Time-Consistent Taxation with Heterogeneous Agents

Jang, Youngsoo

March 2023

Time-Consistent Taxation with Heterogeneous Agents*

Youngsoo Jang†

March 2023

Abstract

I study optimal taxes and transfers along transition paths in an incomplete markets model with uninsurable risk, wherein individuals play a dynamic game with successive governments that lack the ability to commit to future policies. I characterize and solve for Markov-perfect equilibria in this dynamic game. I find that the government balances two types of externalities: income redistribution externalities through transfers and pecuniary externalities caused by changes in the factor composition of income. Commitment affects how the government balances the two types of externalities along the transition path. Quantitative analysis with a calibrated economy shows that the government with commitment substantially increases taxes and transfers early in the transition and maintains them thereafter. By doing so, the government attains front-loaded positive externalities from reduced income inequality and favorable factor price changes for low-income individuals while placing negative externalities from stagnant income redistribution and unfavorable factor price changes for low-income individuals in the long run. Without commitment, this equilibrium is not credible because the government disregards the upfront welfare gains and balances the two types of externalities in a forward-looking manner in each period.

JEL classification: E61, H21.

Keywords: Markov-Perfect Policy, Income Taxation, Incomplete Markets, Heterogeneous Agents

---

*I have benefited from helpful comments by Adrien Auclert, Juan Carlos Conesa, Begoña Domínguez, Timothy Kam, Tatyana Koreshkova, Dirk Krueger, Burhan Kuruscu, Virgiliu Midrigan, José-Víctor Ríos-Rull, Gianluca Violante, Minchul Yum, Kai Zhao, the seminar participants at the Australian National University, the University of Queensland, the University of Connecticut, Korea University, Korea Labor Institute, the University of Melbourne, and the session participants at the VAMS workshop in Sydney and the Jinan IESR Quantitative Macro Workshop. Discussions with Ji-Woong Moon were constructive at all stages of this project. I also thank Yunho Cho, Pantelis Kazakis, Qian Li, and Hoonsuk Park for their generous help. All remaining errors are mine.

†School of Economics, University of Queensland, Colin Clark Building (#39), Blair Dr, St Lucia, QLD 4072, Australia. E-mail: youngsoo.jang@uq.edu.au

1 Introduction

Government policies lack commitment because political procedures sequentially determine policy execution. Previous research has analyzed the strategic behavior of successive governments and found that this lack of commitment can result in significant differences in policy implications (Kydland and Prescott, 1977; Calvo, 1978; Barro and Gordon, 1983; Lucas and Stokey, 1983; Chari and Kehoe, 1990; Athey, Atkeson and Kehoe, 2005; Klein, Krusell and Ríos-Rull, 2008). Despite many government policies addressing redistribution and its welfare effects, relatively few studies have investigated the distributional implications of government commitment. This paper fills this gap by examining a dynamic game between individuals and successive governments in the standard incomplete markets model with uninsurable income risk (the Bewley-Huggett-Aiyagari model). In this context, I investigate how optimal taxes and transfers along transition paths differ according to government commitment technologies.

Specifically, I compare two economies: one with the time-inconsistent optimal income taxes and transfers with a Ramsey planner (with commitment) and one with the time-consistent optimal taxes and transfers (without commitment). In the economy with the time-inconsistent optimal policy, because the government can commit to all future tax and transfer policies, it chooses a sequence of income taxes that maximizes the utilitarian welfare function along the transition path. By contrast, in the time-consistent optimal case, the government can only set a tax rate and transfers for the next period and cannot commit to them thereafter. Thus, the government sequentially chooses a tax/transfer policy that maximizes the utilitarian welfare function under this commitment constraint, and this action continues perpetually. To concentrate on the role of government commitment rather than various government fiscal policy instruments, I assume a simple government financing rule in the benchmark case. The government balances its budget in each period, levying a flat tax from labor and capital income and redistributing its revenue to households through lump-sum transfers after covering a predetermined level of government spending.[1]

This paper, to my knowledge, makes the first contribution to the literature by theoretically characterizing the equilibria of this dynamic game with incomplete financial markets. Using the generalized Euler equation (GEE) approach, as proposed by Klein et al. (2008), I characterize the Markov-perfect equilibria (MPEs) by computing the first-order condition (FOC) of the government's policy decision. The analysis reveals that the government does not consider the direct impact of individual decisions on consumption, saving, and labor supply on welfare, as these terms are offset in the government's optimal condition through the envelope theorem. Instead, the government focuses on two types of externalities while setting taxes: (i) income redistribution externalities through changes in transfers and (ii) pecuniary externalities due to changes in the factor composition of individual income. For instance, when the income tax rate increases, it raises

[1] The assumption on the tax base is relaxed later.

the after-tax income for low-income individuals but lowers it for high-income individuals. This results in positive externalities for low-income individuals by improving their welfare but negative externalities for high-income individuals by worsening their welfare. Meanwhile, this tax increase reduces overall savings and aggregate capital, resulting in an increase in the equilibrium interest rate and a decrease in the equilibrium market wage due to general equilibrium effects—pecuniary externalities. These changes impact individual welfare differently, with low-income individuals, who have a higher proportion of labor income, being worse off (negative externalities), while high-income individuals, who have a higher proportion of capital income, are better off (positive externalities). The government must account for these two types of externalities and their differing impacts on individuals while making its policy decisions.

Commitment technologies determine how to strike a balance between the two types of externalities along the equilibrium path. With commitment, the Ramsey planner balances the two types of externalities over the entire time horizon, resulting in time-inconsistent outcomes. When making changes in tax rates and transfers at a specific point in time, the government with commitment takes into account the effects not only on the current and future economy, but also on the past. For instance, if the government has already achieved significant income redistribution through high taxes and transfers up until a certain point, and then decides to change the taxes and transfers, the Ramsey planner will consider the impact of past income redistribution on welfare. This policy, which is desired at time 0, is not optimal when evaluated in a forward-looking manner. On the other hand, without commitment, the government implements time-consistent policies by balancing the two types of externalities in a forward-looking manner at each period, without considering the welfare impacts from the past income redistribution. This leads to suboptimal outcomes over the entire time horizon, as the government only takes into account the effects of its policies on the current and future economy at each point in time.

The second contribution of the paper is to provide a numerical solution algorithm for the dynamic game in the incomplete markets model. Note that although I have used this method to quantify my theoretical findings, it is not limited to the examples presented in this paper. It can be applied to other general games with incomplete financial markets, including economies with general optimal policies or political procedures. The general structure of this game, as Krusell et al. (1996); Krusell and Ríos-Rull (1999) show, is complex because it involves the interplay of three equilibrium objects: individual decisions, the aggregate law of motion for the household distribution, and the endogenous government policy function, all of which must be consistent with each other. This becomes more challenging with one-shot deviations required to solve sub-perfect Nash equilibria because the deviations produce a sequence of taxes/transfers that are off-the-equilibrium paths—alternative paths of the economy that are not selected in the equilibrium but still need to be defined to shape the equilibrium. To address these computational issues, the paper draws on the backward induction method of Reiter (2010) and makes modifications to suit the characteristics of the MPE, considering the existence of off-the-equilibrium paths, while still retaining its computational benefits.[2] I use this numerical method for quantitative analysis. I calibrate the model to the U.S. economy and take it as the starting point during its transition. I solve the Ramsey problem (time-inconsistent policy) using a numerical approach from Dyrda and Pedroni (2022) and the time-consistent case using my own numerical solution method. The results of both cases are then compared along the transition path.

In quantitative analysis, I find that commitment makes substantial differences in the aggregate economy, inequality, and welfare. The Ramsey planner, with its time-inconsistent optimal policy, chooses more substantial income taxes than the government with the time-consistent optimal policy over the entire transition path. Compared to the calibrated initial economy, the Ramsey planner gradually increases income taxes by 16 percentage points. However, with a lack of commitment, the optimal income tax rate rapidly increases by 2 percentage points. This gap in tax policies results in differences in the size of transfers. The time-inconsistent optimal income tax economy generates larger transfers than the time-consistent optimal income tax economy. The ratio of transfers-to-initial GDP in the case with commitment increases by 9.2 percentage points, but that in the case without commitment increases by 4.7 percentage points. These differences in taxes and transfers result in diverse dynamics in the economy and distributions. The economy managed by the Ramsey planner is less efficient but more equal due to the imposition of higher income taxes and larger transfers. Aggregate consumption, capital, and output are higher in the time-consistent optimal policy scenario, but their inequality is lower in the time-inconsistent scenario. The welfare gain, as measured by the consumption equivalent variation, is greater in the time-inconsistent optimal income tax scenario (+2.19%) than in the time-consistent scenario (+0.57%).

To shed light on the economic reasoning behind differences in government policy decisions, I apply my theoretical findings to the quantitative results. Through this analysis, I find that the Ramsey planner, with its commitment, achieves front-loaded positive externalities through both reduced income inequality and changes in the factor composition of income. To understand this result, note that the government is inclined to represent the interests of low-income individuals more, as the externalities are weighted by the marginal utility of consumption. On the one hand, the committed scenario, which results in a more substantial increase in taxes and transfers, rapidly reduces after-tax income inequality in the early stages of the transition, leading to positive externalities for low-income individuals. On the other hand, this policy change induces an increase in the market wage and a decrease in the market interest rate in the early phase of the transition because the adjustment of aggregate capital is slower than that of aggregate labor. Therefore, the two types of externalities become positive in the early phase of the transition, leading the government

[2] Section 5 explains the key ideas and Appendix A demonstrates each step of the algorithm in detail.

to attain up-front welfare gains.

In the long run, the government with commitment bears welfare losses from mitigated income redistribution externalities and negative pecuniary externalities caused by changes in the factor composition of income unfavorable to low-income individuals. The higher taxes and transfers in the case with commitment lead to a higher market interest rate and a lower market wage in the long run, due to a lower capital-to-labor ratio. These price changes act as negative externalities for low-income individuals whose income composition is biased toward labor. Additionally, since transfers are already substantial, welfare gains from reducing income inequality (i.e., income redistribution externalities) become insignificant. Therefore, the theoretical findings suggest that the Ramsey planner balances these two types of externalities over the entire time horizon by allocating upfront positive externalities from reduced income inequality and factor price changes favorable to low-income individuals while putting negative externalities coming from unfavorable factor price changes for low-income individuals and stagnant income redistribution later on. The significant welfare gain (+2.19%) implies that this method of balancing throughout the entire time horizon is effective when commitment is available.

Without commitment, the previously discussed strategy is not credible. Suppose that the government without commitment finds itself in the long-run equilibrium of the committed scenario. Then, the government disregards the upfront welfare gains and balances the two types of externalities in a forward-looking manner in each period. In other words, the government perceives this economy as one where negative pecuniary externalities from changes in the factor composition of income outweigh mitigated externalities from income redistribution through transfers because income redistribution becomes stagnant. As a result, the government decides to reduce taxes and transfers in the next period. Households will rationally anticipate this decision and increase their savings and labor supply, which increases the correlation between individual labor productivity and labor supply and savings, leading to a deviation from the long-run equilibrium in the committed scenario. The quantitative results demonstrate that this way of making policy decisions by the government results in smaller welfare gains (+0.57%). I conduct the same quantitative exercises, but changes the tax base. I find that the outcomes are consistent with those under proportional income taxes.[3] These findings suggest that government commitment plays a crucial role in making policy decisions, which can induce a quantitatively significant impact on the overall economy, inequality, and welfare.

Related Literature: This paper belongs to the stream of macroeconomic literature that examines

[3] Specifically, I compare economies with and without commitment when only labor income taxes vary over time. I cannot compare capital income tax cases because time-consistent capital income taxes significantly shrink the economy. As a result, the calibrated exogenous government spending cannot be covered by the tax revenue in the government budget.

the implications of time-inconsistent features for government policies, following the seminal study of Kydland and Prescott (1977). A branch of this literature examines the impact of government commitment technologies in MPE that depends on the fundamental economic state variables (Cohen and Michel, 1988; Krusell et al., 1996; Krusell and Ríos-Rull, 1999; Klein and Ríos-Rull, 2003; Klein et al., 2008; Azzimonti, 2011; Song et al., 2012; Bassetto et al., 2020; Laczó and Rossi, 2020).[4] This paper is closely aligned with this stream, relying on the concept of MPE. Like this paper, these studies with Markov-perfect policies have found that government commitment has significant impacts on the optimal design of fiscal and public policies. However, this paper distinguishes itself by examining how the interaction between government commitment and individual heterogeneity affects optimal policy design through the incomplete financial markets channel. Corbae et al. (2009) has a similarity with this paper in the sense that their paper studies the government's decisions on income taxation in response to wage inequality, by comparing a series of economies with different commitment technologies in an incomplete markets model. However, whereas their study limits its scope to quantifying long-run equilibria without offering a theoretical framework, this paper provides a theoretical characterization of the MPE and analyzes quantifiable results during the transition.

This paper is also related to macroeconomic studies on constrained efficiency in dynamic general equilibrium models with incomplete markets that focus on distributive externalities through wages and interest rates. For example, Davila, Hong, Krusell and Ríos-Rull (2012) analyzes constrained efficiency in Aiyagari's (1994) model; Dávila and Korinek (2018) examines the implications of constrained efficiency with collateral constraints; Park (2018) conducts a similar analysis with human capital; and Itskhoki and Moll (2019) characterizes the optimal distribution of development policies between workers and entrepreneurs in an economy with financial constraints. While distributive externalities play a crucial role in these studies, the mechanism differs from that in the present paper, which examines a decentralized economy with a government rather than a centralized economy with a social planner. In a centralized economy, constrained efficiency can be achieved through consideration of both dynamic allocations of consumption at the individual level and distributive externalities at the aggregate level. In this paper, the government endogenously responds to efficient individual decisions made in competitive equilibrium, considering only the externalities and ignoring the direct welfare impacts of individual decisions, which cancel out in the government's optimal condition through the envelope theorem.

The solution method in this paper is a non-negligible, independent contribution to the literature. Broadly, two types of methods are often used to solve macroeconomic models with MPE. The

[4] There is also another branch of this literature focusing on overcoming the time-inconsistent issues through the use of rules (Barro and Gordon, 1983; Rogoff, 1985; Athey et al., 2005) a richer range of policy instruments (Lucas and Stokey, 1983; Debortoli et al., 2017, 2021), and reputational equilibria (Chari and Kehoe, 1990; Domínguez, 2007a,b).

first is the Klein, Krusell and Ríos-Rull's (2008) approach, which is a solution method using the GEE. This method is accurate and efficient but not general enough to handle heterogeneous-agent models with incomplete markets. The method in this paper is a solution method applicable to heterogeneous-agent models. The other approach is the Krusell and Smith's (1998) method, which applicable to heterogeneous agent models. For example, Corbae et al. (2009) used this approach in their heterogeneous agent economy. However, this simulation-based method is computationally costly because economies without commitment would have more than one aggregate law of motion (e.g., the law of motion for the distributions and the endogenous tax policy function). This structure increases the computational burden in an exponential manner. Additionally, in some cases, the endogenous policy function can be highly non-linear, which poses a challenge for capturing it accurately using the parameterized law of motion in the Krusell and Smith's (1998) method. The approach used in this paper is an efficient non-simulation-based solution approach that captures the non-linearity in a non-parametric way as in Reiter (2010).

The remainder of this paper proceeds as follows. Section 2 presents the model and defines the equilibrium. Section 3 characterizes the MPE by using the GEE. Section 4 explains the core ideas of the numerical solution algorithm. Section 5 describes the calibration strategy. Section 6 presents the results of the policy analysis. Section 7 concludes this paper. Finally, Appendix A demonstrates the full details of the numerical solution algorithm.

2 Model

The quantitative model here builds upon the canonical model of Aiyagari (1994), incorporating wealth effects of labor supply. In this model, given a tax policy rule, heterogeneous households make decisions on consumption, savings, and labor supply at the intensive margin, as in standard incomplete markets models. A notable difference from standard models is the manner in which the tax/transfer policy is set. It is determined endogenously, according to the government's commitment technology. In an equilibrium state, the tax/transfer policy, individual decisions, and the evolution of the distribution are all consistent with one another, given the government's commitment constraint.

2.1 Environment

The model economy is populated by a continuum of infinitely lived households. Their preferences follow

$$E\left[\sum_{t=0}^{\infty}\beta^t u(c_t, 1-n_t)\right] \tag{1}$$

where c_t is consumption, $n_t \in [0,1]$ is labor supply in period t ($(1-n_t)$ refers to leisure), and β is the discount factor. Preferences are represented by

$$u(c_t, 1-n_t) = \frac{{c_t}^{1-\sigma}}{1-\sigma} + B\,\frac{(1-n_t)^{1-1/\chi}}{1-1/\chi} \tag{2}$$

where σ is the coefficient of relative risk aversion, B is the utility of leisure, and χ is the Frisch elasticity of labor supply. Note that the preferences here capture the wealth effects of labor supply. Such wealth effects are crucial for a welfare analysis, closely related to efficiency loss. An increase in transfers, for example, decreases the overall labor supply, shrinking the size of the aggregate economy and playing a role in reducing welfare.[5]

The representative firm produces output with a constant returns to scale. The firm's technology is represented by

$$Y_t = zF(K_t, N_t) = zK_t^{\theta}N_t^{1-\theta} \tag{3}$$

where z is the total factor productivity (TFP), K_t is the quantity of aggregate capital, N_t is the quantity of aggregate labor, and θ is the capital income share. Capital depreciates at the rate of δ each period.

In each period, households confront an uninsurable, idiosyncratic shock ϵ_t to their wage that follows an AR-1 process:

$$\log(\epsilon_{t+1}) = \rho_\epsilon \log(\epsilon_t) + \eta_{t+1}^{\epsilon} \tag{4}$$

where $\eta_{t+1}^{\epsilon} \sim N(0, \sigma_\epsilon^2)$. Using the method in Rouwenhorst (1995), I approximate the AR-1 process as a finite-state Markov chain with transition probabilities $\pi_{\epsilon,\epsilon'}$ from state ϵ to state ϵ'. Households earn $w_t\epsilon_t n_t$ as their labor income where w_t is the market equilibrium wage. They can self insure through assets a_t. Such households have capital income of as much as $r_t a_t$ where r_t is the equilibrium risk-free interest rate.

[5] By contrast, Corbae et al. (2009) employed the preferences in Greenwood, Hercowitz and Huffman (1988) that lack the wealth effects of labor supply to mitigate the computational burden.

The government obtains its tax revenue by levying taxes on household capital and labor income at proportional flat tax rate, τ_t.[6] Given its tax revenue, the government covers a constant government spending G, and the rest is used for lump-sum transfers $T_t \geq 0$. The government runs a balanced budget each period:

$$G + T_t = \tau_t \left[r_t K_t + w_t N_t \right] \text{ where } T_t \geq 0. \tag{5}$$

Note that T_t are always non-negative following the literature, where lump-sum taxes are not considered.

2.2 Competitive Equilibrium, Exogenous Policy

In this section, I define the competitive equilibrium, given an exogenous tax and transfer policy. I start with a setting to address time-inconsistent policies. To describe problems with commitment (the Ramsey problem), household dynamic problems need to be represented in a sequential manner. At the beginning of each period, households differ from one another in asset holdings a and labor productivity ϵ. $\mu_t(a, \epsilon)$ denotes the distribution of households in period t. Given a sequence of prices $\{r_t, w_t\}_{t=0}^{\infty}$, income taxes $\{\tau_t\}_{t=0}^{\infty}$, and lump-sum transfers $\{T_t\}_{t=0}^{\infty}$, households in period t solve

$$v_t(a, \epsilon) = \max_{c_t,\, a_{t+1},\, n_t} u(c_t(a, \epsilon), 1 - n_t(a, \epsilon)) + \beta \sum_{\epsilon_{t+1}} \pi_{\epsilon_t, \epsilon_{t+1}} v_{t+1}(a_{t+1}(a, \epsilon), \epsilon_{t+1}) \tag{6}$$

such that

$$c_t + a_{t+1} = (1 - \tau_t) w_t \epsilon_t n_t + (1 + r_t(1 - \tau_t))a + T_t.$$

Definition 2.2.1. ***Sequential Competitive Equilibrium (SCE), given a Sequence of Taxes***
Given G, an initial distribution $\mu_0(\cdot)$, and a sequence of income taxes and transfers $\{\tau_t, T_t\}_{t=0}^{\infty}$, an SCE is a sequence of prices $\{w_t, r_t\}_{t=0}^{\infty}$, a sequence of allocations $\{c_t, n_t, a_{t+1}, K_t, N_t\}_{t=0}^{\infty}$, a sequence of value functions $\{v_t(\cdot)\}_{t=0}^{\infty}$, and a sequence of distributions over the state space $\{\mu_t(\cdot)\}_{t=1}^{\infty}$, such that for all t

(i) Given $\{\tau_t, T_t\}_{t=0}^{\infty}$ and $\{w_t, r_t\}_{t=0}^{\infty}$, the decision rules $a_{t+1}(a, \epsilon)$ and $n_t(a, \epsilon)$ solve the household problem in (6), and $v_t(a, \epsilon)$ is the associated value function.

[6] In a later section, I relax this assumption.

(ii) The representative agent firm engages in competitive pricing:

$$w_t = (1-\theta) z \left(\frac{K_t}{N_t} \right)^{\theta} \tag{7}$$

$$r_t = \theta z \left(\frac{K_t}{N_t} \right)^{\theta - 1} - \delta. \tag{8}$$

(iii) The factor markets clear:

$$K_t = \int a \, \mu_t(\mathbf{d}(a \times \epsilon)) \tag{9}$$

$$N_t = \int \epsilon \, n_t(a, \epsilon) \, \mu_t(\mathbf{d}(a \times \epsilon)) \tag{10}$$

(iv) The government budget constraint (5) is satisfied.

(v) Let $\mathcal{B}(A \times E)$ denote the Borel σ-algebra on $A \times E$. For any $B \in \mathcal{B}(A \times E)$, the sequence of distributions over individual $\{\mu_t(\cdot)\}_{t=1}^{\infty}$ satisfies

$$\mu_{t+1}(B) = \int_{\{(a,\epsilon) | (a_{t+1}(a,\epsilon), \epsilon_{t+1}) \in B\}} \sum_{\epsilon_{t+1}} \pi_{\epsilon, \epsilon_{t+1}} \mu_t(\mathbf{d}(a \times \epsilon)). \tag{11}$$

In contrast, to handle problems without commitment, it is convenient to present the household dynamic problems in a recursive manner. In addition to the individual state variables a and ϵ, there are two aggregate state variables, including the distribution of households $\mu(a, \epsilon)$ over a and ϵ and income tax τ.[7] A variable with a prime symbol denotes its value in the next period.

Let $v(a, \epsilon; \mu, \tau)$ denote the value of households associated with a state of $(a, \epsilon; \mu, \tau)$. They solve

$$v(a, \epsilon; \mu, \tau) = \max_{c>0,\, a' \geq \underline{a},\, 0 \leq n \leq 1} \left[\frac{c^{1-\sigma}}{1-\sigma} + B \, \frac{(1-n)^{1-1/\chi}}{1-1/\chi} + \beta \sum_{\epsilon'} \pi_{\epsilon, \epsilon'} v(a', \epsilon'; \mu', \tau') \right] \tag{12}$$

such that

$$c + a' = (1-\tau) \, w(\mu) \, \epsilon \, n + (1 + r(\mu)(1-\tau)) \, a + T$$

$$\tau' = \Psi(\mu, \tau)$$

$$\mu' = \Gamma(\mu, \tau, \tau') = \Gamma(\mu, \tau, \Psi(\mu, \tau))$$

where $\underline{a} \leq 0$ is a borrowing limit, $\tau' = \Psi(\mu, \tau)$ is the perceived law of motion of taxes, and

[7] Note that either τ or T is a state variable because once one of them chosen, the other is fixed in the balanced government budget, $G + T = \tau[rK + wN]$. I here choose τ as a state variable.

$\mu' = \Gamma(\mu, \tau, \tau')$ is the law of motion for the distribution over households. Note that households here solve the above problem given an exogenous tax policy function $\tau' = \Psi(\mu, \tau)$.

Definition 2.2.2. ***Recursive Competitive Equilibrium (RCE), given a Law of Motion for Tax.*** Given G and $\Psi(\mu, \tau)$, an RCE is a set of prices $\{w(\mu), r(\mu)\}$, a set of decision rules for households $g^a(a, \epsilon; \mu, \tau)$ and $g^n(a, \epsilon; \mu, \tau)$, a value function $v(a, \epsilon; \mu, \tau)$, a distribution of households $\mu(a, \epsilon)$ over the state space, and the law of motion for the distribution of households $\Gamma(\mu, \tau, \Psi(\mu, \tau))$ such that

(i) Given $\{w(\mu), r(\mu)\}$, the decision rules $a' = g^a(a, \epsilon; \mu, \tau)$ and $n = g^n(a, \epsilon; \mu, \tau)$ solve the household problem in (12), and $v(a, \epsilon; \mu, \tau)$ is the associated value function.

(ii) The representative agent firm engages in competitive pricing:

$$w(\mu) = (1-\theta)z\left(\frac{K}{N}\right)^{\theta} \tag{13}$$

$$r(\mu) = \theta z\left(\frac{K}{N}\right)^{\theta-1} - \delta. \tag{14}$$

(iii) The factor markets clear:

$$K = \int a\, \mu(\mathbf{d}(a \times \epsilon)) \tag{15}$$

$$N = \int \epsilon\, g^n(a, \epsilon; \mu, \tau)\, \mu(\mathbf{d}(a \times \epsilon)) \tag{16}$$

(iv) The government budget constraint (5) is satisfied.

(v) The law of motion for the distribution of households $\mu' = \Gamma(\mu, \tau, \Psi(\mu, \tau))$ is consistent with individual decision rules and the stochastic process of ϵ.

2.3 Competitive Equilibrium, Endogenous Policy

In this section, I define competitive equilibria where the income tax is endogenously determined. I model the tax choice in two ways: the time-inconsistent optimal income tax with commitment (Ramsey problem) and the time-consistent optimal income tax without commitment. I begin with the Ramsey problem.

Definition 2.3.1. ***The Ramsey Problem:***
An SEC with the Time-inconsistent Optimal Income Tax with Commitment

Given μ_0, the government chooses $\{\tau_t\}_{t=0}^{\infty}$ such that

$$\{\tau_t\}_{t=0}^{\infty} = \underset{\{\tilde{\tau}_t\}_{t=0}^{\infty}}{\operatorname{argmax}} \int E_0 \sum_{\hat{t}=0}^{\infty} \beta^{\hat{t}} u\big(c_{\hat{t}}^*(a,\epsilon|\{\tilde{\tau}_{\hat{t}}\}_{t=0}^{\infty}), 1 - n_{\hat{t}}^*(a,\epsilon|\{\tilde{\tau}_t\}_{t=0}^{\infty})\big)\mu_0(\mathbf{d}(a\times\epsilon))$$

where $\big(c_{\hat{t}}^*(a,\epsilon|\{\tau_t\}_{t=0}^{\infty}), n_{\hat{t}}^*(a,\epsilon|\{\tau_t\}_{t=0}^{\infty})\}_{\hat{t}=0}^{\infty}\big)$ is an allocation in Definition (2.2.1) in period $\hat{t}$, given $\{\tilde{\tau}_t\}_{t=0}^{\infty}$.

Note that the consumption and labor decisions at time t, (c_t^*, n_t^*), are affected not only by the policy in this period but also by a sequence of income taxes. Therefore, the current decisions are influenced by past and future taxes, which leads to the time-inconsistency issue.

For the time-consistent case, I have employed the definition in Krusell and Ríos-Rull (1999); Klein and Ríos-Rull (2003).

Definition 2.3.2. ***An RCE with the Time-consistent Optimal Income Tax without Commitment.***

(i) A set of functions $\{w(\cdot), r(\cdot), g^a(\cdot), g^n(\cdot), v(\cdot), \Gamma(\cdot)\}$ satisfy Definition (2.2.2).

(ii) For each (μ,τ), the government chooses $\tau^{WO}(\mu,\tau)$ such that

$$\tau^{WO}(\mu,\tau) = \underset{\tilde{\tau}'}{\operatorname{argmax}} \int \hat{V}(a,\epsilon;\mu,\tau,\tilde{\tau}')\mu(\mathbf{d}(a\times\epsilon)) \tag{17}$$

where

$$\hat{V}(a,\epsilon;\mu,\tau,\tilde{\tau}') = \max_{c>0,\, a'\geq \underline{a},\, 0\leq n\leq 1} \left[\frac{c^{1-\sigma}}{1-\sigma} + B\,\frac{(1-n)^{1-1/\chi}}{1-1/\chi} + \beta\sum_{\epsilon'} \pi_{\epsilon,\epsilon'} v(a',\epsilon';\mu',\tilde{\tau}')\right]$$

such that

$$c + a' = (1-\tau)\,w(\mu)\,\epsilon\, n + (1 + r(\mu)(1-\tau))\,a + T$$

$$\tau' = \tilde{\tau}',\ \text{and thereafter } \tau'' = \Psi(\mu', \tau' = \tilde{\tau}') \tag{18}$$

$$\mu' = \Gamma(\mu,\tau,\tilde{\tau}),\ \text{and thereafter } \mu'' = \Gamma(\mu', \tilde{\tau}, \tau'' = \Psi(\mu', \tau' = \tilde{\tau}')) \tag{19}$$

(iii) $a' = \hat{g}_a(a,\epsilon;\mu,\tilde{\tau} : \tilde{\tau}')$ and $n = \hat{g}_n(a,\epsilon;\mu,\tilde{\tau} : \tilde{\tau}')$ solve (17) at prices that clear markets and satusfy the government budget constraint, and Γ is consistent with individual decisions and the stochastic process of ϵ.

(iv) For each (μ,τ), the policy outcome function satisfies $\Psi(\mu,\tau) = \tau^{WO}(\mu,\tau)$.

Note that the government's solution to the problem is consistent with the sub-perfect Nash equilibrium obtained with the one-shot deviation principle. In the economy with the optimal income tax without commitment, the government implements a time-consistent optimal policy as in

Klein and Ríos-Rull (2003); Corbae et al. (2009): a tax rate that is sequentially chosen only for the next period while maximizing its utilitarian welfare under this commitment constraint. The government cannot commit to the future tax rate from the period after the next period. Thus, if the chosen tax rate $\tilde{\tau}'$ deviates from the equilibrium tax policy function $\Psi(\cdot)$, future tax rates will follow the equilibrium tax policy function $\Psi(\cdot)$ because of the lack of commitment. (18) presents such dynamics. The law of motion for the distribution of households $\Gamma(\cdot)$ has to capture all the changes in the evolution of distributions caused by the deviation of the income tax from the equilibrium tax function, as shown in (19). In equilibrium, for each aggregate state (μ, τ), the government's choice of the tax rate, $\tau^{WO}(\mu, \tau)$, should be equal to the equilibrium tax function $\psi(\mu, \tau)$, as presented in (iv).

3 Characterization of the Equilibria

3.1 The Case without Commitment

Despite the definitions provided earlier on the government's decision-making process for the income tax, the underlying economic trade-offs that inform these decisions can be challenging to observe. This section employs the generalized-Euler equation (GEE) approach, introduced by Klein et al. (2008), to characterize the equilibria, thereby illuminating the economic trade-offs considered by the government in making decisions on tax and transfer policies.

The GEE approach provides insight into the economic forces driving the policymaker's decision through its FOC. This condition can be derived by utilizing the Benveniste-Scheinkman condition, also known as the envelope condition, to eliminate terms related to the partial derivative of the value function. In this section, I will first analyze the case without commitment and then proceed to the case with commitment. To determine the FOC of the government, I take the partial derivative of the government's value, represented by $\hat{V}$, with respect to the next period's income tax rate, represented by $\tilde{\tau}'$, in its vicinity of the equilibrium value τ':

$$
\begin{aligned}
0 =& \frac{d}{d\tilde{\tau}'}\Big|_{\tilde{\tau}'=\tau'} \int \hat{V}(a, \epsilon; \mu, \tau, \tilde{\tau}')\mu(d(a \times \epsilon)) \\
=& \int \frac{d}{d\tilde{\tau}'}\Big|_{\tilde{\tau}'=\tau'} \\
& \Big[u((1-\tau)\, w(\mu)\epsilon\, \tilde{g}^n(a, \epsilon; \mu, \tau, \tilde{\tau}') + (1 + r(\mu)(1-\tau))a + T - \tilde{g}^a(a, \epsilon; \mu, \tau, \tilde{\tau}'), 1 - \tilde{g}^n(a, \epsilon; \mu, \tau, \tilde{\tau}')) \\
& + \beta \sum_{\epsilon'} \pi_{\epsilon,\epsilon'} v(\tilde{g}^a(a, \epsilon; \mu, \tau, \tilde{\tau}'), \epsilon'; \mu' = \Gamma(\mu, \tau, \tilde{\tau}'), \tilde{\tau}') \Big] \mu(\mathbf{d}(a \times \epsilon))
\end{aligned} \tag{20}
$$

Note that the tilde over g^a and g^n means that the deviation of $\tilde{\tau}'$ from its equilibrium value τ' makes the decision rules for assets and labor supply different from those in equilibrium.

An obscure part in computing the FOC (20) is the derivative of v with respect to μ'. Let m_q denote the $q-th$ moment of μ. I assume that $Q \in \mathbb{N}$ exists such that $\{m_q\}_{q=1}^{Q}$ is a sufficient statistic of μ. This assumption allows me to replace μ with $\{m_q\}_{q=1}^{Q}$ in the value function. Then, the FOC (20) is given by:

$$
\begin{aligned}
0 = \int \Bigg[& u_c(c,1-n) \cdot \Big((1-\tau)w(m_1)\epsilon \frac{\partial \tilde{g}^n(a,\epsilon;\{m_q\}_{q=1}^{Q},\tau,\tau')}{\partial \tau'} - \frac{\partial \tilde{g}^a(a,\epsilon;\{m_q\}_{q=1}^{Q},\tau,\tau')}{\partial \tau'} \Big) \\
& - u_n(c,1-n) \cdot \frac{\partial \tilde{g}^n(a,\epsilon;\{m_q\}_{q=1}^{Q},\tau,\tau')}{\partial \tau'} \\
& + \beta \sum_{\epsilon'} \pi_{\epsilon,\epsilon'} \Bigg\{ \frac{\partial v(a',\epsilon',\{m'_q\}_{q=1}^{Q},\tau')}{\partial a'} \cdot \frac{\partial \tilde{g}^a(a,\epsilon;\{m_q\}_{q=1}^{Q},\tau,\tau')}{\partial \tau'} \\
& + \sum_{q=1}^{Q} \Big(\frac{\partial v(a',\epsilon',\{m'_q\}_{q=1}^{Q},\tau')}{\partial m'_q} \cdot \frac{dm'_q}{d\tau'} \Big) + \frac{\partial v(a',\epsilon',\{m'_q\}_{q=1}^{Q},\tau')}{\partial \tau'} \Bigg\} \Bigg] \mu(\mathbf{d}(a \times \epsilon)). \qquad (21)
\end{aligned}
$$

where $u_c(c)$ is the derivative of u in c and m_1 is the first moment of μ. Note that the first moment is sufficient to determine w and r. I will eliminate the derivative terms of the value $\frac{\partial v}{\partial a'}$, $\frac{\partial v}{\partial m'_q}$, and $\frac{\partial v}{\partial \tau'}$ by using the Benveniste-Scheinkman condition and then interpret the economic logic behind the equations. First, $\frac{\partial v}{\partial a'}$ is given by:

$$
\frac{\partial v(a,\epsilon;\{m_q\}_{q=1}^{Q},\tau)}{\partial a} = u_c(c,1-n)(1+r(m_1)(1-\tau)). \qquad (22)
$$

Similarly, with $\frac{\partial T}{\partial \tau} = rk + wN$, $\frac{\partial v}{\partial \tau}$ is given by:

$$
\begin{aligned}
\frac{\partial v(a,\epsilon;\{m_q\}_{q=1}^{Q},\tau)}{\partial \tau} = & u_c(c,1-n)\Big(w(m_1)(N - \epsilon \cdot g^n(a,\epsilon;\{m_q\}_{q=1}^{Q},\tau) + r(m_1)(K-a) \Big) \\
& + \omega(a,\epsilon,\{m_q\}_{q=1}^{Q},\tau) \cdot \frac{\partial g^a(a,\epsilon;\{m_q\}_{q=1}^{Q},\tau)}{\partial \tau} \\
& + \zeta(a,\epsilon;\{m_q\}_{q=1}^{Q},\tau) \cdot \frac{\partial g^n(a,\epsilon;\{m_q\}_{q=1}^{Q},\tau)}{\partial \tau} \\
& + \beta \sum_{\epsilon'} \pi_{\epsilon,\epsilon'} \Bigg\{ \sum_{q=1}^{Q} \Big(\frac{\partial v(a',\epsilon';\tau',\{m'_q\}_{q=1}^{Q})}{\partial m'_q} \cdot \frac{dm'_q}{d\tau} \Big) \\
& + \frac{\partial v(a',\epsilon';\tau',\{m'_q\}_{q=1}^{Q})}{\partial \tau'} \cdot \frac{\partial \Psi(\tau,\mu)}{\partial \tau} \Bigg\} \\
\text{where } \; & \zeta(a,\epsilon;\{m_q\}_{q=1}^{Q},\tau) = -\big(u_c(c,1-n) \cdot (1-\tau)w(m_1)\epsilon + u_n(c,1-n) \big) \\
& \omega(a,\epsilon;\{m_q\}_{q=1}^{Q},\tau) = -u_c(c,1-n) + \beta(1+r(m'_1)(1-\tau')) \sum_{\epsilon'} \pi_{\epsilon,\epsilon'} u_c(c',1-n').
\end{aligned}
\qquad (23)
$$

ξ and ω represent wedges in the Euler equation for consumption and the FOC for optimal leisure choice, respectively.

Note that $u_c(c, 1-n)(w(N-\epsilon\cdot g^n)+r(K-a))$ implies an individual welfare change driven by income redistribution via changes in taxes/transfers. This term is one of the key economic forces in characterizing the MPE. Let χ denote this term in the subsequent discussion:

$$\chi(a,\epsilon;\{m_q\}_{q=1}^{Q},\tau) = u_c(c,1-n)\Big(w(m_1)(N-\epsilon\cdot g^n(a,\epsilon;\{m_q\}_{q=1}^{Q},\tau)+r(m_1)(K-a)\Big). \quad (24)$$

χ represents *income redistribution externalities via transfers*. χ indicates a type of externalities because individuals do not take into account how changes in taxes and transfers affect their disposable income, which is determined by variations in the government budget constraint in competitive equilibrium. χ is also related to income redistribution via transfers because they measure the difference between before- and after-tax income, the extent of which varies across individuals. Individuals with a lower pre-tax income will have more income after receiving transfers, while those with a higher pre-tax income will have less due to heaver taxes. Thus, if an individual's effective labor ϵg^n and asset holdings a are below the average values (N and K), χ becomes a positive externality for that individual. In contrast, if an individual's effective labor and assets are above average, χ becomes a negative externality.

The next step is to eliminate $\frac{\partial v(a',\epsilon';\{m_q'\}_{q=1}^{Q},\tau')}{\partial m_q'}$. It is difficult to determine the required value of Q to obtain sufficient statistics for μ. For the purposes of this analysis, I assume that $Q=1$, which means that $m_1 = K$ is sufficient for capturing the evolution of the distribution, as in Krusell and Smith (1998). An alternative interpretation of this assumption is that the government only considers changes in future prices and not higher moments of the future distribution when determining income taxes. This assumption enables a further characterization of the MPE.

With the Benveniste-Scheinkman condition, $\frac{\partial v}{\partial K}$ is given by:

$$\begin{aligned}\frac{\partial v(a,\epsilon;K,\tau)}{\partial K} =& u_c(c,1-n)\Big((1-\tau)(f_{NK}(K)\epsilon\cdot g^n(a,\epsilon;K,\tau)+f_{KK}(K)a)+\frac{\partial T}{\partial K}\Big)\\ &+\zeta(a,\epsilon;K,\tau)\cdot\frac{g^n(a,\epsilon;K,\tau)}{\partial K}+\omega(a,\epsilon;K,\tau)\cdot\frac{g^a(a,\epsilon;K,\tau)}{\partial K}\\ &+\beta\sum_{\epsilon'}\pi_{\epsilon,\epsilon'}\Big\{\frac{\partial v(a',\epsilon_j';K',\tau')}{\partial K'}\cdot\frac{\partial\Gamma(K,\tau,\tau')}{\partial K}+\frac{\partial v(a',\epsilon';K',\tau')}{\partial\tau'}\cdot\frac{\partial\Psi(K,\tau)}{\partial K}\Big\}. \quad (25)\end{aligned}$$

Note that $u_c(c,1-n)\big((1-\tau)(f_{NK}(K)\epsilon+f_{KK}(K)a)+\frac{\partial T}{\partial K}\big)$ implies an individual welfare change driven by variations in the factor composition between capital and labor income following an increase in the current tax τ. As discussed in Davila et al. (2012), this effect differs across individuals and depends on the composition of their income. To clarify how this effect is linked to the factor composition of individual income, I proceed with further steps following Davila et al. (2012). Because f is homogeneous of degree 1, $f_{KK}(K,N)K + fKN(K,N)N = 0$. In addition, because

$T = \tau(rK + wN) - G = \tau(f_K K + f_N N) - G$, $\frac{\partial T}{\partial K} = f_K(K)\tau$. Then, with these conditions, $\frac{\partial v}{\partial K}$ is given by:

$$\begin{aligned}\frac{\partial v(a,\epsilon;K,\tau)}{\partial K} &= u_c(c,1-n)\big((1-\tau)(-\frac{\epsilon \cdot g^n(a,\epsilon;K,\tau)}{N} + \frac{a}{K})f_{KK}(K)K + f_K(K)\tau\big) \\ &+ \zeta(a,\epsilon;K,\tau)\cdot\frac{g^n(a,\epsilon;K,\tau)}{\partial K} + \omega(a,\epsilon;K,\tau)\cdot\frac{g^a(a,\epsilon;K,\tau)}{\partial K} \\ &+ \beta\sum_{\epsilon'}\pi_{\epsilon,\epsilon'}\left\{\frac{\partial v(a',\epsilon'_j;K',\tau')}{\partial K'}\cdot\frac{\partial\Gamma(K,\tau,\tau')}{\partial K} + \frac{\partial v(a',\epsilon';K',\tau')}{\partial \tau'}\cdot\frac{\partial\Psi(K,\tau)}{\partial K}\right\} \end{aligned} \tag{26}$$

The first term is another important economic force in characterizing the MPE. For further discussion, I refer to this term as ϕ:

$$\phi(a,\epsilon;K,\tau) = u_c(c,1-n)\big((1-\tau)(-\frac{\epsilon\cdot g^n(a,\epsilon;K,\tau)}{N} + \frac{a}{K})f_{KK}(K)K + f_K(K)\tau\big). \tag{27}$$

ϕ represents pecuniary externalities caused by changes in the factor composition of income. It measures the welfare change following a shift of the factor composition of income between r and w, which is driven by general equilibrium effects after an increase in K. ϕ is considered a type of externality because individuals take a sequence of factor prices as given in competitive equilibrium and do not consider the possibility that these prices may vary due to endogenous government policies and their impacts on welfare. Another feature of ϕ is that the extent of pecuniary externalities differs across individuals based on their factor composition of income, as noted by Davila et al. (2012). Note that because $f_{KK}(K)K < 0$, $f_K(K)\tau > 0$, the sign of this term highly depends on the value of $(-\frac{\epsilon\cdot g^n}{N} + \frac{a}{K})$. For example, suppose that there is an increase in K. In this case, if $\frac{\epsilon\cdot g^n}{N}$ is substantially larger than $\frac{a}{K}$, indicating that the factor income is biased toward labor (the factor composition of income of the consumption-poor), then $\frac{\partial v}{\partial K}$ is positive because w increases while r is reduced in general equilibrium. In contrast, the consumption-rich, whose factor income is biased more toward capital, are more likely to experience a loss in welfare due to a decline in r reducing their income.

Next, I substitute (22), (23), and (26) into the derivative of the government value in $\tilde{\tau}'$, $\frac{\partial \hat{v}}{\partial \tilde{\tau}'}$, in the FOC (21) and eliminate the partial derivatives of the future values by substituting them out. Additionally, I simplify the notation by employing sequential terms. I refer to $y_{i,t}$ as the variable of y for individual i in period t.

Then, $\frac{\partial \hat{V}_{i,t}}{\partial \tau_{t+1}}$ is given by:

$$\frac{\partial \hat{V}_{i,t}}{\partial \tau_{t+1}} = E_{i,t}\left[\sum_{s=1}^{\infty}\beta^s\cdot\Big(\underbrace{\phi_{i,t+s}}_{\text{Pecuniary Externalities}}\cdot\underbrace{\frac{\Delta K_{t+s}}{\Delta\tau_{t+1}}}_{\text{Efficiency Effect}} + \underbrace{\chi_{i,t+s}}_{\text{Income Redist. Externalities}}\cdot\underbrace{\frac{\Delta\tau_{t+s}}{\Delta\tau_{t+1}}}_{\text{Policy Scale Effect}}\Big)\right] \tag{28}$$

where $E_{i,t}$ is the conditional expectation for individual i at time t; $\frac{\Delta K_{t+s}}{\Delta \tau_{t+1}}$ is the total variation of aggregate capital K in period $t+s$ caused by a tax rate change at $t+1$; and $\frac{\Delta \tau_{t+s}}{\Delta \tau_{t+1}}$ is the total variation in the taxes in period $t+s$ caused by a tax rate change at $t+1$.[8] A notable feature is that the wedges related to individual decisions disappear above. More precisely, the product terms between these wedges and the derivative of individual decisions, such as $\omega \cdot \frac{\partial g^a}{\partial K}$ and $\zeta \cdot \frac{\partial g^n}{\partial K}$, are offset through the envelope theorem. This canceling out implies that the government does not consider the direct impact of distortions that are involved with individual decisions on welfare. The government recognizes that given a set of policies, individuals make optimal decisions on consumption, saving, and labor supply in competitive equilibrium; therefore, there is no room for varying individual welfare via this channel.

Instead, the government considers the two types of externalities that are heterogeneous across individuals—pecuniary externalities and income redistribution externalities—along the transition path. Pecuniary externalities work through changes in aggregate capital. For example, if the government increases τ_{t+1}, efficiency is reduced because aggregate capital K falls below the initial level for a period of time ($\frac{\Delta K_{t+s}}{\Delta \tau_{t+1}} < 0$). These changes in K along the transition path induce variations in the factor composition of income between r and w over time. In this case, while w decreases, r increases over the transition path. Income redistribution externalities work through changes in transfers. For instance, if the economy converges to the long-run equilibrium in a mean-reverting way, its tax rate τ remains above the initial level for a time once τ_{t+1} increases ($\frac{\Delta \tau_{t+s}}{\Delta \tau_{t+1}} > 0$). These increases in τ imply more transfers along the transition path, influencing income redistribution externalities.

Finally, I substitute (28) into the FOC (21). Then, the government-optimal condition is given by:

$$\underbrace{-\int E_{i,t}\left[\sum_{s=1}^{\infty} \beta^s \cdot \phi_{i,t+s} \cdot \frac{\Delta K_{t+s}}{\Delta \tau_{t+1}}\right] \mu_t(\mathbf{d}(a_{i,t} \times \epsilon_{i,t}))}_{\text{Aggregate Pecuniary Externalities}} = \underbrace{\int E_{i,t}\left[\sum_{s=1}^{\infty} \beta^s \cdot \chi_{i,t+s} \cdot \frac{\Delta \tau_{t+s}}{\Delta \tau_{t+1}}\right] \mu_t(\mathbf{d}(a_{i,t} \times \epsilon_{i,t}))}_{\text{Aggregate Income Redist. Externalities}} \tag{29}$$

where

$$\phi_{i,\tilde{t}} = u_c(c_{i,\tilde{t}}, 1-n_{i,\tilde{t}})\big((1-\tau_{\tilde{t}})(-\frac{\epsilon \cdot g^n_{i,\tilde{t}}}{N_{\tilde{t}}} + \frac{a_{i,\tilde{t}}}{K_{\tilde{t}}})f_{KK}(K_{\tilde{t}})K_{\tilde{t}} + f_K(K_{\tilde{t}})\tau_{\tilde{t}}\big)$$

$$\chi_{i,\tilde{t}} = u_c(c_{\tilde{t}}, 1-n_{\tilde{t}})\bigg(w_{\tilde{t}}(N_{\tilde{t}} - \epsilon_{\tilde{t}} \cdot g^n_{i,\tilde{t}} + r_{\tilde{t}}(K_{\tilde{t}} - a_{i,\tilde{t}})\bigg).$$

The above government-optimal condition implies that for each period t, when deciding τ, the government without commitment strikes a balance between the two types of externalities at the

[8]The precise definitions are presented in Appendix C.

aggregate level. For further discussion, I refer to the left-hand side as *aggregate pecuniary externalities* because the absolute value of this term indicates the sum of the present discounted value of pecuniary externalities caused by variations in the factor composition of income over all individuals. Analogously, I refer to the right-hand side as *aggregate income redistribution externalities* because this term means the sum of the present discounted value of income redistribution externalities caused by changes in transfers over all individuals.

The above optimal condition (29) reveals what the government takes into account in its policy decisions. First, the government places greater weight on the interests of the consumption-poor because the two types of externalities ϕ and χ are weighted with the marginal utility of consumption. Additionally, the distribution of consumption is right-skewed, leading the government to further consider the interests of the consumption-poor. This government attention to the consumption-poor allows me to regard its policy decisions as a cost-benefit analysis of the consumption-poor for taxes and transfers. Second, the consumption-poor experience different welfare changes from the two types of externalities, as mentioned previously with (24) and (27) . Suppose that the government decides to increase τ. While income redistribution externalities χ improve welfare for the consumption-poor by increasing their after-tax income via transfers, pecuniary externalities ϕ are negative for the consumption-poor—whose income is more biased toward labor—because the tax change reduces w and increases r in general equilibrium. Third, for the consumption-poor, pecuniary externalities ϕ are less than or equal to $f_k(K)\tau$ ($\phi \geq f_k(K)\tau$) and income redistribution externalities are non negative ($\chi \geq 0$). Finally, as incomes for the consumption-poor become more equally distributed $(a, \epsilon \cdot g^n) \to (K, N)$, pecuniary externalities ϕ converge to $f_K(K)\tau$ from a negative value, and their income redistribution externalities χ converge to 0 from a positive value. These findings leads to the following proposition.

Proposition 1. *For each period t, the government without commitment makes a policy decision as follows:*

1. *When aggregate pecuniary externalities are* ***greater*** *than aggregate income redistribution externalities, the government* ***reduces*** τ_{t+1}.

2. *When aggregate pecuniary externalities are* ***less*** *than aggregate income redistribution externalities, the government* ***raises*** τ_{t+1}.

3. *As the income of the consumption-poor approaches the average, the government* ***reduces*** τ_{t+1}.

The first and second statements posit that pecuniary externalities can be seen as the consumption-poor's marginal cost of increasing τ_{t+1} and income redistribution externalities can be regarded as their marginal benefit. Accordingly, the government decides to increase (decrease) its tax rate when the marginal cost (i.e., aggregate pecuniary externalities) is greater (less) than the marginal benefit (i.e., aggregate income redistribution externalities). The third statement is closely related to how χ and ϕ change as the income of the consumption-poor approaches the average. Specifically, the marginal benefit of increasing τ_{t+1} for the consumption-poor, ϕ, converges to 0 from a positive value as their income approaches the average. Additionally, as their income approaches the average, in their marginal cost of raising τ_{t+1}, χ, the impact of $f_k(K)\tau$ becomes more pronounced. Therefore, the most effective way to minimize losses from χ is to reduce τ when the income of the consumption-poor approaches the average. This insight is distilled into the third statement.

3.2 Comparison with Constrained Efficiency

Another interesting investigation for the optimal conditions (29) is to compare this to the planner's optimal condition in Davila et al. (2012) because pecuniary externalities are thoroughly investigated in their study. The planner's optimal condition in Davila et al. (2012) is given by:

$$\omega_{i,t} + \beta \int E_{i,t}[\phi_{t+1}]\mu(\mathbf{d}(a \times \epsilon)) = 0 \tag{30}$$

where

$$\omega_{i,t} = -u'(c_{i,t}) + \beta(1 + r(K_{t+1})) \sum_{\epsilon'} \pi_{\epsilon,\epsilon'} u'(c_{t+1}) \tag{31}$$

$$\phi_{i,t} = u'(c_{i,t})\big((-\frac{\epsilon_{i,t}}{L_t} + \frac{a_{i,t}}{K_t})f_{KK}(K_t)K_t\big) \tag{32}$$

In contrast with the government's optimal conditions (29) in this paper, the consumption Euler equation part in Davila et al. (2012) does not need to be null. What matters for the social planner is to satisfy this optimal condition considering the pecuniary externality and distortions embedded in the consumption Euler equation. This distinction arises because of different assumptions between the two economies. The Davila, Hong, Krusell and Ríos-Rull's (2012) economy is centralized: the social planner can manipulate individual saving decisions while preserving the constraints caused by incomplete markets and uninsurable idiosyncratic income risk. This centralized economy assumption makes the consumption Euler equation non-zero. However, the economy in my paper is decentralized. Although the government exists and endogenously determines taxes, individuals optimally choose consumption and saving; therefore, the government has no room for improvement regarding individual consumption dynamic allocations.

3.3 Comparison with the Case with Commitment

Regarding the case with commitment (the Ramsey problem), I assume that the government has already taken an optimal sequence of taxes/transfers and now considers a change in its tax rate at time $t+1$. Then, repeating the previous calculations leads me to obtain the government's optimal condition as follows:

$$\underbrace{-\int E_{i,0}\left[\sum_{s=1}^{\infty}\beta^{s}\cdot\phi_{i,s}\cdot\frac{\Delta K_{s}}{\Delta\tau_{t+1}}\right]\mu_{0}(\mathbf{d}(a_{i,0}\times\epsilon_{i,0}))}_{\text{Aggregate Pecuniary Externalities at } t=0}=\underbrace{\int E_{i,0}\left[\sum_{s=1}^{\infty}\beta^{s}\cdot\chi_{i,s}\cdot\frac{\Delta\tau_{s}}{\Delta\tau_{t+1}}\right]\mu_{0}(\mathbf{d}(a_{i,0}\times\epsilon_{i,0}))}_{\text{Aggregate Income Redist. Externalities at } t=0}. \tag{33}$$

In comparison with (29), the government's optimal condition (33) highlights the role of commitment technologies. Although the government always balances the two types of aggregate externalities in both cases, commitment makes a difference in how the balance is struck. With commitment, the government (the Ramsey planner) balances the externalities by taking conditional expectations at time 0 over the entire time horizon. In contrast, as (29) shows, a lack of commitment causes the government to take a conditional expectation in each period. Because of this difference, commitment leads the government to consider the effect of taxes and transfers in a backward manner. The optimal condition (33) suggests that the government considers how a tax change in period $t+1$ affects not only the current and future economy but also the past. In contrast, as (29) shows, a lack of commitment causes the government to consider the effect of policies only on the current and future economy without taking into account their effect on the past. These findings suggest that the government with commitment makes policy decisions that are optimal at time 0 but not necessarily desirable when evaluated in a forward-looking manner, thereby leading to time-inconsistent outcomes.

The comparison above clearly shows the qualitative differences in the government's policy decisions with and without commitment. However, it does not provide a quantitative assessment of the magnitude of these differences, which is an important question to address. In the following section, I propose a numerical solution method to quantitatively analyze these differences.

4 Numerical Solution Algorithm

Here, I focus on conveying the key ideas of the numerical solution algorithm. Appendix A demonstrates each step of the algorithm with details.

Although the characterization of the MPE in the previous section helps us better understand the government's decisions on policies, it is not very useful in numerically computing the equilibrium because of its sequential feature. Basically, solving the model entails a substantial computational burden. The law of motion for the distribution of households $\Gamma(\cdot)$ has to be consistent with individual decisions. Additionally, because the labor supply is endogenous with wealth effects, the two factor markets—K and N—must clear. Furthermore, perhaps the most challenging part is finding the equilibrium policy function $\Psi(\cdot)$ that must be consistent with individual decisions and the law of motion for the distribution of households. That is, three equilibrium objects—specifically, individual decisions, $g^n(\cdot)$ and $g^a(\cdot)$, the law of motion for the distribution, $\Gamma(\cdot)$, and the policy function, $\Psi(\cdot)$—interact and have to be consistent with one another in an MPE.

I address the above computational issues by taking ideas from the backward induction method of Reiter (2010). This study introduced a non-simulation-based solution method to solve an incomplete markets economy with aggregate uncertainty. As in Krusell and Smith's (1998), the Reiter's (2010) approach also reduces the dimension of distributions in the law of motion $\Gamma(\cdot)$ to some finite moments of the distribution, and they are defined across the aggregate finite grid points. However, the way of finding $\Gamma(\cdot)$ differs substantially between the two methods. In Krusell and Smith (1998), their algorithm repeatedly simulates the model economy through the inner and outer loops. In the inner loops, the value is solved given a perceived law of motion for the distribution of households, and the law of motion is updated after a simulation in the outer loop. This procedure is repeated until the perceived law of motion is equal to the updated one.

By contrast, the backward induction method of Reiter (2010) does not simulate the economy to update the law of motion for the distribution of households $\Gamma(\cdot)$; rather, this is updated while solving for the value given a set of proxy distributions across the aggregate finite grid points. Given a proxy distribution, finding the law of motion for the distribution of households $\Gamma(\cdot)$ is feasible by using the moment-consistent conditions. For example, individual decision rules for assets allow me to obtain the information (e.g., the mean or variance) on the aggregate capital in the next period. A simulation step is followed not to update the law of motion for the distribution of households $\Gamma(\cdot)$ but to update a set of proxy distributions across the finite nodes in the aggregate state. Simulations are required much less in Reiter (2010) than in Krusell and Smith (1998), which improves the computational efficiency of the backward induction method. Additionally, with these proxy distributions, the backward induction method allows me to approximate not only the aggregate law of motion for the distribution $\Gamma(\cdot)$ but also the tax policy function $\Psi(\cdot)$. This is feasible because, with the value function, these endogenous tax functions can be directly obtained by solving (17).

However, I wish to clarify that I cannot directly apply the Reiter's (2010) method to the model in this paper because of the presence of off-equilibrium paths after one-shot deviations that are required to find the sub-perfect Nash equilibrium. In the incomplete markets economy with aggregate uncertainty, for which the Reiter's (2010) method is originally designed, the distribution of aggregate shocks (TFP) Z is ergodic. Thus, all the aggregate states Z are not measure zero. With a positive probability, all the states in Z are realized on the equilibrium path. However, cases in the MPE do not have this property. For example, in the economy without commitment, the government chooses a tax rate by comparing one-time deviation policies, as in (17). Some tax paths will not be reached on the equilibrium path but the corresponding value needs defining to solve the problem that the government confronts.

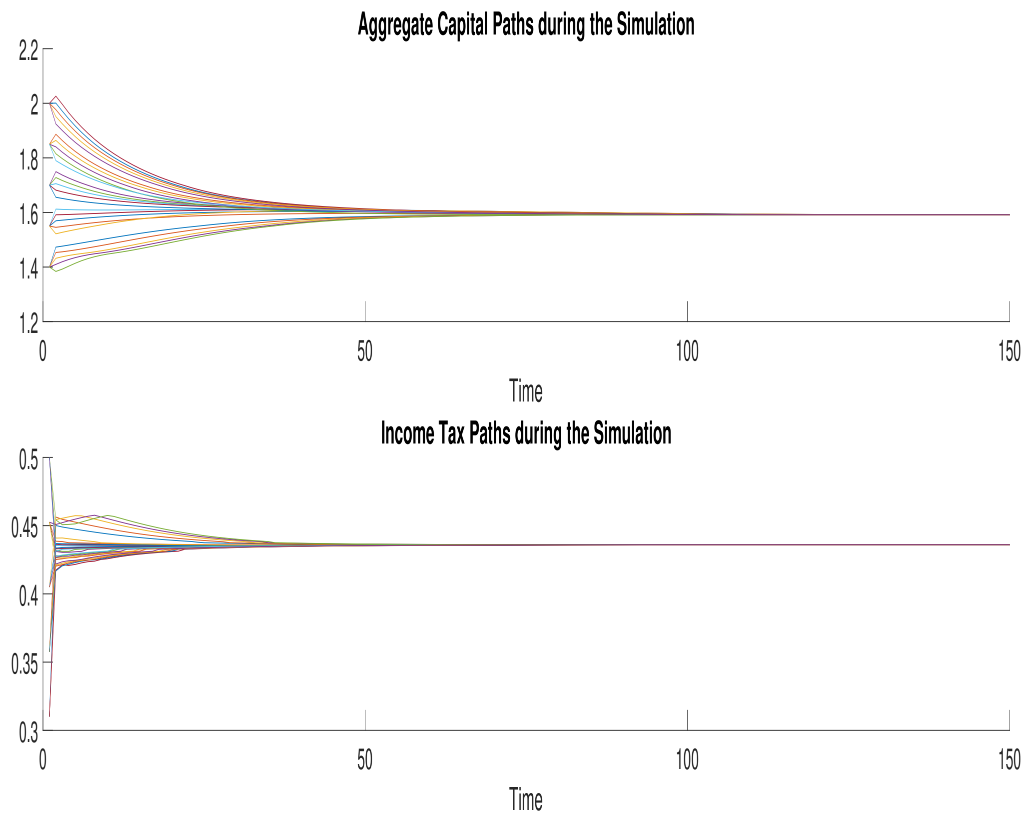

Figure 1: Transitions from off the Equilibrium to the Equilibrium

To cope with this issue, I make three variations to the original backward induction method of Reiter (2010). First, as mentioned above, I approximate not only the aggregate law of motion for the distribution of households but also the endogenous tax policy function. I find these mappings in a nonparametric way as in Reiter (2010). Second, I arrange distributions for all types of off the equilibrium paths by taking the initial distribution of the simulations as the previous proxy distribution for each finite grid point of the aggregate state. Figure 1 shows various transitions from off the equilibrium to the steady-state equilibrium in the case without commitment. Finally, I modify the way of constructing the reference distributions, which is required to update the proxy distributions in Reiter (2010), by reflecting the features of the MPE, in which how many times a tax rate off-the-equilibrium takes place is unknown before simulation. Appendix A demonstrates the full details of the solution method, in addition to its performances in efficiency and accuracy.

Because of these somewhat complex variations in the Reiter's (2010) method, one might consider simply using the Krusell and Smith's (1998) method to solve this model. However, their approach is less efficient in addressing this class of models in MPE. First, finding the two aggregate laws of motion–Γ and Ψ—is computationally very costly when using this simulation-based solution method. When this method is employed to solve the economy in this paper, this process is the same as adding another outer loop to the outer loop in the Krusell and Smith's (1998) original algorithm, thereby exponentially increasing the computational burden. Second, the parametric assumption of the Krusell and Smith's (1998) approach acts as a barrier because the equilibrium tax function $\Psi(\cdot)$ could be severely non-linear in the aggregate state. The parametric assumption works well when the law of motion for household distributions $\Gamma(\cdot)$ is close to linear. However, it is possible for $\Psi(\cdot)$ to be severely non-linear and the method in this paper can cope with this non-linearity.

Regarding the case with commitment (the Ramsey problem), I employ the approach in Dyrda and Pedroni (2022) by parameterizing the transition path of income taxes as follows:

$$\tau_t = \left(\sum_{i=0}^{m_{x0}} \alpha_i^x P_i(t) \right) \exp(\lambda t) + (1 - \exp(-\lambda t)) \left(\sum_{j=0}^{m_{xF}} \beta_j^x P_j(t) \right), \quad t \leq t_F \tag{34}$$

where $\{P_i(t)\}_{i=1}^{m_{x0}}$ and $\{P_j(t)\}_{j=0}^{m_{xF}}$ are families of the Chebyshev polynomial; m_{x0} and m_{xF} are the orders of the polynomial approximation for the short- and long run dynamics, respectively; $\{\alpha_i^x\}_{i=0}^{m_{x0}}$ and $\{\beta_j^x\}_{j=0}^{m_{xF}}$ are weights on the consecutive elements of the family; and λ controls the convergence rate of the fiscal instrument. This setting assumes that the economy has the long-run steady state at the latest in period t_F. I first choose $m_{x0} = m_{xF} = 2$ and $t_F = 250$. Then, I seek $\{\alpha_0^x, \alpha_1^x, \alpha_2^x, \beta_0^x, \beta_1^x, \beta_2^x, \lambda\}$ that maximize the welfare function of the utilitarian government at time 0.[9]

5 Calibration

I calibrate the model to capture the features of the U.S. economy. I divide the parameters into two groups. The first set of parameters requires solving the stationary distribution of the model to match the moments generated by the model with their empirical counterparts. The other set of the parameters is determined outside the model. I take the values of these parameters from the macroeconomic literature and policies.

[9]The inclusion of lump-sum transfers prevents the non-existence of a Ramsey steady state, which is examined in Straub and Werning (2020). Further details are provided in Dyrda and Pedroni (2022).

Table 1: Parameter Values of the Baseline Economy

	Description (Target)	Value
β	Discount factor ($K/Y = 3$)	0.951
B	Utility of leisure (AVG Wrk Hrs = 1/3)	3.803
σ	Relative risk aversion	2
χ	Frisch elasticity of labor supply	0.75
$\underline{a}$	Borrowing constraint	0
z	Total Factor Productivity	1
θ	Capital income share	0.36
δ	Depreciation rate	0.08
ρ_ϵ	Persistence of wage shocks	0.955
σ_ϵ	STD of wage shocks	0.20
G	Government spending	$G/Y = 0.19$
τ	AVG income tax	0.31

Table 1 displays the parameters. I internally calibrate two parameters: the discount factor β and the utility of leisure B. β is selected to match a capital-to-output ratio of 3, and B is chosen to reproduce an average of hours worked of 8 hours a day. The other parameters are determined outside the model. The coefficient of relative risk aversion is set to 2. The Frisch elasticity of labor supply χ is taken to be 0.75. I set the borrowing constraint as $\underline{a} = 0$. The TFP z is set as 1, and the capital income share θ is chosen to reproduce the empirical finding that the share of capital income is 0.36. The annual depreciation rate δ is 8 percent. The persistence of wage shocks ρ_ϵ is set to be 0.955, and the standard deviation of wage shocks σ_ϵ is taken as 0.20. The values of ρ_ϵ and σ_ϵ lie in the range of those frequently used in the literature. Government spending G is set so that the fraction of government spending out of GDP is equal to 0.19. The flat income tax rate is chosen as 0.31 in the baseline economy, implying the ratio of transfers to GDP to be 0.046, which is closer to its empirical counterpart of, 0.044.[10]

6 Results

In this section, I quantitatively explore how commitment technologies affect the aggregate economy, inequality, and welfare. For this, I compare the economy with the time-consistent optimal policy to the economy with the time-inconsistent optimal policy with the Ramsey planner. I assume that the initial economy begins at the calibrated steady-state through all the exercises and compare their equilibrium results along the transition path.

[10] I take the value from Jang et al. (2021) that excludes Social Security and Medicare in their calculation to reflect the lack of a lifetime structure.

6.1 Time-Consistent versus Time-Inconsistent Policy: Income Tax

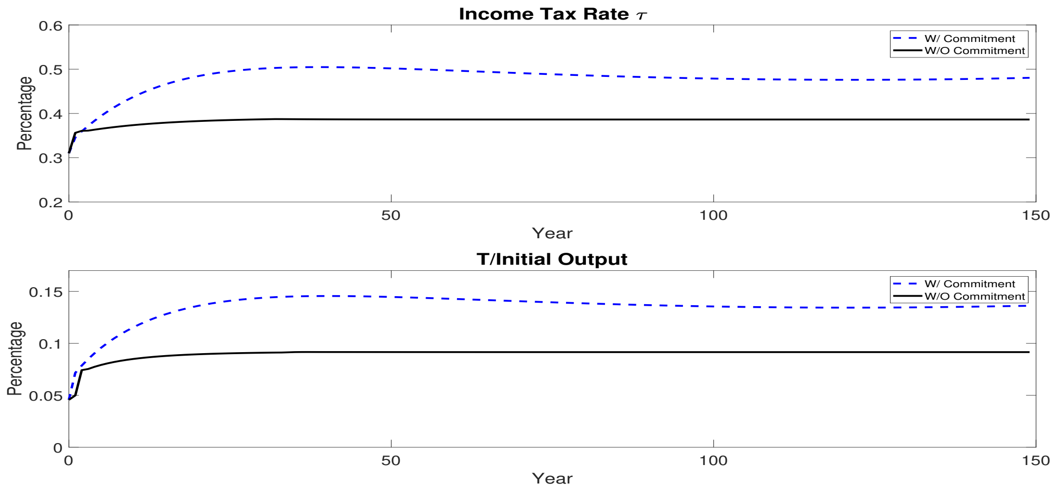

Figure 2: Time-Consistent and Time-Inconsistent Policies: Taxes/Transfers Transition Paths

Figure 2 shows the time-consistent and time-inconsistent optimal taxes and the implied ratio of transfers to the initial output. The top panel of Figure 2 implies that the government with commitment chooses more substantial income taxes than the government with the time-consistent optimal policy over the entire transition path. The Ramsey planner initially increases income taxes by 16 percentage points and then reduces them thereafter. However, without commitment, the government with the optimal income tax raises taxes by only 2 percentage points. This gap in tax policies results in differences in the size of transfers. The time-inconsistent optimal income tax economy generates larger transfers than the time-consistent optimal income tax economy. The ratio of transfers to the initial GDP in the case with commitment gradually increases by 9.2 percentage points, but in the case without commitment, it only increases by 4.7 percentage points.

Table 2: Welfare Outcomes According to Commitment (Income Taxes)

Welfare (CEV)	Time-Inconsistency	Time-Consistency
OPT INC TAX	+2.19%	+0.57%

This distinction in income taxes creates different welfare consequences. Table 2 shows that welfare, measured by the consumption equivalent variation (CEV) of the utilitarian welfare function, is significantly higher in the case with the time-inconsistent optimal tax. The time-inconsistent optimal tax improves welfare by 2.19 percent, while the time-consistent optimal tax improves welfare by 0.57 percent. To understand this disparity in welfare consequences, I examine how differently the inputs of the social welfare function vary over time according to the availability of commitment. Note that welfare increases when the overall level of consumption and leisure increases, and their inequality is reduced.

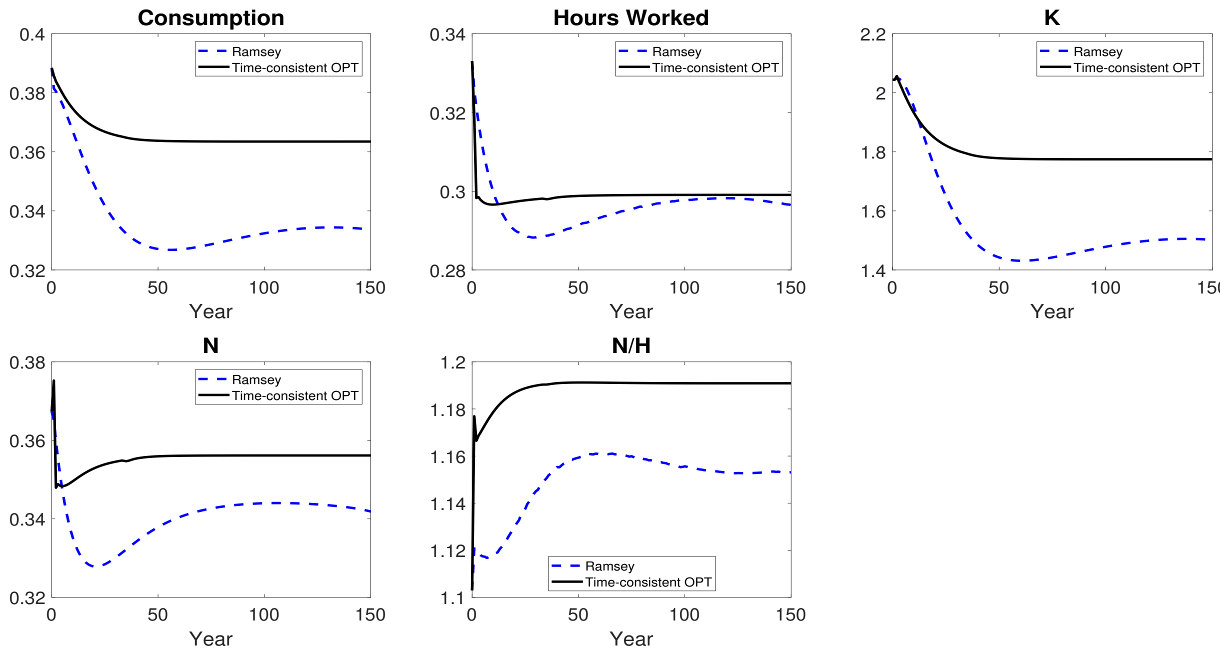

Figure 3: Time-Consistent and Time-Inconsistent Income Tax: Aggregate Outcomes

Figure 3 illustrates the changes in the levels of the aggregate variables. The figure suggests that the case without commitment generates more efficient outcomes. All the aggregate variables in the economy are larger with the time-consistent optimal policy than with the time-inconsistent optimal policy. This result may seem obvious because lower taxes in the time-consistent case cause fewer distortions. However, this finding might be unclear when examining the welfare consequences. Despite the fact that consumption is substantially larger in the case with the time-consistent policy, the gaps in hours worked are not significant and do not offer a complete understanding of welfare implications

Figure 4 illustrates the inequalities in consumption, hours worked, wealth, and after-tax income. The figure suggests that more significant welfare improvements in the case with the time-inconsistent optimal policy are driven mainly by larger reductions in inequalities in consumption and leisure. Although consumption inequality, measured by the Gini coefficient, decreases by less than 5 percent with the time-consistent optimal policy, it is reduced by approximately 10 percent with the time-inconsistent optimal tax. Similarly, inequality in hours worked is also reduced more in the case with the time-inconsistent optimal policy. These findings imply that the commitment instrument allows the Ramsey planner to better manage the evolution of inequality, leading to a better welfare outcome. However, this explanation does not provide a clear economic logic behind

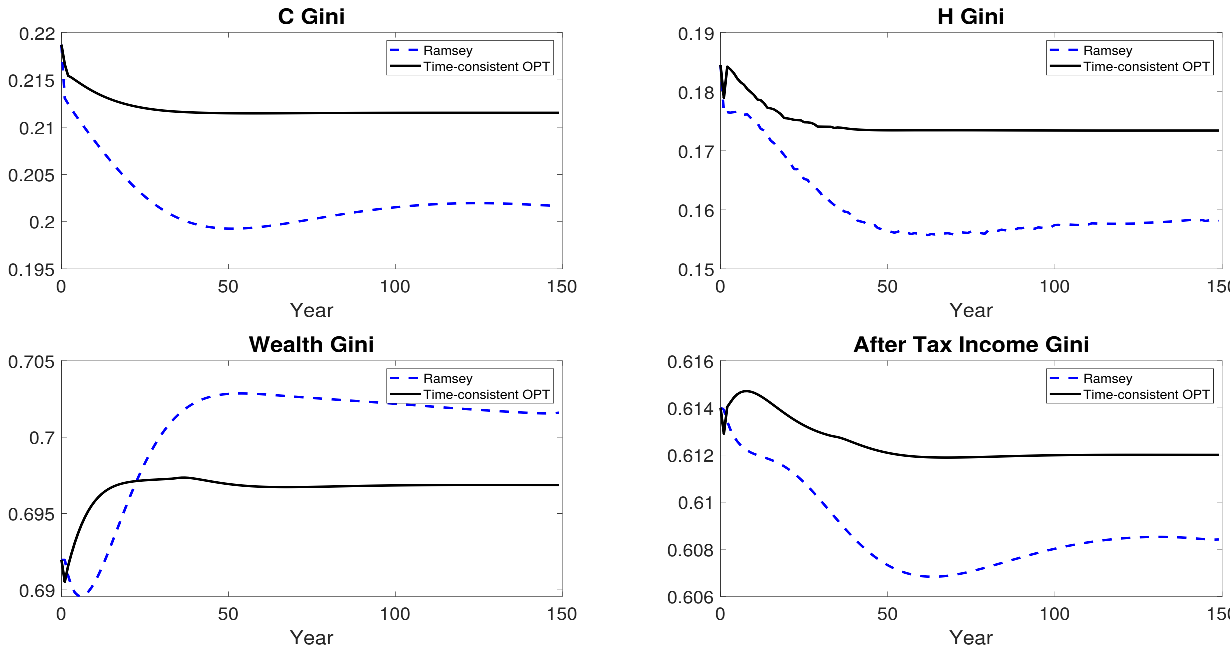

Figure 4: Time-Consistent and Time-Inconsistent Income Taxes: Distributional Outcomes

the quantitative outcomes. To better understand the economic logic behind these differences, I employ the theoretical findings in the previous chapter to interpret these quantitative outcomes.

Figure 5 shows the dynamics of the factor prices w and r depending on the availability of commitment. This figure suggests that the government with commitment (the Ramsey planner) benefits from pecuniary externalities in the early phase of the transition by exploiting differences in the speed of adjustments between K and N. Of course, in the long run, the increased tax rate in the case with commitment reduces the ratio of K to N, leading to a decrease in w and an increase in r. However, during their adjustment, the ratio of K to N increases in the early phase of the transition because the adjustment of K is slower than that of N, leading to a rise in w and a reduction in r. These price changes help improve the welfare of the consumption-poor, who are better represented by the government, through pecuniary externalities. Additionally, as shown in the bottom-right panel of Figure 4, the government with commitment receives front-loaded welfare gains due to the rapid reduction in inequality in after-tax income. Therefore, the government with commitment achieves up-front welfare gains through both types of externalities.

Figure 5 also indicates that the government with commitment endures welfare losses in the long run. In the long run, w is lower and r is higher than their initial values, and these changes become negative pecuniary externalities for the consumption-poor. Additionally, in the long run, as Figure (33) shows, there is no further reduction in after-tax income inequality, leading to minuscule welfare gains from income redistribution externalities. Recall that, as (33) shows, the sum of all

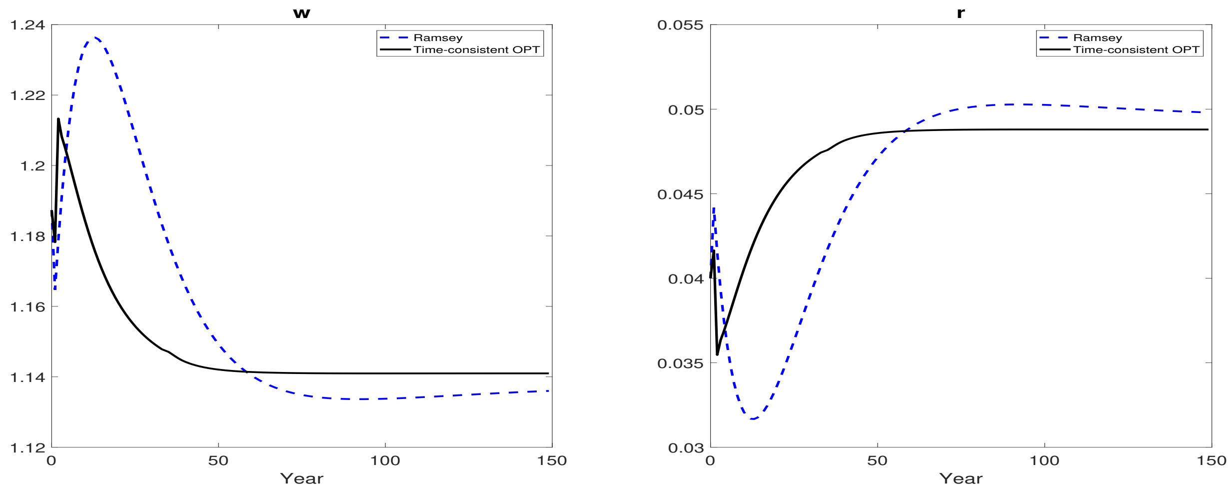

Figure 5: Time-Consistent and Time-Inconsistent Income Taxes: Dynamics of w and r

pecuniary externalities is equal to that of all income redistribution externalities. These findings imply that the government with commitment takes front-loaded welfare gains via positive income redistribution externalities and pecuniary externalities while enduring welfare losses from small income redistribution externalities and negative pecuniary externalities.

Note that the above strategy is not credible without commitment. Suppose that the government without commitment is in the long-run equilibrium of the economy with commitment. As shown in (29), a lack of commitment causes the government to measure the costs and benefits of changing a tax rate in a forward-looking manner. Since front-loaded welfare gains are disregarded, the marginal cost of raising the tax rate, which equals the aggregate pecuniary externalities, is greater than that of the initial economy, and the marginal benefit, which equals the aggregate income redistribution externalities, is close to zero. Proposition 1 indicates that the government, in this case, will decide to reduce its tax rate in the next period. Individuals rationally anticipate the government's motive to reduce its tax rate in the next period and preemptively increase their savings. As a result, the long run economy with commitment is no longer sustainable.

6.2 Time-Consistent versus Time-Inconsistent Policy by Tax Base

Figure 6 displays the time-consistent and time-inconsistent optimal taxes and transfers based on the tax base. The top (bottom) panels illustrate the outcomes when labor (capital) income taxes are permitted to vary over time while the capital (labor) income tax is held at the calibrated level of 0.31. In this section, I concentrate on the cases where labor income taxes change over time rather than the cases where capital income taxes change over time. The bottom panels demonstrate that when capital income tax is the only instrument available to the government, the time-consistent

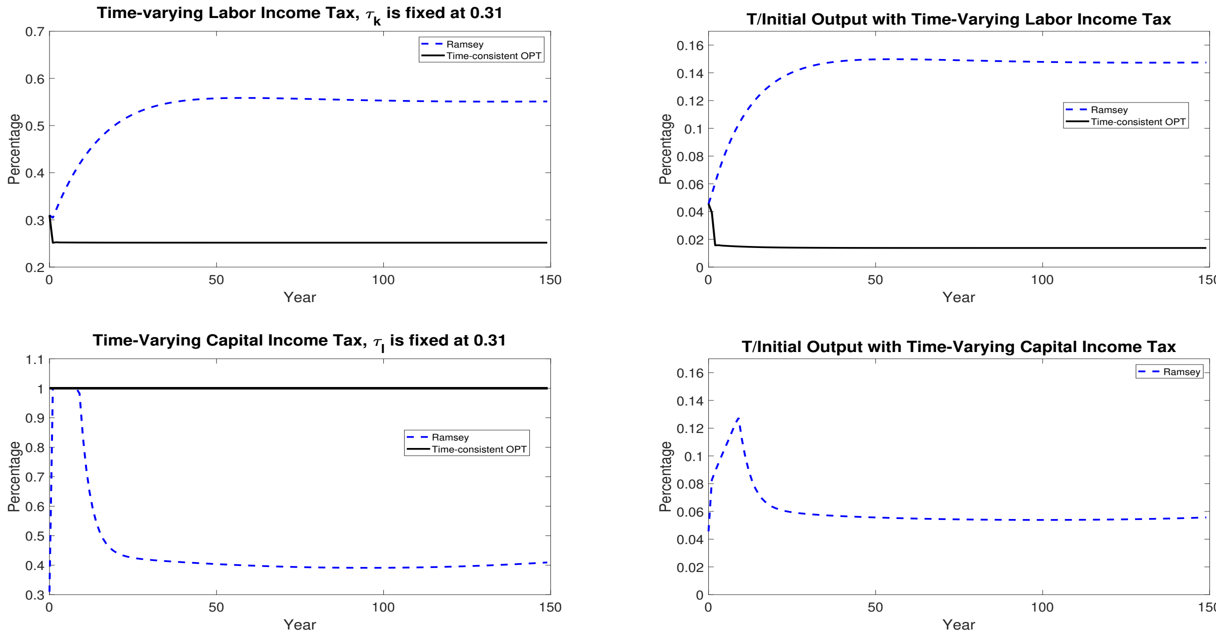

Figure 6: Time-Consistent and Time-Inconsistent Policies: Taxes/Transfers by Tax Base

government raises it to the maximum rate of 100 percent. This extremely high capital income tax has been analyzed theoretically in previous studies, such as Chari and Kehoe (1990). These studies have shown that the time-consistent government makes the capital income tax confiscatory, resulting in no savings. In my quantitative exercise, time-consistent capital income taxes reduce the size of the economy to such an extent that the tax revenue cannot cover the exogenous government spending in its budget.[11] The results of the time-inconsistent capital income tax economy are presented in Appendix B, which are consistent with those in Dyrda and Pedroni (2022).

Table 3: Welfare Outcomes According to Commitment (Labor Income Taxes)

Welfare (CEV)	Time-Inconsistency	Time-Consistency
OPT Labor INC TAX (τ_k is given by 0.31)	+2.20%	-1.24%

Based on the top panels of Figure 6, the economies with labor income taxes deliver similar messages as in the cases with proportional income taxes. The government with commitment raises its labor income taxes early in the transition, resulting in more significant transfers. However, without commitment, the government reduces its labor income taxes, resulting in lower transfers. Table 3 shows that the welfare consequences are also similar to the cases with proportional income taxes:

[11]To have sustainable capital income tax rates, the size of government spending might need to be endogenous by assuming households to value it, as in Klein et al. (2008). The current setting does not have this component.

the time-inconsistent government brings more significant welfare improvements to the economy than the time-consistent government.

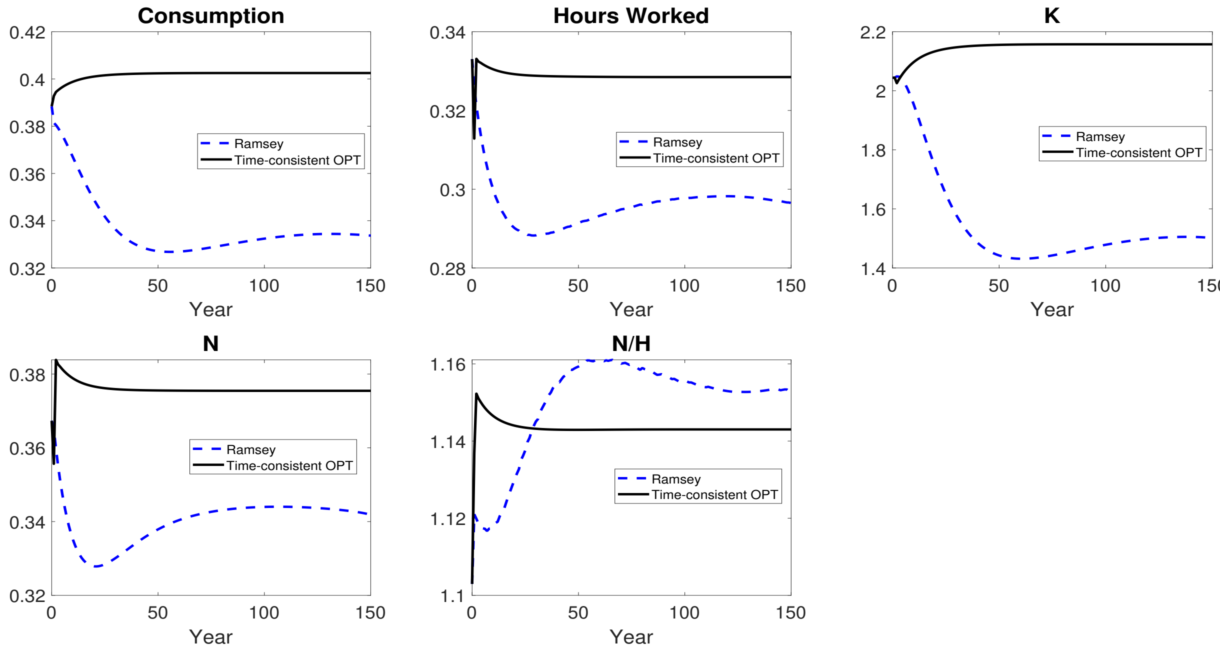

Figure 7: Time-consistent and Time-inconsistent Labor Income Taxes: Aggregate Outcomes

Figure 7 shows that heavier labor income taxes in the case with commitment lead to a less efficient economy than without commitment, which is similar to the results in the cases with proportional income taxes. In the economy with commitment, aggregate consumption, hours worked, capital, and effective labor are lower than those in the case without commitment because heavier labor income taxes in the case with commitment lead to a greater loss in efficiency, causing the economy to shrink. Although the dynamics of the effective labor to hours worked ratio exhibit a different pattern, this gap is due to the larger differences in transfers driven by heavier labor income taxes inducing labor market selection for the case with commitment. The dynamics of overall aggregate variables are similar to those in the cases with proportional income taxes.

Figure 8 shows that the dynamics of inequality play a crucial role in the welfare outcomes in both cases. The economy with commitment experiences more substantial transfers, which result in reduced inequality in after-tax income. This reduction in inequality leads to more equally distributed consumption and hours worked, ultimately improving welfare. In contrast, without commitment, the government reduces taxes and transfers, leading to more significant inequality in after-tax income. This increase in inequality worsens welfare in the economy. Overall, these findings are consistent with those in the cases with proportional income taxes. To obtain a better understanding, I apply the theoretical findings to the quantitative results.

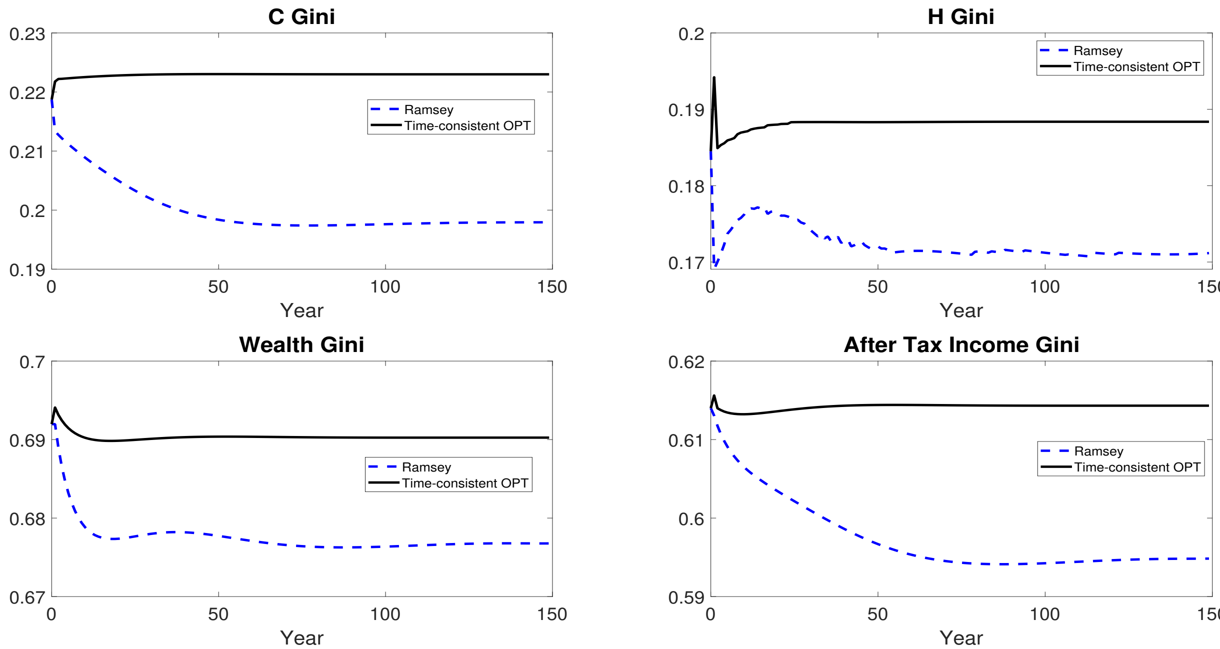

Figure 8: Time-consistent and Time-inconsistent Labor Income Taxes: Distributional Outcomes

Note that the government with commitment obtains positive externalities from reduced income inequality, as the bottom-right panel of Figure 8 shows. Furthermore, Figure 9 indicates that it also obtains positive pecuniary externalities from changes in the factor composition of income—increases in w and decreases in r—in the early phase of the transition. These price changes seem inconsistent with the outcomes after increasing labor income taxes, which implies the opposite changes due to an increase in the capital-to-labor ratio in the long run. Nonetheless, as Figure 7 shows, increases in the labor income tax reduce the aggregate capital and labor supply in the early phase of the transition. Additionally, because the speed of adjustment in capital is slower than that in labor, the government obtains front-loaded positive pecuniary externalities from increases in w and decreases in r, along with positive externalities from reduced income inequality. However, in the long run, the government with commitment endures welfare losses from negative pecuniary externalities and small income redistribution externalities. My theoretical findings imply that the government with commitment balances the two types of externalities by allocating positive externalities in the early phase of the transition while putting negative externalities afterward.

Without commitment, this strategy is not credible. Suppose that, as in the case with proportional income taxes, the government without commitment is in the long-run equilibrium of the economy with commitment. This government ignores the upfront welfare changes and compares the two types of externalities in a forward-looking manner. From the government's perspective, negative externalities from the factor composition of income are more significant than small ex-

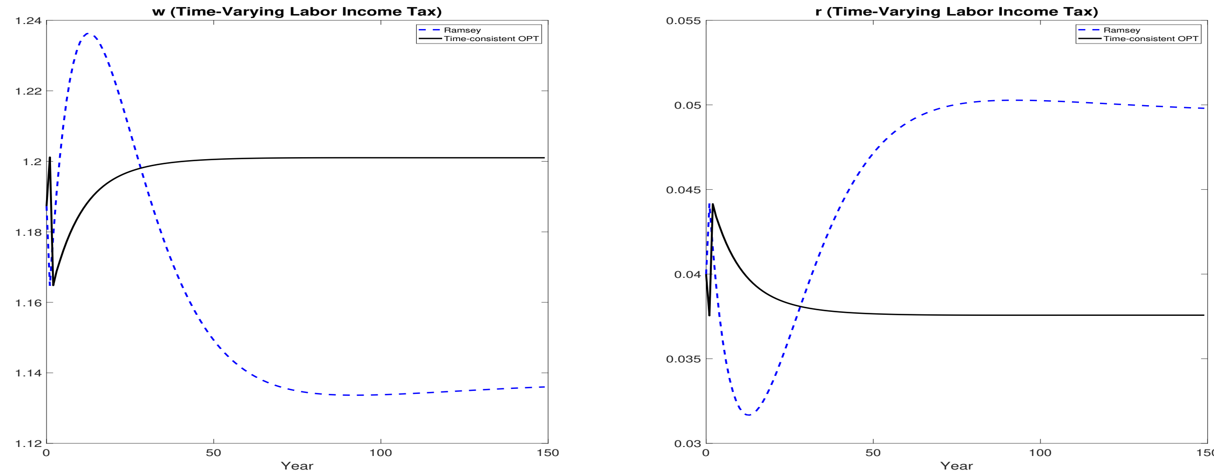

Figure 9: Time-Consistent and Time-Inconsistent Labor Income Taxes: Dynamics of w and r

ternalities from stagnant changes in income inequality. Therefore, the government is willing to rebalance the two types of externalities by reducing labor income taxes because doing so will alleviate negative externalities from lower wages and spare more income under the lack of commitment. Households rationally anticipate the government's incentive and increase labor supply and savings. Consequently, the long-run equilibrium in the case of commitment is not maintained.

7 Conclusion

This paper examines how the availability of government commitment affects the macroeconomy, inequality, and welfare. I characterize the MPE with heterogeneous agents via the GEE; develop a numerical solution method for this game; and apply this method to the standard incomplete markets model with uninsurable idiosyncratic income risk. I find that commitment affects how the government addresses the two types of externalities: income redistribution externalities through transfers and pecuniary externalities via changes in the factor composition of income. The magnitude of this qualitative difference is measured using a quantitative method in a calibrated economy. The quantitative analysis shows that commitment has significant impacts on the government's tax/transfer decisions along the equilibrium path. As a result, these different policies create disparities in individual decisions, which generate differences in the aggregate economy, inequality, and welfare.

Note that the solution method itself could provide many opportunities for studying unexplored research topics. Given the fundamental feature of Reiter (2010), this solution method can be compatible with aggregate uncertainty. This research direction makes it possible to revisit questions on fiscal policies according to the political and commitment structure. Another exciting application of the method is addressing the interactions between policies and life-cycle dimensions. The Kim's (2021) method makes this direction reachable. She extends the Reiter's (2010) backward induction method to solve an overlapping generations model. Such analyses are deferred to future work.

References

Aiyagari, S Rao, "Uninsured idiosyncratic risk and aggregate saving," *The Quarterly Journal of Economics*, 1994, *109* (3), 659–684.

Athey, Susan, Andrew Atkeson, and Patrick J Kehoe, "The optimal degree of discretion in monetary policy," *Econometrica*, 2005, *73* (5), 1431–1475.

Azzimonti, Marina, "Barriers to investment in polarized societies," *American Economic Review*, 2011, *101* (5), 2182–2204.

Barro, Robert J and David B Gordon, "Rules, discretion and reputation in a model of monetary policy," *Journal of monetary economics*, 1983, *12* (1), 101–121.

Bassetto, Marco, Zhen Huo, and José-Víctor Ríos-Rull, "Institution Building without Commitment," *Working Paper*, 2020.

Calvo, Guillermo A, "On the Time Consistency of Optimal Policy in a Monetary Economy," *Econometrica*, 1978, *46* (6), 1411–1428.

Carroll, Christopher D, "The method of endogenous gridpoints for solving dynamic stochastic optimization problems," *Economics letters*, 2006, *91* (3), 312–320.

Chari, Varadarajan V and Patrick J Kehoe, "Sustainable plans," *Journal of political economy*, 1990, *98* (4), 783–802.

Cohen, Daniel and Philippe Michel, "How should control theory be used to calculate a time-consistent government policy?," *The Review of Economic Studies*, 1988, *55* (2), 263–274.

Corbae, Dean, Pablo D'Erasmo, and Burhanettin Kuruscu, "Politico-economic consequences of rising wage inequality," *Journal of Monetary Economics*, 2009, *56* (1), 43–61.

Dávila, Eduardo and Anton Korinek, "Pecuniary externalities in economies with financial frictions," *The Review of Economic Studies*, 2018, *85* (1), 352–395.

Davila, Julio, Jay H Hong, Per Krusell, and José-Víctor Ríos-Rull, "Constrained efficiency in the neoclassical growth model with uninsurable idiosyncratic shocks," *Econometrica*, 2012, *80* (6), 2431–2467.

Debortoli, Davide, Ricardo Nunes, and Pierre Yared, "Optimal time-consistent government debt maturity," *The Quarterly Journal of Economics*, 2017, *132* (1), 55–102.

_, _, and _, "Optimal fiscal policy without commitment: Revisiting Lucas-Stokey," *Journal of Political Economy*, 2021, *129* (5), 1640–1665.

Domínguez, Begoña, "On the time-consistency of optimal capital taxes," *Journal of Monetary Economics*, 2007, *54* (3), 686–705.

_ , "Public debt and optimal taxes without commitment," *Journal of Economic Theory*, 2007, *135* (1), 159–170.

Dyrda, Sebastian and Marcelo Pedroni, "Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Income Risk," *The Review of Economic Studies*, 2022, *90* (2), 744–780.

Greenwood, Jeremy, Zvi Hercowitz, and Gregory W Huffman, "Investment, capacity utilization, and the real business cycle," *The American Economic Review*, 1988, pp. 402–417.

Haan, Wouter J Den, "Assessing the accuracy of the aggregate law of motion in models with heterogeneous agents," *Journal of Economic Dynamics and Control*, 2010, *34* (1), 79–99.

Itskhoki, Oleg and Benjamin Moll, "Optimal development policies with financial frictions," *Econometrica*, 2019, *87* (1), 139–173.

Jang, Youngsoo, Takeki Sunakawa, and Minchul Yum, "Heterogeneity, Transfer Progressivity and Business Cycles," 2021.

Kim, Heejeong, "Inequality, disaster risk, and the great recession," *Review of Economic Dynamics*, 2021.

Klein, Paul and José-Víctor Ríos-Rull, "Time-consistent optimal fiscal policy," *International Economic Review*, 2003, *44* (4), 1217–1245.

_ **, Per Krusell, and José-Víctor Ríos-Rull**, "Time-consistent public policy," *The Review of Economic Studies*, 2008, *75* (3), 789–808.

Krusell, Per and Anthony A Smith Jr, "Income and wealth heterogeneity in the macroeconomy," *Journal of political Economy*, 1998, *106* (5), 867–896.

_ **and José-Víctor Ríos-Rull**, "On the size of US government: political economy in the neoclassical growth model," *American Economic Review*, 1999, *89* (5), 1156–1181.

_ **, Vincenzo Quadrini, and José-Víctor Ríos-Rull**, "Are consumption taxes really better than income taxes?," *Journal of monetary Economics*, 1996, *37* (3), 475–503.

Kydland, Finn E and Edward C Prescott, "Rules rather than discretion: The inconsistency of optimal plans," *Journal of political economy*, 1977, *85* (3), 473–491.

Laczó, Sarolta and Raffaele Rossi, "Time-consistent consumption taxation," *Journal of Monetary Economics*, 2020, *114*, 194–220.

Lucas, Robert and Nancy Stokey, "Optimal fiscal and monetary policy in an economy without capital," *Journal of Monetary Economics*, 1983, *12* (1), 55–93.

Park, Yena, "Constrained efficiency in a human capital model," *American Economic Journal: Macroeconomics*, 2018, *10* (3), 179–214.

Reiter, Michael, "Recursive computation of heterogeneous agent models," *manuscript, Universitat Pompeu Fabra*, 2002, pp. 28–35.

__ , “Solving the incomplete markets model with aggregate uncertainty by backward induction,” *Journal of Economic Dynamics and Control*, 2010, *34* (1), 28–35.

Rogoff, Kenneth, “The optimal degree of commitment to an intermediate monetary target,” *The quarterly journal of economics*, 1985, *100* (4), 1169–1189.

Rouwenhorst, K Geert, “Asset Pricing Implications of Equilibrium Business Cycle Models. In: Cooley, T.F. (Ed.),” in “Frontiers of business cycle research,” Princeton University Press, 1995, pp. 294–330.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti, “Rotten parents and disciplined children: A politico-economic theory of public expenditure and debt,” *Econometrica*, 2012, *80* (6), 2785–2803.

Straub, Ludwig and Iván Werning, “Positive long-run capital taxation: Chamley-Judd revisited,” *American Economic Review*, 2020, *110* (1), 86–119.

Appendix A Numerical Solution Algorithm

Solving the Markov-Perfect Equilibria (MPE) of consecutive governments entails heavy computational burdens with heterogeneous agents. As in standard macroeconomic heterogeneous agent models, individual decisions should be consistent with the aggregate law of motion for the distribution of agents. On top of that, the aggregate tax policy function must be compatible with individual decisions and the aggregate law of motion for the distribution of agents. In other words, these three equilibrium objects—individual decisions, the law of motion for the distribution, and the tax policy function—have to be consistent with each other in the Markov-perfect equilibrium.

I address this computational issue by taking ideas from the Backward Induction method of Reiter (2010). This method discretizes the aggregate state into finite grid points. For each aggregate grid point, the Backward Induction algorithm allows updating the aggregate law of motion while solving the decision rules thanks to the existence of the proxy distribution. This means that for each aggregate grid point, the backward induction algorithm would make it possible to approximate not only the aggregate law of motion for the distribution; but also the tax policy function consistent with the choice of government lacking commitment. With the value function, this endogenous tax policy outcome can be directly obtained when the proxy distribution is explicitly available.

Unfortunately, the original Reiter's (2010) method cannot directly be applied to the MPE models because the existence of off-the-equilibrium paths makes it challenging to arrange the proxy distribution. In the model of Krusell and Smith (1998), for which the Reiter's (2010) method is originally designed, the distribution of TFP shocks Z is stationary, thus all the aggregate states Z are not measure zero. With a positive probability, all the states Z are realized on the equilibrium path. However, the MPE economy does not have this property. Let us think about a case in a stationary distribution. In this equilibrium, this optimal policy is obtained by comparing among one-time deviation policies. Some tax paths would not be reached at all on the equilibrium path.

I have three variations from the original backward induction method. First, I have to approximate not only the aggregate law of motion for distributions but also the tax policy function that is endogenous. I find these mappings in a non-parametric way, as in Reiter (2010). Second, I arrange distributions for all types of off the equilibrium paths, taking the initial distribution of the simulations as the previous proxy distribution for each aggregate state. Finally, I modify the construction of the reference distributions in Reiter (2002, 2010), reflecting the features of economies in the MPE wherein how many times a policy off-equilibrium takes place is unknown before simulations.

Here, I show how to apply the algorithm to the case with proportional income taxes. Note that I solve all the value functions in the following steps with the Endogenous Grid Method of Carroll (2006).

A.1 Notation and Sketch of the Solution Method

The aggregate law of motion Γ and the tax policy function Ψ are evolved with the distribution μ that is an infinite dimensional equilibrium object, and thus it not not feasible in computations. To handle this issue, the Backward Induction method replaces μ with m, a set of moments from the distribution and discretize it. Here, I take the mean of the distribution and discretize it, $M = \{m_1, \cdots, m_{N_m}\}$. Furthermore, I discretze the tax policy, $T = \{\tau_1, \cdots, \tau_{N_\tau}\}$. This setting allows me to define the aggregate law motion and the tax policy function on each grid (m_{i_m}, τ_{i_τ}) such that $m' = G(m_{i_m}, \tau_{i_\tau}, \tau')$ where $\tau' = P(m_{i_m}, \tau_{i_\tau})$. Note that G and P do not rely on a parametric law.

Across a grid of aggregate states (m_{i_m}, τ_{i_τ}), each point selecting a proxy distribution, the Backward Induction method simultaneously solves for households' decision rules and an intratemporally consistent end-of-period distribution. This implies a future approximate aggregate state consistent with households' expectation ($m' = G(m_{i_m}, \tau_{i_\tau}, \tau')$). Likewise, the backward induction can find the tax policy function that is consistent with the choice of the government, by using household's value functions and the proxy distribution ($\tau^m = \tau' = P(m_{i_m}, \tau_{i_\tau})$). Theses mappings imply that G interacts with P. Given P, first, I find G during the iteration of value functions, and then update P with the value function and proxy distribution. I repeat this until P is convergent.

Given a distribution over individual states at each aggregate grid point (m_{i_m}, τ_{i_τ}), my goal is to obtain the law of motion for households distribution G and the tax policy function P that are intratemporally consistent with the end-of-period distribution and the choice of the government. Explicitly,

$$m' = G(m_{i_m}, \tau_{i_\tau}, \tau') \tag{35}$$

$$\tau' = P(m_{i_m}, \tau_{i_\tau}) \tag{36}$$

$$\tau' = \tau^m(m_{i_m}, \tau_{i_\tau}) \tag{37}$$

$$w = W(m_{i_m}, \tau_{i_\tau}) \tag{38}$$

$$T = TR(m_{i_m}, \tau_{i_\tau}) \tag{39}$$

(35) is to approximate Γ, (36) is to do Ψ, (37) is for the choice of the government, (38) is the mapping for the market wage, and (39) is the mapping for transfers.

The backward induction method explicitly computes G, P, τ^m, W, and T, given a set of proxy distributions before the simulation step. An issue is that computing $G(m_{i_m}, \tau_{i_\tau}, \tau')$ in solving the value is costly because it depends on τ' not only on the equilibrium path but also off the equilibrium path. To address this issue, I reduce $G(m_{i_m}, \tau_{i_\tau}, \tau')$ into $\tilde{G}(m_{i_m}, \tau_{i_\tau}) = G(m_{i_m}, \tau_{i_\tau}, P(m_{i_m}, \tau_{i_\tau}))$ while solving the value function; retrieve $G(m_{i_m}, \tau_{i_\tau}, \tau')$ with the converged value function and the proxy distribution. Note that $G(m_{i_m}, \tau_{i_\tau}, \tau')$ must also satisfy an intratemporal consistency.

A.2 Computing the Aggregate Mappings given a Set of Proxy Distributions

(1) Given $v^n(a,\epsilon;m,\tau)$ and $\tau' = P^q(m,\tau)$, where $n = 1,2,\cdots$ and $q = 1,2,\cdots$ denote the rounds of iteration, at grid (m_{i_m},τ_{i_τ}), where $i_m = 1,\cdots,N_m$ and $i_\tau = 1,\cdots,N_\tau$ are grid indexes, solve for intratemporally consistent m'.

a) Guess m' . Using v^n and P^q, solve for $a' = g_a^{n+1}(a,\epsilon;m_{i_m},\tau_{i_\tau})$ and $n = g_n^{n+1}(a,\epsilon;m_{i_m},\tau_{i_\tau})$ using

$$v^{n+1}(a,\epsilon;m_{i_m},\tau_{i_\tau}) = \max_{c,a',n} u(c,1-n) + \beta\sum_{\epsilon'}\pi_{\epsilon,\epsilon'}v^n(a',\epsilon',m',\tau') \tag{40}$$

such that

$$c + a' = (1-\tau_{i_\tau})w(m_{i_m},\tau_{i_\tau})\epsilon n + (1+(1-\tau_{i_\tau})r(m_{i_m},\tau_{i_\tau}))a + T(m_{i_m},\tau_{i_\tau})$$

$$\tau' = P^q(m_{i_m},\tau_{i_\tau})$$

b) Using the proxy distribution, $\mu(a,\epsilon;m_{i_m},\tau_{i_\tau})$, compute the distribution consistent with capital stock in the end of period $\tilde{m}'$, wage $\tilde{w}$, and transfers $\tilde{T}$.

$$\tilde{m}' = \int g_a^{n+1}(a,\epsilon;m_{i_m},\tau_{i_\tau})\mu(\mathbf{d}(a\times\epsilon);m_{i_m},\tau_{i_\tau}) \tag{41}$$

$$\tilde{w} = (1-\theta)\left(\frac{m_i}{N}\right)^{\theta} \tag{42}$$

$$\tilde{T} = \tau_{i_\tau}(r(m_{i_m},\tau_{i_\tau})m_i + w(m_{i_m},\tau_{i_\tau})N) \tag{43}$$

where

$$N = \int g_n^{n+1}(a,\epsilon;m_{i_m},\tau_{i_\tau})\epsilon\,\mu(\mathbf{d}(a\times\epsilon);m_{i_m},\tau_{i_\tau})$$

c) If $\max\left\{|\tilde{m}'-m'|,|\tilde{w}-w|,|\tilde{T}-T|\right\}$ >precision, update m', w, and T; set $r = \theta\left(\frac{w}{1-\theta}\right)^{\frac{\theta-1}{\theta}} - \delta$; and return to a).

(2) Having found $m' = \tilde{G}^q(m_{i_m},\tau_{i_\tau})$, $w = W^q(m_{i_m},\tau_{i_\tau})$, and $T = TR^q(m_{i_m},\tau_{i_\tau})$, use (40) to define $v^{n+1}(a,\epsilon;m,\tau)$ consistent with $v^n(a',\epsilon';G^q(m_{i_m},\tau_{i_\tau}),P^q(m_{i_m},\tau_{i_\tau}))$. If $||v^{n+1} - v^n|| >$ precision, $n = n+1$ and return to (1).

(3) For each aggregate grid $(m_{i_m},\tau_{i_\tau},\tau'_{i_\tau})$, retrieve $G^q(m_{i_m},\tau_{i_\tau},\tau'_{i_\tau})$ by solving for intratempollay consistent $\hat{m}'$.

a) For each $(m_{i_m}, \tau_{i_\tau}, \tau'_{i_\tau})$, guess $\hat{m}'$. With v^∞, solve for $a' = \hat{g}_a(a, \epsilon; m_{i_m}, \tau_{i_\tau}, \tau'_{i_\tau})$ and $n = \hat{g}_n(a, \epsilon; m_{i_m}, \tau_{i_\tau}, \tau'_{i_\tau})$ using

$$\hat{v}(a, \epsilon; m_{i_m}, \tau_{i_\tau}, \tau'_{i_\tau}) = \max_{c,a',n} u(c, 1-n) + \beta \sum_{\epsilon'} \pi_{\epsilon,\epsilon'} v^\infty(a', \epsilon', m', \tau'_{i_\tau})$$

such that

$$c + a' = (1 - \tau_{i_\tau})\hat{w}(m_{i_m}, \tau_{i_\tau}, \tau'_{i_\tau})\epsilon n + (1 + (1 - \tau_{i_\tau})\hat{r}(m_{i_m}, \tau_{i_\tau}, \tau'_{i_\tau}))a + \hat{T}$$

b) For each $(m_{i_m}, \tau_{i_\tau}, \tau'_{i_\tau})$, using the proxy distribution, $\mu(a, \epsilon; m_{i_m}, \tau_{i_\tau})$, compute the distribution consistent with the end of period aggregate capital stock.

$$\tilde{m}' = \int \hat{g}_a(a, \epsilon; m_{i_m}, \tau_{i_\tau}, \tau'_{i_\tau})\mu(\mathbf{d}(a \times \epsilon); m_{i_m}, \tau_{i_\tau})$$

$$\tilde{w} = (1-\theta)\left(\frac{m_i}{N}\right)^\theta$$

$$\tilde{T} = \tau_{i_\tau}(\hat{r}m_i + \hat{w}N)$$

where

$$N = \int \hat{g}_n(a, \epsilon; m_{i_m}, \tau_{i_\tau}, \tau'_{i_\tau})\epsilon\, \mu(\mathbf{d}(a \times \epsilon); m_{i_m}, \tau_{i_\tau})$$

c) If $\max\left\{|\tilde{m}' - \hat{m}'|, |\tilde{w} - \hat{w}|, |\tilde{T} - \hat{T}|\right\} >$ precision, update $\hat{m}', \hat{w}$, and $\hat{T}$; set $\hat{r} = \theta\left(\frac{\hat{w}}{1-\theta}\right)^{\frac{\theta-1}{\theta}} - \delta$; and return to a).

(4) Having found $m' = G^q(m_{i_m}, \tau_{i_\tau}, \tau'_{i_\tau})$, keep $G^q(m_{i_m}, \tau_{i_\tau}, \tau'_{i_\tau})$. Note that here is no update of the value.

(5) For each aggregate grid (m_{i_m}, τ_{i_τ}), find $\tau^{m,q}(m_{i_m}, \tau_{i_\tau})$.

a) Given $(a, \epsilon; m_{i_m}, \tau_{i_\tau})$, using $\hat{v}(a, \epsilon; m_{i_m}, \tau_{i_\tau}, \tau'_{i_\tau})$ in (3) - a), solve $\Psi^q(m, \tau)$ as follows:

$$\Psi^q(m_{i_m}, \tau_{i_\tau}) = \underset{\tilde{\tau}'}{\operatorname{argmax}}\, \hat{V}(m_{i_m}, \tau_{i_\tau}, \tilde{\tau}') = \int \hat{v}(a, \epsilon; m_{i_m}, \tau_{i_\tau}, \tilde{\tau}')\mu(\mathbf{d}(a \times \epsilon); m_{i_m}, \tau_{i_\tau}) \tag{44}$$

The golden section search is used to find $\Psi^q(m_{i_m}, \tau_{i_\tau})$ with a cubic spline for $\hat{V}$ over τ'.

b) For each aggregate grid (m_{i_m}, τ_{i_τ}), if $P^q(m_{i_m}, \tau_{i_\tau}) = \Psi^q(m_{i_m}, \tau_{i_\tau})$, G^q and P^q are the solutions, given the proxy distribution. Then, go to the next step. Otherwise, they are not the solutions. Take $P^{q+1} = \omega \cdot P^q + (1-\omega) \cdot \Psi^q$, and go back to (1).

A.3 Constructing the Reference Distributions

Until now, I have solved G and P for a given set of proxy distributions. In the following step, I will simulate the economy and update the distribution selection function, as in Reiter (2002, 2010); but, the simulation step in this paper is substantially different from that in his method. He addresses Krusell and Smith (1998) model where aggregate uncertainty exists. Thus, what matters in his papers is to obtain the Ergodic set that is not affected by the initial distribution.

However, in economies without government commitment, it is important to obtain not only distributions on the equilibrium path but also those off the equilibrium path. For example, let us think of an economy without commitment in the stationary equilibrium. Then, there will be a unique value of $\tau^* = P(m^*, \tau^*)$ and $m^* = G(m^*, \tau^*, \tau^*)$. In this case, I may not know the value of other alternatives because this economy has nothing but the unique equilibrium path. This difficulty might lead the previous studies to employ local solution methods in solving this type of the MPE. By constrast, my approach is a global solution method, which means I need proxy distributions over all types of off the equilibrium paths.

To reserve distributions off the equilibrium path, I use the proxy distributions in the previous step as the initial distribution for the simulation. For each (m_{i_m}, τ_{i_τ}), I run a simulations for T periods from the proxy distribution $\mu_0 = \mu(a, \epsilon; m_{i_m}, \tau_{i_\tau})$, implying the number of simulations is $N_m \times N_\tau$ and that of simulation outcomes is $T \times N_m \times N_\tau$. Note that any type of (m_{i_m}, τ_{i_τ}) will be observed at least once in the simulations. For each (m_{i_m}, τ_{i_τ}), using $\mu_0 = \mu(a, \epsilon; m_{i_m}, \tau_{i_\tau})$ and v^∞ from the previous step, I simulate the economy in a forward manner. I compute the market cleared w_t and r_t and transfers T_t satisfying the government budget condition for each simulation period $t = 1, \cdots, T$. In addition, I solve the government's problem τ_t^m for each simulation period $t = 1, \cdots, T$ with the $m' = G(m_{i_m}, \tau_{i_\tau}, \tau'_{i_\tau})$ obtained in the previous step.

I gather all the simulated distributions and rearrange the index as $\tilde{t} = 1, \cdots, T \times N_m \times N_\tau$. In creating the reference distributions from the simulation, I need a measure of distance for the moments of a distribution. For (m, τ), define an inverse norm

$$d((m_0, \tau_0), (m_1, \tau_1)) = (m_0 - m_1)^{-4} + (\tau_0 - \tau_1)^{-4} \tag{45}$$

In contrast to an economy with uncertainty, the initial simulation results should be preserved, having to be used to construct the reference distributions off the equilibrium path (non-Ergodic set). For each t, when (m_t, τ_t) with $m_t \in [m_k, m_{k+1})$ and $\tau_t \in [\tau_s, \tau_{s+1})$,

$$\begin{aligned}
d_1(m_k, \tau_s) &= d_1(m_k, \tau_s) + (m_t - m_k)^{-4} + (\tau_t - \tau_s)^{-4} \\
d_1(m_{k+1}, \tau_s) &= d_1(m_{k+1}, \tau_s) + (m_t - m_{k+1})^{-4} + (\tau_t - \tau_s)^{-4} \\
d_1(m_k, \tau_{s+1}) &= d_1(m_k, \tau_{s+1}) + (m_t - m_k)^{-4} + (\tau_t - \tau_{s+1})^{-4} \\
d_1(m_{k+1}, \tau_{s+1}) &= d_1(m_{k+1}, \tau_{s+1}) + (m_t - m_{k+1})^{-4} + (\tau_t - \tau_{s+1})^{-4}
\end{aligned}$$

Above m_k (τ_s) is the k-th (s-th) grid point for m (τ). Note that distances between a given node and non-adjacent moments are not taken into account, which is different from the corresponding step in Reiter (2002, 2010).

I construct the reference distributions for each (m_{i_m}, τ_{i_τ}) using the above, when $(m_{\tilde{t}}, \tau_{\tilde{t}}) \in \Big([m_{i_m}, m_{i_m+1}), [\tau_{i_\tau}, \tau_{i_\tau+1})\Big)$,

$$\mu^r(a, \epsilon; m_{i_m}, \tau_{i_\tau}) = \sum_{\tilde{t}=1}^{T \times N_m \times N_\tau} \frac{d((m_{i_m}, \tau_{i_\tau}), (m_{\tilde{t}}, \tau_{\tilde{t}}))}{d_1(m_{i_m}, \tau_{i_\tau})} \mu_{\tilde{t}}(a, \epsilon). \tag{46}$$

Each reference distribution is a weighted sum of distributions over the simulation only when simulated moments are adjacent to a given pair of grid points (m_{i_m}, τ_{i_τ}). Since the simulation moments are not on an Ergodic set, this should be considered.

I arrange the finite grid, which is the distribution support, as explicit. The distribution over (a, ϵ) used below size $(N_a \times N_\epsilon)$ with $\epsilon \in E = \{\epsilon_1, \cdots, \epsilon_{N_\epsilon}\}$ and $a \in A = \{a_1, \cdots, a_{N_a}\}$. I represent $\mu^r(a, \epsilon; m_{i_m}, \tau_{i_\tau})$ using $\mu^r_{i_a, i_\epsilon}(i_m, i_\tau)$, indexing $(a_{i_a}, \epsilon_{i_\epsilon})$ over $A \times E$ for (m_{i_m}, τ_{i_τ}). The moment of a reference distribution, $\sum_{i_\epsilon}^{N_\epsilon} \mu^r_{i_a, i_\epsilon}(i_m, i_\tau) a_{i_a}$, will not be consistent with m_{i_m}. However, the proxy distribution at (i_m, i_τ) will have this property.

A.4 Updating the Proxy Distributions

Following Reiter (2002, 2010), for each aggregate grid (i_m, i_τ), I solve for μ_{i_a, i_ϵ}, the proxy distribution, as the solution to a problem that minimizes the distance to the reference distribution while imposing that each type of sums to its reference value and moment consistency.

$$\min_{\{\mu_{i_a,i_\epsilon}\}_{i_a=1,i_\epsilon=1}^{N_a,N_\epsilon}} \sum_{i_a=1}^{N_a}\sum_{i_\epsilon=1}^{N_\epsilon}\left(\mu_{i_a,i_\epsilon}-\mu^r_{i_a,i_\epsilon}(i_m,i_\tau)\right)^2 \tag{47}$$

$$\sum_{i_a=1}^{N_a}\mu_{i_a,1_\epsilon}=\sum_{i_a=1}^{N_a}\mu^r_{i_a,i_\epsilon}(i_m,i_\tau) \text{ for } i=1,\cdots,N_\epsilon \tag{48}$$

$$\sum_{i_\epsilon=1}^{N_\epsilon}\sum_{i_a=1}^{N_a}\mu_{i_a,i_\epsilon}\cdot a_{i_a}=m_{i_m} \tag{49}$$

$$\mu_{i_a,i_\epsilon}\geq 0 \tag{50}$$

The first-order condition for μ_{i_a,i_ϵ} with λ_i as the LaGrange multiplier for (48) and ω the multiplier (49) is

$$2(\mu_{i_a,i_\epsilon}-\mu^r_{i_a,i_\epsilon}(i_m,i_\tau))-\lambda_i-\omega\cdot a_{i_a}=0 \tag{51}$$

If I ignore the non-negative constraints for probabilities in (50), I have N_ϵ constraint in (48). 1 constraint in (49) and $N_a\times N_\epsilon$ first-order conditions in (50). These are a system of $N_a\times N_\epsilon+N_\epsilon+1$ linear equations in $\left(\{\mu_{i_a,i_\epsilon}\}_{i_a=1,i_\epsilon=1}^{N_a,N_\epsilon},\{\lambda_{i_\epsilon}\}_{i_\epsilon}^{N_\epsilon},\omega\right)$.

I construct a column vector $\mathbf{x}$. The first block of $\mathbf{x}$ are the stack of the elements from the proxy distribution, such that $\mathbf{x}(j)=\mu_{i_a,i_\epsilon}$ where $j=(i_\epsilon-1)\times N_a+i_a$. Next are the N_ϵ multipliers λ_i, followed by one multiplier ω. I solve for $\mathbf{x}$ using a system of linear equations, $\mathbf{Ax}=\mathbf{b}$ in Figure 10. The non-zero element of $\mathbf{A}$ and $\mathbf{b}$ are described here. The coefficients for μ_{i_a,i_ϵ} are entered into $\mathbf{A}$ as

$$\mathbf{A}((i_\epsilon-1)\times N_a+i_a,(i_\epsilon-1)\times N_a+i_a)=2 \tag{52}$$

$$\mathbf{A}(N_\epsilon\times N_a+i_\epsilon,(i_\epsilon-1)\times N_a+i_a)=1 \text{ for } i_\epsilon=1,\cdots,N_\epsilon \tag{53}$$

$$\mathbf{A}(N_\epsilon\times N_a+N_\epsilon+1,(i_\epsilon-1)\times N_a+i_a))=a_{i_a}. \tag{54}$$

The coefficient for λ_i are entered in $\mathbf{A}$, for $i_\epsilon=1,\cdots N_\epsilon$ and $i_a=1,\cdots,N_a$, as

$$\mathbf{A}((i_\epsilon-1)\times N_a+i_a,N_\epsilon\times N_a+i_\epsilon)=-1 \tag{55}$$

The coefficients for ω sets the following elements of $\mathbf{A}$, for $i_\epsilon=1,\cdots N_\epsilon$ and $i_a=1,\cdots,N_a$,

$$\mathbf{A}((i_\epsilon-1)\times N_a+i_a,N_\epsilon\times N_a+N_\epsilon+1)=-a_{i_a}. \tag{56}$$

The elements of $\mathbf{b}$ are, for $i_\epsilon = 1, \cdots N_\epsilon$ and $i_a = 1, \cdots, N_a$,

$$\mathbf{b}((i_\epsilon - 1) \times N_a + i_a) = 2\mu^r_{i_a, i_\epsilon}(i_m, i_\tau) \tag{57}$$

$$\mathbf{b}(N_\epsilon \times N_a + i_\epsilon) = \sum_{i_a=1}^{N_a} \mu^r_{i_a, i_\epsilon}(i_m, i_\tau) \tag{58}$$

$$\mathbf{b}(N_\epsilon \times N_a + N_\epsilon + 1) = m_{i_m}. \tag{59}$$

I solve $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ iteratively using an active set method corresponding to probabilities that are not set to 0.

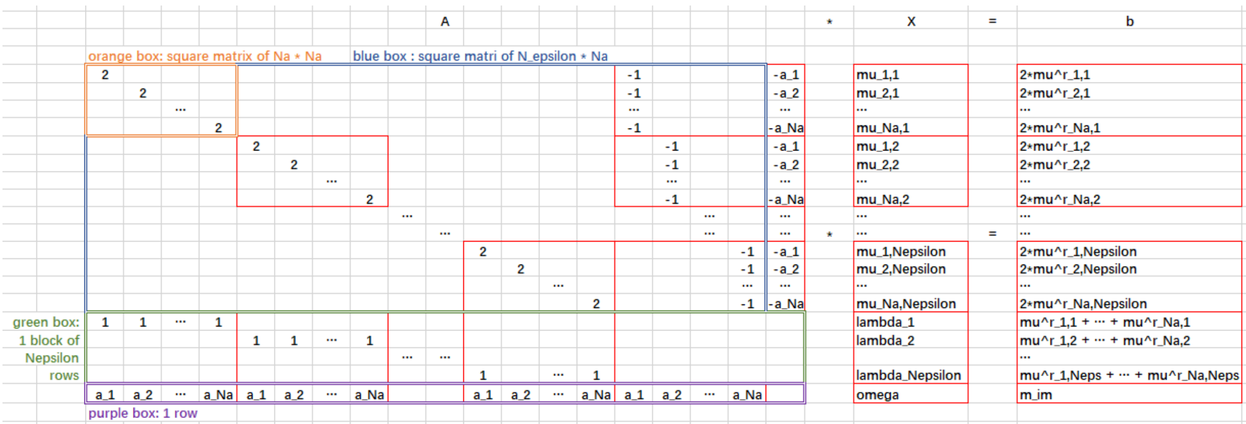

Figure 10: $\mathbf{A} \times \mathbf{x} = \mathbf{b}$

To solve the linear system, I use the active set approach to non-negative constraints in Reiter (2002, 2010). If any of the first $N_\epsilon \times N_a$ elements of $\mathbf{x}$ are negative, the constraint $\mu_{i_a, 1_\epsilon} \geq 0$ has been violated for some $(i_\epsilon - 1)N_a + i_a = j \in J_0$ where

$$J_0 = \{j | 1 \leq j \leq N_a \times N_\epsilon \text{ and } \mathbf{x}(j) < 0\}. \tag{60}$$

For some $O > 0$, set the most negative O elements indexed in J_0 to 0, $\mu_{i_a, i_\epsilon} = 0$. Remove the $j - th$ row and column of A along with the $j - th$ element of $\mathbf{b}$. Solve the reduced system with O less rows. If any of the $N_\epsilon \times (N_a - O)$ elements are negative, again discard the most negative O. I repeat this procedure until the most negative elements of $\mathbf{x}$ is larger than a precision level. This iteratively implements the non-negativity of probabilities (60).

Table 4 shows the setting of the grids in this paper. With this setting, I continue to repeat the whole steps above until no improvement in accuracy statistic proposed by Den Haan (2010). I find that the mean errors on the equilibrium path are sufficiently small (considerably less than 0.6% for all cases) and the mean errors over transitions from off the equilibrium to the equilibrium are also reasonably small (not exceeding 0.6% for all cases). Furthermore, the method is substantially

Table 4: Setting for Computation

	num. of nodes	Description
N_a	400(400)	asset (distribution)
N_ϵ	10	persistence wage process
N_m	5	aggregate capital (aggregate)
N_τ	7	income tax (aggregate)

efficient in a usual personal computer.

Table 5: Accuracy and Efficiency of the Solution Method

	OPT w/o Commitment
Run time	11.1 min
DH of m at EQ	0.394%
DH of w at EQ	0.048%
DH of τ at EQ	0.153%
AVG(DH) of m	0.668%
AVG(DH) of w	0.251%
AVG(DH) of τ	0.129%
MAX(DH) of m	2.133%
MAX(DH) of w	0.949%
MAX(DH) of τ	0.415%

$AVG(\cdot)$ and $MAX(\cdot)$ are computed with all of the results both on and off the equilibrium paths.
Processor: i7-10770 @ 2.9GHz, RAM: 16GB

Appendix B Ramsey Problem with Capital Income Tax

The results are consistent with those in Dyrda and Pedroni (2022). This front-loaded capital income tax is to quickly reduce overall inequality. Afterward, the Ramsey planner balances between inequality and redistribution over the transition path.

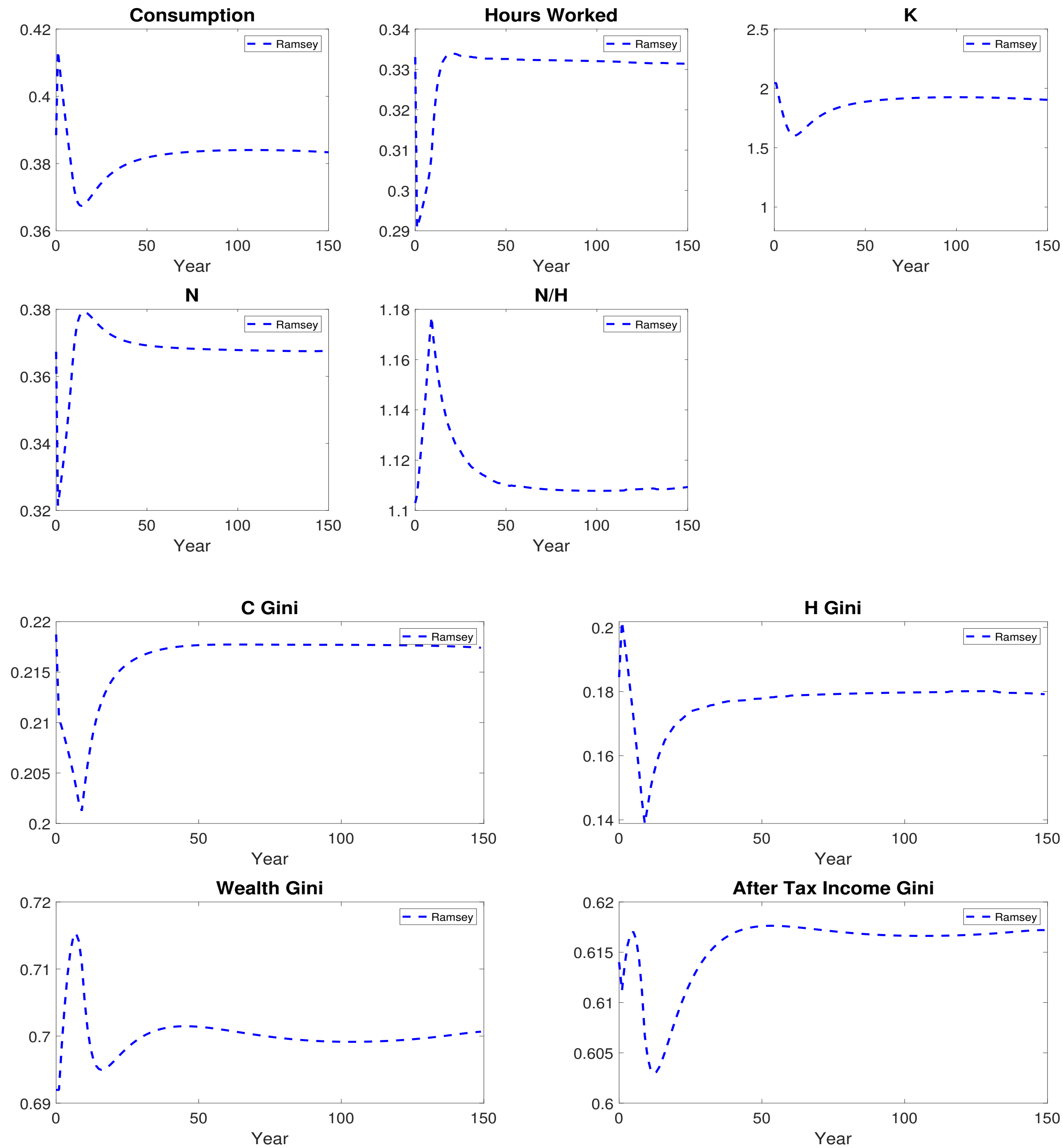

Figure 11: Results with Capital Income Tax in the Ramsey Problem

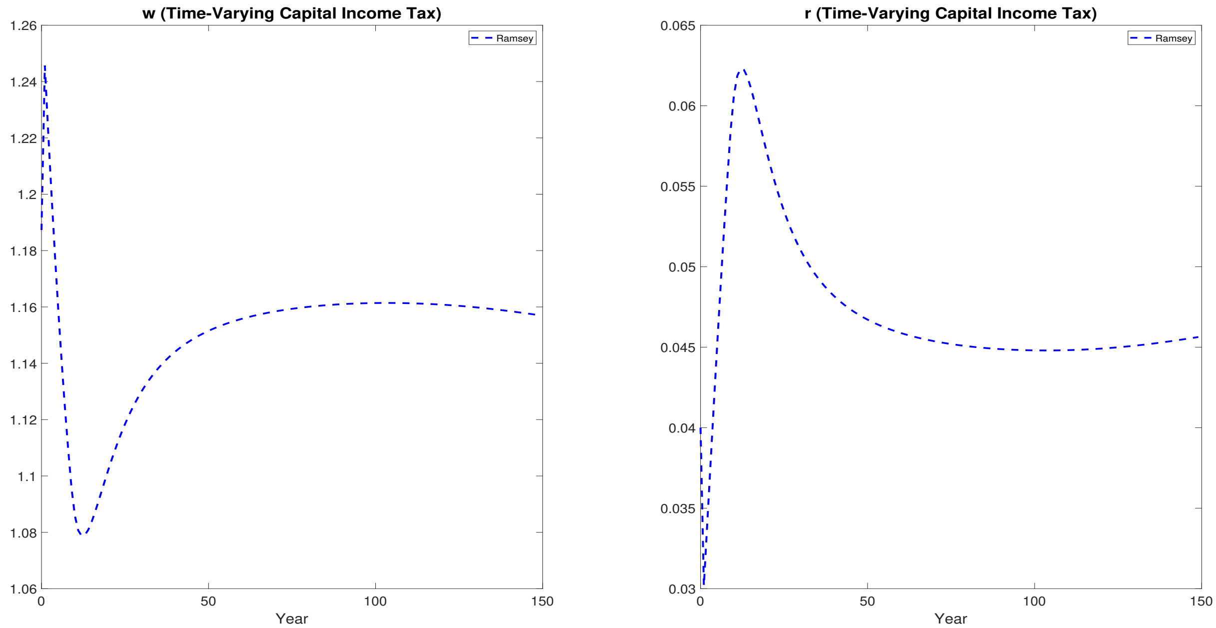

Figure 12: Time-Inconsistent Capital Income Taxes: Dynamics of w and r

Appendix C Definition of $\frac{\Delta K_{t+s}}{\Delta \tau_{t+1}}$ and $\frac{\Delta \tau_{t+s}}{\Delta \tau_{t+1}}$

$$\frac{\Delta K_{t+s}}{\Delta \tau_{t+1}} = \frac{\partial K_{t+1}}{\partial \tau_{t+1}} \cdot \frac{\Xi \Gamma_{t+s-1}}{\Xi K_{t+1}} + \frac{\Xi \Gamma_{t+s-1}}{\Xi \tau_{t+1}}$$
$$\frac{\Delta \tau_{t+s}}{\Delta \tau_{t+1}} = \frac{\partial K_{t+1}}{\partial \tau_{t+1}} \cdot \frac{\Xi \Psi_{t+s-1}}{\Xi K_{t+1}} + \frac{\Xi \Psi_{t+s-1}}{\Xi \tau_{t+1}}$$

where

$$\frac{\Xi \Gamma_{t+s-1}}{\Xi K_t} = F^K_{t+s}(F^K_{t+s-1}, G^K_{t+s-1}) = \begin{cases} \Gamma_{K_t} & \text{if s=1} \\ \Gamma_{K_{t+s-1}} F^K_{t+s-1} + \Gamma_{\tau_{t+s-1}} G^K_{t+s-1} & \text{if } s \geq 2 \end{cases}$$

$$\frac{\Xi \Psi_{t+s-1}}{\Xi K_t} = G^K_{t+s}(F^K_{t+s-1}, G^K_{t+s-1}) = \begin{cases} \Psi_{K_t} & \text{if s=1} \\ \Psi_{K_{t+s-1}} F^K_{t+s-1} + \Psi_{\tau_{t+s-1}} G^K_{t+s-1} & \text{if } s \geq 2 \end{cases}$$

$$\frac{\Xi \Gamma_{t+s-1}}{\Xi \tau_t} = F^\tau_{t+s}(F^\tau_{t+s-1}, G^\tau_{t+s-1}) = \begin{cases} \Gamma_{\tau_t} & \text{if s=1} \\ \Gamma_{K_{t+s-1}} F^\tau_{t+s-1} + \Gamma_{\tau_{t+s-1}} G^\tau_{t+s-1} & \text{if } s \geq 2 \end{cases}$$

$$\frac{\Xi \Psi_{t+s-1}}{\Xi \tau_t} = G^\tau_{t+s}(F^\tau_{t+s-1}, G^\tau_{t+s-1}) = \begin{cases} \Psi_{\tau_t} & \text{if s=1} \\ \Psi_{K_{t+s-1}} F^\tau_{t+s-1} + \Psi_{\tau_{t+s-1}} G^\tau_{t+s-1} & \text{if } s \geq 2 \end{cases}$$