Year: 2012

Time consistent optimal fiscal policy over the business cycle

Feng, Zhigang

Abstract: This paper examines a dynamic stochastic economy with a benevolent government that cannot commit to future policies. Following Phelan and Stacchetti (2001), we consider sequential sustainable equilibria (SSE). We numerically solve for the set of equilibrium payoffs, and investigate whether the time consistency problem of capital income tax is quantitatively important. For a realistically calibrated economy, we find that the optimal sustainable capital income tax rate is pro-cyclical and close to zero on average, while the labor income tax is countercyclical. Moreover, the welfare cost of no commitment is very small (0:22%) when compared with the Ramsey allocation. We also find that the best sustainable equilibrium outcome may achieve substantially higher social welfare than the Markov-perfect equilibrium as considered by Klein, Krusell and Rios-Rull (2008).

Time Consistent Optimal Fiscal Policy over the Business Cycle

Zhigang Feng[*][†]

Department of Banking and Finance

University of Zurich

March 28, 2012

Abstract

This paper examines a dynamic stochastic economy with a benevolent government that cannot commit to future policies. Following Phelan and Stacchetti (2001), we consider sequential sustainable equilibria (SSE). We numerically solve for the set of equilibrium payoffs, and investigate whether the time consistency problem of capital income tax is quantitatively important. For a realistically calibrated economy, we find that the optimal sustainable capital income tax rate is pro-cyclical and close to zero on average, while the labor income tax is counter-cyclical. Moreover, the welfare cost of no commitment is very small (0.22%) when compared with the Ramsey allocation. We also find that the best sustainable equilibrium outcome may achieve substantially higher social welfare than the Markov-perfect equilibrium as considered by Klein, Krusell and Rios-Rull (2008).

Keywords: Optimal Fiscal Policy, Business Cycle, Recursive Game Theory
Time Consistency.

JEL Codes: E61, E62, H21, H62, H63

1 Introduction

This paper studies the optimal tax policy in environments without a commitment technology that can bind the actions of future governments. We consider a stochastic growth economy with endogenous government spending. We describe the economy as a dynamic game between a benevolent government and a continuum of representative households. Following Phelan and Stacchetti (2001), we focus on the whole set of sequential sustainable equilibria (SSE). A novel numerical algorithm based on the approximation technique of Feng et. al. (2011) has been developed. Using this method, we numerically solve for the set of equilibrium payoffs, and investigate whether the time consistency problem of capital income tax is quantitatively important. For a realistically calibrated economy, we find that the optimal sustainable capital income tax rate is pro-cyclical and close to zero on average, while the labor income tax is counter-cyclical. Moreover, the welfare cost of no commitment is very small (0.22%) when compared with the Ramsey allocation. It has been shown that Ramsey policy with zero capital income tax is sustainable when the discounting is sufficiently low and the labor supply is sufficiently inelastic. The intuition of this result is closely related to Folk Theorem. In equilibrium, the government's temptation to deviate to a higher capital tax and lower labor tax in the first period is deterred by a lower continuation value for the government.[1] When the discount factor is low, the former effect dominates, and the government cannot resist the temptation. On the other hand, if the labor supply is sufficiently elastic, it will be attractive for the government to raise capital tax. Furthermore, we show that all equilibrium sequences are Markovian.[2] In particular, we find that the worst SSE is a Markov perfect equilibrium. We also find that the best sustainable equilibrium outcome may achieve substantially higher social welfare than the Markov-perfect equilibrium as considered by Klein, Krusell and Rios-Rull (2008).

There is a well-established literature that explores the properties of optimal taxation in the case when the government has access to a commitment technology. A celebrated result of Chamley (1986) and Judd (1985) showed that capital income should not be taxed in a deterministic steady state. Their results have been extended to the stochastic growth model [see Zhu (1992), Chari et al. (1994, 1995)] and to models with balanced government budget constraints [see Stockman (2000)]. The intuition for a

---

[1] This lower continuation value arises as a consequence of the household's belief about higher future tax rate, which discourages investment and reduces output.

[2] Chang (1998) first discovered this property in a monetary model without commitment.

zero capital tax comes from the fact that a capital tax is equivalent to an increasing tax on consumption goods in future periods, which violates the Atkinson and Stiglitz's (1976) result on uniform taxation.

One of the main assumptions underlying the above results is that a government can commit to a sequence of tax policies. However, it is generally recognized that the government usually does not have access to a perfect commitment technology.[3] In this case, as the seminal paper of Kydland and Prescott (1977) pointed out, a benevolent government would choose to deviate from the prescribed sequence of taxes. To illustrate the point, suppose that the policymakers have chosen a Ramsey policy for today and for the future, which prescribes a heavy tax rate on initial capital stock and a light tax rate on future capital to stimulate investment. However policymakers may choose to renege on the original plan when the future arrives. The reason is that once capital is accumulated it is sunk, and taxing capital is no longer distortionary. The government would choose higher capital taxes, since the investment decision has already been made, and a light tax rate on labor income to encourage work effort. The private agents anticipate this incentive to renege and rationally choose lower levels of investment when there is no commitment device to tie the hands of future policymakers. The missing commitment technology makes the Ramsey policy time inconsistent. It is also inappropriate to find the optimal policy using optimal control theory, which requires that the private agent's expectation about future government actions are time invariant.

A substantial amount of work has been done in order to find the ways to overcome time consistency problems [see Lucas and Stokey (1983), Persson, Persson, and Svensson (1987), Chari and Kehoe (1990)]. However, it remains open question how severe the time consistency problem is.

To answer this question, this paper builds on the literature that uses recursive game theory to study the time consistent policies. In an early contribution, Chari and Kehoe (1990) advocate sequential sustainable equilibrium (SSE) as a natural solution concept since we want to assure that the government maintains the sequential rationality and the economy remains in a competitive equilibrium. Phelan and Stacchetti (2001) blend the dynamic programming tools developed by Abreu, Pearce, and Stacchetti (1990) for repeated games with the recursive methods provided by Kydland and Prescott (1980) for the Ramsey problem with commitment to find the sustainable equilibrium value

---

[3]Romer and Romer (2010) identify that there are 104 legislated tax changes between 1947 and 2006 in the U.S.

3

set of a Ramsey tax model without commitment. Similar tools have been developed by Sleet (1997) for overlapping generation economy and by Chang (1998) for monetary economy. We complement this literature by providing an alternative numerical algorithm. Our algorithm is inspired by the key observation that the boundaries of the equilibrium payoff set can be recursively characterized. This observation simplifies the computation and permits a variety of experiments.

This paper is most closely related to Fernandez-Villaverde and Tsyvinski (2001). In an incomplete working paper, they try to solve a similar model by using the method of Phelan and Stacchetti (2001). While they follow the approximation strategy proposed by Judd, Yeltekin and Conklin (2000) and Sleet and Yeltekin (2003), we adapt a recent numerical algorithm developed by Feng et al. (2011). Their algorithm exhibits good convergence properties and accuracy properties for the simulated moments.

Our paper is related to Klein and Rios-Rull (2003), Klein, Krusell and Rios-Rull (2008), Martin (2009), Ortigueira (2006). They study the properties of Markov perfect equilibrium for the same class of economies like ours. The strategies that support the Markov perfect equilibrium are simple to describe. The drawback of their approach is that they limit the tax policies to differentiable strategies. It is not clear how much is lost by looking at this subset of equilibria. Furthermore we show that the worst SSE is a Markov perfect equilibrium.

Our analysis is also related to Benhabib and Rusticchini (1997) and Marcet and Marimon (1994). They provide an alternative method to solve policy games without commitment. They use the techniques of optimal control in which they explicitly impose additional constraints on the standard optimal tax problem such that a government does not deviate from the prescribed sequence of taxes. Similar to Chari and Kehoe (1990), their methods requires that the worst punishment to the deviating government can be easily determined.

Our research complements the existing literature on several fronts. With respect to the papers on Markov perfect equilibria, we study how much better the government can do by allowing for history depend strategies. With respect to the literature that studies the time consistent policies using optimal control, we provide a constructive procedure to compute the equilibrium payoff of deviation. Finally with respect to the literature on SSE, we offer quantitative assessments of the time consistency problem of capital and labor income taxes in a stochastic setting. We also provide an alternative numerical algorithm for the computation of the set of equilibrium payoffs.

In this paper, we provide analysis for two sets of economies. The first set of economies is calibrated to match the main features of the postwar U.S. economy. In the model economy, a stochastic shock to production technology mimics the TFP process. The representative infinitely lived households maximize life time expected utility by choosing a sequence of consumption, leisure and investment taking prices and tax sequence as given. Firms maximize profits using a neoclassical production technology. The government finances the consumption of public goods through labor and capital income taxes. The commitment technology is not available in the sense that the government sets the tax policy period by period. The results of our analysis can be summarized as follows. For a set of realistically calibrated parameter values, the Ramsey allocation cannot be supported by the credibility mechanism. However, the tax rate on capital income is close to zero on average. Numerical simulations indicate that the optimal sustainable capital income tax is pro-cyclical while labor tax is counter-cyclical. This is different from the Ramsey policy. The Ramsey government utilizes a counter-cyclical capital income tax to smooth government spendings and pro-cyclical labor tax to smooth the household's after tax labor income. When there is no commitment, the government cannot use capital income tax as a shock absorber, since the unexpected change in capital tax rate will distort the private agent's belief about future policies. In terms of welfare effect, we find that the cost of commitment is in the order of 0.22% in consumption equivalence when compared with the Ramsey allocation.

The second set is designed to compare the best SSE with the Markov perfect equilibrium. To facilitate the comparison with Klein, Krusell and Rios-Rull (2008), we abstract from uncertainty and consider the same set of economies as in their analysis. It has been found that by limiting the government's strategy to a differentiable one, as in Markov perfect equilibrium, the overall welfare may be reduced by 5.6% in consumption equivalence.

This paper abstracts from several important features of the data. First, for computational reasons, we don't consider government debt. It is not clear to us how restrictive this abstraction is. In an environment without government default, it has been shown that public debt may serve as a substitution for commitment [see Lucas and Stocky (1983), Dominguez (2007)]. If default is an option for the governments, they are tempted to default on their debts [see Prescott (1977), Chari and Kehoe (1993), Stockman (2004), Dominguez (2010)]. It is an interesting avenue to explore how gov-

ernment default affect the properties of optimal taxation, which is beyond the scope of the current research. Second, we abstract from any heterogeneity of agents. Different types of agents may raise interesting political economy issues that deserve careful study. However, it should be noted that our algorithm is built upon Feng et, al (2011), whose analysis allows for heterogeneous agents. It should be straightforward to extend our method to an environment with different types of government and consumers. For computational reasons, we opt for a setup with representation household.

The paper proceeds as follows. Section 2 is devoted to explaining the economic environment. Section 3 discusses how we can describe the problem as a dynamic game between the government and households. The equilibrium concept will be discussed in this section. Section 4 outlines some details how to characterize the set of sustainable equilibria and find the set by a recursive procedure. Section 5 explains the numerical implementation of this procedure. Sections 4 and 5 are rather technical and the reader can skip them without any loss of the economic intuition. The economy is calibrated to match certain features of the U.S. economy in section 6. Section 7 presents our findings and section 8 concludes. Some proofs and further computational details are documented in the appendix.

2 Economic Environment

The economy is populated by a measure one of representative, infinitely lived households and a competitively behaving firm. The primitive characteristics of the economy are defined by a stationary Markov chain. Time and uncertainty are represented by a countably infinite tree Σ. Each node of the tree, $\sigma \in \Sigma$, is a finite history of shocks $\sigma = s^t = (s_0, s_1, \ldots, s_t)$ for a given initial shock s_0. The process of shocks (s_t) is assumed to be a Markov chain with finite support S. The probability of each of these histories is given by $\pi(s^t)$. With this probability we can define conditional probabilities of future events given a particular history s^t that we denote by $\pi(s_{t+1}|s^t)$.

2.1 The firm

In each time period $t = 0, 1, 2, \ldots$, the firm produces the final good, $y(s^t)$, using capital, $k(s^{t-1})$ and labor, $l(s^t)$ through a neoclassical production function:

$$y\left(s^t\right) = A\left(s_t\right) F\left(k\left(s^{t-1}\right), l\left(s^t\right)\right) \tag{1}$$

Note that the current technology A is a function of current period shock.

The competitive firm sets the rental rates for its inputs equal to their marginal productivities.

$$r\left(s^t\right) = A(s_t)F_k\left(k\left(s^{t-1}\right), l\left(s^t\right)\right) \tag{2}$$

$$w\left(s^t\right) = A(s_t)F_l\left(k\left(s^{t-1}\right), l\left(s^t\right)\right) \tag{3}$$

The final good can be used for private consumption $c\left(s^t\right)$, public consumption $g\left(s^t\right)$, and as an investment good $i\left(s^t\right)$. The law of motion for capital $k\left(s^t\right)$ is given by:

$$k\left(s^t\right) = i\left(s^t\right) + (1-\delta)k\left(s^{t-1}\right) \tag{4}$$

where δ is the depreciation rate on capital and where we impose $i\left(s^t\right) \geq 0$.

2.2   Household

The representative household derives utility from a sequence of consumption and leisure by a time-separable utility function:

$$\sum_{t=0}^{\infty}\sum_{s^t} \beta^t \pi(s^t)u(c(s^t), 1 - l(s^t)) \tag{5}$$

where $\beta \in (0,1)$ is the discount factor, $l(s^t) \in [0,1)$ is the labor supply and the period utility function u is strictly concave, differentiable and satisfies the Inada condition.

The households choose a sequence of consumption and labor supply $\{c(s^t), l(s^t)\}$ to maximize their life-time expected utility as defined by equation (5) subject to the budget constraint

$$c(s^t) + i(s^t) = (1 - \tau_l(s^t))w(s^t)l(s^t) + (1 - \tau_k(s^t))\left(r(s^t) - \delta\right)k(s^{t-1}) \tag{6}$$

$$k\left(s_0\right) > 0 \tag{7}$$

where $\tau_k\left(s^t\right)$ is a proportional tax on capital gain and $\tau_l\left(s^t\right)$ a proportional tax on labor income, which are set by the current government as explained in the following subsection. $k\left(s_0\right)$ is the initial endowment of capital of households.

2.3 Government

The government is benevolent in the sense that it seeks to maximize social welfare:

$$\sum_{r=t}^{\infty} \sum_{s^r} \beta^{r-t} \pi(s^r) \left[u(c(s^r), 1 - l(s^r)) + G(g(s^r)) \right] \tag{8}$$

where G is a function of the public consumption $g(s^t)$, which is financed through taxes on labor income and capital gain.[4] A commitment technology is not available in the economy. Hence the government can only set tax rates on capital and labor incomes for the current period. In each period, the government can vary its decisions depending on the history of government policies up to the point at which the decision is made. The government consumption $g(s^t)$ is determined so as to balance the budget:

$$g(s^t) = \tau_l(s^t)w(s^t)l(s^t) + \tau_k(s^t) \left[r(s^t) - \delta \right] k(s^{t-1}) \tag{9}$$

We restrict the choice of tax rate from which the government can choose to the interval $T = \{[\underline{\tau}, \bar{\tau}]\}$.[5]

2.4 Competitive equilibrium

Although we are concerned with the case where the government cannot commit to an arbitrarily specified tax policy, along the equilibrium path, the households act as if the government had established such a policy. As argued by Phelan and Stacchetti (2001), the anonymity of each individual implies that their individual actions do not affect the actions of the government or other individuals and, in turn, they behave competitively. Therefore it is convenient to characterize the competitive equilibrium of the dynamic economy in which a tax policy is arbitrarily specified. We say that a tax policy $\tau = \{(\tau_k(s^t), \tau_l(s^t))\}_{t=0}^{\infty}$ is feasible if $(\tau_k(s^t), \tau_l(s^t)) \in [\underline{\tau}, \bar{\tau}]^2$ for all $t \geq 0$.

Definition 1 *Let $\Upsilon \{k_0, s_0, \tau\}$ be the economy starting with initial capital k_0 and shock s_0, and the tax policy is given by τ. A competitive equilibrium for $\Upsilon \{k_0, s_0, \tau\}$ is a sequence of allocation $(l(s^t), c(s^t), k(s^t))$ and prices $(r(s^t), w(s^t))$ such that (1) given prices and tax rates, consumers solve their problem; (2) input prices equate the marginal*

---

[4]It is equivalent to add the last term in the household's utility function, since each household takes the public spending as exogenously given. Our choice is meant to facilitate the analysis below.

[5]We need impose an upper bound on the tax rate to show the equivalence of the recursive and sequential problem for the households. The lower bound on the tax rate is used to bound the initial value correspondence.

productivities; (3) the government satisfies its budget constraint period by period, and (4) markets clear.

In equilibrium, the behavior of the household is governed by the following first order conditions:

$$u_c(s^t) - \beta \mathbb{E} \left\{ \left[1 + \left(1 - \tau_k(s^{t+1}) \right) \left(r(s^{t+1}) - \delta \right) \right] u_c(s^{t+1}) \right\} = 0 \tag{10}$$

$$u_l(s^t) - \left(1 - \tau_l(s^t) \right) w(s^t) u_c(s^t) = 0 \tag{11}$$

where u_c, u_l denote to the marginal utility with respect to consumption and leisure, and $\mathbb{E}$ is the expectation operator that yields the expected future value of a variable given the history of shock s^t. Equation (10) is the inter-temporal optimality condition where we take expectation over all possible realization of the shocks that can follow the current state. Equation (11) is the static optimality relation that equates the marginal utility from consumption with marginal utility from leisure.

We can reformulate the household's problem recursively following the strategy suggested by Kydland and Prescott (1980) [see Marcet and Marimon (1994), Phelan and Stacchetti (2001) and Feng et al. (2011) for recent applications]. We define a new variable:

$$m(s^t) := u_c(s^t) \left[1 + \left(1 - \tau_k(s^t) \right) \left(r(s^t) - \delta \right) \right] \tag{12}$$

which is the shadow value of investment.

For any s^t, $k(s^t)$, $\{m(s^{t+1})\}$ and an arbitrary specified tax policy τ, the household solves the following problem:

$$\max_{c(s^t), k(s^t)} u \left(c(s^t), l(s^t) \right) + \beta \mathbb{E} \left\{ m(s^{t+1}) k(s^t) \right\} \tag{13}$$

subject to the budget constraint and law of motion of capital as described in equations (6) and (4). The recursive problem is equivalent to the sequence problem as defined in section 2.2 when the transversality condition

$$\lim_{t \to \infty} \sum_{t=0}^{\infty} \sum_{s^t} \beta^t \pi(s^t) m(s^{t+1}) k(s^t) = 0 \tag{14}$$

is satisfied. This is shown in the following proposition.

Proposition 1 *The recursive and the sequence problem for the household are equiva-*

9

lent.

Proof.

See the Appendix.

∎

From now on we denote x and x_+ as the current and future value of the variable x. Now we can summarize the competitive equilibrium conditions for a given tax policy τ as follows.

Definition 2 *Let* $\Upsilon^S \{k, s, (\tau_k, \tau_l), \{m_+\}\}$ *be the static (one-period) economy where each household has an initial capital stock k, the current state is s, the anticipated shadow value of investment is given by the vector $\{m_+\}$ and the government imposes tax (τ_k, τ_l) for current period. $\{c, l, k_+, r, w, g\}$ is a competitive equilibrium for $\Upsilon^S \{k, s, (\tau_k, \tau_l), \{m_+\}\}$ and denoted by $\{c, l, k_+, r, w, g\} \in \mathbf{CE}^S \{k, s, (\tau_k, \tau_l), \{m_+\}\}$ if and only if the following conditions are satisfied: (1) input prices equates the marginal productivities; (2) the government satisfies its budget constraint; (3) market clears; and (4) $u_c(s) = \beta \mathbb{E} \{m_+\}$, $u_l(s) = (1 - \tau_l) w(s) u_c(s)$.*

The following Lemma allows us to think of the original economy as a sequence of static economy with endogenously changing state variables and exogenous stochastic shocks.

Lemma 3 *Given a feasible tax policy $\tau = \{(\tau_{k,t}, \tau_{l,t})\}_{t=0}^{\infty}$, an initial capital stock k_0 and shock s_0, suppose that the sequence $\{c(s^t), l(s^t), k(s^t), r(s^t), w(s^t), g(s^t)\}_{t=0}^{\infty}$ is such that for each t,*

$$\left\{c(s^t), l(s^t), k(s^t), r(s^t), w(s^t), g(s^t)\right\} \in \mathbf{CE}^S \left\{k(s^t), s_t, (\tau_{k,t}, \tau_{l,t}), \{m(s^{t+1})\}\right\},$$

where

$$m(s^{t+1}) := u_c(s^{t+1}) \left[1 + (1 - \tau_{k,t+1}) \left(r(s^{t+1}) - \delta\right)\right],$$

Then $\{c(s^t), l(s^t), k(s^t), r(s^t), w(s^t), g(s^t)\}_{t=0}^{\infty}$ constitutes a competitive equilibrium for $\Upsilon \{k_0, s_0, \tau\}$.

As it will get clear soon, the introduction of m and the equivalence result above will help us characterize the game, as we describe in the following section, recursively.

3 The Dynamic Game

In this section we first describe the game associated with our environment. We then explain the equilibrium concept adapted in this study.

3.1 The game

The timing protocol is as follows.[6] At the beginning of every period t, the shock s_t is revealed, and the government sets the tax rates for the current period $\{\tau_k(s^t), \tau_l(s^t)\} \in [\underline{\tau}, \bar{\tau}]^2$. Then the households simultaneously choose their labor input $l(s^t)$. Given the aggregate capital, which is chosen in the previous period, and labor input, the market determines prices for the capital and labor. Finally, each household independently allocates the after-tax income between consumption and investment for production in the following period. The government uses the totality of tax revenue to finance the consumption of public good.

We use $\Gamma(k_0, s_0)$ to denote the dynamic game between the government and the households when all the households are endowed with initial capital of k_0 and shock of s_0. We denote the public history of the game by $\zeta^t = (\zeta_0, \zeta_1, \ldots, \zeta_t)$, where $\zeta_t = (s_t, \tau_k(s^t), \tau_l(s^t), c_t, l_t, k_t)$, that is, the history of shocks, government policies and representative household's choices. In this paper we study symmetric strategy profiles as in Phelan and Stacchetti (2001), where all households choose the same actions along the equilibrium path. A symmetric strategy profile for $\Gamma(k_0, s_0)$ is a pair of strategies $\sigma = (\sigma_C, \sigma_G)$. The strategy for the government σ_G is a measurable function that maps the publicly observed history ζ^{t-1} and current shock s_t into tax rates for period t, namely $(\tau_k(s^t), \tau_l(s^t)) = \sigma_G(\zeta^{t-1}, s_t)$. The strategy for the household σ_C specifies c_t, l_t, and k_t as a function of expanded history $(\zeta^{t-1}, \tau_k(s^t), \tau_l(s^t))$, that is, $(c_t, l_t, k_t) = \sigma_C(\zeta^{t-1}, \tau_k(s^t), \tau_l(s^t))$. We denote the set of all symmetric strategy profiles for $\Gamma(k_0, s_0)$ by $\sum(k_0, s_0) = \sum_C(k_0, s_0) \times \sum_G(k_0, s_0)$, where $\sum_C(k_0, s_0)$ represents the set of strategies for the households and $\sum_G(k_0, s_0)$ the set of strategies for the government.

3.2 Sustainable equilibrium

The solution concept we use in this study is sustainable equilibrium originated by Chari and Kehoe (2001). To facilitate the definition of sustainable equilibrium, we

---

[6]Figure 2 in appendix provides a graphical representation of the timing.

define the value of a strategy σ (for the government):

$$\Phi_G(k_0, s_0, \sigma) := h = \sum_{t=0}^{\infty} \beta^t \mathbb{E}\left[u(c_t, 1 - l_t) + G(g_t)\right]. \tag{15}$$

As we argued in section 2.4, the household is able to make its optimal choices when they have correctly anticipated the value of $\{m_+\}$, from which the household can formulate the current and future prices for labor, capital and consumption.[7] Therefore we define the value of m as the "payoff" to the household:

$$\Phi_C(k_0, s_0, \sigma) := m = u_c(c_t, 1 - l_t)\left[1 + (1 - \tau_{k,t})\left(r(s^t) - \delta\right)\right]. \tag{16}$$

We define a sustainable equilibrium as follows:

Definition 4 *A symmetric strategy profile σ of the game $\Gamma(k_0, s_0)$ is a sustainable equilibrium if for any $t \geq 0$ and history ζ^t: (1) $\Phi_G\left(k_t, s_t, \sigma|\zeta^{t-1}\right) \geq \Phi_G\left(k_t, s_t, (\sigma|\zeta^{t-1}, \gamma)\right)$ for any strategy $\gamma \in \sum_G(k_t, s_t)$, where $k_t \in \sigma_C\left(\zeta^{t-1}, \gamma\right)$; (2) $\{c_t, l_t, k_{t+1}, r_t, w_t, g_t\}_{t=0}^{\infty}$ is a competitive equilibrium for $\Upsilon\{k_0, s_0, \tau\}$ where $\tau := \{\tau_{k,t}, \tau_{l,t}\}_{t=0}^{\infty}$, $(\tau_{k,t}, \tau_{l,t}) \in \sigma_G\left(\zeta^{t-1}, s_t\right)$, and $(c_t, l_t, k_{t+1}) \in \sigma_G\left(\zeta^t\right)$.*

The first condition says that the continuation payoff for the government following strategy σ is higher than the payoff from any deviation to a different strategy. The second condition requires that the households always respond to the government strategy with decisions that imply a competitive equilibrium since this is the situation compatible with feasibility and optimality. It should be noted that there exist unique values of (h, m) associated with any symmetric strategy profile σ of the game $\Gamma(k_0, s_0)$ as defined by (15) and (16).

4   Self-Generation

In this section, we apply the Abreu, Pearce and Stacchetti's (1990) method to our dynamic game. We follow closely the work of Phelan and Stacchetti (2001) and we only explain in detail the differences between their application and ours. The key observation is that the continuation value (for both the government and the household)

---

[7]$\{m_+\}$ denotes the vector of m_+ under different shocks s_+. It has summarized what the other households will do in all future periods and contingencies.

from the next period on can summarize all relevant information about the future, which is enough for them to choose the optimal strategies for the current period.

Hence, instead of finding the strategy functions that are defined over the history, which is a formidable task, we focus on the characterization of the equilibrium value correspondence as defined below. We define an equilibrium value correspondence $\mathbf{V}(k, s)$ that maps the value of the states (k, s) into the set of possible payoffs associated with a strategy profile σ that constitutes a sustainable equilibrium. That is,

$$\mathbf{V}(k, s) := \{(m, h) \,|\, \sigma \text{ is a sustainable equilibrium for } \Gamma\,(k_0, s_0)\}\,.$$

In order to recursively characterize the equilibrium value correspondence $\mathbf{V}(k, s)$, we need an operator $\mathbb{B}$, which maps the space $\mathbf{A}$ of all value correspondence into itself, where an arbitrary value correspondence is defined as $\mathbf{W}(k, s) : \mathbf{R}_+ \times \mathbf{S} \to \mathbf{R}^2$. For given values of (k, s), the correspondence is a set of (m, h) generated by a sustainable equilibrium strategy profile, where m is the shadow value of investment and h is the government's payoff as defined by equations (15) and (16). Before we proceed with operator $\mathbb{B}$, we first explain the concepts of consistency and admissibility with respect to $\mathbf{W}$. These will be used to define the operator.

Definition 5 *A vector $\psi = (\tau_k, \tau_l, c, l, k_+, \{(m_+, h_+)\})$ is consistent with respect to the value correspondence $\mathbf{W}$ at (k, s) if*

$$(c, l, k_+, F_k(k, l), F_l(k, l), g) \in \mathbf{CE}^S \{k, s, (\tau_k, \tau_l), \{m_+\}\}\,,$$

for $(m(k, s, \psi), h(k, s, \psi)) \in \mathbf{W}(k, s)$, and $(m_+, h_+) \in \mathbf{W}(k_+, s_+)$, where m is the shadow value of investment and h is the value of the vector ψ

$$m(k, s, \psi) := u_c(c, 1 - l)[1 + (1 - \tau_k)(r - \delta)], \tag{17}$$

$$h(k, s, \psi) := u\,(c, 1 - l) + G(g) + \beta\mathbb{E}h_+. \tag{18}$$

In words, consistency guarantees that the vector ψ delivers an allocation that is optimal for the household and satisfies feasibility. It also requires that the promised continuation values (m_+, h_+) belong to the same equilibrium value correspondence as the implied (m, h) do.

Definition 6 *The vector ψ is admissible with respect to $\mathbf{W}$ if it is consistent with*

respect to $\mathbf{W}$ *at* (k, s) *and*

$$h(k, s, \psi) \geq h(k, s, \psi'). \tag{19}$$

This condition says that the government can not increase its payoff by announcing any unexpected tax rate τ' other than τ. Therefore, admissibility guarantees that the government does not have any incentive to deviate.

Definition 7 *For a value correspondence* $\mathbf{W}$*, we define*

$$\mathbb{B}(\mathbf{W})(k, s) = \{(m, h) \,|\, \psi \text{ is admissible with respect to } \mathbf{W} \text{ at } (k, s)\}.$$

In the following, we adapt the results of Abreu, Pearce and Stacchetti (1990) to our game. It should be noted that these results are straightforward extensions of those in Phelan and Stacchetti (2001) for our environment with stochastic shock. Therefore we only list the main results and refer the readers to their paper for proofs.

1. If $\mathbf{W} \subseteq \mathbb{B}(\mathbf{W})$ then $\mathbb{B}(\mathbf{W}) \subseteq \mathbf{V}$.

2. $\mathbf{V}$ is compact and the largest value-correspondence $\mathbf{W}$ such that $\mathbf{W} = \mathbb{B}(\mathbf{W})$.

3. $\mathbb{B}(\cdot)$ is monotone and preserves compactness.

4. If we define $\mathbf{W}_{n+1} = \mathbb{B}(\mathbf{W}_n)$ for all $n \geq 0$, and $\mathbf{V} \subset \mathbf{W}_0$, then $\mathbf{W}_\infty = \mathbf{V}$.

Result 1 has been called as self-generation. From the definition of the sequential equilibrium value correspondence, it is straightforward to see that $\mathbf{V} \subseteq \mathbb{B}(\mathbf{V})$. Together with result 1, it is fairly easy to reach the second result. Results 3 and 4 will be used in the next section to approximate the equilibrium value correspondence.

5 Computation of Equilibria and Recovering Strategies

In this section, we first detail the numerical algorithm that is used to compute the value correspondence. We then explain how to build strategies that support a given pair of payoffs in that correspondence.

5.1 Alternative operator

Computing the mapping $\mathbb{B}$ amounts to find a set $\mathbb{B}(\mathbf{W})$, the set of pairs of (m, h) that can be "enforced" today:

$$\mathbb{B}(\mathbf{W})(k, s) = \{(m, h) | \exists (\tau_k, \tau_l), (c, l, k_+, g, w, r) \text{ and } (m_+, h_+) \in \mathbf{W}(k_+, s_+), \text{ for all } s_+ \succ s\}$$

such that

$$m = u_c \cdot [1 + r(1 - \tau_k)] \tag{20}$$

$$h = u(c, l) + G(g) + \beta \mathbb{E} h_+ \tag{21}$$

$$(m, h) \in \mathbf{W}(k, s) \tag{22}$$

$$h \geq \left[u(c', l') + G(g) + \beta \mathbb{E} h'_+ \mid \left(m'_+, h'_+ \right) \right], \forall \left(m'_+, h'_+ \right) \in \mathbf{W}(k_+, s_+) \tag{23}$$

$$u_c(c, l) - \beta \mathbb{E} \{ m_+ \} = 0 \tag{24}$$

$$(1 - \tau_l) w u_c - u_l = 0 \tag{25}$$

$$\tau_k, \tau_l \in [\underline{\tau}, \bar{\tau}] \tag{26}$$

The constraints (20, 21, 22) are called "regeneration constraints," while (23) is an "incentive" constraint. (24, 25) is necessary to ensure that the continuation of a sustainable plan after any deviation is consistent with a competitive equilibrium.

As pointed out by Chang (1998), computing $\mathbb{B}(\mathbf{W})$ given $\mathbf{W}$ is complicated in particular by the presence of the constraint (23). Chang (1998) suggested an alternative operator to circumvent this complication in the context of finding time-consistent monetary policy, while Phelan and Stacchetti (2001) developed a similar operator in a production economy like ours.

The basic idea of a simpler approach is the following. On the one hand, the government does not need to evaluate the consequences of all possible actions, it only needs to consider the payoff associated with the "best" deviation. On the other hand, and as in Abreu, Pearce and Stacchetti (1990), if the government chooses to deviate, this is then followed by the worst available punishment.

In line with Chang (1998), we redefine the operator $\mathbb{B}: \mathbf{A} \to \mathbf{A}$ by replacing (23) with the following condition,

$$h \geq \tilde{h}(k, s) \tag{27}$$

where $\tilde{h}(k, s)$ is the worst possible payoff for the government when it announces un-expected tax rate τ'. As in Phelan and Stacchetti (2001), we only consider extreme punishments. Therefore, we define $\tilde{h}(k, s)$ as

$$\tilde{h}(k, s) = \max_{\tau_k, \tau_l} \left\{ \min_{c, l, k_+, (m_+, h_+) \in \mathbf{W}} [u(c, l) + G(g) + \beta \mathbb{E} h_+] \right\}$$

such that

$$[c, l, k_+, g, w, r] \in \mathbf{CE}^S \{ (k, s, \tau_k, \tau_l, \{m_+\}) \text{ for all } s_+ \succ s \}$$

Chang (1998) showed that conditions (23) and (27) are equivalent in the sense that the repeated application of operator $\mathbb{B}$, with either (23) or (27), yields a decreasing sequence of sets that converges to the same fixed point $\mathbf{V}$. Phelan and Stacchetti (2001) define their operator using condition (27).

Even though we can simplify the computation by substituting condition (23) with (27), it is not a trivial matter as we need to operate over set. Phelan and Stacchetti (2001) adapt the innovative approximation technique developed by Judd, Yeltekin and Conklin (2003). Instead we apply the numerical method developed by Feng et al. (2011) to our economy. Their method is able to approximate convex-valued value correspondence and has good convergence property as stated in Theorem 4.1 in their paper.

We should emphasize that we compute the upper and lower boundaries of $\mathbf{W}(k, s)$, which are represented by the following functions in line with Phelan and Stacchetti (2001)

$$\bar{h}(k, s, m) := \max_h \{ h | (m, h) \in \mathbf{W}(k, s) \}, \tag{28}$$

$$\underline{h}(k, s, m) := \min_h \{ h | (m, h) \in \mathbf{W}(k, s) \}. \tag{29}$$

As observed by Phelan and Stacchetti (2001), the lowest value in $\mathbf{W}(k, s)$ yields the value of the worst punishment for the government $\tilde{h}(k, s) = \min_m \underline{h}(k, s, m)$, which corresponds to the equilibrium when the government plays the trigger strategy. It also should be noted that the highest value $\max_m \bar{h}(k, s, m)$ in $\mathbf{W}(k, s)$ corresponds to the equilibrium that the government obtains the maximum payoff (playing the best strategy). We are interested in the boundaries since we learn a lot about the equilibrium set from looking at extremes and learning about the best and trigger strategies conveys a lot of information about all the other strategies. To this end, we find an outer

approximation of $\mathbf{W}$: $\hat{\mathbf{W}}(k,s) = \left\{ (m,h) \,|\, h \in \left[\underline{h}(k,s,m), \bar{h}(k,s,m) \right] \right\}$. The following proposition will be useful for the computation of the boundaries. Before we proceed, we need two assumptions.

(A1) There exists at most one set of (c, l, k_+) that solves the following equation system at given (k, s) and $\{\tau_k, \tau_l, m\}$,

$$m = u_c(c, l) \left[1 + (1 - \tau^k)(r - \delta) \right] \tag{30}$$

$$u_l(c, l) + (1 - \tau_l) w u_c(c, l) = 0 \tag{31}$$

$$\left[1 + (1 - \tau_k)(r - \delta) \right] k + (1 - \tau_l) w l - c - k_+ = 0 \tag{32}$$

(A2) $\mathbf{W}$ is convex-valued at given $\{k, s, m\}$.

Assumption A1 rules out some commonly used utility function. However, the following pair of utility and production functions satisfy this assumption.

$$u(c, l) = \frac{c^{1-\sigma_1}}{1 - \sigma_1} + \gamma_l \frac{(1-l)^{1-\sigma_2}}{1 - \sigma_2} \tag{33}$$

$$F(k, l) = A k^\alpha l^{1-\alpha}. \tag{34}$$

It should be noted that the equation system (30-32) will degenerate into a single non-linear equation in term of l.[8] In terms of assumption A2, we use a randomization device similar to Phelan and Stacchetti (2001) to make it applicable. However, the randomization device that we introduced does not require a synchronization between the government's and the household's moves and beliefs. Instead, it only convexifies the payoff set of the government at given value of (k, s, m), which we should interpret as the government plays a mixed strategy.

Proposition 2 *If the equilibrium value correspondence $\mathbf{V}(k,s)$ satisfies assumption (A1, A2), then for all $(m, h) \in \mathbf{V}(k, s)$,*

$$\bar{h}(k, s, m) = \max_{\tau_k, \tau_l} u(c, l) + G(g) + \beta \mathbb{E} \bar{h}(k_+, s_+, m_+) \tag{35}$$

$$\underline{h}(k, s, m) = \max_{\tau_k, \tau_l} u(c, l) + G(g) + \beta \mathbb{E} \underline{h}(k_+, s_+, m_+) \tag{36}$$

$$\tilde{h}(k, s) = \min_m \underline{h}(k, s, m). \tag{37}$$

---

[8]For the preference and production function we used, the above equation system can be converted into the following equation in terms of l: $A k^\theta = B(1-\delta) l^\theta + \left[A k^\theta + B(1 - \tau_k) \theta k^{\theta-1} \right] l$, where $A = m(1-\tau_l)(1-\theta)$, $B = (1-\beta)(1-\alpha_p)(1-\alpha_c)$. When $\{k, s, m, \tau_k, \tau_l\}$ are given, the right hand side of the equation is monotone increasing in $l \in [0, 1]$. Therefore, there exists at most one solution $l^* \in [0, 1]$.

Proof.

See the Appendix.

∎

This result suggests that we can focus on the boundaries of the value correspondence if our primary interest is about the best and trigger strategies. This will greatly contain the computational cost. The above observation inspired us to define operator $\mathbb{F}$.[9] The following result shows that this operator has good convergence property. The repeated application of this operator generates a sequence of sets that converges to the equilibrium value correspondence $\mathbf{V}$.

Theorem 8 *Assume the equilibrium value correspondence $\mathbf{V}$ satisfies (A1, A2). Let $\hat{\mathbf{W}}_0$ be a convex-valued correspondence such that $\hat{\mathbf{W}}_0 \supset \mathbf{V}$. Let $\hat{\mathbf{W}}_n = \mathbb{F}(\hat{\mathbf{W}}_{n-1})$. Then $\lim_{n\to\infty} \hat{\mathbf{W}}_n = \mathbf{V}$.*

Proof.

See the Appendix.

∎

5.2 Recovering strategy

In this section, we outline how to find the strategy that corresponds to a given pair of payoffs belongs to the equilibrium value correspondence. Note that for better exposition, we abstract from uncertainty and we will focus on the strategy that yields the highest payoff for the government. This procedure can be generalized to find strategies supporting any point belongs to the equilibrium payoff set.

1. $t = 0, k_0$ is given, we find the highest possible value of $h_0 = \sup\{h|(m_0, h_0) \in \mathbf{W}^*(k_0)\}$ and its corresponding m_0. Then we find the government's tax policy that can support (m_0, h_0). More specifically, we find $(\tau_{k,0}, \tau_{l,0})$ such that

$$u(c_0, l_0) + G(g_0) + \beta h_1 = h_0 \qquad (38)$$

---

[9] The definition of operator $\mathbb{F}$ and its numerical implementation can be found in the appendix.

where $h_1 = \bar{h}(k_1, m_1)$, $m_1 = \frac{u_c(c_0, l_0)}{\beta}$ and we can find the values of (c_0, l_0, k_1) by solving the following equation system

$$m_0 - u_c(c_0, l_0) \cdot [1 + r_0(1 - \tau_{k,0})] = 0 \qquad (39)$$
$$u_l(c_0, l_0) - (1 - \tau_{l,0})w_0 \cdot u_c(c_0, l_0) = 0 \qquad (40)$$
$$(1 - \tau_{l,0})w_0 l_0 + (1 - \tau_{k,0})r_0 k_0 + (1 - \delta)k_0 - (c_0 + k_1) = 0. \qquad (41)$$

when the values of $(\tau_{k,0}, \tau_{l,0}, m_0)$ are given. Therefore the above problem is well-defined in terms of $(\tau_{k,0}, \tau_{l,0}, m_0, h_0)$.

2. $t = 1$, k_1, m_1, h_1 are given by the solution in step 1. Now we find the government policies $(\tau_{k,1}, \tau_{l,1})$ such that

$$u(c_1, l_1) + G(g_1) + \beta h_2 = h_1 \qquad (42)$$

in the same spirit of step 1.

3. We repeat step 2 for $t \to T$.

It should be noted that the construction above reveals that any sustainable outcome has essentially a Markovian structure in the sense that, $k(s^t), \tau_k(s^t), \tau_l(s^t)$ and $m(s^t), h(s^t)$ only depend on history ζ^{t-1} through $m(s^{t-1}), h(s^{t-1})$. More important, we found that the worst SSE has the following property.

Proposition 3 *For the game* $\Gamma(k_0, s_0)$, *the strategy profile*
$\{(\tau_k(s^t), \tau_l(s^t)), (c(s^t), l(s^t), k(s^t))\}_{t=0}^{\infty}$ *that has an equilibrium payoff for the government of* $\tilde{h}(k_0, s_0)$ *has the following property: there exist measurable functions* ϑ^{τ_k}, ϑ^{τ_l}, ϑ^{k} *and* ϑ^{l} *such that* $\tau_k(s^t) = \vartheta^{\tau_k}(k_{t-1}, s_t)$, $\tau_l(s^t) = \vartheta^{\tau_l}(k_{t-1}, s_t)$, $k(s^t) = \vartheta^{k}(k_{t-1}, s_t)$, $l(s^t) = \vartheta^{l}(k_{t-1}, s_t)$ *for all* $t \geq 0$.

Proof.
    See the Appendix.

    ∎

This proposition essentially says that the worst SSE is a Markov perfect equilibrium. Whenever the government announces an unexpected tax policy, the private agent's belief about the future policies will be distorted. In order to provide enough incentive

for the government to follow the policy recommendation, the households would rather to undermine their expectation about the future to a point such that their choices will yields the lowest payoff for the government for any announced unexpected tax changes. The government understands this and takes it into account before any announcement of tax rates. The advantage of this expectation is that it provides the strongest incentive for the government to follow the policy recommendation. In this equilibrium, the strategy profile is not history dependent since the household's belief has been manipulated to a point that the government's best response is to play the same strategy over time.

6 Calibration

To facilitate comparison with the Markov perfect equilibrium as considered in Klein and Rios-Rull (2003), we adopt this parametrization. The utility function is of the CRRA form, and technology is Cobb-Douglas:

$$u(c,l) = (1-\beta)(1-\alpha_p)\frac{[c^{\alpha_c}(1-l)^{1-\alpha_c}]^{1-\sigma}}{1-\sigma} \tag{43}$$

$$F(K,L) = A(s)K^{\alpha}L^{1-\alpha} \tag{44}$$

We set $\sigma = 1$ to satisfy the assumption A1.[10] The preference of the government from public consumption is given by

$$G(g) = (1-\beta)\alpha_p \ln g \tag{45}$$

The TFP process is calibrated so as to match the variance and auto-correlation reported in Prescott (1986). To offer insight in the effect of uncertainty on the credibility of the trigger-type strategy, we also offer results for the deterministic model where $A(s) = 1$ for all $s \in S$. The rest of the parameter values are fixed such that certain moments in the stationary distribution of the Ramsey allocation match with the real-world statistics. α is picked to match labor income share of national product, β to generate a pre-tax interest rate of around 4% [see McGrattan and Prescott (2000) for a justification of this number based on their measure of the return on capital and

---

[10]This yields a log utility: $u(c,l) = (1-\beta)(1-\alpha_p)[\alpha_c \log c + (1-\alpha_c)\log(1-l)]$.

on the risk-free rate of inflation-protected U.S. Treasury bonds]. δ is chosen to match a capital-output ratio of around 3, α_p to get a share of government consumption around 20% and α_c to get hours worked to be around 25% of total time. Finally we set an upper limit and lower bound on the tax rates that equate to the highest and lowest tax rates observed in a simulated Ramsey economy to facilitate the comparison with the Ramsey policy. This can be viewed as an institutional constraint. We summarize the calibration values in Table 1.

Table 1: Parameter values for the baseline economy

Parameter Values					
β	σ	δ	α	α_p	α_c
0.96	1	0.08	0.36	0.13	0.30
$A \in \{0.976, 1.024\}$			$\pi^{11} = \pi^{22} = 0.946$		

7   Quantitative Analysis

In this section, we first present the results for the stochastic economy in section 7.1. This section starts with the description of the economy with full commitment. This is followed by a quantitative characterization of optimal fiscal policy when there is no commitment device available in the economy. We then compare the outcomes of these two economies. We also discuss the property of the worst equilibrium that supports the best one. Section 7.2 is devoted to understanding what matters to the results. In section 7.3, we compare the best sustainable equilibrium with Markov-perfect equilibrium.

7.1   Optimal taxation over the business cycle

7.1.1   Commitment

Table 2 presents the key finding of the economy with full commitment.[11] The main properties of this economy can be summarized as follows:

1. The capital income tax rate is close to zero on average;

---

[11]The model with full commitment has been solved using the primal approach as in Stockman (2001). The author thanks David Stockman for sharing his codes and clarifying some technical questions.

2. The labor income tax rate is very high in the sense that the lion share of tax revenue comes from the labor income tax;

3. The capital income tax rate is much more volatile than the labor income tax rate;

4. The capital income tax rate is counter-cyclical, while the labor income tax rate is pro-cyclical.

The properties that we found are in line with the findings for the baseline model and the economy with only a technology shock described in Chari et al. (1994). The Ramsey government takes considerations of the distortion of capital taxation. It eliminates such distortion by setting the long-run capital income tax rate to be zero. In terms of business cycle property, the optimal labor tax is pro-cyclical while the optimal capital income tax is counter-cyclical. A pro-cyclical labor income tax helps to smooth the household's after-tax wage income. The counter-cyclical capital income tax provides an efficient means of absorbing shocks to the government's budget, which vary with the size of the tax base over the business cycle. The reason for capital income taxation acting as the shock absorber has been made very clear by Judd (1993). The efficiency cost of such tax rate change over the business cycle is low because the short-run supply elasticity of capital is very small.

Table 2: Business cycle properties of the economy with full commitment

Statistics	Mean	Std.dev.	AutoCorr.	Corr. w/ GDP	Corr. w/ s
Cap. inc. tax	0.002	0.034	0.694	-0.665	-0.748
Lab. inc. tax	0.397	0.002	0.653	0.364	0.409
Output	0.452	0.018	0.935	1.000	0.993
Consumption	0.230	0.008	0.987	0.937	0.890
Hours	0.245	0.003	0.819	0.551	0.635
Capital stock	1.337	0.056	0.992	0.838	0.770

7.1.2 No Commitment

For the economy without commitment, we find the strategies that support the best equilibrium outcome and report the key statistics in table 3. The main findings are listed as follows:

1. The capital income tax rate is very low on average;

2. The labor income tax rate is very high;

3. The capital income tax rate is more volatile than the labor income tax rate. However, the labor income tax rate is more volatile than in the economy with commitment;

4. The capital income tax rate is slightly pro-cyclical, while the labor income tax rate is counter-cyclical.

The average tax rates on capital and labor are very similar to those described by Chari et al. (1994) and Stockman (2001) for the economies with full commitment. Even though the capital income is distortion free in the short term, the government considers the fact that a positive tax on capital will alter the household's belief about the future policy. Depending on the current tax rate, the household's beliefs will lead to certain choices that yield sufficiently low payoff for the government. This effect deters the government's incentive to explore the inelastic aspect of capital supply. In equilibrium, the government sets the average capital tax rate to be 4.7% and reaches the highest social welfare. Furthermore, the consensus on the positive capital income tax and its distortion effect on public beliefs makes is unlikely to use capital tax as a shock absorber. In fact, the capital tax rate is only slightly positively correlated with technology shocks, with a correlation coefficient of merely 0.057, compared with -0.748 for the economy with commitment. Instead the government chooses labor income taxes as a means to absorb shocks to its budget. Consequently the tax rate on labor income exhibits higher volatility, with a standard deviation of 1.4%, compared with 0.2% in the Ramsey policy.

Table 3: Business cycle properties of the economy without commitment

Statistics	Mean	Std.dev	AutoCorr.	Corr. w/ GDP	Corr. w/ s
Cap. inc. tax	0.047	0.054	0.783	0.090	0.057
Lab. inc. tax	0.387	0.014	0.780	-0.043	-0.043
Output	0.455	0.021	0.906	1.000	0.961
Consumption	0.232	0.009	0.923	0.730	0.735
Hours	0.248	0.005	0.526	0.647	0.574
Capital stock	1.331	0.066	0.990	0.794	0.696

These findings are in sharp contrast with the patterns suggested by Klein and Rios-Rull (2003), who analyze the Markov perfect equilibrium for a similar economy without commitment. They find that the average capital income tax rate is 65% and the tax rate on labor income is 12%. Proposition 3 shows that the worst SSE is a Markov perfect equilibrium. When we compare the worst SSE with what Klein and Rios-Rull (2003) found, we can see some similarities. We suspect that what they found corresponds to the worst SSE.[12]

7.1.3   The implication of no commitment

To further understand why commitment makes a difference, Figure 1 presents the dynamic adjustment of an economy in which the government switches from full commitment to no commitment. We assume away the stochastic shocks for comparison. In this experiment, the economy starts from the long-run situation under commitment. The government loses commitment at period zero. The figure shows the outcome path and the period length is one year.

We can see that the government chooses to impose a very high tax rate on the capital income and decreases the labor income tax when the commitment is lost. This happens because the capital income is not distortionary to the current government. The capital stock starts to fall and the hours worked increases. After the first few periods, the capital income tax becomes lower and labor income tax increases over time. This pattern is very close to what Phelan and Stacchetti (2001) found in their numerical example. However it is different from what Klein and Rios-Rull (2003) reported. They suggest that the capital income tax rate keeps increasing and the labor income tax rate decreases monotonically after the government lost commitment.

Next we look at the welfare cost of no commitment. In line with the methodology as explained by Conesa and Krueger (1999), this paper measures the welfare loss of a particular tax policy, compared with Ramsey, by computing the consumption equivalent variation (CEV). We quantify the welfare effect of a given policy by asking how much consumption has to be increased in each state and date in order to equate expected utilities in Ramsey allocation while leaving leisure and public expenditure unchanged.

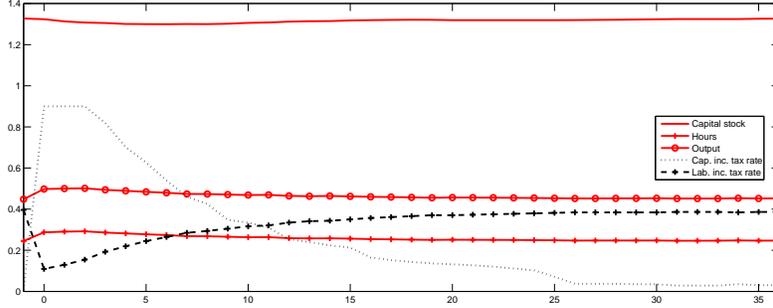[12]See the discussion in section 7.2 below.

24

Figure 1: Simulated outcome path after the government loses commitment

More specifically, we are finding an x such that

$$\sum_{t=0}^{\infty} \beta \mathbb{E}_0 \left[u((1+x)\hat{c}_t, 1 - \hat{l}_t) + G(\hat{g}_t) \right] = \sum_{t=0}^{\infty} \beta \mathbb{E}_0 \left[u(c_t^R, 1 - l_t^R) + G(g_t^R) \right]$$

where $\left\{ \hat{c}_t, \hat{l}_t, \hat{g}_t \right\}$ denotes the equilibrium allocation of the ad hoc tax policy and $\left\{ c_t^R, l_t^R, g_t^R \right\}$ represents the Ramsey allocation. The welfare loss of no commitment is merely 0.22% in consumption equivalent.

The welfare gains of the full commitment mainly come from zero tax rate on capital. The gain from the large initial tax on capital income is nearly zero. It seems to be a contradiction with Chari, Christiano and Kehoe (1994). However we argue that our findings are consistent with theirs. In Chari, Christiano and Kehoe (1994), the government can eliminate the distortion of capital income tax by setting a large initial tax and a zero rate thereafter. This channel only works through the issue of government debt. For computational reasons, we assume that the government does not have access to debt issuance. Consequently, the Ramsey government can only utilize the long run zero tax rate on capital. Stockman (2001) investigates the effects of a balanced-budget restriction on the Ramsey allocation. It has been found that there is a big welfare gain associated with switching from the Ramsey with a fixed debt, which is similar to the Ramsey government in our study, to the Ramsey with no constraint on debt issuance. However it is not necessary to conclude that we will observe a bigger welfare gain than what we found here when the government debt is introduced. This is because the time consistency problem may go away as the government debt functions as the central commitment device among governments [see Domínguez (2007)].

7.2 Relation with Ramsey and Markov perfect

In order to understand the difference between looking at SSE and Markov perfect equilibrium, we study the same set of economies considered by Klein, Krusell and Rios-Rull (2008). We look at the steady states of the economy under four different benevolent governments that we label Pareto, Ramsey, Markov and SSE. Both Pareto and Ramsey economies have commitment technology, while commitment is not available in either Markov or SSE economies. In Pareto economy, the government has access to lump-sum taxation. In the rest three economies, the government is restricted by a period-by-period balanced-budget constraint and to the use of distortion taxation. Markov government only uses differentiable strategies and SSE economy find the best equilibrium outcome through the reputation mechanism. Table 5 reports the steady-state allocations of these four economies under two different tax regimes.

Table 4: Simulations of the deterministic economy

Economy	Govt.	y	k/y	c/y	g/y	c/g	l	τ	$welfare$
τ_l only	Pareto	1.000	2.959	0.509	0.254	2.005	0.350	—	-
	Ramsey	0.700	2.959	0.509	0.254	2.005	0.245	0.397	100
	Markov	0.719	2.959	0.573	0.190	3.017	0.252	0.297	99.23
	Best SSE	0.697	2.959	0.500	0.264	1.895	0.244	0.412	99.81
τ_k only	Pareto	1.000	2.959	0.509	0.254	2.005	0.350	—	—
	Ramsey	0.588	1.734	0.712	0.149	4.779	0.278	0.673	100
	Markov	0.478	1.149	0.688	0.220	3.123	0.285	0.821	81.04
	Best SSE	0.581	1.693	0.710	0.154	4.611	0.279	0.686	96.27

Welfare gain from looking at SSE instead of Markov perfect equilibrium will be 0.5% measured by CEV when there is only labor income tax available. The number will increase to 15% when the government can only tax capital income.

We also consider a case where the government can use both capital and labor income taxation. We then build up the trigger strategy that supports the best equilibrium outcome. Martin (2009) considers the same economy as we do and study the Markov perfect equilibria. He provides conditions under which there exist interior Markov perfect equilibrium. It has been shown that such equilibrium does not exist in the above model with the parameter values we considered. Furthermore, he finds that there exists a corner solution Markov perfect equilibrium which presents zero labor tax. He also suggests that this equilibrium is the worst SSE. Our finding, reported in Table 5, is consistent with Martin (2009). The differences in the tax rates are due to

the fact that we impose a positive lower bound on the tax rate. We also do not limit the government to differentiable strategy.

Table 5: Best and Worst SSE, compared with Markov perfect

Govt.	k/y	c/y	g/y	c/g	l	τ_k	τ_l
Ramsey	2.959	0.509	0.254	2.005	0.245	0.000	0.397
Best SSE	2.896	0.513	0.255	2.009	0.247	0.058	0.387
Markov Perfect	1.150	0.688	0.220	3.125	0.285	0.612	0.000
Worst SSE	1.303	0.681	0.215	3.188	0.283	0.789	0.010

7.3   What matters

In equilibrium, the government's temptation to deviate to a higher capital tax and lower labor tax in the first period is deterred by a lower continuation value for the government. In the short term, the supply of capital is inelastic with respect to the after-tax rate of return. The government can raise extra revenue and improve the current period social welfare without any substantial efficiency loss. However, this tax changes will distort the household's investment decision from the following period, which yields lower payoffs to the government. When the discount factor β is low, the former effect dominates, and the government cannot resist the temptation. On the other hand, as the worst sustainable equilibrium yields lower continuation value for the government, in turn, make the deviation more painful and deter the government's temptation to raise the capital tax. In what follows, we vary the value of discount factor β to see how the optimal taxation change. We can see that the tax rate on capital increase as the value of β decreases.

Table 6: What matters - discount factor?

β	τ_k	τ_l
0.96	4.5%	34.5%
0.86	32.2%	29.8%
0.70	34.5%	27.2%
0.00	92.4%	0.00%

We also examine how the labor supply elasticity affect the sustainability of the Ramsey allocation. In what follows, we use a preference that allows for variation in the labor supply elasticity: $u(c, l) = \log c - \gamma \frac{l^{1+\eta}}{1+\eta}$. We keep the value of β constant while

adjust the value of η. The table below suggests that the steady state capital income tax increases with the labor supply elasticity. The incentive for the government to raise capital tax comes from the benefit of reducing the labor supply distortion. If the labor supply response inertly to the reduced tax rate, then it becomes less attractive for the government to deviate. We can conclude from these two experiment that Ramsey policy with zero capital income tax is sustainable when the discounting is sufficiently low and the labor supply is sufficiently inelastic.

Table 7: What matters - labor supply elasticity?

η	τ_k	τ_l
0.67	5.5%	37.3%
1.00	7.2%	36.5%
2.00	7.7%	36.8%
3.13	12.0%	35.4%

8 Concluding Remarks

In this paper, we have extended the analysis of Phelan and Stacchetti (2001) into a stochastic setting. By looking at the best sustainable equilibrium, we have found that the welfare loss of no commitment, compared with Ramsey allocation, is very small when we use a set of parameter values compatible with the standard macro literature. The average tax rate on the capital income is quite close to zero and pro-cyclical, while the labor income tax assumes most of the tax burden and present counter-cyclical movement. To support the best outcome, we need a trigger strategy that presents an average capital income tax rate of 40%, and an average labor income tax rate of 17.5%, which result a welfare loss of 7% in consumption equivalence. This seems to be plausible.

We also compare the best SSE outcome with Markov-perfect equilibrium. Numerical simulations indicate that the welfare loss can be as large as 5.6% when we limit ourselves to this subset of the equilibria.

It should be noted that our results hold for the class of economies with certain preferences and a set of parameter values due to the difficulties in proving a key result for general case. However, we argue that the specification of preferences is compatible with the existence of a balanced growth path. More important, it satisfies most important properties related to uniform commodity taxation [see Chari and Kehoe (1999)].

In this paper, we abstract from several important features of the data. First, for computational reasons, we maintain the assumption that the government does not have access to debt issuance. Domínguez (2010) studies the time consistent government debt while assuming away capital and all uncertainty. She found that an additional limit on the government's budget deficit can be used to support better equilibria. It remains an open question how the introduction of government debt can affect the property of optimal taxation. To understand what the optimal fiscal policies look like when the government can issue debt but cannot commit the policy beyond the current period will help us to understand important policy questions, such as should the government increase tax rate or lift the debt limit when the economy is on the blink of default. This is particularly relevant given the current fiscal woes and sovereign debt crisis that many countries are facing. Second we abstract from any heterogeneity of agents. Different types of agents may raise interesting political economy issues that deserve careful study. Third the numerical method developed here face some computational challenges, even with the help of high-performance, parallel computing. It will certainly improve the efficiency of the algorithm through a better understanding of the construction of the equilibrium payoff set [c.f. Abreu and Sannikov (2011)]. We leave these questions for further research.

References

Abreu, D., D. Pearce and E. Stacchetti (1990). "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring." Econometrica, 58, 1041-1063

Abreu, D., Y. Sannikov (2011). "An Algorithm for Two Player Repeated Games with Perfect Monitoring." Manuscript.

Athey, S., A. Atkenson and P. J. Kehoe (2005). "The Optimal Degree of Discretion in Monetary Policy." Econometrica, 73, 1431-1475.

Atkinson, A. B. and J. E. Stiglitz (1976). "The Design of Tax Structure: Direct vs. Indirect Taxation." Journal of Public Economics 6, 55-75.

Benhabib, J . and A. Rusticchini (1997). "Optimal Taxes without Commitment." Journal of Economic Theory 77 (1997), 231-59.

Chamley, C. (1986). "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives." Econometrica, 54, 607-622.

Chang, R. (1998). "Credible Monetary Policy in an Infinite Horizon Model: Recursive Approaches." Journal of Economic Theory, 81, 431-61.

Chari, V. V. (1988). "Time Consistency and Optimal Policy Design." Quarterly Review of Federal Reserve Bank of Minneapolis, 12, 4, 17-31.

Chari, V. V., L. J.Christiano and P. J. Kehoe (1994). "Optimal Fiscal Policy in a Business Cycle Model." Journal of Political Economy, 102, 617-652.

Chari, V. V., L. J. Christiano and P. J. Kehoe (1995). "Policy Analysis in Business Cycle Models." in T. F Cooley, ed., Frontiers of Business Cycle Research, Chapter 12, Princeton University Press.

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

Chari, V. V. and P. J. Kehoe (1999). "Optimal Fiscal and Monetary Policy." Federal Reserve Bank of Minneapolis Staff Report, 251.

Conesa, J. C. and B. Domínguez (2010). "Social Security As a Commitment Device." Manuscript.

Conesa, J. C., S. Kitao, and D. Krueger (2009). "Taxing Capital? Not a Bad Idea after All!." American Economic Review, 99, 25-48.

Correia, I. H. (1996). "Should Capital Income be Taxed in the Steady State?." Journal of Public Economics, 60, 147-151.

Domínguez, B. (2007). "Public Debt and Optimal Taxes without Commitment." Journal of Economic Theory, 135, 159-170, July 2007.

Domínguez, B. (2010). "The Time-Consistency of Government Debt and Institutional Restrictions on the Level of Debt." Manuscript.

Feng, Z., J. Miao, A. Peralta-Alva and M. Santos (2011). "Computing Nonoptimal Dynamic Competitive Equilibria." Manuscript.

Fernández-Villaverde, J., A. Tsyvinski (2002), "Optimal Fiscal Policy in a Business Cycle Model without Commitment." Manuscript.

Judd, K., (1985). "Redistributive Taxation in a Simple Perfect Foresight Model." Journal of Public Economics 28, 59-83.

Judd, K., (1993). "Optimal Taxation in Dynamic Stochastic Economies: Theory and Evidence." Hoover Institute, Manuscript.

Judd, K., S. Yeltekin and J. Conklin (2003). "Computing Supergame Equilibria." Econometrica, 71, 1239-1254.

Klein, P., P. Krusell and J.-V. Ríos-Rull (2003). "Time-Consistent Optimal Fiscal Policy." International Economic Review, Vol. 44, No. 4, pp. 1217-1245.

Klein, P., P. Krusell and J.-V. Ríos-Rull (2008). "Time-Consistent Public Policy." Review of Economic Studies, 75, 789-808.

Kydland, F. E. and E. C. Prescott (1977). "Rules rather than Discretion: The Inconsistency of Optimal Plans." Journal of Political Economy, 85, 473-491.

Kydland, F. E. and E. C. Prescott (1980). "Dynamic Optimal Taxation, Rational Expectations and Optimal Control." Journal of Economic Dynamics and Control, 2, 79-91.

Lucas, R., E. J. and N. L. Stocky (1983). "Optimal Fiscal and Monetary Policy in an Economy without Capital." Journal of Monetary Economics 12, 55-93.

Marcet, A. and R. Marimon (2011). "Recursive Contracts." manuscript, Universitat Pomepeu Fabra.

Martin, F. M. (2009). "Markov-perfect capital and labor taxes." Journal of Economic Dynamics and Control, 34, Issue 3, 503-521.

McGrattan, E. R. and E. C. Prescott (2000). "Is the stock market overvalued?." Quarterly Review Federal Reserve Bank of Minneapolis, issue Fall, pages 20-40.

Ortigueira, S. (2006). "Markov-Perfect Optimal Taxation." Review of Economic Dynamics, January 2006, Vol. 9, pages 153-178.

Persson, M., T. Persson and L. Svensson (1987). "Time Consistency of Fiscal and Monetary Policy." Econometrica, 55, 1419-31.

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

Prescott, E. C. (1986). "Theory ahead of Business-Bycle Measurement." Carnegie-Rochester Conference Series on Public Policy, Elsevier, vol. 25(1), pages 11-44, January.

Romer, C. D. and D. H. Romer (2010). "The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks." American Economic Review, vol. 100(3), pages 763-801, June.

Sleet, C. and S. Yeltekin (2002). "On the Computation of Value Correspondence." Working paper, Hoover Institution.

Stockman, D. (2001). "Balance-Budget Rules: Welfare Loss and Optimal Policies." Review of Economic Dynamics 4 (2001), 438-59.

Zhu, X. (1992). "Optimal Fiscal Policy in a Stochastic Growth Model," Journal of Economic Theory 58 (1992), 250-89.

9 Appendix

9.1 Proofs

Proof of Proposition 1:

It is difficult to prove that this proposition holds in general cases. Here we show that the transversality condition (14) holds for the following pair of preference and production functions:$u(c,l) = \frac{c^{1-\sigma}}{1-\sigma} + \nu(l)$, $F(k,l) = Ak^\alpha l^{1-\alpha}$, where $\nu : [0,1) \to \mathbf{R}$ is a decreasing, concave function, with $\lim_{l\to 1} \nu(l) = -\infty$. We also abstract from all uncertainties to simplify the exposition. If we plug the definition of $m(s^{t+1})$ into the first order condition in terms of $k(s^t)$ of the recursive problem, we get the exactly same equation (10). Next we show that the transversality condition holds.

First, there exist $\underline{k} > 0$ and $\bar{k} > 0$ such that for all $k_0 \in [\underline{k}, \bar{k}]$, and feasible tax policy τ, each competitive equilibrium of $\Upsilon\{k_0, \tau\}$ satisfies $k_{t+1} \in [\underline{k}, \bar{k}]$ for all $t \geq 0$ by applying an argument similar to Lemma 2 in Phelan and Stacchetti (2001). Second, we show that there exists $\bar{m} < \infty$ such that $m_t \leq \bar{m}$. Consider the situation where the capital stock of the economy hit the lower bound $\underline{k}$ and the government levies maximum tax rates $\bar{\tau} < 1$ on capital income and labor income. The Inada condition $\lim_{c\to 0} u_c(c, l) = \infty$ and $\lim_{l\to 0} u_l(0, l) = 0$ imply that the household will be better off by spending a strictly positive amount of time $\underline{l} > 0$ in working so that they can obtain some income to finance a positive consumption. The household's income will be $(1-\bar{\tau}_k)(F_k(\underline{k}, \underline{l})-\delta)\underline{k} + (1-\bar{\tau}_l)F_l(\underline{k}, \underline{l})\underline{l} > 0$. The first order condition (11) implies that the optimal consumption will be $\left[\frac{(1-\bar{\tau}_l)F_l(\underline{k},\underline{l})}{\nu_l(\underline{l})}\right]^{\frac{1}{\sigma}} > 0$, which yields the lower bound for c_t. Therefore m_t is bounded by $\bar{m} = [1 + (1 - \bar{\tau}_k)(F_k(\underline{k}, \underline{l}) - \delta)]\underline{c}^{-\sigma} = [1 + (1 - \bar{\tau}_k)(F_k(\underline{k}, \underline{l}) - \delta)]\frac{\nu_l(\underline{l})}{(1-\bar{\tau}_l)F_l(\underline{k},\underline{l})} > 0$. Finally, we have $\lim_{t\to\infty} \beta^t m_t k_t \leq \lim_{t\to\infty} \beta^t \bar{m}\bar{k} = 0$.
∎

Proof of Proposition 2:

By definition, $\bar{h}(k, s, m)$ is the maximum value of h at given (k, s, m). Therefore,

$$\bar{h}(k, s, m) = \max_{\tau^k,\tau^n} \max_{m_+,h_+} u(c, l) + G(g) + \beta\mathbb{E}h(k_+, s_+m_+)$$

$$= \max_{\tau^k,\tau^n} u(c, l) + G(g) + \max_{\tau^k,\tau^n} \max_{m_+,h_+} \beta\mathbb{E}h(k_+, s_+m_+)$$

$$= \max_{\tau^k,\tau^n} u(c, l) + G(g) + \beta\mathbb{E}\bar{h}(k_+, , s_+m_+).$$

where the first equality follows from the definition of $\bar{h}(k, s, m)$, the second equality follows from the fact that there exists at most one pair of (c, l, k_+) consistent with $\{(\tau^k, \tau^n), m\}$ at given (k, s). The last equality uses the definition of $\bar{h}$.

A similar argument applies to $\underline{h}(k, s, m)$. A few comments go as follow. First, $\underline{h}(k, s, m) = \max_{\tau_k, \tau_l} \min_{m_+, h_+} u(c, l) + G(g) + \beta \mathbb{E} h_+$. Secondly, it should be noted that the value of $u(c, l) + G(g) + \beta \mathbb{E} \underline{h}(k_+, s_+, m_+)$ at given $\{\tau_k, \tau_l\}$ may be smaller than $\tilde{h}(k, s)$, which says that the incentive constraint is not satisfied when the government has the worst continuation value. When this happens, the government need a higher continuation value so that the incentive constraint is satisfied. However, the corresponding payoff for the present government cannot be higher than $\tilde{h}(k, s)$. This is because only the *minimization* operates when $\{\tau_k, \tau_l\}$ is given. There always exists $h_+ \in \left[\underline{h}(k_+, s_+, m_+), \bar{h}(k_+, s_+, m_+)\right]$ to bind the incentive constraint when the worst continuation value breaks the incentive constraint. Otherwise m should not belong to the equilibrium set.

$\tilde{h}(k)$ is the payoff of the worst, it must be in the lower boundary $\underline{h}(k, m)$. Because it is the worst of all, it must be equal to $\min_m \underline{h}(k, m)$.

■

Proof of Theorem 8:

First we show that the sequence of $\left\{\hat{\mathbf{W}}_n\right\}$ is decreasing and $\hat{\mathbf{W}}_n \supseteq \hat{\mathbf{W}}_{n+1}$. We claim that the upper boundary of $\left\{\hat{\mathbf{W}}_n\right\}$ is decreasing. This is because $\bar{h}^1(k, s, m)$ is defined as $\max_{\tau_k, \tau_l} u(c, l) + G(g) + \beta \mathbb{E} \bar{h}^0(k_+, s_+, m_+)$ such that $\psi = (\tau_k, \tau_l, c, l, k_+, g, w, r, \{m_+, h_+\})$ is admissible with respect to $\hat{\mathbf{W}}_0$ at (k, s). The admissibility of the vector ψ implies that $\left(m, \bar{h}^1(k, s, m)\right) \in \hat{\mathbf{W}}_0(k, s)$. Therefore $\bar{h}^1(k, s, m) \leq \max\left\{h \mid (m, h) \in \hat{\mathbf{W}}_0(k, s)\right\} = \bar{h}^0(k, s, m)$. Similarly we have $\underline{h}^1(k, s, m) \geq \underline{h}^0(k, s, m)$. The same argument holds for $\hat{\mathbf{W}}_n(k, s)$.

Since the sequence is decreasing, it has a limit $\hat{\mathbf{W}}_\infty$. Proposition 2 implies that $\mathbb{F}(\mathbf{V}) = \mathbf{V}$. By a simple limit argument we have $\lim_{n \to \infty} \hat{\mathbf{W}}_\infty = \mathbf{V}$.

■

Proof of Proposition 3:

We prove this proposition by construction. For an arbitrary initial condition (k_0, s_0), the game $\Gamma(k_0, s_0)$ has strategy profile which yields a period zero payoff of $\tilde{h}(k_0, s_0)$ defined as

$$\tilde{h}(k_0, s_0) = \max_{\tau_{k,0}, \tau_{l,0}} \left\{\min_{c_0, l_0, k_1, (m_1, h_1) \in \mathbf{W}} \left[u(c_0, l_0) + G(g_0) + \beta \mathbb{E} h_1\right]\right\}$$

34

such that

$$[c_0, l_0, k_1, g_0, w_0, r_0] \in \mathbf{CE}^S \{(k_0, s_0, \tau_{k,0}, \tau_{l,0}, \{m_1\}) \text{ for all } s_1 \succ s_0\}.$$

Therefore, we have the period zero tax rates and households' choice, namely $\{\tau_{k,0}, \tau_{l,0}, c_0, l_0, k_1\}$. By the definition of $\tilde{h}(k_0, s_0)$, $h_1 = \tilde{h}(k_1, s_1)$, otherwise we have $\tilde{h}'(k_0, s_0) < \tilde{h}(k_0, s_0)$. This implies that we can find the period one tax rates and households' choices as we did for the initial period. Notice that there exist $\underline{k} > 0$ and $\bar{k} > 0$ such that for all $k_0 \in [\underline{k}, \bar{k}]$, and feasible tax policy τ, each competitive equilibrium of $\Upsilon\{k_0, \tau\}$ satisfies $k_{t+1} \in [\underline{k}, \bar{k}]$ for all $t \geq 0$. Hence, we can find a time invariant mapping from (k, s) to $\{\tau_k, \tau_l, c, l, k_+\}$ such that for $k \in [\underline{k}, \bar{k}]$, we have $k_+ \in [\underline{k}, \bar{k}]$. ∎

9.2 Definition of operator $\mathbb{F}$

Definition 9 *For any convex-valued set* $\hat{\mathbf{W}} = \{(m, h) | h \in [\underline{h}^0(k, s, m), \bar{h}^0(k, s, m)]\}$, *we define operator* $\mathbb{F}$ *as follows:*

$$\mathbb{F}(\hat{\mathbf{W}})(k, s) = \{(m, h) | h \in [\underline{h}^1(k, s, m), \bar{h}^1(k, s, m)]\}$$

where

$$\bar{h}^1(k, s, m) = \max_{\tau_k, \tau_l} u(c, l) + G(g) + \beta \mathbb{E}\bar{h}^0(k_+, s_+, m_+) \tag{46}$$

$$\underline{h}^1(k, s, m) = \max \left\{ \max_{\tau_k, \tau_l} u(c, l) + G(g) + \beta \mathbb{E}\underline{h}^0(k_+, s_+, m_+), \tilde{h}(k, s) \right\} \tag{47}$$

$$\tilde{h}(k, s) = \max_{\tau_k, \tau_l} \left\{ \min_{c, l, k_+ \{m_+\}} u(c, l) + G(g) + \beta \mathbb{E}\underline{h}^0(k_+, s_+, m_+) \right\} \tag{48}$$

such that the vector $(\tau_k, \tau_l, c, l, k_+, g, w, r, \{m_+, h_+\})$ *is admissible with respect to* $\hat{\mathbf{W}}$ *at* (k, s). *We also define* $\underline{h}(k, s, m) = -\infty, \bar{h}(k, s, m) = +\infty$ *if there is no such vector exists.*

The equation (47) is slightly different from (36). This outer maximum is used to take care of the case where the incentive constraint always violates given (k, s). This may happen when the initial value of $\underline{h}^0$ is sufficient low.

9.3 Numerical implementation of operator $\mathbb{F}$

Below we briefly explain our algorithm and we refer to Feng, et. al. (2011) for details in terms of approximating compact-valued set. The numerical method proceeds as follows.

Let $\mathbf{K} \times \mathbf{S} \times \mathbf{M} \times \mathbf{H}$ denotes the space of all equilibrium state vectors (k, s, m, h). First we define a grid $\hat{\mathbf{K}} = \{k^{i_1}\}_{i_1=1}^{N_k}$. We assume that $\mathbf{S}$ is finite and $\mathbf{S} = \{s^{i_2}\}_{i_2=1}^{N_s}$. After this discretization, instead of a correspondence $\mathbf{W} : \mathbf{K} \times \mathbf{S} \to \mathbf{M} \times \mathbf{H}$ we have $\hat{\mathbf{W}} : \hat{\mathbf{K}} \times \mathbf{S} \to \mathbf{M} \times \mathbf{H}$. It is equivalent to think about this correspondence as $N_k \times N_s$ set $\hat{\mathbf{W}}(k^{i_1}, s^{i_2})$, where $k^{i_1} \in \hat{\mathbf{K}}$, and $s^{i_2} \in \mathbf{S}$. Notice, $\hat{\mathbf{W}}$ approximates $\mathbf{W}$ well as N_k goes to ∞.

Our algorithm start with an initial guess $\mathbf{W}^0(k, s) = \{(m, h(k, s, m))\}$ and a per-determined tolerance $\epsilon > 0$.

- Step 1-1: given (k, s), pick $(m, h) \in \mathbf{W}^0(k, s)$. We store the pair of (m, h) in $\Omega(k, s)$ if there exists $(\tau_k, \tau_l) \in T$ and $(m_+, h_+) \in \mathbf{W}^0(k_+, s_+)$ such that

$$h = u(c, l) + G(g) + \beta \mathbb{E} h_+ \tag{49}$$

$$u_c(c, l) - \beta \mathbb{E} m_+ = 0 \tag{50}$$

  where (c, l, k_+) are determined as solutions to the following equations

$$m - u_c(c, l) \cdot [1 + (1 - \tau_k)r] = 0 \tag{51}$$
$$u_l(c, l) - (1 - \tau_l)w \cdot u_c(c, l) = 0 \tag{52}$$
$$(1 - \tau_l)wl + (1 - \tau_k)rk + k - (c + k_+) = 0. \tag{53}$$

- Step 1-2: given (k, s), and $\Omega(k, s)$, denote $\Omega^m := \{m | (m, h) \in \Omega(k, s)\}$, we define

$$\bar{h}^1(k, s, m) = \max_{\tau_k, \tau_l} \max_{(m_+, h_+) \in \mathbf{W}^0} u(c, l) + G(g) + \beta \mathbb{E} \bar{h}^0(k_+, s_+, m_+) \tag{54}$$

$$\underline{h}^1(k, s, m) = \max_{\tau_k, \tau_l} \min_{(m_+, h_+) \in \mathbf{W}^0} u(c, l) + G(g) + \beta \mathbb{E} \underline{h}^0(k_+, s_+, m_+), \tilde{h}^0(k, s) \tag{55}$$

for all $m \in \Omega^m$. Otherwise we set

$$\bar{h}^1(k, s, m) = +\infty \tag{56}$$

$$\underline{h}^1(k, s, m) = -\infty \tag{57}$$

and finally,

$$\tilde{h}^1(k, s) = \min_{m \in \Omega^m} \underline{h}^1(k, s, m). \tag{58}$$

- Step 2: we define $\mathbf{W}^1(k, s) = \{(m, h) | m \in \mathbf{W}_m^1(k, s), h \in \left[\underline{h}^1(k, s, m), \bar{h}^1(k, s, m) \right] \}$.

- Step 3: we set $\mathbf{W}^* = \mathbf{W}^1$ if $\| \mathbf{W}^1 - \mathbf{W}^0 \| < \epsilon$, otherwise, we set $\mathbf{W}^0 = \mathbf{W}^1$ and restart from step 1.
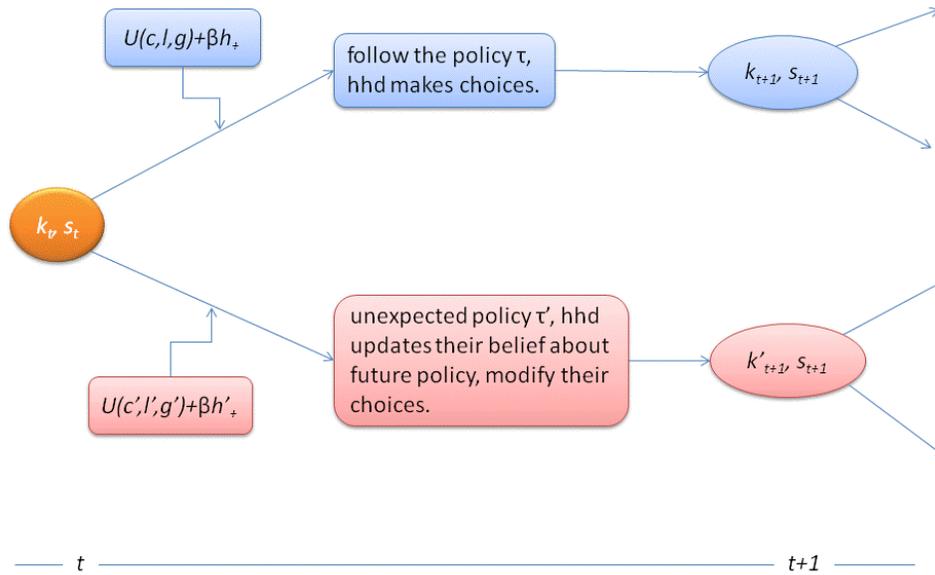
9.4 Figures

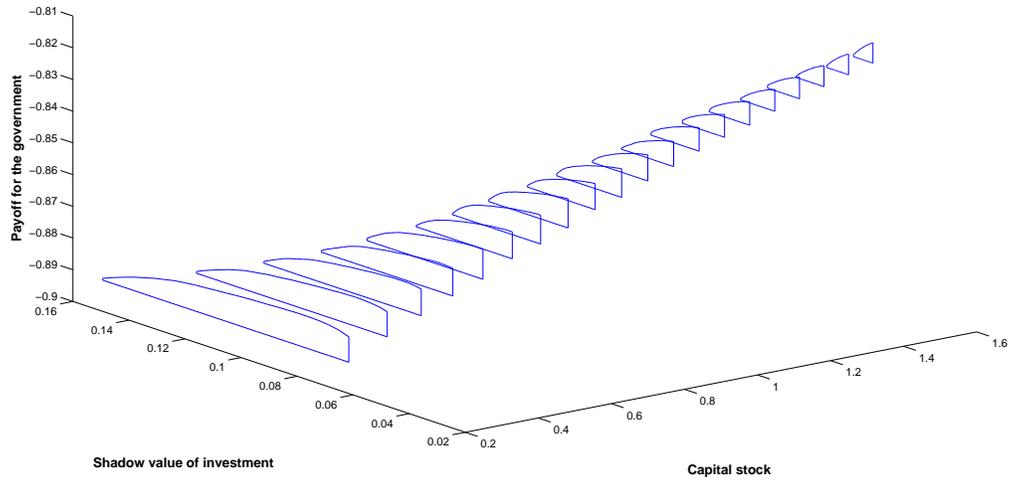Figure 2: Timing of the game between the government and households

Figure 3: The equilibrium value correspondence for deterministic economy