\tau_1}\frac{1-\gamma}{1+\gamma}$$

because

$$\frac{3-\gamma}{\gamma+1}>\frac{1-\gamma}{\gamma+1}.$$

Therefore, if $\partial^2 U_{tr}/(\partial\tau_1)^2<0$,

$$\frac{\beta_{tr}}{\beta_{op}}\tau_2^*>\frac{3-\gamma}{(1-\tau_1)(\gamma+1)}+1>\frac{2}{1+\gamma}+\frac{\tau_1}{1-\tau_1}\frac{1-\gamma}{1+\gamma}$$

implies $\partial U_{tr}/\partial\tau_1<0$.

7.1.3 Relationship between $A_2\left(\beta_{op},\gamma\right)$, β_{tr}/β_{op} and $(\gamma+1)/\gamma$

We can rewrite A_2 as

$$\begin{aligned}A_2\left(\beta_{op},\gamma\right) &= \frac{1-\gamma}{\left[\beta_{op}/\left(1-\tau_{UC}^*\right)\right]^{1/(\gamma+1)}\left(1-\tau_{UC}^*\right)}+\gamma+1\\ &= (\gamma+1)\left[1+\frac{1-\gamma}{\gamma}\frac{1}{\left[\beta_{op}/\left(1-\tau_{UC}^*\right)\right]^{1/(\gamma+1)}}\right].\end{aligned}$$

It follows immediately that $A_2\left(\beta_{op},\gamma\right)\geqslant(\gamma+1)/\gamma$ if and only if $\beta_{op}\leqslant(1-\tau_{UC}^*)$.

Denote A as $\beta_{op}/(1-\tau_{UC}^*)$. $A\geqslant 1$ when $\beta_{op}\geqslant 1-\tau_{UC}^*$. We can rewrite A_2 and $1/\beta_{op}$ as:

$$\begin{aligned}A_2 &= (\gamma+1)\left[1+\frac{1-\gamma}{\gamma}A^{-1/(\gamma+1)}\right]\\ 1/\beta_{op} &= A^{-1}(1-\tau_{UC}^*)^{-1}=A^{-1}\frac{\gamma+1}{\gamma},\end{aligned}$$

and their ratio is

$$\begin{aligned}\frac{A_2}{1/\beta_{op}} &= \left[1+\frac{1-\gamma}{\gamma}A^{-1/(\gamma+1)}\right]A\gamma\\ &= A\gamma+(1-\gamma)A^{\gamma/(\gamma+1)}\\ &= A^{\gamma/(\gamma+1)}+\gamma\left(A-A^{\gamma/(\gamma+1)}\right).\end{aligned}$$

Since $\gamma/(\gamma+1)<\gamma<1$ and $A\geqslant 1$, the ratio is then greater than 1. Therefore, $A_2\geqslant 1/\beta_{op}\geqslant\beta_{tr}/\beta_{op}$.

7.1.4 Limits of $\tilde{H}_1$ and $\hat{H}_1$ when $\rho_1 \to 1$

When we view $\tilde{H}_1$ and $\hat{H}_1$ as the cutoffs in β_{tr}/β_{op} and are functions of ρ_1, β_{op} and γ, we can express $\tilde{H}_1$ and $\hat{H}_1$ as the following:

$$\begin{aligned}
\tilde{H}_1 &= \left(\frac{1}{1-\rho_1^\gamma} - 1\right)\left[\rho^{-\gamma/(\gamma+1)}\left[\beta_{op}\left(1-\tau_{UC}^*\right)\right]^{-1/(\gamma+1)} - \left(1+\frac{\tau_{UC}^*}{\beta_{op}}\right)\right]/\tau_{UC}^* \\
\hat{H}_1 &= \frac{1}{1-\rho_1^\gamma}\left[\left(1+\frac{\tau_{UC}^*}{\beta_{op}}\right) - \rho^{\gamma/(\gamma+1)}\left[\beta_{op}\left(1-\tau_{UC}^*\right)\right]^{-1/(\gamma+1)}\right]/\tau_{UC}^*
\end{aligned}$$

If we can show that

$$\left[\beta_{op}\left(1-\tau_{UC}^*\right)\right]^{-1/(\gamma+1)} \leqslant 1+\frac{\tau_{UC}^*}{\beta_{op}}$$

Then immediately $\lim_{\rho_1\to 1}\tilde{H}_1 = -\infty$ and $\lim_{\rho_1\to 1}\hat{H}_1 = +\infty$ since $\left(1-\rho_1^\gamma\right)^{-1} \to +\infty$.

To prove the inequality above holds, again denote A as $\beta_{op}/\left(1-\tau_{UC}^*\right)$. We can rewrite the inequality equivalently as:

$$\begin{aligned}
A^{-1/(\gamma+1)}\left(1-\tau_{UC}^*\right)^{-1} &\leqslant 1+\frac{\tau_{UC}^*}{1-\tau_{UC}^*}\frac{1}{A} \\
A^{\gamma/(\gamma+1)} &\leqslant A\frac{\gamma}{\gamma+1}+\frac{1}{\gamma+1}
\end{aligned}$$

The right-hand-side is a linear function of A, and the left-hand-side is a concave function of A since $\gamma/\left(\gamma+1\right) < \gamma < 1$. The two sides equal when $A = 1$, and the right-hand-side is tangential to the left-hand-side at $A = 1$. Therefore, the inequality must hold.