Federal Reserve Bank of Minneapolis
Research Department Staff Report 283

January 2001

Public Trust and Government Betrayal*

Christopher Phelan

Federal Reserve Bank of Minneapolis

ABSTRACT

This paper presents a simple model of government reputation which captures two characteristics of policy outcomes in less developed countries: governments which betray public trust do so erratically, and, after a betrayal, public trust is regained only gradually.

*This paper was formerly titled "Randomly Bad Government." The author has benefited greatly from discussions with Andrew Atkeson, V.V. Chari, Harold Cole, Patrick Kehoe, and Narayana Kocherlakota, and from the research assistance of Aleh Tsyvinski. The views expressed herein are the those of the author and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

1. Introduction

Government policy in developed countries is boring. It is relatively stable and predictable, and, for the most part (at least relative to less developed countries), promises made are promises kept. Governments keep their promises despite the fact that policymakers face a well-known *time-consistency problem*. That is, it is seldom in the short-run best interest of a government to keep capital taxes low, honor its debt obligations, or inflate the currency only by the expected amount.[1]

Much of the theory on credible government policy concerns itself precisely with accounting for this ability of governments to make and keep promises. In *trigger models* (such as in Chari and Kehoe 1990), good outcomes correspond to a particular subset of equilibria of a game with multiple equilibria. In these good equilibria, households trust the government and the government does not betray this trust because a deviation by the government causes a reversion to a worse equilibrium. With an infinite period model and sufficiently little discounting, such a threat induces the government not to deviate.

In *reputation models* (such as in Barro and Gordon 1983, Celentani and Pesendorfer 1996, and Cole and Kehoe 1998), good outcomes occur both in finite period models and in infinite period models without explicit history-dependent (or trigger) strategies. In such models, the type of government is unobserved by households and the government's reputation is the household sector's belief (its Bayesian posterior) that the government is of a particular type. If the government is possibly an honest (or irrational) type that simply cannot betray the trust of households, then a betrayal destroys the belief that the government is possibly honest. This loss of doubt regarding the type of government can be a sufficient inducement

[1] The seminal paper is Kydland and Prescott (1977).

for governments of all types (honest or not) to act in a trustworthy manner.

A difficulty of both trigger and reputation models, however, is their ability to account for bad outcomes. Standard reputation models (specifically those in which government type is permanent) overexplain good outcomes in that they predict good outcomes will always occur.[2] Trigger models, on the other hand, allow for bad equilibrium outcomes (and, in fact, rely on them) but generally miss key characteristics of bad outcomes.

Bad outcomes tend to have two characteristics which do not easily match existing models. One is that bad outcomes tend to be associated with unpredictable government policy. That is, for long periods of time, exchange rates, tax policies, or monetary policies are stable, and then the government freezes all bank accounts (Brazil), massively devalues the currency in the foreign exchange market (Argentina and many others), declares the fiat currency valueless (Russia), defaults on debt (too many to mention), or massively inflates the currency (again, too many to mention). The other characteristic most models tend to miss is that after such episodes, trust (in the form of money or debt holdings or capital accumulation) is rebuilt only gradually.

In contrast, I present here a very simple model which, in spirit, can capture the most basic characteristics of bad policy outcomes. In my model, a government can be either good, which means it must tax output at a low enough rate to make production worthwhile, or bad, which means it has the option of either taxing at this low rate or confiscating all output. Government type cannot be directly observed by households, making the model close to the reputation models of Kreps and Wilson (1982), Milgrom and Roberts (1982),

[2] In a finite period reputation model with permanent types, bad outcomes can occur toward the end of time. However, in the limit as the number of dates approaches infinity, or in the Markov perfect equilibrium of the infinitely repeated game, the good outcome always occurs.

Barro and Gordon (1983), and their successors.[3] When government type is permanent, this model has a result in line with those in these earlier papers—as long as there is not too much discounting, in the Markov perfect equilibrium, both good and bad governments act in a trustworthy manner. This holds for any probability that the government is good, as long as the probability is positive.

The surprising result of this paper is that if government type can change (specifically, if it follows a Markov process), then no matter how small the transition probabilities are and regardless of the rate of discount, the unique outcome of a Markov perfect equilibrium has the bad government following a mixed strategy. Bad governments do not routinely act in a trustworthy or an untrustworthy manner, but instead randomize regarding whether to betray the households. Thus, policy is unpredictable. Further, the equilibrium has the property that trust is rebuilt only gradually. The percentage of households which produce in equilibrium is directly related to how long it has been since a confiscation of output by the government.

The intuition behind these results is straightforward. If very few or no households trust the government, a bad government has little or nothing to gain by betraying the trust it has been granted. Further, if bad governments always betray trust, as long as some probability exists that the government truly is good, a bad government can earn a reputation as a good government by acting good for one date. This enhanced reputation is valuable to the bad government, and thus bad governments always betraying trust is not an equilibrium. However, if, in a proposed Markov perfect equilibrium, bad governments always act like good governments, then there is no punishment to betraying. It is better to betray today and

[3] In particular, Cukierman and Meltzer (1986), Cole, English and Dow (1995), and Celentani and Pesendorfer (1996) deal with reputation in a government policy setting.

follow the equilibrium from tomorrow on rather than follow the equilibrium from today on. Thus, any Markov perfect equilibrium must entail at least some mixing.

This mixing implies interesting dynamics. Specifically, if bad governments mix and good governments always act in a trustworthy manner, then observing good behavior by the government causes households to gradually increase their Bayesian posterior that the government is good. Further, for a bad government to value a good reputation (a necessary condition to get the bad government to be willing to mix) a higher percentage of households must produce, the higher their Bayesian posterior. Thus, trust increases gradually in equilibrium.

In Section 2, I present the model. In Section 3, I define Markov strategies and Markov equilibria. In Section 4, I consider the special cases in which government type is common knowledge and in which government type is private but fixed through time. In Section 5, I assume government type follows a nondegenerate Markov process. Here, the main results of the paper are proved. In Section 6, I characterize the Markov perfect equilibrium of the limiting economy where the probability that the government is good goes to zero. In Section 7, I consider non-Markov equilibria, and in Section 8, I conclude.

2. The Model

Consider the following simple game. A continuum of households faces a sequence of governments which can be of type *good* or *bad*. The government's type is not directly observable by households. At each date $t = 0, \ldots, \infty$, households move first, simultaneously to each other. Each household can produce at cost c a good with value q or not produce. The government and other households observe the measure (or fraction) μ of households which produce, but the action of any particular household is private to that household. After households move, the government moves. A bad government can tax output at an exogenous rate $\tau < 1$ or confiscate all output. A good government has no choice to make. It always sets the tax to τ. If measure μ of households produce, a bad government's static payoff is $\mu\tau q$ if it taxes at rate τ and μq if it confiscates all output. A good government's payoffs are not defined because it never makes a choice. A household which does not produce receives a static payoff of zero regardless of the play of the government. A household which produces receives a payoff of $(1-\tau)q - c > 0$ if the government taxes at rate τ and a payoff of $-c < 0$ if the government confiscates all output. These assumptions ensure that a household should produce if it anticipates that the government will tax at rate τ and should not produce if it anticipates confiscation.

At date $t = 0$ the government is good with probability $\rho_0 \geq 0$. At the start of each date, a bad government is replaced by a good government with probability $\epsilon \geq 0$. Alternatively, a good government is replaced by a bad government with probability $\delta \geq 0$. Government death and rebirth is not observed by and cannot be directly communicated to households.[4] Both households and the bad government discount at the rate $0 < \beta < 1$. The only additional restriction on the parameters $(\tau, q, c, \rho_0, \epsilon, \delta)$ is

$$\frac{c}{(1-\tau)\, q} < 1 - \delta - \epsilon. \tag{A1}$$

Since $(1-\tau)q > c$, assumption (A1) requires that the transition probabilities δ and ϵ be sufficiently small. This assumption is made to ensure that, in equilibrium, if a government does not confiscate for a sufficient number of dates, it is trusted enough to ensure that

[4] This assumption plays a major role in the firm reputation models of Tadelis (1999) and Mailath and Samuelson (2000).

all households produce. If there is too much mixing between types (δ and ϵ are too high), a history of not confiscating isn't sufficiently informative to ensure that all households produce.

No restrictions are put on the discount rate β (other than $0 < \beta < 1$), and no restrictions are put on the invariant (or long-run) probability that the government is of type good, $\epsilon/(\epsilon + \delta)$. Thus, ρ_0 and ϵ can be set arbitrarily close to zero, which implies that this model can be made (in a sense) arbitrarily close to a model in which it is common knowledge that the government is bad.

3. Markov Strategies and Markov Equilibria

Define Markov strategies relative to the state variable ρ—the households' posterior probability that the government is of type good. A *Markov strategy* is a specification of $\mu(\rho)$, the measure of households which produce as a function of ρ, and $\pi(\mu, \rho)$, the probability that a bad government confiscates as a function of μ and ρ.

Since households are anonymous, they cannot individually affect the play of the government or the future values of ρ and thus cannot individually affect the future play of the game. Whether a household should produce depends solely on whether the probability that the government confiscates at the current date is at or below a cutoff value, π^*. In particular, define π^* such that

$$(1 - \pi^*)(1 - \tau)q - c = 0,$$

which implies that

$$0 < \pi^* = 1 - \frac{c}{(1 - \tau)\, q} < 1.$$

Households are said to be optimizing if for all ρ, $\mu(\rho) > 0$ implies that $(1 - \rho)\pi(\mu(\rho), \rho) \leq \pi^*$

and $\mu(\rho) < 1$ implies that $(1-\rho)\pi(\mu(\rho),\rho) \geq \pi^*$. (Together, these inequalities imply that if $0 < \mu(\rho) < 1$, then $(1-\rho)\pi(\mu(\rho),\rho) = \pi^*$.)

Note that if ρ is sufficiently high (or households are sufficiently confident that they are facing a good government), households should produce regardless of the probability that a bad government confiscates. Specifically, if $\rho > 1-\pi^*$, then $(1-\rho)\pi < \pi^*$ for all $\pi \in [0,1]$; thus, household optimization implies that $\mu(\rho) = 1$. Thus, while π^* is the *cutoff probability of confiscation*, $\rho^* \equiv 1-\pi^*$ can be considered the *cutoff posterior*.

Unlike households, a bad government can affect the future play of the game and thus cares how it affects future values of ρ. If a government confiscates at date t, it must have been the bad type. Given the government was bad at date t, the posterior at date $t+1$ is $\rho' = \epsilon$, the probability that a bad government is replaced by a good government. If the government does not confiscate at date t, Bayes' rule defines the new posterior by

$$\rho'(\rho,\pi) = (1-\delta)\frac{\rho}{\rho+(1-\rho)(1-\pi)} + \epsilon\ (1 - \frac{\rho}{\rho+(1-\rho)(1-\pi)}).$$

This function is strictly increasing in π. In particular, $\rho'(\rho,1) = 1-\delta$, the highest possible value for ρ. If households expect a bad government to confiscate with probability one, a bad government can achieve the highest possible reputation by not confiscating.

Let $V(\rho)$ denote the expected lifetime payoff to a bad government associated with strategy (μ,π). Recursively,

$$\begin{aligned} V(\rho) &= \pi(\mu(\rho),\rho)\Big[q\mu(\rho) + \beta(1-\epsilon)V(\epsilon)\Big] \\ &\quad +(1-\pi(\mu(\rho))\Big[\tau q\mu(\rho) + \beta(1-\epsilon)V(\rho'(\rho,\pi(\mu(\rho),\rho))\Big]. \end{aligned}$$

A Markov strategy is said to respect government optimization if and only if for all

(μ, ρ) such that $\pi(\mu, \rho) > 0$, confiscating is weakly preferred to not confiscating, or

$$q\mu + \beta(1-\epsilon)V(\epsilon) \geq \tau q\mu + \beta(1-\epsilon)V(\rho'(\rho, \pi(\mu, \rho))),$$

and for all (μ, ρ) such that $\pi(\mu, \rho) < 1$, not confiscating is weakly preferred to confiscating, or

$$q\mu + \beta(1-\epsilon)V(\epsilon) \leq \tau q\mu + \beta(1-\epsilon)V(\rho'(\rho, \pi(\mu, \rho))).$$

(Together, these inequalities imply that if $0 < \pi(\mu, \rho) < 1$, the bad government must be indifferent between confiscating and not confiscating, and thus the above inequalities must hold as an equality.)

A Markov strategy is said to be a Markov perfect equilibrium if it respects both household and bad government optimization.

4. Special Cases

Assume, for the moment, that $\rho_0 = \epsilon = 0$, or that after all histories, it is common knowledge that the government is the bad type. In this case, a Markov perfect equilibrium is simply the fraction of households which produce μ and a probability that the bad government confiscates $\pi(\mu)$, and an implied value

$$V = \pi(\mu)[q\mu + \beta(1-\epsilon)V] + (1-\pi(\mu))[\tau q\mu + \beta(1-\epsilon)V].$$

Government optimization requires (for all μ) that if $\pi(\mu) < 1$, $q\mu + \beta V \leq \tau q\mu + \beta V$. Since $\tau < 1$, the latter requires that $\pi(\mu) = 1$ for all $\mu > 0$. Household optimization then implies that $\mu = 0$, leaving $\mu = 0$ for all dates as the unique outcome of a Markov perfect equilibrium.[5]

[5]There is a continuum of Markov equilibria for this case, defined by $\mu = 0$, $\pi(\mu) = 1$ for $\mu > 0$ and $\pi(0) \geq \pi^*$, but all have the same outcome path.

This is not surprising. Without history-dependent strategies, only repetition of the one-shot equilibrium is possible.

Next assume that $\epsilon = \delta = 0$ but that $\rho_0 > 0$. This assumption brings the model more in line with those in most papers regarding reputation in game theory (Kreps and Wilson 1982, Milgrom and Roberts 1982, Celentani and Pesendorfer 1996, among others). That is, government type is permanent (governments are never replaced) but is not observed by households. Here, similar to the chain-store paradox papers, if the government ever confiscates, then its type is known forever ($\rho = 0$). Given this, $V(0) = 0$ because the subgame following a confiscation is identical to the case considered above. If $\tau q/(1-\beta) > q$ (or imitating the good type forever is preferred to the one-shot gain from confiscating all output), then this game has a unique Markov perfect equilibrium outcome in which for all $\rho > 0$, $\mu(\rho) = 1$, $\pi(1, \rho) = 0$ (and thus $\rho' = \rho$), and $V(\rho) = \tau q/(1-\beta)$. This is, again, not surprising given the earlier work on reputation.

5. The General Case

Assume that $\epsilon > 0$ and $\delta > 0$. That is, assume that at the start of each date, governments have a positive probability of dying and being reborn as the other type. For simplicity, also assume that $\rho_0 = \epsilon$, implying that the game starts as if there were a confiscation at the preceding date. (The case in which $\rho_0 \neq \epsilon$ is treated later.) I show here that there is a unique Markov perfect equilibrium which always has the same structure. First, I simply assert and describe this equilibrium strategy. Next, I show that this strategy satisfies household optimization and bad government optimization if the bad government weakly wishes to confiscate if all households produce. Following this, I show that no other Markov perfect equilibrium exists, and in the process, I show that a bad government strictly wishes to confiscate if all households produce. Finally, I consider $\rho_0 \neq \epsilon$.

Let $(\hat{\mu}, \hat{\pi})$ denote the unique Markov perfect equilibrium. The first asserted characteristic of this equilibrium is that if $\rho < \rho^*$, the bad government randomizes to make households indifferent to producing or not, or $(1-\rho)\hat{\pi}(\hat{\mu}(\rho), \rho) = \pi^*$. When $\pi = \pi^*/(1-\rho)$, the function $\rho'(\rho, \pi)$ simplifies to

$$\rho^{'*}(\rho) = \rho(1 - \delta - \epsilon)/\rho^* + \epsilon.$$

This function is linear, has a positive intercept, and has a slope greater than one. Thus, starting from ϵ, successive application of $\rho^{'*}(\rho)$ steps above ρ^* in a finite number of steps (denoted N). That is, under $\hat{\pi}$, if fewer than N dates have passed since the last confiscation, $\rho < \rho^*$ and $\rho \geq \rho^*$ otherwise.

Let $\hat{\rho}_i$, $i \in \{0, 1, \ldots, N\}$ denote the value of ρ induced by $\hat{\pi}$ if i consecutive dates have passed without a confiscation, and let $\hat{\mu}_i = \hat{\mu}(\hat{\rho}_i)$ and $\hat{V}_i = V(\hat{\rho}_i)$. Since $0 < \pi^*/(1-\rho) < 1$ for $\rho < \rho^*$, for $(\hat{\mu}, \hat{\pi})$ to be a Markov perfect equilibrium, the bad government must be indifferent between confiscating and not confiscating for $i < N$. This implies that for $i \in \{0, \ldots, N-1\}$,

$$\hat{V}_i = q\hat{\mu}_i + \beta(1-\epsilon)\hat{V}_0$$

$$\hat{V}_i = \tau q\hat{\mu}_i + \beta(1-\epsilon)\hat{V}_{i+1}.$$

This is a sequence of $2N$ equations and $2N+1$ unknowns with full rank.

The second asserted characteristic of the unique equilibrium is that if $\rho \geq \rho^*$, $\hat{\mu}(\rho) = 1$ and $\hat{\pi}(1, \rho) = 1$. That is, all households produce, and the bad government confiscates with probability one. This completes the description of government behavior and adds to the above system the equation

$$\hat{V}_N = q + \beta(1-\epsilon)\hat{V}_0.$$

This system of linear equations has a unique solution where the vector $\{\hat{\mu}_0, \hat{\mu}_1, \ldots, \hat{\mu}_{N-1}\}$ describes household play for $\rho < \rho^*$, completing the description of household play.

Is this an equilibrium? Given the strategy of the bad government, household optimization is immediate. By construction, households are indifferent between producing or not if $\rho \leq \rho^*$ and strictly prefer to produce when $\rho > \rho^*$. Again, by construction, a bad government is indifferent between confiscating or not for $\rho < \rho^*$. Thus, government optimization is satisfied if the bad government weakly prefers to confiscate when $\rho \geq \rho^*$, which is yet unproved. This is shown as a consequence of proving that no other Markov perfect equilibrium exists, to which I now proceed.

Let (μ, π) denote an arbitrary Markov perfect equilibrium. Analogously, let ρ_i, $i \in \{0, 1, \ldots, \infty\}$ denote the value of ρ induced by π if i consecutive dates have passed without a confiscation, and let $\mu_i = \mu(\rho_i)$ and $V_i = V(\rho_i)$. The first step is to show that in any Markov perfect equilibrium, $\mu(\epsilon) = \mu_0 > 0$. This eliminates, among other things, a Markov equilibrium in which no household ever produces because of fear that the bad government will confiscate with probability one. This cannot be an equilibrium outcome because by deviating, the government can costlessly earn a higher ρ and ensure a higher payoff.

LEMMA 1. $\mu_0 > 0$.

Proof. If $\mu_0 = 0$, then household optimization implies that $\pi_0 \geq \pi^*$. For a bad government to confiscate with positive probability, its payoff must be that of confiscating with certainty;

that is, $V_0 = \beta(1-\epsilon)V_0$ or $V_0 = 0$. Consider the following deviation strategy: never confiscate until μ is positive. Either this strategy has a positive payoff or $\mu = 0$ forever. For $\mu_i = 0$, $\pi_i \geq \pi^*/(1-\rho_i)$. Since $\rho'(\rho, \pi)$ is increasing in π, $\rho_i \geq \hat{\rho}_i$. Thus, $\rho_N > \rho^*$ and household optimization implies that $\mu_N = 1$. ∎

An immediate implication of Lemma 1 is that since $V_0 \geq q\mu_0 + \beta(1-\epsilon)V_0$, $V_0 > 0$. The next result shows that in any Markov perfect equilibrium, $\mu(\epsilon) = \mu_0 < 1$. This result eliminates, among other things, a strategy in which all households always produce because neither type of government will confiscate. This cannot be an equilibrium because it implies no punishment for a deviating government.

Lemma 2. $\mu_0 < 1$.

Proof. If $\mu_0 = 1$, then $V_0 \geq q + \beta(1-\epsilon)V_0$ or $V_0 \geq q/(1-\beta(1-\epsilon))$. This lifetime payoff is possible only if $\mu = 1$ at every date and $\pi = 1$ at every date. This outcome is inconsistent with household optimization. ∎

Lemmas 1 and 2 imply that $0 < \mu_0 < 1$ and thus that $\pi_0 = \hat{\pi}_0 = \pi^*/(1-\epsilon)$, which implies that $\rho_1 = \hat{\rho}_1$. The next two lemmas establish an induction argument to show $0 < \mu_i < 1$ for all $i < N$.

Lemma 3. For $i \geq 1$, if $q\mu_{i-1} + \beta(1-\epsilon)V_0 = \tau q\mu_{i-1} + \beta(1-\epsilon)V_i$, then $\mu_i > 0$.

Proof. If $\mu_i = 0$, then household optimization implies that $\pi_i \geq \pi^*/(1-\rho_i) > 0$. This implies that $V_i = \beta(1-\epsilon)V_0$, or $V_i < V_0$ since $V_0 > 0$. Next, since $q\mu_{i-1} + \beta(1-\epsilon)V_0 = \tau q\mu_{i-1} + \beta(1-\epsilon)V_i$ and $\mu_{i-1} \geq 0$, $V_i \geq V_0$. ∎

Lemma 4. For $i \geq 1$, if $\mu_{i-1} < 1$, $\rho_i < \rho^*$ and $q\mu_{i-1} + \beta(1-\epsilon)V_0 = \tau q\mu_{i-1} + \beta(1-\epsilon)V_i$, then $\mu_i < 1$.

Proof. If $\mu_i = 1$ and $\rho_i < \rho^*$, then $\pi_i \leq \pi^*/(1-\rho_i) < 1$. This implies that $V_i = \tau q + \beta(1-\epsilon)V_{i+1} \geq q + \beta(1-\epsilon)V_0$, or

$$V_{i+1} \geq \frac{q(1-\tau)}{\beta(1-\epsilon)} + V_0.$$

Since $\tau q \mu_{i-1} + \beta(1-\epsilon)V_i = q\mu_{i-1} + \beta(1-\epsilon)V_0$,

$$V_i = \frac{q(1-\tau)\mu_{i-1}}{\beta(1-\epsilon)} + V_0,$$

and thus $V_{i+1} > V_i$. The fact that $V_{i+1} > V_i$ implies that $V_{i+1} > q + \beta(1-\epsilon)V_0$. Thus, $\mu_{i+1} = 1$ and $V_{i+1} = \tau q + \beta(1-\epsilon)V_{i+2}$. Since $V_i = \tau q + \beta(1-\epsilon)V_{i+1}$, $V_{i+1} > V_i$ implies that $V_{i+2} > V_{i+1}$. However, I can continue in this manner, getting $\mu_{i+n} = 1$ for all $n \geq 1$ and $V_{i+n} > q + \beta(1-\epsilon)V_0$, so that the bad government never confiscates. But the bad government never confiscating implies $V_{i+n} = \tau q/(1-\beta(1-\epsilon))$ for all $n \geq 0$, contradicting $V_{i+1} > V_i$. ∎

Since $0 < \mu_0 < 1$, $q\mu_0 + \beta(1-\epsilon)V_0 = \tau q \mu_0 + \beta(1-\epsilon)V_1$ and thus Lemmas 3 and 4 establish that $0 < \mu_1 < 1$ and $\rho_2 = \hat{\rho}_2$. If I continue, this implies for $i \leq N$, $\rho_i = \hat{\rho}_i$ and for $i \leq N-1$, $0 < \mu_i < 1$. By definition, $\mu_N = 1$. Lemma 5 establishes that a bad government after N consecutive nonconfiscations sets $\pi = 1$.

LEMMA 5. For $i \geq N$, $\pi_i = 1$, $\mu_i = 1$, and $V_i = q + \beta(1-\epsilon)V_0 > \tau q/(1-\beta(1-\epsilon))$.

Proof. Suppose that $\tau q + \beta(1-\epsilon)V_{N+1} \geq q + \beta(1-\epsilon)V_0$, or

$$V_{N+1} \geq \frac{q(1-\tau)}{\beta(1-\epsilon)} + V_0.$$

Since $0 < \mu_{N-1} < 1$, $\tau q \mu_{N-1} + \beta(1-\epsilon)V_N = q\mu_{N-1} + \beta(1-\epsilon)V_0$, or

$$V_N = \frac{q(1-\tau)\mu_{i-1}}{\beta(1-\epsilon)} + V_0,$$

and thus $V_{N+1} > V_N$. The fact that $V_{N+1} > V_N$ implies that $V_{N+1} > q + \beta(1-\epsilon)V_0$. Thus, $\mu_{N+1} = 1$ and $V_{N+1} = \tau q + \beta(1-\epsilon)V_{N+2}$. Since $V_N = \tau q + \beta(1-\epsilon)V_{N+1}$, $V_{N+1} > V_N$ implies that $V_{N+2} > V_{N+1}$. As in the preceding proof, I can continue in this manner, getting $\mu_{N+n} = 1$ for all $n \geq 1$ and $V_{N+n} > q + \beta(1-\epsilon)V_0$, so that the bad government never confiscates. But given this, $V_{N+n} = \tau q/(1-\beta(1-\epsilon))$ for all $n \geq 0$, contradicting $V_{N+1} > V_N$.

Thus, $\tau q + \beta(1-\epsilon)V_{N+1} < q + \beta(1-\epsilon)V_0$. This implies that $\pi_N = 1$ and $\rho_{N+1} = 1 - \delta > \rho^*$; thus, $\mu_{N+1} = 1$. The same logic implies that $\tau q + \beta(1-\epsilon)V_{N+2} < q + \beta(1-\epsilon)V_0$; thus, by induction, $\pi_i = 1$, $\mu_i = 1$, and $V_i = V_0 + \beta(1-\epsilon)V_0$ for all $i \geq N$. Since for $i \geq N$, $V_i > \tau q + \beta(1-\epsilon)V_i$, $V_i > \tau q/(1-\beta(1-\epsilon))$. ∎

Thus, I have established that the earlier asserted equilibrium satisfies government optimization and is unique. The fact that μ_i and V_i are strictly monotonic for $i \leq N$ follows quickly.

Lemma 6. For all $(i, j) \in \{0, \ldots, N\}^2$, such that $i < j$, $\mu_i < \mu_j$, and $V_i < V_j$.

Proof. The fact that the government is mixing between confiscating and not confiscating when $\rho = \epsilon$ implies that

$$q\mu_0 + \beta(1-\epsilon)V_0 = \tau q\mu_0 + \beta(1-\epsilon)V_1.$$

Since $\mu_0 > 0$ and $\tau < 1$, this implies that $V_1 > V_0$. Given that

$$V_0 = q\mu_0 + \beta(1-\epsilon)V_0$$

and

$$V_1 = q\mu_1 + \beta(1-\epsilon)V_0,$$

the fact that $V_1 > V_0$ implies that $\mu_1 > \mu_0$. For all $i < N$, government mixing from ρ_i implies that

$$q\mu_i + \beta(1-\epsilon)V_0 = \tau q\mu_i + \beta(1-\epsilon)V_{i+1},$$

or, rearranging,

$$V_{i+1} = V_0 + \frac{q(1-\tau)\mu_i}{\beta(1-\epsilon)}.$$

Since $\mu_1 > \mu_0$, $V_2 > V_1$. As above, I can use this to show that $\mu_2 > \mu_1$, and so on. ∎

This completes the characterization of the case where $\rho_0 = \epsilon$. Consider initial ρ values other than $\rho = \epsilon$, but requiring $\rho_0 \in [0,1)$. Here, I simply assert and verify the equilibrium. Let $\Gamma_N = [\rho^*, 1)$, $\Gamma_{N-1} = \{\rho|\rho'^*(\rho) \in \Gamma_N\}$, $\Gamma_{N-2} = \{\rho|\rho'^*(\rho) \in \Gamma_{N-1}\}$, ..., $\Gamma_0 = \{\rho|\rho'^*(\rho) \in \Gamma_1\}$. Next, let $\Gamma_{-1} = \{\rho|\rho'^*(\rho) \notin \Gamma_1 \cup \ldots \cup \Gamma_N\}$. By construction, $\rho_i \in \Gamma_i$ ($i \in \{0, \ldots, N\}$), and $\Gamma_{-1}, \Gamma_0, \ldots, \Gamma_N$ is a partition of $[0,1)$.

For each ρ in Γ_i, $i \geq 0$, let $\mu(\rho) = \mu_i$, $\pi(\rho) = \pi^*/(1-\rho)$, and $V(\rho) = V_i$ for $\rho \in \Gamma_i$. Thus, $\mu(\rho)$ and $V(\rho)$ are step functions. This specification satisfies household optimization because $(1-\rho)\pi(\rho) = \pi^*$, and thus households are indifferent. To show that it satisfies (bad) government optimization, consider $\rho \in \Gamma_N$. Deviating by not confiscating delivers $\tau q + \beta(1-\epsilon)V_N < q + \beta(1-\epsilon)V_0 = V_N$. For $\rho \in \Gamma_{N-1}$, for the government to be willing to randomize, $V(\rho)$, $\mu(\rho)$ must satisfy

$$V(\rho) = \tau q\mu(\rho) + \beta(1-\epsilon)V_N$$

and

$$V(\rho) = q\mu(\rho) + \beta(1-\epsilon)V_0.$$

These two linear equations are uniquely solved by $V(\rho) = V_{N-1}$ and $\mu(\rho) = \mu_{N-1}$. I can continue in this fashion to $\rho \in \Gamma_0$.

This leaves $\rho \in \Gamma_{-1}$, which consists of ρ sufficiently small such that $\rho'^*(\rho) \in \Gamma_0$. (This includes $\rho = 0$.) For $\rho \in \Gamma_{-1}$, $\mu(\rho) = 0$ and $V(\rho) = \beta(1-\epsilon)V_0$. This holds because if $\mu(\rho) > 0$, then it must be the case that $0 \leq \pi(\mu(\rho), \rho) \leq \pi^*/(1-\rho)$. This in turn implies that $\epsilon \leq \rho'(\rho, \pi(\mu(\rho), \rho)) \leq \rho'^*(\rho)$. Since both ϵ and $\rho'^*(\rho)$ are elements of Γ_0, $\rho'(\rho, \pi(\mu(\rho), \rho)) \in \Gamma_0$. Further, if the bad government confiscates, $\rho' = \epsilon \in \Gamma_0$. Thus, for $\rho \in \Gamma_{-1}$, the continuation for the bad government is the same whether it confiscates or not, and thus $\mu(\rho) = 0$ is necessary for bad government optimization. For households to be willing to set $\mu(\rho) = 0$, one needs $\pi(0, \rho) \geq \pi^*/(1-\rho)$. For the bad government to be willing to confiscate given the current period payoff is zero whether it confiscates or not, one needs $\rho'(\rho, \pi(0, \rho)) \in \Gamma_0$. Setting $\pi(0, \rho) = \pi^*/(1-\rho)$ accomplishes this, as does a neighborhood above this value. Thus, for $\rho_0 \in \Gamma_{-1}$, there is not a unique Markov perfect equilibrium but, nevertheless, a unique outcome in terms of $\mu(\rho_0)$ and $V(\rho_0)$.

6. Limits and Discontinuities

In Section 4, I established that for $\rho_0 > 0$, the Markov perfect equilibrium when $\epsilon = \delta = 0$ has $\mu = 1$ and $\pi = 0$ at all dates along the equilibrium path and delivers the value $\tau q/(1-\beta)$, as long as there is not too much discounting.[6] This holds as well when $0 < \delta < 1$. Since when $\rho_0 = 0$, $\epsilon = 0$, and $0 \leq \delta < 1$, the Markov perfect equilibrium has $\mu = 0$ and $\pi = 1$ at all dates along the equilibrium path and delivers a value of zero, there is a discontinuity at $\rho_0 = 0$, ($\lim_{\rho_0 \to 0} V(\rho_0) \neq V(0)$). This discontinuity is not a new result. It could be considered

[6] By this, I mean $\tau q/(1-\beta) > q$ or the lifetime value of a bad government receiving τq forever exceeds the one-time payoff from confiscating q.

the main point of the standard reputation models of Kreps and Wilson (1982) and Milgrom and Roberts (1982). In Section 5, I showed that this continuity disappears when $\epsilon > 0$. That is, $\lim_{\rho_0 \to 0} V(\rho_0) = V(0) = \beta(1-\epsilon)V(\epsilon)$.

But having $\epsilon > 0$ when bringing ρ_0 to zero is somewhat unfair. When $\epsilon = 0$, considering arbitrarily small ρ_0 brings the economy with uncertainty about government type arbitrarily close to one in which government type at all dates is known with certainty. When $\epsilon > 0$, a small value of ρ_0 no longer implies the model is close to one in which the government type is always known with certainty.

Instead, consider a sequence of economies where both ρ_0 and ϵ go to zero. In particular, while formally possible, it makes little sense to consider $\rho_0 < \epsilon$ because this implies the government starts with a worse reputation than is possible at any point later in the game. Given this, a natural sequence of economies which converges to the common knowledge benchmark is where $\epsilon \to 0$ and $\rho_0 = \epsilon$ (the worst possible continuation reputation) at each point in the sequence. I next show that this sequence has a limit where $\mu(\rho_0)$ remains interior and $V(\rho_0) > 0$.

To this end, define $N(\epsilon)$ as the smallest integer such that $\rho'^{*N(\epsilon)}(\epsilon) > \rho^*$. The following limiting results are obtained: First, as ϵ goes to zero, $N(\epsilon)$ goes to infinity. That is, after a confiscation, it takes an arbitrarily large number of consecutive nonconfiscations for all households to produce. Second, as $\epsilon \to 0$, $\pi(\epsilon) \to \pi^*$ (which is interior), and (as stated earlier) $\mu(\epsilon)$ remains interior and $V(\epsilon) > 0$. Since the posterior ρ evolves according to $\rho'^*(\rho) = \rho(1-\delta-\epsilon)/\rho^* + \epsilon$, this implies that as $\rho_0 = \epsilon \to 0$, ρ is almost always very near zero. Thus, μ is almost always approximately equal to $\lim_{\epsilon \to 0} \mu(\epsilon)$, and π is almost always approximately equal to π^*. These results are proved in the following lemma.

Lemma 7. For given values of $(\beta, \tau, q, c, \rho_0, \delta)$, $\lim_{\epsilon \to 0} N(\epsilon) = \infty$, $\lim_{\epsilon \to 0} \pi(\epsilon) = \pi^*$, $0 < \lim_{\epsilon \to 0} \mu(\epsilon) < 1$, and $\lim_{\epsilon \to 0} V(\epsilon) > 0$.

Proof. The function $\rho^{'*}(\rho) = \rho(1-\delta-\epsilon)/\rho^* + \epsilon$ implies that as $\epsilon \to 0$, $\rho^{'*}(\rho) = \rho(1-\delta)/\rho^*$. Since $\rho^* < 1-\delta$, this function is simply a constant greater than one multiplying ρ, implying that $\lim_{\epsilon \to 0} N(\epsilon) = \infty$. Next, since $\pi(\rho) = \pi^*/(1-\rho)$, $\lim_{\epsilon \to 0} \pi^*/(1-\epsilon) = \pi^*$. From Lemma 5,

$$q + \beta(1-\epsilon)V(\epsilon) > \frac{\tau q}{1-\beta(1-\epsilon)}.$$

This implies that

$$\lim_{\epsilon \to 0} V(\epsilon) \geq \frac{q}{\beta}\Big(\frac{\tau}{1-\beta} - 1\Big).$$

Since $\tau q/(1-\beta) > q$, this is positive. Finally, since $V(\epsilon) = q\mu(\epsilon) + \beta(1-\epsilon)V(\epsilon)$, the preceding inequality implies that

$$\lim_{\epsilon \to 0} \mu(\epsilon) \geq \frac{1}{\beta}\Big(\tau - (1-\beta)\Big),$$

completing the proof. ∎

7. Other Equilibria

The Markov perfect equilibrium examined in this paper is not the unique equilibrium. Consider the following history-dependent strategy starting from $\rho_0 = \epsilon$: For the first $N-1$ dates (where N is as defined above), $\mu = 0$ and $\pi(\rho_i) = \pi^*/(1-\rho_i)$ (where for all i, ρ_i is the same as in the Markov perfect equilibrium). From date N on, regardless of the play of the government from dates 1 to $N-1$, $\mu = 1$ and $\pi = 0$. If the government confiscates at date N or later, this strategy starts over. This strategy always satisfies household optimization and satisfies the optimization of the bad government if $(1-\beta^N(1-\epsilon)^N)\tau q/(1-\beta(1-\epsilon)) \geq q$. Further, as $\epsilon \to 0$, because $N(\epsilon) \to \infty$, this equilibrium has a value which converges to zero, the value of the Markov perfect equilibrium and worst equilibrium when $\rho_0 = \epsilon = 0$. Thus while there is a discontinuity in the value of the Markov perfect equilibrium at $\rho_0 = \epsilon = 0$, there is no discontinuity at this point regarding the value of the worst equilibrium.

8. Conclusion

I have presented a simple model in which in the unique Markov Perfect equilibrium, bad governments do not always act in an untrustworthy manner, but instead randomize. While not proved here, the logic that a bad government always acting in an untrustworthy manner cannot be a Markov perfect equilibrium should generalize to other models. In Chari and Kehoe (1990), the Markov perfect equilibrium (which is also the worst equilibrium) has no household ever investing because the benevolent government will always confiscate whatever investment is made. However, if as in my model, there is always a positive probability that the government simply cannot confiscate all investment, then by deviating and not confiscating, a government can cheaply acquire a reputation as the type which cannot confiscate. The same should hold in models of monetary growth, debt repudiation, and capital taxation, with or without a benevolent government.

References

Barro, Robert J., and David B. Gordon. 1983. Rules, discretion, and reputation in a model of monetary policy. *Journal of Monetary Economics* 12 (July): 101–21.

Celentani, Marco, and Wolfgang Pesendorfer. 1996. Reputation in dynamic games. *Journal of Economic Theory* 70 (July): 109–32.

Chari, V.V., and Patrick J. Kehoe. 1990. Sustainable plans. *Journal of Political Economy* 98 (August): 783–802.

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 (May): 365–85.

Cole, Harold L., and Patrick J. Kehoe. 1998. Models of sovereign debt: Partial versus general reputations. *International Economic Review* 39 (February): 55–70.

Cukierman, Alex, and Allan H. Meltzer. 1986. A theory of ambiguity, credibility, and inflation under discretion and asymmetric information. *Econometrica* 54 (September): 1099–128.

Kreps, David M., and Robert Wilson. 1982. Reputation and imperfect information. *Journal of Economic Theory* 27 (August): 253–79.

Kydland, Finn E., and Edward C. Prescott. 1977. Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy* 84 (June): 473–91.

Mailath, George J., and Larry Samuelson. 2000. Who wants a good reputation? *Review of Economic Studies,* forthcoming.

Milgrom, Paul, and John Roberts. 1982. Predation, reputation, and entry deterrence.

Journal of Economic Theory 27 (August): 280–312.

Tadelis, Steven. 1999. What's in a name? Reputation as a tradeable asset. *American Economic Review* 89 (June): 548–63.