Heterogeneous Beliefs and Optimal Taxation*

Facundo Piguillem†

University of Minnesota
and FRB Minneapolis

Anderson L. Schneider

University of Minnesota
and FRB Minneapolis

Job Market Paper

March 8, 2009

Abstract

Although time consistency problems play an important role in public policy, game theoretical models in macroeconomics seem to indicate the opposite. Due to the complexity of this kind of models, it is commonly assumed that information is complete and perfect. In turn, this assumption becomes the key element that allows agents to coordinate perfectly to punish the government if it does not do what private agents want. As a result, a wide range of feasible payoffs can be sustained as equilibrium, including the best payoff under commitment. Since this approach is widely used for normative purposes a natural question emerges: are the above results robust to small variations in information? This paper analyzes an investment taxation problem in an economy with incomplete information. Specifically, we study an environment with the following main characteristics: 1) the aggregate productivity (fundamental) is stochastic, 2) only the government observes it and; 3) every agent privately receives a noisy signal about the fundamental. The first characteristic implies that the best policy (tax on investment) with commitment is state contingent. The second and third characteristics make the information incomplete. In particular, agents have different information sets, and therefore different beliefs, about the true state of the economy. As a result, independently of the accuracy of the signal, incomplete information reduces the set of equilibrium payoffs. First, we show that any policy that depends solely on the fundamental cannot be an equilibrium. Second, the best equilibrium policy is independent of the fundamental. Finally, for any discount factor strictly smaller than one and for any size of the noise, the best equilibrium is inefficient.

*The authors are grateful to Larry Jones for his support and encouragement. We would also like to thank V.V. Chari, Chris Phelan and Hakki Yazici for long lasting discussions. Remaining mistakes are ours. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

†Contact Piguillem: Department of Economics, University of Minnesota, 4-101 Hanson Hall, 1925 Fourth Street South, Minneapolis, MN 55455. Email: `fpiguill@econ.umn.edu`

1 Introduction

Since the seminal work of Kydland and Prescott (1977), a great deal of research has been devoted to the study of time inconsistency problems. In short, an optimal policy under commitment is time inconsistent if its continuation plan is not optimal. In another words, governments or policy makers have the incentive to break their promises when re-optimizing ex post, giving rise to inefficient outcomes. These types of problems have been studied in a wide variety of policy settings but the most common ones are related to capital taxation, optimal monetary policy and default decisions. In their original work, Kydland and Prescott (1977) propose that the way to avoid these problems is to tie the hands of policy makers after the fact by forcing them to use rules (set ex ante with commitment) rather than allowing them discretion (set ex post, without commitment). The difficulty with this approach is that it undermines the capability of the government to react optimally when there are changes in the fundamentals.

Following up on this idea, Chari and Kehoe (1990) and (1993) showed that if the policy maker is patient enough, the optimal policy under commitment is sustainable even when the government is endowed with full discretion. Their argument relies on reputational considerations borrowed from game theory. A key factor in the result is the assumption that information is complete and perfect. As a consequence of this assumption, if after some history the government is found to have deviated from the optimal policy, every agent would observe it and all of them would, in the best equilibrium, coordinate to punish the government. Notice that to achieve this kind of equilibrium it is important not only that all agents know that the government has deviated but also that all agents know that everyone knows, that everyone knows and so on–the usual recursion due to common knowledge. Therefore, coordination is not only possible but also perfect. Because of their simplicity and tractability these kinds of models have become the dominant tool to analyze environments without commitment in macroeconomics. Thus, although time inconsistency problems play an important role in public policy, reputational considerations can be used as a way of solving the problem without the need of institutional reforms designed to mimic commitment.

This paper analyzes time inconsistency problems in economies with incomplete information. Specifically, we study an environment similar to Chari and Kehoe (1990) in which an investment decision must be made by private agents before the government sets a capital income tax. In that setting, if there is no commitment and a finite horizon, governments (even benevolent ones) will set capital tax rates too high. They show however, that in an infinitely repeated setting, trigger strategies can enforce the commitment outcome when agents are sufficiently patient. We deviate from the Chari and Kehoe example in two important ways. First, we assume that aggregate productivity in the economy (the fundamental) is stochastic. This by itself does not change the Chari and Kehoe result - if agents are patient enough, the full commitment outcome can be supported. Second, we assume that agents do not share the same information. Specifically, we assume that the government sees the true aggregate state, but that this is not observed by private agents. Rather, each agent privately receives a different signal (payoff relevant) about the aggregate state of the economy. The signals can be made arbitrarily close to the aggregate state.

Here, agents have different information sets, and therefore different beliefs, about the fundamentals. In this environment the optimal policy under commitment depends solely on the state of the fundamental. However, if the government deviates agents cannot be certain if what happened is actually a deviation or is the optimal reaction to a change in the fundamentals. In addition, agents do not know other agents' beliefs. As a result, independent of the accuracy of the signal, incomplete information reduces the set of equilibrium payoffs.

First, we show that strategy profiles for the government that depend solely on the fundamentals along the equilibrium path cannot be equilibrium profiles. In particular, there is no strategy profile in the repeated game that delivers the best allocation under commitment, regardless the punishment prescription off the equilibrium path. Second, we show that when government's private shock takes on two values and agents are patient enough the best equilibrium can be achieved with strategy profiles that depend only on public histories. In another words, the best equilibrium is a policy independent of the fundamentals. Finally, for any discount factor strictly smaller than one the best equilibrium is inefficient.

The first result is in line with the literature about repeated games with private imperfect monitoring and a finite number of players (e.g. Mailath and Samuelson (2006)). However, the reasoning behind the argument is slightly different. In games with private imperfect monitoring problems arise when players do not know exactly in which state they are (on or off the equilibrium path) and therefore they are not able to coordinate to punish each other off the equilibrium path. In our environment, the government always knows with certainty everything that has happened while the agents have different beliefs about past histories of the fundamental. If the agents trusted the government when using a strategy that depends only on its private information the government would defect and no agent would be able detect the deviation. Thus, any equilibrium strategy has to depend on some object that is fully observed by the agents. Since in our environment the only variable that is perfectly observed by every player is the tax on investment (the action space for the government) any equilibrium strategy has to depend on past taxes. Moreover, if the strategy for the government depends only on the history of taxes then the results about environments with perfect and complete information carry over entirely. The fact that the government's strategy depends on a history that is perfectly observed by every player allows the full coordination of the agents to punish the government if it deviates. Although the size of the set of public equilibria depends on the size of the discount factor, in the economy analyzed here this kind of strategy generates payoffs that are uniformly bounded away from the best one (the best payoff under commitment).

Would equilibrium strategies that depend on both public and private information increase the payoff? The answer is no. Given that agents have no way to foresee the fundamental, the game becomes one of repeated adverse selection. Thus, when an equilibrium strategy depends on private information punishments happen with positive probability on the equilibrium path. The punishment takes the form of a smaller continuation payoff after those actions that are especially tempting. Nonetheless, the "punishment cost" could be compensated with a larger present payoff fitting the present action to the realization of the fundamental. Unlike the usual environments in game theory, here the agent and the principal have the same payoff functions. Therefore, optimal punishments along the Pareto frontier arbitrarily close to the optimal average welfare are not available. In the language of contract theory, every punishment for the agent (the government) would hurt the principal as well (the agents in this economy). Since the punishment cost is always at least as large as the gain from discretion, the best equilibrium implies a policy that is independent of the fundamental.

Regarding the literature about this topic, to the best of our knowledge there are three closely related papers: Sleet (2001), Athey, Atkeson, and Kehoe (2005) and Sleet and Yeltekin (2006). The first paper considers the problem of a monetary authority that receives a private signal about the true state of the economy and both households and firms have the same information set. They show that under some conditions the optimal policy with commitment is an equilibrium, while in other cases the monetary authority chooses not to use the private signal. The second paper, again in an optimal monetary policy context, considers an environment where agents have the same information sets and only the policy maker observes the (random) true state of the economy. They find that if the time inconsistency problem is "severe" the optimal policy is independent of the true state of the economy; otherwise some dependency is allowed. Sleet and Yeltekin (2006) also analyze an economy with government debt in which private agents have the same information sets, but the government privately observes a taste shock related to the public good consumption. They find that the interaction between informational frictions and the possibility of debt repudiation yields more persistence in both taxes and debt (in the best sustainable allocation) when compared to the benchmark economy (with full commitment and complete information).

This paper differs from the above in the following ways.First, agents have different information sets. Second, in the above papers the information privately known by the government does not directly affect either the payoff or the feasibility set of the agents. Consequently, agents cannot extract from their information sets any useful information about the signals received by the government. In this paper, agents can foresee in an arbitrarily precise way the signal received by the government. Therefore, their results can be viewed as the limit case of the economy studied here. Finally, all the papers mentioned before analyze Public

Perfect Equilibria. That is, the strategy space of the government is constrained to include only the last realization of its private information. This paper extends the result to the unconstrained strategy space.

The paper proceeds as follows: Section 2 describes the environment. Section 3 defines and characterizes the equilibria. In Sections 4 and 5 we characterize the best equilibrium under commitment. Section 6 shows the inefficiency result. Section 7 describes an alternative environment for which the results go through. The last section concludes.

2 The Economy

2.1 Uncertainty

We consider a repeated game with a benevolent government and a continuum of households indexed by $i \in I = [0, 1]$. Time is discrete and indexed by $t = 0, 1, 2, ...$ At the the beginning of every period t, the outcome $R_t \in \Upsilon$ of a random variable $\hat{R}_t$ is realized. The set Υ has cardinality equal to N. Let $\bar{R} \equiv max_{\hat{R}\in\Upsilon}\{R\}$. The outcome R_t is observed only by the government. Each $\hat{R}_t$ is distributed i.i.d. over time with probability distribution P. The process $\{\hat{R}_t\}_{t=0}^{\infty}$ is independent of any choices made by the government or households.

In each period, conditional on the realization of $\hat{R}_t$, each individual privately observes a draw $y_t^i \in [1, \overline{y}]$ of a random variable Y from the probability density function $f(y|R_t)$ and distribution function $F(y|R_t)$. The value $\overline{y}$ may be infinite. Conditional on the realization of the aggregate state, the individual shocks are i.i.d. across agents. We assume a version of the Law of Large Numbers with a continuum of random variables relying on the construction of Sun (2006). We also shall impose the following.

Assumption 1. *Stochastic processes.*

1. *For all* $R, R' \in \Upsilon$, $F(y|R) > F(y|R')$ *if* $R' > R$. *(First order stochastically dominance)*

2. $f(y|R) > 0$ *for all* $R \in \Upsilon$ *and almost all* $y \in [1, \overline{y}]$. *(Full support)*

3. *For all* $R \in \Upsilon$, $\int_1^{\overline{y}} y f(y|R) dy = R$

The first condition assures that given individual actions, aggregate output and aggregate investment are strictly increasing in the realizations of the aggregate state. The second condition is a technical assumption, standard in the literature, that prevents dealing with zero probability events. The full support assumption makes sure that everything can happen in every period, independently of the realization of $\hat{R}$. Finally, the third assumption is just for simplicity and to save on notation.

2.2 Stage Game: Actions and Payoffs

There are three goods in the economy in each point in time: two private goods, consumption and investment, and a public good. In each period every agent receives a physical endowment of $\omega > 0$ units of the private good. After observing the individual shock y^i, each household chooses investment $x_t^i(y^i) \in [0, \omega]$. Returns on investment in period t are agent-specific, and given by the draw y_t^i. Given individual investment decisions, the realized aggregate output, $Y(R_t) = \int_0^1 \int_1^{\overline{y}} y_t^i x_t^i(y_t^i) f(y_t^i|R_t) dy_t^i di$ and aggregate investment $X(R_t) = \int_0^1 \int_1^{\overline{y}} x_t^i(y_t^i) f(y_t^i|R_t) dy_t^i di$ are only observed by the government. Next, the government chooses a tax on investment $\tau_t \in [0, 1]$. Agents are risk neutral in the private good. If household i invests x_t^i and the government sets taxes equal to τ, then its individual consumption is given by $c^i = (1 - \tau) y^i x^i + (\omega - x^i)$.

There is a technology that automatically transforms aggregate output into a public good g. If the aggregate state is R_t and the government sets tax on investment τ_t then the amount of public good provided, as a function of the aggregate state, is given by:

$$g_t(R_t) = \tau_t Y_t(R_t) \tag{1}$$

We assume that agents do not observe the provided amount of the public good. This is a strong assumption that greatly simplifies the analysis and the exposition of the main results. As we explain in Section 7 there is an alternative environment where at the end of each period agents do observe g_t but with a noise that is uncorrelated with y^i. The following results are still true when this noise approaches zero. We interpret this assumption as a situation in which even when agents can observe the final action of the government, they disagree in the amount of resources needed to achieve that result.

Preferences are separable between the private and the public good. If household i gets a draw y^i, invest x^i, the tax on investment is τ and aggregate output is Y, its payoff is given by:

$$u(y^i, \tau, Y, x^i) = [(1-\tau)y^i x^i + (\omega - x^i) + v(\tau Y)] \tag{2}$$

where $v : \mathbb{R}_+ \mapsto \mathbb{R}$ is twice continuously differentiable and strictly concave function.

Given the realization of R_t, the profile of individual investment functions and tax τ_t, government's payoff is given by:

$$W(R_t, \tau_t; x^i_{i\in I}) = \int_0^1 \int_1^{\overline{y}} u(y^i_t, \tau_t, Y_t, x^i_t) f(y^i_t|R_t) dy^i_t di$$

We abuse notation writing the above function as in (3) below. We can do this because given any profile of individual investment functions the payoff for the government depends only on the aggregate values for investment and output. In addition, as we show later since agents are ex-ante identical, the decision function for all agents are equal.

$$W(R_t, \tau_t, X_t, Y_t) = \int_1^{\overline{y}} u(y^i_t, \tau_t, Y_t, x^i_t) f(y^i_t|R_t) dy^i_t \tag{3}$$

The following sequence of events summarizes the information structure of the stage game:

1. R is realized;

2. y^i is drawn for each i;

3. Each individual chooses x^i;

4. Government privately observes R_t and sets τ;

5. The public good is produced according to (1);

6. Consumption is realized.

In order to simplify the future analysis, we shall assume that the marginal value of provision of the public good is higher than the marginal value of individual consumption in any possible state.

Assumption 2 (Optimality of full taxation). $v'(\bar{R}) > 1$.

To understand this condition, take any period t after which all the agents have chosen x_t^i, and therefore both R_t and y_t^i have been realized. Let X_t be the aggregate investment at time t. Then, using the government budget constraint, the one period utility for the government is given by:

$$W(R_t, \tau_t, X_t, Y_t) = (1-\beta)[(1-\tau_t)Y_t + (\omega - X_t) + v(\tau_t Y_t)]$$

If the government increases taxes slightly, say by $d\tau_t$, the benefit of increasing the public good is given by $Y_t(R_t)v'(\tau_t Y_t(R_t))d\tau_t$, while the loss in private consumption is given by $Y_t(R_t)d\tau_t$. Combining these two effects, the government has incentives to increase the current tax as long as $Y_t(R_t)[v'(\tau_t Y_t(R_t)) - 1] > 0$, which is guaranteed by Assumption 2. This assumption implies that in every state the optimal deviation for the government is to set the tax in 100%.

3 Perfect Bayesian Equilibrium

In the repeated game, a public history is a collection of variables that have been observed by all the players. At the beginning of period t, a public history is $h^{P,t} \equiv \{\tau_0, ..., \tau_{t-1}\}$. Let the set of all public histories $h^{P,t}$ at time t be given by $H^{P,t}$. In contrast, a history for the government at time t consists only of the observed outcomes by the government. In terms of outcome paths, because households are competitive there is no loss of generality if we define private histories for the government which do not include the aggregates [1]. In this way, let $h^{g,t} \equiv \{h_0^g, ..., h_t^g\} \in H^{g,t}$ with $h_s^g = \{\tau_{s-1}, R_{s-1}\}$ if $t > 1$ and $H^{g,0} = \emptyset$. A history for individual i is given by $h^{i,t} \equiv \{h_0^i, ..., h_t^i\} \in H^{i,t}$ with $h_s^i = \{\tau_{s-1}, y_{s-1}^i\}$ if $s > 1$ and $H^{i,0} = \emptyset$.

The information sets for an individual player i at t correspond to all histories of the game $h^t \equiv \{h_0, ..., h_t\} \in H^t$ with $h_s = \{\tau_{s-1}, R_{s-1}, \{y_{s-1}^i\}_{i \in I}\}$ that are consistent with her own history at time t.

We restrict the analysis to pure strategies for the government. A pure strategy for the government is a sequence $\{\sigma_{G,t}\}_{t=0}^{\infty}$ with $\sigma_{G,t} = H^{g,t} \times \Upsilon \to [0,1]$. A strategy for an agent $i \in I$ is given by $\{\sigma_{i,t}\}_{t=0}^{\infty}$ with $\sigma_{i,t} : H^{i,t} \times [1, \bar{y}] \to [0, \omega]$. Both $\sigma_{G,t}$ and $\sigma_{i,t}$ are assumed to be measurable functions.

In order to consider any kind of perfection, given the informational restrictions, individual agents have to form beliefs over their information sets. We have opted to analyze Perfect Bayesian equilibria. Let $\mu(\cdot|\tilde{h}^{i,t}, y^i)$ be the probability distribution over histories $\hat{h}^{g,t} \in H^{g,t}$ consistent with individual history $\tilde{h}^{i,t}$. Let Σ_G be the set of possible strategy profiles for the government and Σ be the set of possible strategy profiles $\sigma = (\sigma_G, \{\sigma_i\}_{i \in I})$. A strategy profile $\sigma \in \Sigma$ induces, after any history $h^t \in H^t$, a continuation profile $(\sigma_G|_{h^{g,t}}, \{\sigma_i|_{h^{i,t}}\}_{i \in I}) \in \Sigma$.

Given the risk neutrality assumption and the fact that agents are both anonymous and atomistic, optimality for the individuals can be reduced to a simple rule. Given $\sigma_G \in \Sigma_G$, let $E_{\sigma_G}(\tau|h^{i,t}, y^i)$ be the conditional expectation that a household with history $h^{i,t}$ and idiosyncratic return y^i has about the random variable $\sigma_{G,t}$. This function, for each individual i, is measurable with respect to the sigma-algebra generated by his individual histories. A household with history $h^{i,t}$ and idiosyncratic return y^i will invest a positive amount only if the expected marginal return on investment is positive:

[1]For a detailed explanation of this reasoning, see Chari and Kehoe (1990)

$$x^*(h^{i,t}, y^i) = \begin{cases} \omega & \text{if } y^i E_{\sigma_g}(1-\tau|h^{i,t}, y^i) - 1 > 0 \\ [0,\omega] & \text{if } y^i E_{\sigma_g}(1-\tau|h^{i,t}, y^i) - 1 = 0 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

Given a profile $\sigma \in \Sigma$ and a sequence of belief profiles $\mu \equiv \{(\mu(\cdot|h^{i,t}, y^i))_{i\in I}\}_{t=0}^{\infty}$, expected payoffs for the players are naturally defined from the stochastic outcomes that the strategies induce. The payoff for the government at time zero in the repeated game is given by:

$$V(\sigma) = (1-\beta)E\left[\sum_{t=0}^{\infty} \beta^t W(R_t, X_t, \tau_t, Y_t)\right]$$

Notice that, since the government observes all the aggregates in the game, it does not need to form beliefs about individuals' actions. It only needs to take into consideration that its own actions affect individuals' beliefs.

Definition 1. *A pair (σ, μ) consisting of strategy profiles and belief profiles is a Perfect Bayesian Equilibrium if:*

(i) Given $\{\mu^i\}_{i\in I}$, $\sigma_{i,t}(h^{i,t}, y^i) = x^*(h^{i,t}, y)\ \forall i \in I, h^{i,t} \in H^{i,t}, \forall y^i \in [1, \bar{y}]$;

(ii) $V(\sigma_G|_{h^{g,t}}, \{\sigma_i\}_{i\in I}) \geq V(\widetilde{\sigma}, \{\sigma_i\}_i)\ \forall \widetilde{\sigma} \in \Sigma_G, h^{g,t} \in H^{g,t}, \forall R$;

(iii) Beliefs are given by Bayes' rule whenever possible.

Conditions (i) and (ii), respectively, require that, given beliefs, the government's and the individual's continuation strategies be best responses to each other after any history. Households' deviations cannot be detected and therefore at each period they maximize, given their beliefs, utility from private consumption. Regarding individuals' decisions about investment, we shall assume the following:

Assumption 3 (**Monotone Likelihood Ratio**). *For all* $R_H, R_L \in \Upsilon$ *and for all* $\hat{y}, y \in [1, \bar{y}]$ *we have that* $\frac{f(\hat{y}|R_L)}{f(y|R_L)} \leq \frac{f(\hat{y}|R_H)}{f(y|R_H)}$ *if* $\hat{y} \geq y$ *and* $R_H > R_L$.

Assumption 4 (**Analytical pdf**). *For each* $R \in \Upsilon$, $f(.|R) : [1, \bar{y}] \to \mathbb{R}$ *is analytic.*

A function defined on the real line is analytic if it is equal to its Taylor expansion. The role of Assumption 4 is to guarantee that for any $\sigma \in \Sigma$, the set of individuals indifferent between investing or not has Lesbegue measure zero. This assumption is not crucial for the main result of this paper but simplifies the characterization of the individual decisions. Without it, is not even clear that aggregate investment, and therefore government's revenues, are decreasing in individual taxes. Analytical functions are fairly common among continuous and differentiable distributions.[2]

Lemma 1. *Under Assumption 4, for any* $\sigma \in \Sigma$ *and any* $h^t \in H^t$, *the set of individuals for which* $y^i E_{\sigma_g}(1 - \tau|h^{i,t}, y^i) - 1 = 0$ *in (4) has Lesbegue measure zero.*

Proof: In the appendix.

4 Ramsey Equilibrium

Before proceeding with the characterization of Perfect Bayesian equilibria, we first consider the benchmark case in which the government has a commitment technology that it is used to bind itself to a tax policy $\sigma_G : \Upsilon \to [0, 1]$ in each period. When such technology is available, the static nature of the government's problem allows us to restrict the analysis to an one-period game. Following the literature, we call it the Ramsey game. The introduction of a commitment technology can be formalized by changing the timing of the one shot Bayesian game. The Ramsey game that we analyze evolves as follows:

[2]Fox example, Normal, Uniform, Exponential, Gamma, Beta and Pareto are analytic. Among the distributions satisfying the monotone likelihood ratio are the Normal, Exponential, Uniform and Beta.

0. The government sets a tax policy $\sigma_G : \Upsilon_R \to [0,1]$;

1. R is realized;

2. y^i is drawn for each i;

3. Each individual chooses x^i;

4. Government observes both R and aggregate output Y and it sets τ according to $\sigma_G(\cdot)$;

5. Public good is produced according to equation (1);

6. Consumption of both goods is realized.

There are two differences with the stage game in Section 2. First, the government sets a tax policy $\sigma_G(\cdot)$ before observing R. Second, only after aggregate output is realized government learns about the realization of R and sets the investment tax according to $\sigma_G(\cdot)$. This specific choice about the sequence of events implies the existence of a strategy profile in the Bayesian game that attains, despite incentive questions, the outcome in the Ramsey game. At the same time it prevents any discussion about communication issues.

Given a a tax policy σ_G, let $E_{\sigma_G}(\tau|y^i)$ be the conditional expectation that a household with draw y^i has about the random variable σ_G .

Definition 2. *The Ramsey equilibrium is a function* $\sigma_G : \Upsilon_R \to [0,1]$ *and, for each* $i \in [0,1]$, *a function* $\sigma_i : [1, \overline{y}] \to [0, \omega]$ *such that:*

a) σ_G *maximizes* $\int_\Upsilon W(R,\tau,Y)dP(R)$ *given* $\sigma_i|_{i\in I}$.

b) $$\sigma_i^*(y) = \begin{cases} \omega & \text{if } y \cdot E_{\sigma_G}(1-\tau|y) - 1 > 0 \\ [0,\omega] & \text{if } y \cdot E_{\sigma_G}(1-\tau|y) - 1 = 0 \\ 0 & \text{otherwise} \end{cases}$$

It turns out to be difficult to characterize in a precise way the Ramsey policy with a general function $v(.)$. All what we can say is that $g(R) < g(R')$ if $R < R'$, that is, the government spending is higher in more productive states.[3] Since aggregate output, $Y(R)$, is increasing in R as well, the magnitude of the taxes is unclear. The properties and relative size of taxes would be very different depending on the shape of $v(.)$, even when it is assumed that this function is strictly concave and twice differentiable. For that reason in the next Lemma we assume that $v(.)$ is linear. That is, $v(x) \equiv b \cdot x$. Of course, because of Assumption 2, $b > 1$. When that is the case the characterization is intuitive and straightforward. Moreover, it highlights the main complications related to the individual investment decisions. The next lemma characterizes the equilibrium of the Ramsey game.

Lemma 2. *Under Assumptions 2-4 if $v(.)$ linear, then the Ramsey equilibrium is given by:*

1. $(1 - \sigma_G^*(\bar{R}))\sigma_G^*(R) = 0$ for all $R \in \Upsilon$, $R \neq \bar{R}$
2. $\sigma_i^*(y^i)$ is given by item b) in the definition of the Ramsey equilibrium.

Proof: In the appendix.

Lemma 2 states that the solution to the Ramsey game with linear utility is in a corner. By Assumption 2, taxes being either zero or one in all states cannot be a solution. Moreover, the government taxes a positive amount in the highest aggregate state. This happens because taxing in the highest state is always less costly than taxing in a lower state. For any given average tax, it is always possible to increase the payoff by increasing the tax in the highest state and reducing the tax in a lower state in such a way that the average tax remains the same. The proof of the lemma exploits this idea. It is worth to comment the role of Assumptions 3-4 in the lemma. If the function $H(y;\tau) = y \cdot E_{\sigma_G}(1 - \tau|y) - 1$ had the single crossing property[4], then the characterization of the individual decisions would be very

[3]Here the difficulty is similar as in the Mirrleesian literature, where is not known in general whether or not workers with higher productivity work more than low productivity workers.

[4]By single crossing we mean the following: there exists $y^* \geq 1$ such that $\sigma_i(y) = \omega$ if $y \geq y^*$ and $\sigma_i(y) = 0$ otherwise.

simple. This is not always the case for any possible tax function $\tau : \Upsilon \to [0,1]$. In the proof of Lemma 2 we exploit those assumptions in order to handle the more general case in which there are multiple agents indifferent between investing or not.

For future reference, we define the Ramsey strategy profile $(\sigma_G^*, \sigma_i^*) \in \Sigma$ as the repetition of the Ramsey equilibrium. Formally, for all $h^{g,t} \in H^{g,t}$ and $\hat{R} \in \Upsilon$, let $\sigma_{G,t}^*(h^{g,t}, \hat{R}) = \sigma_G^*(\hat{R})$. Regarding household's strategies, for all $h^{i,t} \in H^{i,t}$ and $y^i \in [1, \overline{y}]$, let $\sigma_{i,t}(h^{i,t}, y^i) = \sigma_i^*(y^i)$ in Lemma 2. Beliefs are given by $\mu(\hat{R}|\hat{y}^i) = \frac{f(\hat{y}^i|\hat{R})P(\hat{R})}{\sum_{R\in\Upsilon} f(\hat{y}^i|R)P(R)}$.

5 Unattainability of the Ramsey Outcome

In this section we show that there is no strategy profile $(\sigma_G, \sigma_i) \in \Sigma$, together with some belief system, that yields the outcome path of the Ramsey equilibrium. If the Ramsey outcome were an equilibrium, it should be the case that **on the equilibrium path** the government was playing actions that depend only on the current shock R_t. But then, the government would have profitable deviations. For instance, every time that the government is supposed to play the lowest tax prescribed by the equilibrium strategy it could choose the highest tax consistent with equilibrium behavior. Because agents cannot be certain about the real value of R this deviation would be undetectable. This result remains true regardless the punishment prescription off the equilibrium path. In order to state proposition 1, given a pair of strategies $\sigma = (\sigma_g, (\sigma_i)_{i\in I}) \in \Sigma$ and a belief profile μ, the outcome path is denoted by a sequence of stochastic functions $Z = \{(x_t^i(\sigma), c_t^i(\sigma))_{i\in I}, g_t(\sigma), \tau_t(\sigma)\}_{t=0}^{\infty}$. As usual, the outcome path is defined as the induced outcome starting from the initial history $H^{g,0}$.

In what follows, let $Z^* = \{\{x_t^i(\sigma^*), c_t^i(\sigma^*)\}_{i\in I}, g_t(\sigma^*), \tau_t(\sigma^*)\}$ denote the outcome path of the Ramsey allocation, where σ^* was defined at the end of the last section.

Proposition 1. *Under assumptions 2 and 4 there is no belief system μ and $\sigma \in \Sigma$ that generates Z^* on the equilibrium path.*

Proof: Suppose not, that is, there is a pair of strategy profiles $\hat{\sigma} \in \Sigma$ and a belief system μ such that Z^* is an equilibrium outcome.

Let $S = \left\{ \tau \in [0,1] : \tau = \sigma^*(R_t), R_t \in \Upsilon \right\}$ and define:

$$H^{*,i} = \{h^{i,t} \in H^i : \tau_s \in S, \forall s \leq t-1\}$$

Notice that $H^{*,i}$ is the set of possible histories on the equilibrium path for household i. In addition, because of the full support assumption, all individual histories in $H^{*,i}$ have non-zero measure.

Given $\hat{\sigma} \in \Sigma$, let $\mu(\cdot|h^{i,t}, y_t^i)$ be the induced probability distribution over $H^{g,t} \times \Upsilon$ given the history $(h^{i,t}, y_t^i)$. In the appendix we show that, for all $h^{i,t} \in H^{*,i}$, the belief system should follow the following updating:

$$\mu(h^{g,t}, R_t|\tilde{h}^{i,t}, y_t^i) = P(R_t|y_t^i)\frac{\mu(h^{g,t-1}|\tilde{h}^{i,t-1})}{\sum_{\hat{h}^{g,t-1}} \mu(\hat{h}^{g,t-1}|\tilde{h}^{i,t-1})}$$

By recursive calculations the above expression can be reduced to

$$\mu(h^{g,t}, R_t|\tilde{h}^{i,t}, y_t^i) = f(y_t^i|R_t)\prod_{s=0}^{t-1}\frac{f(y_s^i|R_s)}{\sum_{\hat{R}_s} f(y_s^i|\hat{R}_s)} \qquad (5)$$

That is, belief are unaffected by government actions. In addition, $\mu(\hat{\sigma}(h^{g,t}, R_t)|h^{i,t}, y_t^i) = \mu(\hat{\sigma}(h^{g,t}, R_t)|\tilde{h}^{i,t}, y_t^i)$ for all $h^{i,t}, \tilde{h}^{i,t} \in H^{*,i}$ and all $y_t^i \in [1, \overline{y}]$. Therefore, equation (4) implies that $\hat{x}^i(h^{i,t}, y_t^i) = x^i(\tilde{h}^{i,t}, y_t^i)$ for all $h^{i,t}, \tilde{h}^{i,t} \in H^{*,i}$ and all $y_t^i \in [1, \overline{y}]$. Take some period t and $R' \in \Upsilon$ such that $\hat{\sigma}(h^{g,t}, R') < \max_{\tau \in S} \tau$. Then consider the following one shot deviation by part of the government:

$$\tilde{\sigma}_G(h^{g,s}, R) = \begin{cases} \tau^D \equiv \max_{\tau \in S} \tau & \text{if } s = t \text{ and R=R'} \\ \hat{\sigma}(h^{g,s}, R) & \text{otherwise} \end{cases}$$

Following history $(h^{g,t}, R')$, the equilibrium strategy generates a government's payoff of:

$$(1-\beta)W(R_t, \hat{\sigma}(h^{g,t}, R'), X(h^{g,t}), Y(h^{g,t})) + \beta V(\hat{\sigma}|_{\{h^{g,t}, R', \hat{\sigma}(h^{g,t}, R')\}}, (\hat{\sigma}_i|_{\{h^{i,t}, y_t^i, \hat{\sigma}(h^{g,t}, R')\}})_{i \in I}) \quad (6)$$

where $X(h^{g,t})$ and $Y(h^{g,t})$ are the aggregates following the outcome path after history $h^{g,t}$.[5]

The one shot deviation strategy generates a payoff of:

$$(1-\beta)W(R_t, \tau^D, X(h^{g,t}), Y(h^{g,t})) + \beta V(\hat{\sigma}|_{\{h^{g,t}, R', \tau^D\}}, (\hat{\sigma}_i|_{\{h^{i,t}, y_t^i, \tau^D\}})_{i\in I}) \tag{7}$$

Since $\tau^D \in S$ it follows that $\{h^{i,t}, \tau^D, y_t^i\} \in H^{*,i}$ and therefore $\hat{\sigma}^i|_{\{h^{i,t}, \hat{\sigma}(h^{g,t}, R'), y_t^i\}} = \hat{\sigma}^i|_{\{h^{i,t}, \tau^D, y_t^i\}}$ for all $i \in I$. In addition, $\hat{\sigma}|_{\{h^{g,t}, R', \hat{\sigma}(h^{g,t}, R')\}} = \tilde{\sigma}|_{\{h^{g,t}, R', \tau^D\}}$ by construction. Thus, the continuation payoffs are equal in both (6) and (7). By assumption 2, $\hat{\sigma}(h^{g,t}, R') < \tau^D$ implies that $W(R_t, \tau^D, X(h^{g,t}), Y(h^{g,t})) > W(R_t, \hat{\sigma}(h^{g,t}, R'), X(h^{g,t}), Y(h^{g,t}))$. Hence the deviation is profitable, a contradiction ∎

Assumption 4 plays a role in proposition 1 only to the extent that it guarantees that a measure zero of agents is indifferent between investing or not. A crucial feature that prevents any strategy achieving the Ramsey outcome is the fact that, otherwise, the individual strategies do not depend on history on the equilibrium path. The government then takes advantage of this situation, defecting whenever possible.

Although the Ramsey payoff cannot be attained, one may wonder if it can be approached arbitrarily close for high enough discounting. The answer to this question is negative under some circumstances, and we will elaborate it in the next section. Notice that if indeed there is a strategy profile than can approach the (repeated) payoff of the Ramsey equilibrium, such profile requires some coordination among agents. In another words, a positive measure of agents should have strategies depending on public histories. If only a measure zero of agents could coordinate any punishment that they could used would have no effect on the government's payoff.

Proposition 1 can actually be made stronger. As its proof makes clear, there is nothing special to the Ramsey outcome other than **on the equilibrium path** taxes are independent of history [6] and stochastic.

[5] Since individual decisions are measurable with respect to their information sets, the aggregates cannot depend on the actual realization of R_t after history $h^{g,t}$.

[6] Or in another words, the outcome is just the repetition of the static Ramsey equilibrium.

In order to formalize the above statement, consider the following class of strategies;

Definition 3. *Let* $R^t = (R_{0,\ldots,}R_t) \in \Upsilon^t$ *be a history of the aggregate state. A strategy profile* $\sigma_G \in \Sigma_G$ *is* ***purely private*** *if for all* t, $h^{g,t}, \hat{h}^{g,t} \in H^{g,t}$ $\sigma_{G,t}(h^{g,t}, R_t) = \sigma_{G,t}(\hat{h}^{g,t}, \hat{R}_t)$ *if* $R^t = \hat{R}^t$.

Definition 4. *A strategy profile* $\sigma_G \in \Sigma_G$ *is* ***non-trivially purely private*** *if*

1. *It is purely private;*
2. *There exist* t *and* $R^t, \hat{R}^t \in \Upsilon^t$ *such that* $R^t \neq \hat{R}^t$ *with* $\sigma_{G,t}(h^{g,t}, R_t) \neq \sigma_{G,t}(\hat{h}^{g,t}, \hat{R}_t)$.

As in Proposition 1 we can define the outcome path of a purely private strategy profile for the government in the usual way. Then, we have the following proposition;

Proposition 2. *Under Assumptions 2-4, there is no belief system* μ *such that the outcome of* ***non-trivially purely private*** *strategy profile for the government is a Perfect Bayesian Equilibrium.*

Proof: In the appendix.

The main steps of the proof are similar to those in Proposition 1. If **on the equilibrium path** the government were following a purely private strategy, there would be more than one tax consistent with equilibrium behavior in at least one period. Then, when is time to choose the low tax, the government could deviate choosing the highest tax prescribe by the strategy. By the same arguments as in Proposition 1 these kinds of deviations are not detectable. The main implication of Proposition 2 is that the problem becomes one of repeated hidden information as if the households had not information whatsoever about the true state of the economy. This fact allow us to use the usual tools for repeated agency problems where the households are the principal and the government is the agent.

6 Best Equilibrium and Inefficiency Result

Proposition 2 implies that in any equilibrium in which the government uses pure strategies, its strategy profile has to condition actions on public information somehow. If the government condition its actions only on public histories (of previous taxes), a wide range of payoffs can be sustained. In this case the private agents can fully coordinate to punish deviations made by the government, switching to the worst equilibrium. In the worst equilibrium, the government indeed uses a strategy profile that only depends on public information. Regardless the history, the government always taxes investment fully. Anticipating this behavior, agents do not save. As a consequence, government's action after any history is also a best response to individual agents' strategies.

Proposition 3 (**Worst Equilibrium**). *The pair of strategies* $\sigma_{G,t}^{worst}(h^{g,t}, R_t) = 1$ *for all* $h^{g,t} \in H^{g,t}$ *and all* $R_t \in \Upsilon$ *together with* $\sigma_{i,t}^{worst}(h^{i,t}, y_t^i) = 0$ *for all* $h^{i,t} \in H^{i,t}$*, all* $y_t^i \in [1, \overline{y}]$ *and all* $i \in I$ *is a Perfect Bayesian Equilibrium. It yields the lowest payoff amongst Perfect Bayesian equilibria.*

Proof: In the appendix.

Denote by V^{worst} the payoff generated by σ^{worst}.

Given the results in Propositions 1 and 2, in this section we analyze a class of government strategies that depends on both public and private histories. We restrict attention to government strategies that condition behavior on public histories and the most recent realization of the private shock, but not on the entire history of private shocks.[7] However, in appendix 9.7 we show that this result is indeed true when strategies are allowed to depend on the whole history of private information. Since the proof and the arguments used are similar we chose to leave the general case for the appendix.

Let $\mathring{\Sigma}$ be the set of strategy profiles that conform with the restriction explained above. A strategy $\sigma \in \mathring{\Sigma}$ induces a stochastic outcome path. Given $\sigma \in \mathring{\Sigma}$, let $\tau^{t-1}(R^{t-1})$ be the

[7] In doing so we follow the same approach of both Sleet and Yeltekin (2006) and Athey, Atkeson, and Kehoe (2005).

t-period public history induced by σ_G and the sequence of shocks R^{t-1}.

Definition 5. *A sequence of functions $\mathcal{A} \equiv \{\tau_t, X_t, Y_t\}_{t=0}^{\infty}$ is the stochastic* ***aggregate outcome*** *induced by a strategy profile $\sigma \in \mathring{\Sigma}$ if:*

(i) $\tau_0(R_0) = \sigma_{G,0}(R_0)$, *and* $\tau_t(R^t) = \sigma_{G,t}(\tau^{t-1}(R^{t-1}), R_t)$;

(ii) $X_t(R^t) = \int_1^{\overline{y}} x_t^i(R^t, y_t^i) f(y_t^i|R_t) dy_t^i$ *and* $Y_t(R^t) = \int_1^{\overline{y}} y_t^i x_t^i(R^t, y_t^i) f(y_t^i|R_t) dy_t^i$, *where* $x_t^i(R^t, y_t^i)$ *is given by:*

$$x_t^i(R^t, y_t^i) = \begin{cases} \omega & \text{if } y_t^i E(1-\tau_t(R^t)|y_t^i) - 1 > 0 \\ [0, \omega] & \text{if } y_t^i E(1-\tau_t(R^t)|y_t^i) - 1 = 0 \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

Notice that in Definition 5 the tax functions are potentially non-stochastic. Moreover, there is no restriction that constrains the tax functions to be time-stationary.

Proposition 4 gives a full characterization of aggregate allocations induced by profiles $\sigma \in \mathring{\Sigma}$. In order to state the proposition, given an stochastic **aggregate outcome** $\mathcal{A}$ of a profile $\sigma \in \mathring{\Sigma}$, let $V(\mathcal{A}) \equiv V(\sigma)$ denote the time zero expected payoff for the government. Similarly, denote by $V(\mathcal{A}|\tau^t(R^t))$ the continuation payoff from $\mathcal{A}$ after the public history $\tau^t(R^t)$. We also define, for a function $\tau_t \subseteq \mathcal{A}$, the set $\tau_t(\tau^{t-1}(R^{t-1}), \Upsilon) \equiv \{\hat{\tau}|\hat{\tau} = \tau_t(\tau^{t-1}(R^{t-1}), R)$ for some $R \in \Upsilon\}$ consisting of all values for taxes that can be assigned in period t through the mapping $\tau_t(\cdot)$. Finally, let $W^d(R_t, X(R^t), Y(R^t))$ the best deviation that the government can achieve after both the shock R_t and investment decisions $X(R^t)$ (together with aggregate output $Y(R^t)$) is realized. By Assumption 2, such deviation sets taxes to be equal 100%. Then we have the following:

Proposition 4. *$\mathcal{A} \equiv \{\tau_t, X_t, Y_t\}_{t=0}^{\infty}$ is the stochastic aggregate outcome of a Perfect Bayesian Equilibrium $\sigma \in \mathring{\Sigma}$ if and only if the following conditions are satisfied:*

(1) $\forall t$, $X_t(R^t) = \int_1^{\overline{y}} x_t^i(R^t, y_t^i) f(y_t^i|R_t) dy_t^i$ *and* $Y_t(R^t) = \int_1^{\overline{y}} y_t^i x_t^i(R^t, y_t^i) f(y_t^i|R_t) dy_t^i$, *where* $x_t^i(R^t, y_t^i)$ *is given by (8);*

(2) $\forall t, \tau^t(R^t),\ V(\mathcal{A}|\tau^t(R^t)) \geq (1-\beta)W^d(R_t, X(R^t), Y(R^t)) + \beta V^{worst}$*;*

(3) $\forall t, \tau^{t-1}(R^{t-1}), R_t, \hat{\tau} \in \tau_t(\tau^{t-1}(R^{t-1}), \Upsilon)$ *:*

$$(1-\beta)W(R_t, \tau_t(\tau^{t-1}(R^{t-1}), R_t), X(R^t), Y(R^t)) + \beta V(\mathcal{A}|\tau^{t-1}(R^{t-1}), \tau_t(\tau^{t-1}(R^{t-1}), R_t)) \geq$$
$$(1-\beta)W(R_t, \hat{\tau}, X(R^t), Y(R^t)) + \beta V(\mathcal{A}|\tau^{t-1}(R^{t-1}), \hat{\tau})$$

Proof: In the appendix.

Condition (2) in the previous proposition comes from the standard reversion to the worst equilibrium in the case that the government deviates from a prescribed action along the path of play. Condition (3) is an incentive compatibility constraint that prevents the government to make profitable deviations along the path by misrepresenting its private information.

Let Λ be the set of aggregate allocations $\mathcal{A}$ that satisfies the conditions in proposition 4. For a given value of the discount factor β, let Ψ_β be the set of equilibrium payoffs of the repeated game that can be supported by some aggregate allocation $\mathcal{A}$:

$$\Psi_\beta = \{v^* : \exists \mathcal{A} \in \Lambda \text{ and } v^* = V(\mathcal{A})\}$$

The next lemma is a standard recursive result implied by the restriction of equilibria within the set $\mathring{\Sigma}$.

Lemma 3. *A payoff* v^* *is supported by the aggregate outcome* $\mathcal{A}$ *if and only if there exists functions* $\{\tau, X, Y, v'\}$ *with* $\tau : \Upsilon \to \mathbb{R}_+$, $X : \Upsilon \to \mathbb{R}_+$, $Y : \Upsilon \to \mathbb{R}_+$ *and* $v' : \Upsilon \to \mathbb{R}_+$ *such that:*

1. $v = \sum_{R \in \Upsilon} P(R)[(1-\beta)W(R, X(R), \tau(R), Y(R)) + \beta v'(R)]$
2. $\forall R \in \Upsilon,\ \ (1-\beta)W(R, X(R), \tau(R), Y(R)) + \beta v'(R) \geq (1-\beta)W^d(R, X(R), Y(R)) + \beta V^{worst}$
3. $\forall R, \hat{R} \in \Upsilon,\ \ (1-\beta)W(R, X(R), \tau(R), Y(R)) + \beta v'(R) \geq$
$$(1-\beta)W(R, X(R), \tau(\hat{R}), Y(R)) + \beta v'(\hat{R})$$

4. $\forall R \in \Upsilon,\;\; v'(R) \in \Psi_\beta$

5. $X(R) = \int_1^{\overline{y}} x^i(y^i) f(y^i|R) dy^i$ *and* $Y(R) = \int_1^{\overline{y}} y^i x^i(y^i) f(y^i|R) dy^i$*, where* $x^i(y^i)$ *is given by:*

$$x^i(y^i) = \begin{cases} \omega & \textit{if } y^i E(1-\tau(R)|y^i) - 1 > 0 \\ [0,\omega] & \textit{if } y^i E(1-\tau(R)|y^i) - 1 = 0 \\ 0 & \textit{otherwise} \end{cases}$$

For a given list of functions $\{\tau, X, Y, v'\}$ satisfying the conditions in lemma 3, let $\mathcal{OP}(\tau, R)$ be the aggregate allocations $(X(R), Y(R))$ generated by the optimal individual decisions according to condition (5) above. In order to save notation, we define $W(\tau(R), Z(\tau, R)) \equiv W(R, X(R), \tau(R), Y(R))$.

Next we restrict attention to the case in which the government's private shock can take only two possible values. From now on, τ and V will be vectors with each entry being a tax (or continuation value) contingent on R. Given both proposition 4 and lemma 3, the best equilibrium solves the following problem:

$$T = \max_{\{\tau, V\}} \sum_{s=L,H} P(R_s)[(1-\beta)W(\tau(R_s), Z(\tau, R_s)) + \beta V(\tau(R_s))] \tag{PR}$$

subject to:

for all $s = L, H$,

(IC-On) $\quad (1-\beta)W(\tau(R_s), Z(\tau, R_s)) + \beta V(\tau(R_s)) \geq (1-\beta)W(\tau(R_{-s}), Z(\tau, R_s)) + \beta V(\tau(R_{-s}))$

(IC-Off) $\quad (1-\beta)W(\tau(R_s), Z(\tau, R_s)) + \beta V(\tau(R_s)) \geq (1-\beta)W(1, Z(\tau, R_s)) + \beta V^{worst}$

(E) $\quad V(\tau(R_s)) \in \Psi_\beta$

(OP) $\quad Z(\tau, R_s) \in \mathcal{OP}(\tau, R_s)$

The constraint (IC-On) is the incentive compatibility constraint on the equilibrium path. That is, if the equilibrium strategy implies that on the equilibrium path two different taxes are used, and since R is virtually not observable, it has to be on the government's best interest to do it so. Constraint (IC-Off), or sustainability constraint, makes sure that the governments does not want to use a tax that is not contemplated on the equilibrium path. This is only possible if the payoff for any equilibrium tax is larger than the best deviation ($\tau = 1$ for any level of investment) plus the worst possible continuation value. The third constraint requires that every continuation value can be implemented as an equilibrium for some strategy pair and belief system. Finally, the last constraint imposes the optimality of households' decision rules.

Next we show that the above problem can be reduced to a more simple static maximization problem. Let $\bar{V} \equiv \sup\{\Psi_\beta\}$ and consider the following static problem,

$$\hat{T} = \max_{\{\tau,V\}} \sum_{s=L,H} P_{R_s}[(1-\beta)W(\tau(R_s), Z(\tau, R_s)) + \beta V(\tau(R_s))] \qquad \text{(PS)}$$

subject to;

$$\text{(IC-ON-S)} \quad (1-\beta)W(\tau(R_s), Z(\tau, R_s)) + \beta V_s \geq (1-\beta)W(\tau(R_{-s}), Z(\tau, R_s)) + \beta V_{-s};\ \forall s = L, H,$$

$$\text{(OP-S)} \quad Z(\tau, R_s) \in \mathcal{OP}(\tau, R_s) \text{ for all } \tau \in [0,1]^2 \text{ and all } s = L, H$$

$$\text{(E-S)} \quad V_s \in [V^{worst}, \bar{V}] \text{ for all } s = L, H$$

The next result states a key relation between problems PS and PR. In what follows, let $\hat{T}(\tau^*, V^*)$ be the value of the problem PS when its solution is given by (τ^*, V^*). $T(\cdot,\cdot)$ is defined in a similar fashion.

Proposition 5. *Let (τ^*, V^*) be a solution to problem PS and (τ^{**}, V^{**}) be the solution to PR. Then, there exists $\beta^* \in (0,1)$ such that for all $\beta \geq \beta^*$,*

i) $\hat{T}(\tau^*, V^*) \geq T(\tau, V)$ for all τ and V.

ii) If $V^* \in \Psi_\beta$ then $\hat{T}(\tau^*, V^*) = T(\tau^{**}, V^{**})$.

Proof: First notice there are two differences between PS and PR. First, the constraint (IC-Off), present in PR, is not included in PS. Second, the constraint (E) in PR requires that the chosen continuation values belong to the equilibrium value set, while PS only requires that the continuation values V_i belong to a convex and compact set. Notice that if β is large enough the constraint (IC-Off) would not be binding and we can disregard it. Thus, from this point of view, both problems are equivalent. But, by construction $\Psi_\beta \subseteq [V^{worst}, \bar{V}]$, therefore since the objective function is the same in both problems, and the feasible set in PR is smaller than in PS, part *i)* of the proposition follows. In addition, part *ii)* is immediate. If the solution to PS is feasible in PR then it must be the case that this value is the maximum ∎

We say that an strategy profile $\sigma \in \Sigma$ is **public** if for all t, $\tau_{t-1} \in [0,1]$, $h^{g,t} \in H^{g,t}$ and $R, \hat{R} \in \Upsilon$ we have that $\sigma_{G,t}(h^{g,t-1}, \tau_{t-1,}R) = \sigma_{G,t}(h^{g,t-1}, \tau_{t-1,}\hat{R})$. The main result of this section shows that, when agents discount the future high enough, the best Perfect Bayesian Equilibrium (within the class of strategies that we consider) is achieved through the use of public strategies by the government. As a byproduct, the best equilibria is inefficient.

Proposition 6. *If $\beta \geq \beta^*$ then, the solution to PS implies $\tau_L = \tau_H = \tau^B$ and $V_L = V_H = \bar{V}$. Moreover,*

1) $\tau^B = argmax_{\{\tau, Z(\tau,R_s) \in \mathcal{OP}(\tau,R_s)\}} \left\{ E_R \big[W(\tau, Z(\tau, R_s)) \big] \right\}$

2) $\bar{V} = E_{R_s} \big[W(\tau^B, Z(\tau^B, R_s)) \big]$

3) $\bar{V} \in \Psi_\beta$.

Proof: The lagrangian for this problem is

$$\begin{aligned} L &= \sum_{s=L,H} P_s[W(\tau_s, Z(\tau, R_s)) + \beta V_s] + \sum_{s=L,H} \lambda_s[W(\tau_s, Z(\tau, R_s)) + \beta(V_s - V_{-s}) - W(\tau_{-s}, Z(\tau, R_s))] \\ &+ \sum_{s=L,H} \gamma_s \beta[\bar{V} - V_s] \end{aligned}$$

First notice that we are not considering the constraint that $V_s \geq V^{worst}$. This is guaranteed for a β large enough. Second, we do not need to consider the case in which both $\gamma_s > 0$. Because $W(.)$ is strictly increasing in τ, given interior aggregate allocations, the (IC-ON-S) would imply $\tau_L = \tau_H$. In the same way, if $\lambda_s > 0$ in both states, then both (IC-ON-S) would be binding, and therefore $\tau_L = \tau_H$. Thus, we only need to consider cases in which only one λ_s and only one γ_s can be strictly positive.

The first order conditions with respect to V_s are,

$$P_L + \lambda_L - \lambda_H - \gamma_L = 0$$

$$P_H + \lambda_H - \lambda_L - \gamma_H = 0$$

If $\gamma_L = \gamma_H = 0$ the equations above imply $P_L = -P_H$ which is not possible. That is, in a best equilibrium, in at least one state, the continuation value has to be a best equilibrium. Thus, we need to consider two cases.

Case 1: Suppose $\gamma_H > 0$ (hence $V_H = \bar{V}$ and $V_L \leq \bar{V}$), then the above equations imply $\lambda_H = P_L + \lambda_L$ or $\lambda_H > \lambda_L$, since at least one multiplier has to be zero, it follows that $\lambda_H > 0$. One can see by using (IC-ON-S) that in this case $\tau_L \geq \tau_H$ (the best equilibrium requires smaller continuation values for larger taxes).

Case 2: $\gamma_L > 0$ ($V_L = \bar{V}$, $V_H \leq \bar{V}$) and $\lambda_L > 0$. Using a similar argument it follows that $\tau_L \leq \tau_H$ in this case.

Consider case 1. Replacing the binding constraints in the objective function the problem becomes:

$$\hat{T} = \max_{\{\tau_L, \tau_H, V_L\}} P_L W(\tau_L, Z(\tau, R_L)) + P_H W(\tau_L, Z(\tau, R_H)) + \beta V_L$$

subject to;

$$W(\tau_H, Z(\tau, R_H)) + \beta\bar{V} \geq W(\tau_L, Z(\tau, R_H)) + \beta V_L$$
$$Z(\tau, R_s) \in \mathcal{OP}(\tau, R_s) \text{ for all } s$$
$$V_L \in [V^{worst}, \bar{V}]$$

But then again, either $V_L = \bar{V}$ or the (IC-ON-S) is binding, in both cases the solution implies $\tau_L = \tau_H$. A similar argument can be used to show that the second candidate solution implies the same result. Therefore, in any case $V_s = \bar{V}$. Then, maximizing the return function (imposing the additional constraint that taxes are equal) gives the first part of the proposition. The second part of the proposition follows from the fact that the maximum value is the summation of the period by period payoff, then if β is large enough this would be an equilibrium. ∎

Proposition 6 shows that any best equilibrium, within the constrained class of strategies that we analyze here, can be implemented with public strategies when agents discount the future high enough. Denote the best public equilibrium $\sigma^{BP} \in \Sigma$. Taxes then are a deterministic function of its own past history. Agents are able to predict perfectly the tax that they would have to pay after investing:

$$\sigma^{BP}_{i,t}(h^{i,t}, y^i) = \begin{cases} \omega & \text{if } y(1-\tau_t(\tau^{t-1})) - 1 > 0 \\ [0,\omega] & \text{if } y(1-\tau_t(\tau^{t-1})) - 1 = 0 \\ 0 & \text{otherwise} \end{cases}$$

The best equilibrium payoff under pure-public strategies for government is given by:

$$V^{BP} = V(\sigma^{BP}) = \max_{\tau_t(\tau^{t-1})} \sum_{t=0}^{\infty} \beta^t E[W(\tau_t, R_t| \left\{\sigma^{BP}_{i,t}\right\}_{i\in I})]$$

The best payoff under commitment, or the payoff of the Ramsey equilibrium, is given by:

$$V(\sigma^*) = \sum_{t=0}^{\infty} \beta^t E[W(\sigma^*(R_t), R_t| \left\{\sigma^*_{i,t}\right\}_{i\in I})]$$

Proposition 7. *There exists* $\hat{\beta} < 1$ *such that for all* $\beta \in [\hat{\beta}, 1)$,$V^{BP} < V(\sigma^*)$.

Proof: In the appendix.

The proposition follows because the Ramsey equilibrium exhibits taxes that are state dependent. Furthermore, the difference between the best one-period payoff under public equilibria and the Ramsey does not depend on the size of the discount factor.

7 About the Observability of Government Spending

The fact that the government spending is not observable by the agents can be rationalized in the following way. Suppose that the end of each period every agent observes a different realization of the government spending g^i with $g^i \sim \xi(g^i|g)$. Assume further that $g^i \in [0, \bar{G}]$ and $\xi(g^i|g) > 0$ for all g and all $g^i \in [0, \bar{G}]$. In order to keep the structure simple we maintain the assumption that individuals cannot observe either the aggregate output or the aggregate investment. However, to deal with this assumption we can assume, as we are doing now, that individuals observe both of them with a noise in the same way as follows.

The new timing would be,

1. R is realized;
2. y^i is drawn for each i;
3. Each individual chooses x^i;
4. Government observes R and sets τ;
5. g is produced by the government;

6. g^i is observed by each agent i and consumption takes place.

Until now we have assumed that agents do not observe the government spending. As explained in Section 2 this assumption was made to avoid the full (and common) observability of R. That is, since agents perfectly observe the tax rate and since on the equilibrium path there is a one to one correspondence between R and Y, the perfect and common observability of g would reintroduced common knowledge in the game. This creates an inconsistency because agents value something that they cannot observe. In the alternative environment we allowed individuals to observe the public spending but in a noisy way. Of course, now agents can extract additional information about R at the end of each period but the signal extraction is not common among agents. That is, for a given tax rate the distribution of individually observed public spending would change with R. In our environment this assumption could be interpreted as different individual perceptions about the cost of providing the public good. In monetary policy environments like Athey et. al. (2005) or Canzoneri (1985) the assumption could be interpreted as the fact that agents "suffer" in different ways different levels of inflation.

The definition of a Ramsey equilibrium in this new environment is equivalent to that in Section 4, but now the welfare function has to be modified to consider the fact that the aggregate utility of the government spending is different than the utility of the aggregate spending. Let σ_g be a parameter of the density function $\xi(g^i|g)$ with the property that $\int_{g-\epsilon}^{g+\epsilon} \xi(g^i|g)dg^i \to 1$ as $\sigma_g \to 0$ for all $\epsilon > 0$. It is easy to see that this new Ramsey policy and the new Ramsey payoff would converge to those in Section 4 as σ_g goes to zero. In addition, the individual beliefs will still follow a law of motion as (11) with $P(R_{t-1}|y^i_{t-1})$ replaced by $P(R_{t-1}|y^i_{t-1}, g^i_{t-1})$ [8]. Therefore, Propositions 1 and 2 remain unchanged. As before, for the equilibrium strategy to improve upon the payoff it must be a function of at least one publicly observed object. As long as R is *i.i.d.* over time and g^i is uncorrelated with r^i, past realizations of g^i carry no information about the future behavior of R. Thus,

[8]However, some measurability issues could arise due to the fact that belief depend on the government's strategy, see Bergin and Bernhardt (1992)-(1995).

conditioning on the past realizations of g^i do not increase the equilibrium payoff set. That is, it is possible to show that Proposition 6 holds for every σ_g. However, if R were not *i.i.d.* over time or g^i were correlated with r^i it could be the case that the optimal policy without commitment is not invariant over time.

8 Conclusion

In this paper we show how small changes in the informational assumptions can have drastic consequences for both the set of equilibrium strategies and the set of equilibrium payoffs in a macro game without commitment. First we show that every equilibrium strategy has to depend on some information that is **publicly** known for every agent in the economy, otherwise no coordination is possible. In addition, when we analyze equilibria that depend on both public and private histories we found that in the best equilibrium the government **does not use** its private information. As a result, for any discount factor strictly smaller than one, the payoff of the best equilibrium without commitment is always strictly smaller than the payoff with commitment. Moreover, this distance does not approach zero as the discount factor approaches one. In this sense, the welfare in a economy without commitment is uniformly bounded away from the efficient one.

The results of this paper support the arguments for strong institutions that tie the hands of policy makers. To endowed governments with full discretion and to impose the right incentives to avoid deviations from optimal policies could be impossible or too costly. In this sense, the original recommendation of Kydland and Prescott (1977) is still valid.

The implication of this paper apparently contradicts the fact that most policies react to the state of the economy. However, according to the interpretation of Section 7 this will not constitute a contradiction. It could be optimal for the government to react to **past** (and publicly known) states of the economy as long as it provides information about future states. What the policy maker looses is the possibility of fine tuning using **not perfectly precise** signals. On the other hand, it is fairly common to find examples of policy makers that have been institutionally banned from the use of discretion, e.g, the implementation of a currency board system. This usually happens when the policy maker has a “bad reputation”, like Argentina in the 90’s. We think that future research on time inconsistency problems should include the possibility for the policy maker of building reputation. This line of research, if successful, will provide a more precise answer to this kind of problems.

References

Athey, S., A. Atkeson, and P. Kehoe (2005): "The Optimal Degree of Discretion in Monetary Policy," *Econometrica*, 73(5), 1431–1475.

Athey, S., K. Bagwell, and C. Sanchirico (2004): "Collusion and price rigidity," *Review of Economic Studies*, 71, 317–349.

Bergin, J., and D. Bernhardt (1992): "Anonymous Sequential Games with Aggregate Uncertainty," *Journal of Mathematical Economics*, 21(6), 543–62.

——— (1995): "Anonymous Sequential Games with Aggregate Uncertainty," *Economic Theory*, 5(3), 461–83.

Chang, R. (1998): "Credible Monetary Policy in an Infinite Horizon Model: Recursive Approaches," *Journal of Economic Theory*, 81(2), 431–461.

Chari, V. V., and P. Kehoe (1993): "Sustainable Plans and Mutual Default," *Review of Economic Studies*, 60(1), 175–95.

Chari, V. V., and P. J. Kehoe (1990): "Sustainable Plans," *Journal of Political Economy*, 98(4), 783–802.

Kydland, F., and E. Prescott (1977): "Rules Rather Than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy*, 85(3), 473–91.

Mailath, G. J., and L. Samuelson (2006): *Repeated Games and Reputations. Long-run relationships*. Oxford University Press.

Phelan, C., and E. Stacchetti (2001): "Sequential Equilibria in a Ramsey Tax Model," *Econometrica*, 69(6), 1491–1518.

Rudin, W. (1987): *Real and Complex Analysis*. McGraw-Hill.

Sleet, C. (2001): "On Credible Monetary Policy and Private Government Information," *Journal of Economic Theory*, 99, 338–376.

Sleet, C., and S. Yeltekin (2006): "Optimal taxation with endogenously incomplete debt markets," *Journal of Economic Theory*, 127(1), 36–73.

Sun, Y. (2006): "The Exact Law of Large Numbers via Fubini Extension and Characterization of Insurable Risks," *Journal of Economic Theory*, 126(1), 31–69.

9 Appendix

9.1 Proof of Lemma 1

The proof basically works out in two steps. First, we show that we can write beliefs recursively. Second, we use a property about the zeros of analytical functions.

9.1.1 Step 1: Beliefs Updating

Given a pure strategy profile $\sigma_G \in \Sigma_G$ and $h^{g,t} = (R_0, \tau_0, ..., R_{t-1}, \tau_{t-1})$, let $\mu(\cdot)$ be the induced probability distribution over histories for the government. Also let $\mu(|h^{i,t}, y_t^i)$ be the respective conditional probability given history $h^{i,t}$, that is, the probability measure used by agent i in (4) to calculate the expected marginal return of investment. With some abuse of notation, for $h^{i,t}$ consistent with $\hat{h}^{g,t}$, let $f(h^{i,t}|\hat{h}^{g,t}) \equiv \prod_{s=0}^{t-1} f(y_s^i|\hat{R}_s)$.

In addition, for $\hat{h}^{g,t}$ consistent with $\tilde{h}^{i,t}$, with some abuse of notation we can use Bayes' theorem to write:

$$\mu(h^{g,t}|\tilde{h}^{i,t}) = \frac{f(\tilde{h}^{i,t}|\hat{h}^{g,t})\mu(h^{g,t})}{\sum_{\hat{h}^{g,t}} f(\tilde{h}^{i,t}|\hat{h}^{g,t})\mu(\hat{h}^{g,t})} \tag{9}$$

The next lemma shows that $\mu(h^{g,t}|\tilde{h}^{i,t})$ can be written recursively.

Remark: The lemma considers the more general case of mixed strategies $\sigma_{G,t} : H^{g,t} \to \Delta(\Upsilon)$.

Lemma 4. *Consider $\sigma_G \in \Sigma_G$. For $\hat{h}^{g,t}$ consistent with $\tilde{h}^{i,t}$ we have that:*

$$\mu(h^{g,t}|\tilde{h}^{i,t}) = P(R_{t-1}|y_{t-1}^i)\frac{\sigma_{G,t-1}(\tau_{t-1}|h^{g,t-1})\mu(h^{g,t-1}|\tilde{h}^{i,t-1})}{\sum_{\tau_{t-1},\hat{h}^{g,t-1}} \sigma_{G,t-1}(\tau_{t-1}|\hat{h}^{g,t-1})\mu(\hat{h}^{g,t-1}|\tilde{h}^{i,t-1})} \tag{10}$$

Proof: First, we have the following:

$$\mu(h^{g,t}) = P(R_{t-1})\sigma_{G,t-1}(\tau_{t-1}|h^{g,t-1})\mu(h^{g,t-1})$$

and

$$f(\tilde{h}^{i,t}|h^{g,t}) = f(y^i_{t-1}|R_{t-1})f(\tilde{h}^{i,t-1}|h^{g,t-1})$$

Using the last two expressions in (9) we get:

$$\mu(h^{g,t}|\tilde{h}^{i,t}) = \frac{f(y^i_{t-1}|R_{t-1})f(\tilde{h}^{i,t-1}|h^{g,t-1})P(R_{t-1})\sigma_{G,t-1}(\tau_{t-1}|h^{g,t-1})\mu(h^{g,t-1})}{\sum_{\hat{h}^{g,t}} f(y^i_{t-1}|R_{t-1})f(\tilde{h}^{i,t-1}|\hat{h}^{g,t-1})P(R_{t-1})\sigma_{G,t-1}(\tau_{t-1}|\hat{h}^{g,t-1})\mu(\hat{h}^{g,t-1})}$$

which yields

$$\mu(h^{g,t}|\tilde{h}^{i,t}) = \frac{f(y^i_{t-1}|R_{t-1})P(R_{t-1})\sigma_{G,t-1}(\tau_{t-1}|h^{g,t-1})f(\tilde{h}^{i,t-1}|h^{g,t-1})\mu(h^{g,t-1})}{\sum_{\hat{R}_{t-1}} f(y^i_{t-1}|\hat{R}_t)P(\hat{R}_{t-1})\sum_{\tau_{t-1},\hat{h}^{g,t-1}} \sigma_{G,t-1}(\tau_{t-1}|\hat{h}^{g,t-1})f(\tilde{h}^{i,t-1}|\hat{h}^{g,t-1})\mu(\hat{h}^{g,t-1})}$$

and

$$\mu(h^{g,t}|\tilde{h}^{i,t}) = P(R_{t-1}|y^i_{t-1})\frac{\sigma_{G,t-1}(\tau_{t-1}|h^{g,t-1})\mu(h^{g,t-1}|\tilde{h}^{i,t-1})}{\sum_{\tau_{t-1},\hat{h}^{g,t-1}} \sigma_{G,t-1}(\tau_{t-1}|\hat{h}^{g,t-1})\mu(\hat{h}^{g,t-1}|\tilde{h}^{i,t-1})} \qquad (11)$$

∎

9.1.2 Step 2: Analytical Functions

Lemma 5. *Suppose that $K(y,R)$ is analytic in y for all $R \in \Upsilon$.*

Define

$$m(y) = \int_{\Upsilon} K(y,R)dP(R)$$

and let

$$C = \{y \in [1,\bar{y}] : m(y) = 0\}$$

Then, C has Lebesgue measure zero on $[1,\bar{y}]$.

Proof: We start with a result about analytic functions. Since analytic functions have power-series expansions about all points in their domain, the set of roots is at most countable. The proof is an adaptation of Theorem 10.18 in Rudin (1987) about holomorphic functions.

Claim 1: Suppose $f : [1, \overline{y}] \to \mathbb{R}$ is analytic. Let $Z(f) = \{x \in \text{int}([1, \overline{y}]) : f(x) = 0\}$. Then either $Z(f) = (1, \overline{y})$ or $Z(f)$ has no limit points in $[1, \overline{y}]$. In the latter case $Z(f)$ is at most countable.

Proof of Claim 1: Let M be the set of limit points of $Z(f)$. Take $x_0 \in Z(f)$. We will argue that either $x_0 \in \text{int(M)}$ or x_0 is an isolated point of $Z(f)$. To see this, notice that:

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$$

for $x \in B_r(x_0) \subseteq [1, \overline{y}]$, where $B_r(x_0)$ is an open ball or radius r around x_0.

Then it follows that either all $a_n = 0$, in which case $B_r(x_0) \subseteq M$ and therefore $x_0 \in \text{int}(M)$, or there exists $\hat{n} > 0$ (since $f(x_0) = 0$) such that $a_{\hat{n}} \neq 0$. In the latter case, define:

$$g(x) = \begin{cases} (x - x_0)^{-\hat{n}} f(x) & \text{for } x \in [1, \overline{y}] \setminus \{x_0\} \\ a_{\hat{n}} & \text{for } x = x_0 \end{cases}$$

Because $g(x_0) \neq 0$ and $g(\cdot)$ is continuous, there exists a neighborhood $B_{\tilde{r}}(x_0)$ of x_0 in which $g(\cdot)$ has no zero, and therefore $f(\cdot)$ has no zero in $B_{\tilde{r}}(x_0)$. Then it follows that x_0 is an isolated point of $Z(f)$. This finish the claim that either $x_0 \in \text{int(M)}$ or x_0 is an isolated point of $Z(f)$.

Next take $x \in M$. Because $f(\cdot)$ is continuous, it follows that $x \in M \subseteq Z(f)$. Then either $x \in \text{int(M)}$ or x is a limit point of M. By the reasoning above, x cannot be a limit point of M, because $x \in Z(f)$ and therefore $x \in \text{int(M)}$ or x is an isolated point of $Z(f)$. It then follows that M is open. If $B = [1, \overline{y}] - M$, then B is open since M is the set of limit points of $Z(f)$. Since $[1, \overline{y}]$ is connected, it cannot be the union of the disjoint open sets M and B. Then either $M = (1, \overline{y})$, in which case $Z(f) = (1, \overline{y})$, or $M = \emptyset$. In the latter case $Z(f)$ has at most finitely many points in each compact subset of $[1, \overline{y}]$. But since $[1, \overline{y}]$ is a countable union of compact sets, $Z(f)$ is at most countable ∎

In order to finish the proof, recall that $m(y) = \int_\Upsilon K(y,R)dP(R)$. Since $K(y,R)$ is analytic, let $K(y,R) = \sum_{n=0}^{\infty} c_n(R)(y-y_0)^n$ for some y_0. Notice that the constants $\{c_n(R)\}_n$ depend on R. Then we have:

$$\begin{aligned} m(y) &= \int_\Upsilon [K(y,R)]\, dP(R) \\ &= \int_\Upsilon \left[\sum_{n=0}^{\infty} c_n(R)(y-y_0)^n\right] dP(R) \\ &= \sum_{R\in\Upsilon} \left[\sum_{n=0}^{\infty} c_n(R)(y-y_0)^n\right] P(R) \\ &= \sum_{n=0}^{\infty} \left[\int_\Upsilon c_n(R)dP(R)\right] (y-y_0)^n \end{aligned}$$

which it is analytic. Then an application of claim 1 implies that set $C = \{y \in [1,\bar{R}] : m(y) = 0\}$ is at most countable, and therefore has Lebesgue measure zero ∎

Given that the composition of analytic functions is itself analytic, a straightforward application of Lemma 5 using (10) in (4) implies that there is a measure zero of agents indifferent between investing or not.

9.2 Proof of Lemma 2

Before we present the proof, consider the following notation. Let $\tau(\hat{R}) \equiv \sigma_G^*(\hat{R})$ and $\tau \equiv \{\tau(R)\}\,|_{R\,\in\Upsilon}$. Given the vector τ, from the individual agent's decision, consider the following function:

$$H(y^i,\tau) = y^i(1 - [E(\tau(R)|y^i)]) - 1 \tag{12}$$

Given assumption 4, the set of agents $i \in I$ such that $H(y^i,\tau) = 0$ is at most countable (see the proof of Lemma 1). Moreover, the set of points y^i such that $H(y^i, \tau) = 0$ is indeed finite as long as there is a state $R \in \Upsilon$ with tax bounded away from the unity. To see this, notice that, for $\hat{y}^i$ high enough, we have $H(\hat{y}^i, \tau) > 0$ whenever $E(\tau(R)|\hat{y}^i) < 1$.

Given the previous reasoning, suppose there are N cutoff points $\{y_i^*\}_{i=1}^N$ satisfying $H(y_i^*, \tau) = 0$. We order them in an ascending order, i.e., $y_i^* \leq y_{i+1}^*$, and let $y_{N+1}^* = \bar{y}$. It is important to keep in mind that since $H(1, \tau) \leq 0$, $\frac{\partial H(y_i^*, \tau)}{\partial y_i^*} > 0$ when i is odd and $\frac{\partial H(y_i^*, \tau)}{\partial y_i^*} < 0$ when i is even.

Notice that, using the implicit function theorem, we have:

$$\begin{aligned} \frac{\partial y_i^*}{\partial \tau(R)} &= \left[\left(1 - E(\tau(R)|y_i^*) - y_i^* \frac{\partial E(\tau(R)|y_i^*)}{\partial y_i^*} \right) (1 - E(\tau(R)|y_i^*)) \right]^{-1} P(R|y_i^*) \quad (13) \\ &= \left(\frac{\partial H}{\partial y_i^*} (1 - E(\tau(R)|y_i^*)) \right)^{-1} P(R|y_i^*) \\ &= J(y_i^*) P(R|y_i^*) \end{aligned}$$

where $J(y_i^*) \equiv \left(\frac{\partial H}{\partial y_i^*} (1 - E(\tau(R)|y_i^*)) \right)^{-1}$.

By the definition of $\{y_i^*\}_{i=1}^N$, the aggregate investment is given by $X(\tau(R), R) = \sum_{i=1}^N \int_{y_i^*}^{y_{i+1}^*} f(y|R) dy$. In a similar fashion aggregate output is $Y(\tau(R), R) = \sum_{i=1}^N \int_{y_i^*}^{y_{i+1}^*} y f(y|R) dy$.

Notice that $\frac{\partial X(\tau(R),R)}{\partial \tau} = \sum_{i=1}^N (-1)^i f(y_i^*|R) \frac{\partial y_i^*}{\partial \tau} < 0$ because $\frac{\partial y_i^*}{\partial \tau} > 0$ when i is odd and $\frac{\partial y_i^*}{\partial \tau} < 0$ when i is even.

In the same way $\frac{\partial Y(\tau(R),R)}{\partial \tau} = \sum_{i=1}^N (-1)^i y_i^* f(y_i^*|R) \frac{\partial y_i^*}{\partial \tau} < 0$ and $\frac{\partial Y(.)}{\partial \tau} - \frac{\partial X(.)}{\partial \tau} < 0$ because $y_i^* \geq 1$ for all i.

We can write the static payoff for the government as:

$$\sum_{R \in \Upsilon} P(R) W(R, \tau(R), \tau(R), Y(\tau(R), R))$$

where

$$W(R, X(\tau(R), R), \tau(R), Y(\tau(R), R)) \equiv (1-\tau(R))Y(\tau(R), R)+(\omega - \tau(R))+b\tau(R)\ Y(\tau(R), R)$$

Towards a contradiction, suppose that the solution has $\tau_{\bar{R}} < 1$ and $\tau_R > 0$ for some $R \in \Upsilon$. Then consider the following perturbation: increase $\tau_{\bar{R}}$ by $d\tau_{\bar{R}} > 0$ and decreases τ_R by $d\tau_R < 0$ such that it keeps y_N^* fixed.

Then, at the solution, the change Δ in payoff should be zero:

$$\begin{aligned}
\Delta &= (b-1)[P(\bar{R})Y(\bar{R})d\tau_{\bar{R}} + P(R)Y(R)d\tau_R] + \\
&\quad \sum_{i=1}^{N}(-1)^i \left[\sum_{\hat{R}\in\Upsilon} P(\hat{R}) \left([(1+\tau(\hat{R})(b-1))y_i^* - 1] f(y_i^*|R) \right) \right] \left(\frac{\partial y_i^*}{\partial \tau_R} d\tau_R + \frac{\partial y_i^*}{\partial \tau_{\bar{R}}} d\tau_{\bar{R}} \right) \\
&= \Delta_1 + \Delta_2
\end{aligned}$$

where

$$\Delta_1 = (b-1)[P(\bar{R})Y(\bar{R})d\tau_{\bar{R}} + P(R)Y(R)d\tau_R]$$

and

$$\Delta_2 = \sum_{i=1}^{N}(-1)^i \left[\sum_{\hat{R}\in\Upsilon} P(\hat{R}) \left([(1+\tau(\hat{R})(b-1))y_i^* - 1] f(y_i^*|R) \right) \right] \left(\frac{\partial y_i^*}{\partial \tau_R} d\tau_R + \frac{\partial y_i^*}{\partial \tau_{\bar{R}}} d\tau_{\bar{R}} \right)$$

Let the perturbation described above satisfies:

$$\frac{\partial y_N^*}{\partial \tau_R} d\tau_R + \frac{\partial y_N^*}{\partial \tau_{\bar{R}}} d\tau_{\bar{R}} = 0$$

Using (13) we get that:

$$d\tau_R = -\frac{P(\bar{R}) f(y_N^* | \bar{R})}{P(R)(f(y_N^* | R)} d\tau_{\bar{R}}$$

For each other y_i^* we have:

$$\frac{\partial y_i^*}{\partial \tau_R} d\tau_R + \frac{\partial y_i^*}{\partial \tau_{\bar{R}}} d\tau_{\bar{R}} = \frac{J(y_i^*) P(\bar{R})}{\sum_{\hat{R}} P(\hat{R}) f(y_i^* | \hat{R})} \left[-\frac{f(y_N^* | \bar{R})}{f(y_N^* | R)} f(y_i^* | R) + f(y_i^* | \bar{R}) \right] d\tau_{\bar{R}}$$

Assumption 3 implies that $-\frac{f(y_N^*|\bar{R})}{f(y_N^*|R)} f(y_i^*|R) + f(y_i^*|\bar{R}) < 0$ since $y_N^* \geq y_i^*$ for all i. Thus, because $d\tau_{\bar{R}} > 0$, $\frac{\partial y_i^*}{\partial \tau_R} d\tau_R + \frac{\partial y_i^*}{\partial \tau_{\bar{R}}} d\tau_{\bar{R}}$ is negative when i is odd and positive when i is even, and therefore $(-1)^i \left(\frac{\partial y_i^*}{\partial \tau_R} d\tau_R + \frac{\partial y_i^*}{\partial \tau_{\bar{R}}} d\tau_{\bar{R}} \right) \geq 0$ for all i.

Therefore $\Delta_2 > 0$. It remains to show that $\Delta_1 > 0$.

$$\begin{aligned}
\Delta_1 &= (b-1)[P(\bar{R}) Y(\bar{R}) d\tau_{\bar{R}} + P(R) Y(R) d\tau_R] \\
&= (b-1) d\tau_{\bar{R}} \left[P(\bar{R}) Y(\bar{R}) - P(R) Y_L \frac{P(\bar{R}) f(y_N^* | \bar{R})}{P(R) f(y_N^* | R)} \right] \\
&= P(\bar{R})(b-1) d\tau_{\bar{R}} \left[Y(\bar{R}) - Y(R) \frac{f(y_N^* | \bar{R})}{f(y_N^* | R)} \right]
\end{aligned}$$

Because $d\tau_{\bar{R}} > 0$, it is sufficient to show that $\frac{Y(\bar{R})}{f(y_N^*|\bar{R})} > \frac{Y(R)}{f(y_N^*|R)}$.Notice that $\frac{Y(\bar{R})}{f(y_N^*|\bar{R})} = \omega \sum_{i=1}^N \int_{y_i^*}^{y_{i+1}^*} y \frac{f(y|\bar{R})}{f(y_N^*|\bar{R})} dy$. and $\frac{Y(R)}{f(y_N^*|R)} = \omega \sum_{i=1}^N \int_{y_i^*}^{y_{i+1}^*} y \frac{f(y|R)}{f(y_N^*|R)} dy$. Each of this variables represents an integral using a normalized probability distribution function $h(y|R') = \frac{f(y|R')}{f(y^*|R')}$ with $h(y^*|R') = 1 \; \forall R' \in \Upsilon$.

Hence $\frac{Y(\bar{R})}{f(y_N^*|\bar{R})} > \frac{Y(R)}{f(y_N^*|R)}$ and $\Delta > 0$, a contradiction ∎

9.3 Proof of Proposition 2

Towards a contradiction, fix $\sigma \in \Xi$ with $\sigma_G \in \Xi_G$ being non-trivially essentially private. Define $S = \left\{ \tau \in [0,1] : \tau = \sigma_{G,t}(h^{g,t}, R_t), \text{ for some } R^t \in \Upsilon^t,\ R^{t-1} \subseteq h^{g,t} \right\}$. The fact that σ_G is non-trivially essentially private makes sure that S is not a singleton. Also define

$$H^{*,i} = \{h^{i,t} \in \cup_{t=0}^{\infty} H^{i,t} : \tau_s \in S, \forall s \leq t-1\}$$

Because of the full support condition in assumption 1, all histories in $H^{*,i}$ have non-zero measure.

Since we are considering pure strategies only, using (10), for all $h^{i,t} \in H^{*,i}$ the belief system is given by:

$$\mu(h^{g,t}, R_t|\tilde{h}^{i,t}, y_t^i) = P(R_t|y_t^i) \frac{\mu(h^{g,t-1}|\tilde{h}^{i,t-1})}{\sum_{\hat{h}^{g,t-1}} \mu(\hat{h}^{g,t-1}|\tilde{h}^{i,t-1})}$$

By recursive calculations the above expression can be reduced to:

$$\mu(h^{g,t}, R_t|\tilde{h}^{i,t}, y_t^i) = f(y_t^i|R_t) \prod_{s=0}^{t-1} \frac{f(y_s^i|R_s)}{\sum_{\hat{R}_s} f(y_s^i|\hat{R}_s)} \tag{14}$$

In this way, when σ_G is non-trivially essentially private, beliefs do not depend on the strategy followed by the government. Thus, given the features of σ_G, $\mu(\sigma_G(h^{g,t}, R_t)|h^{i,t}, y_t^i) = \mu(\sigma_G(h^{g,t}, R_t)|\tilde{h}^{i,t}, y_t^i)$ for all $h^{i,t}, \tilde{h}^{i,t} \in H^{*,i}$ and $h^{g,t}$ consistent with histories in $H^{*,i}$.

Now, consider the following one shot deviation strategy $\tilde{\sigma}_G \in \Sigma_G$. Take any period $t > 0$ with history $h^{g,t}$ such that $\sigma_G(h^{g,t}, R_t) < \bar{\tau}^S \equiv \max_\tau\{S\}$ and consider the following alternative strategy.

$$\tilde{\sigma}_G(h^{g,s}, R_s) = \begin{cases} \sigma_G(h^{g,s}, R_s) & \text{if } s \neq t \\ \bar{\tau}^S & \text{if } s = t \end{cases}$$

Let X_t and Y_t be, respectively, the aggregate output and investment generated by $\{\sigma^i|_{h^{i,t}}\}_{i\in I}$ in time t after R_t is realized. Then $(\sigma|_{h^{g,t}}, \{\sigma^i|_{h^{i,t}}\}) \in \Sigma$ generates the following payoff:

$$V(\sigma_G|_{h^{g,t}}, \{\sigma_i|_{h^{i,t}}\}_{i\in I}) = (1-\beta)W(R_t, X_t, \sigma_{G,t}(h^{g,t}, R_t), Y_t) + \beta V(\sigma_G|_{\{h^{g,t}, R_t, \sigma_G(h^{g,t}, R_t)\ \}}, \sigma_i|_{\{h^{i,t}, \sigma_G(h^{g,t}, R_t)\}})$$

The alternative strategy yields:

$$V(\tilde{\sigma}_G|_{h^{g,t}}, \{\sigma_i|_{h^{i,t}}\}_{i\in I}) = (1-\beta)W(R_t, X_t, \bar{\tau}^S, Y_t) + \beta V(\tilde{\sigma}|_{\{h^{g,t}, R_t, \bar{\tau}^S\}}, \hat{\sigma}^i|_{\{h^{i,t}, \bar{\tau}^S\}})$$

Since $\bar{\tau}^S \in S$, it follows that $\{h^{i,t}, \bar{\tau}^S, y_t^i\} \in H^{*,i}$ and therefore $\hat{\sigma}^i|_{\{h^{i,t}, \sigma_G(h^{g,s}, \tau_s, R_s)\}} = \hat{\sigma}^i|_{\{h^{i,t}, \bar{\tau}^S\}}$ for all $i \in I$.

In addition, from the one shot deviation construction, $\hat{\sigma}|_{\{h^{g,t}, R_t, \sigma_G(h^{g,s}, R_s)\}} = \tilde{\sigma}|_{\{h^{g,t}, R_t, \bar{\tau}^S\}}$. Hence, $V(\sigma_G|_{\{h^{g,t}, R_t, \sigma_G(h^{g,s}, R_s)\ \}}, \sigma_i|_{\{h^{i,t}, \sigma_G(h^{g,s}, R_s)\}}) = V(\tilde{\sigma}|_{\{h^{g,t}, R_t, \bar{\tau}^S\}}, \hat{\sigma}^i|_{\{h^{i,t}, \bar{\tau}^S\}})$. But since $\bar{\tau}^S > \sigma_G(h^{g,t}, R_t)$, from assumption 2 we have that $V(\tilde{\sigma}_G|_{h^{g,t}}, \{\sigma_i|_{h^{i,t}}\}_{i\in I}) > V(\sigma_G|_{h^{g,t}}, \{\sigma_i|_{h^{i,t}}\}_{i\in I})$ ∎

9.4 Proof of Proposition 3

First, take an arbitrarily agent and history $h^{i,t}$. Suppose that every agent is playing according to σ^{worst}. In this situation the government cannot increase the provision of the public good regardless the tax it chooses. Therefore it is weakly optimal for the government to fully tax investment. In this way, condition (i) in definition 1 is met for all histories. Next, suppose the government is playing according to σ_G^{worst}. Then regardless the individual signals, all the agents assign probability one to full taxation and the optimal action is to choose zero investment. This follows straight from (4). Therefore condition (ii) in definition 1 holds and the proposed strategy it is an equilibrium.

The fact that σ^{worst} yields the the worst equilibrium follows from the observation that the level of provision of the public good is at its minimum value under σ^{worst}. Therefore assumption (2) yields the result ∎

9.5 Proof of Proposition 4

Fix $\mathcal{A}$ an aggregate outcome of $\sigma \in \mathring{\Sigma}$. Condition (1) comes from the definition of aggregating investment decisions over individuals, where such decisions are given by (8) along the stochastic outcome path. Condition (2) comes from both proposition 3 and the definition of the best deviation $W^d(R_t, X(R^t), Y(R^t))$, while (3) follows from condition (ii) in the definition of a perfect Bayesian equilibrium along the equilibrium path.

Conversely, take $\mathcal{A} \equiv \{\tau_t, X_t, Y_t\}_{t=0}^{\infty}$ satisfying the conditions above. Then we construct $\sigma \in \mathring{\Sigma}$ such that $\mathcal{A}$ is the induced aggregate outcome. We construct σ_G as follows. Along the equilibrium path, set $\sigma_{G,t}(\tau^{t-1}(R^{t-1}), R_t) = \tau_t(R^t)$. For all other government's history $\hat{h}^{g,t} \in H^{G,t}$ off-path, set $\sigma_{G,t}(\hat{h}^{g,t}) = \sigma_{G,t}^{\text{worst}}(\hat{h}^{g,t})$. Regarding the individual decisions, for taxes on the path, set decisions as (8), and for all other off-path histories $\hat{h}^{i,t} \in H^{i,t}$, set $\sigma_{i,t}(\hat{h}^{i,t})$=$\sigma_{i,t}^{\text{worst}}(\hat{h}^{i,t})$. Beliefs on path are set according to $E(\tau_t|y^*) = \sum_{R\in\Upsilon}\left(\frac{P(R)f(y^*|R)\tau_t(R_t)}{\sum_{R\in\Upsilon}P(R)f(y^*|R)}\right)$, while beliefs off-path assign full measure to taxes being equal to one.

Next, we check that this strategy profile is indeed an equilibrium. First, by condition (1), households are optimizing giving on the path of play. By construction, on off-path histories optimality also holds. Regarding government optimality, condition (3) prevents the government to obtain a profitable deviation along the path of play when it considers a switch to $\hat{\tau} \in \tau_t(\tau^{t-1}(R^{t-1}), \Upsilon)$. For all other taxes $\hat{\tau} \notin \tau_t(\tau^{t-1}(R^{t-1}), \Upsilon)$, switches from the path are prevented by condition (2). All the equilibrium conditions hold off-path because σ^{worst} is also an equilibrium. ∎

9.6 Proof of Proposition 7

Notice that the best payoff that can be achieved with deterministic tax functions is

$$V^{BP} \leq \sum_{t=0}^{\infty} \beta^t \max_{\tau_t} E[W(\tau_t, R_t | \left\{\sigma_{i,t}^{BP}\right\}_{i \in I})]$$

Define,

$$\tau^* = argmax_{\tau_t} E[W(\tau_t, R_t | \left\{\sigma_{i,t}^{BP}\right\}_{i \in I})]$$

And notice that

$$E[W(\tau^*, R_t | \left\{\sigma_{i,t}^{BP}\right\}_{i \in I})] < E[W(\sigma_G^*(R_t), R_t | \left\{\sigma_{i,t}^{*}\right\}_{i \in I})] \tag{15}$$

otherwise it would contradict the Ramsey solution.

Then, from (15), we obtain:

$$V^{BP} \leq \sum_{t=0}^{\infty} \beta^t E[W(\tau^*, R_t | x_{\forall i \in I}^{*i})] < \sum_{t=0}^{\infty} \beta^t E[W(\sigma^*(R_t), R_t | \left\{\sigma_{i,t}^{*}\right\}_{i \in I})] = V(\sigma^*)$$

Since (15) holds regardless the magnitude of β, the result holds for $\beta \in [\hat{\beta}, 1)$, the sufficient condition under which the best equilibrium can be achieved using public strategies ∎

9.7 History dependency

We now show that even when we allow for history dependency the best equilibrium has the property that the tax is independent of the fundamental. The proof is by induction showing that in any period t after any public history $h^{P,t}$ the best equilibrium solves a problem like PS. Let $\tau^B(h_B^{g,t})$ be an equilibrium strategy that generates the best equilibrium payoff, where $h_B^{g,t}$ is constructed recursively as $h_B^{g,t} = \{h_B^{g,t-1}, R_t, \tau^B(h_B^{g,t-1})\}$ with $h_B^{g,0} \in \Upsilon$. In the same way, $h_B^{i,t} = \{h_B^{i,t-1}, y_t^i, \tau^B(h_B^{g,t-1})\}$ and $h_B^{i,0} \in \Upsilon$ and the public history is $h^{P,t} = \{h_B^{P,t-1}, \tau^B(h_B^{g,t-1})\}$

with $h_B^{P,t} = \varnothing$. Finally let $V(h_B^{P,t-1}) = E[V(h_B^{g,t})|R^{t-1}]$ be the ex-ante value of the equilibrium strategy conditional on the realization of R^{t-1}. At period zero the payoff for this strategy is given by:

$$V(\varnothing) = E_{R_1}\left[W(\tau^B(R_1), Z(\tau^B, R_1)) + \beta V(\tau^B(R_1))\right] \tag{16}$$

Lemma 6. *A pair $\{\tau(R_1), V(\tau(R_1))\}$ can be implemented as an equilibrium at time 1 if and only if, for all $m = L, H$ and all $\hat{\tau} \in [0,1]$,*

$$\begin{aligned}
(EQ) \qquad & W(\tau(R_{1,m}), Z(\tau, R_{1,m})) + \beta V(\tau(R_{1,m})) \geq W(\tau(R_{1,-m}), Z(\tau, R_{1,m})) + \beta V(\tau(R_{1,-m})) \\
& W(\tau(R_{1,m}), Z(\tau, R_{1,m})) + \beta V(\tau(R_{1,m})) \geq W(1, Z(\tau, R_{1,m})) + \beta V^{worst} \\
& V(\hat{\tau}) \in \Psi_\beta \text{ for all } \hat{\tau} \in [0,1] \\
& Z(\tau, R_{1,m}) \in \mathcal{OP}(\tau, R_{1,m}) \text{ for all } \tau \in [0,1]^2
\end{aligned}$$

Proof: First, it is straightforward to show that if $\{\tau(R_1), V(\tau(R_1))\}$ can be implemented as an equilibrium it has to satisfies the above set of inequalities. Then, suppose $\{\tau(R_1), V(\tau(R_1))\}$ satisfies (EQ). Since, $V(\tau(R_1)) \in \Psi_\beta \; \forall R_1$, there exist $\tilde{\tau}_{R_1}(h^{g,t}, R_t)$ and $\tilde{\sigma}^i_{R_1}(h^{i,t}, y^i_t)$ for each $R_1 \in \Upsilon$ that together with the belief system $\tilde{\mu}_{R_1}$ are an equilibrium for all $t \geq 2$ and $V(\tau(R_1)) = V(\tilde{\sigma}_G, \tilde{\sigma}^i)$. Let $\tilde{H}^{g,t}_{R_1}$ be the (on the equilibrium path) possible histories generated by $\tilde{\tau}_{R_1}(h^{g,t}, R_t)$ after each R_1, $\tilde{H}^{t,i}$ be the set of possible individual histories generated by generated by $\{\tau(R_1), \tilde{\tau}_{R_1}(h^{g,t}, R_t)\}$ and consider the following pair of strategies and belief system,

$$\sigma_{G,t}(h^{g,t}, R_t) = \begin{cases} \tau(R_1) & \text{if } t = 1, \; \forall R_1 \\ \tilde{\tau}_{R_1}(h^{g,t}, R_t) & \text{if } t > 1 \text{ and } \sigma_{G,1}(R_1) = \tau(R_1) \text{ and } h^{g,t} \in \tilde{H}^{g,t}_{R_1}, \; \forall R_1 \\ 1 & \text{otherwise} \end{cases}$$

$$\sigma^i_t(h^{i,t}, y^i_t) = \begin{cases} x^{i,*}(h^{i,t}, y^i_t) & \text{if } h^{i,t} \in \tilde{H}^{t,i}, \forall t \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\mu(h^{g,t}|h^{i,t}) = \begin{cases} \frac{f(y^i_1|R)}{\sum_{\hat{R}} f(y^i_1|\hat{R})} & \text{if } t = 1, \forall y^i_1 \\ \tilde{\mu}_{R_1}(h^{g,t}|h^{i,t}) & \forall h^{g,t} \in \tilde{H}^{g,t}_{R_1}; \text{ if } t > 1 \text{ and } \tau_1 = \tau(R_1), \forall h^{i,t}\ \forall R_1 \\ 1 & \forall h^{g,t} \notin \tilde{H}^{g,t}_{R_1}; \text{ if } t > 1 \text{ and } \tau_1 \notin \tau(R_1), \forall h^{i,t} \\ 0 & \text{otherwise} \end{cases}$$

The constructed strategies are the usual in repeated games. For the government states: play the function $\tau(R)$ that satisfies (EQ) at period 1, then depending on the realization of R_1 (and therefore of τ_1) pick the appropriated continuation strategy, if you ever deviated from the equilibrium path play $\tau_t = 1$ for ever. The strategy for households states: invest optimally (using the constructed belief system) if the history is consistent with equilibrium behavior by the government, otherwise never invest again. The belief system simply states that at period 1 belief are calculated according to Baye's rule, and then beliefs are the same as in the continuation strategy if the government follows the tax function $\tau(R)$ and put probability one on off of the equilibrium behavior if the government does not set a tax consistent with $\tau(R)$. Notice that this is a well defined probability measure since the complement of $\tilde{H}^{g,t}_{R_1}$ happens with probability zero on the equilibrium. By construction these strategies and belief system are an equilibrium for $t > 1$. In addition, it is easy to see that given the belief system the behavior of the agents is optimal even at period 1. It remains to show that given the belief system the proposed strategy for the government is indeed optimal at period 1. But that follows from the fact that the function $\tau(R_1)$ satisfies the first two inequalities of (EQ). ∎

Assumption 5. *$sup\{\Psi_\beta\} \in \Psi_\beta$.*

The above assumption simply states that the maximum is well defined.

Proposition 8. *The best equilibrium is: for all $t \geq 0$,*

1) $\forall h_B^{g,t}$, $\forall R_t$; $\tau_t^B(h_B^{g,t}, R_t) = \tau^B$; *if* $\tau_s^B = \tau^B$ *for all* $s < t$ *and* $\tau_t^B(h_B^{g,t}, R_t) = 1$ *otherwise*

2) $\forall h_B^{i,t}$, $\forall y_t^i$; $x^i(h_B^{i,t}, y_t^i) = \omega$; *if* $y_t^i \geq \frac{1}{1-\tau^B}$ *and* $\tau_s^B = \tau^B$ *for all* $s < t$; *and* $x^i(h_B^{i,t}, y_t^i) = 0$ *otherwise.*

3) $\forall h_B^{P,t}$, $V(h_B^{P,t}) = \bar{V}$

where τ^B *and* $\bar{V}$ *are defined in Proposition 6*

Proof: The proof is by induction. Notice that at period zero, after the empty history, the best equilibrium payoff can be written as in (16). Then by Lemma 6 the best equilibrium strategy maximizes (16) subject to (EQ). Since $sup\{\Psi_\beta\} \in \Psi_\beta$, by Proposition 6 the solution is $\tau_L = \tau_H = \tau^B$ and continuation values $V_L = V_H$. Because of this, after period one there is only one continuation value and only one on path public history. After this public history and for every realization of R at $t = 1$, the payoff is

$$V(\tau_1^B) = E_{R_2}\big[W(\tau_2^B(\tau_1^B, R_1, R_2), Z(\tau_2^B(\tau_1^B, R_1), R_2)) + \beta V(\tau_2^B(\tau_1^B, R_1, R_2), \tau_1^B)|R_1\big] \quad (17)$$

This equation is similar to (16) and it is a best equilibrium payoff as well. In fact, it is easy to see that we can use the same argument as in period one appealing to Lemma 6 changing only the time index. Thus, after any realization of R_1 if $\tau_2^B(\tau_1^B, R_1, R_2)$ was not a solution to that problem there would be another tax function $\hat{\tau}$ and continuation values $\hat{V}$ that solve it and can be implemented as an equilibrium, generating a value $E_{R_2}\big[E[W(\hat{\tau}, Z(\hat{\tau}, R_2))] + \beta\hat{V}\big] > V(\tau_1^B)$. But this would contradict $V(\tau_1^B)$ being a best equilibrium.

Therefore, at period 2 it must be the case that $\tau_L = \tau_H = \tau^B$ and $V_L = V_H$. In order to conclude the induction, suppose that the statement is true at any period $t > 2$, then it must be the case that

$$V(h^{P,t}) = E_{R_t}\left[W(\tau^B(h_B^{g,t}, R_t), Z(\tau^B(h_B^{g,t}), R_t)) + \beta V(\tau^B(h_B^{g,t}, R_t), h^{P,t})|R^{t-1}\right]$$

Then, the argument is the same as for period 2. The strategy $\tau^B(h_B^{g,t})$ has to solve P0 otherwise there would be a combination of tax function $\hat{\tau}$ and continuation values $\hat{V}$ such that $E[W(\hat{\tau}, Z(\hat{\tau}, R_t))] + \beta\hat{V}] > V(h^t)$. Given that by construction the new strategies are incentive-feasible after period t and that for periods before t the continuation values are independent of the state this policy won't violated any equilibrium requirement. Thus, at period t we obtain $\tau_L = \tau_H = \tau^B$. But if this is true the best tax independent of the state is given by τ^B as defined in Proposition 6, the same as the continuation value. Regarding optimal individual investments, the result is obvious. After any individual history agents know exactly what the tax will be, therefore, given y^i, if the after tax return in investment is positive they invest everything and if the return is negative they do not invest anything (since the set of indifferent agents has measure zero the rule assigned to them is irrelevant).

∎