Taxation without Commitment in a Heterogeneous-Agent Economy

Jang, Youngsoo

March 2022

Taxation without Commitment in a Heterogeneous-Agent Economy*

Youngsoo Jang†

March 2022

Abstract

How do differences in the government's political and commitment structure affect the aggregate economy, inequality, and welfare? I examine this question using a standard incomplete markets model with uninsurable idiosyncratic income risk, wherein a flat tax rate and transfers are endogenously determined according to its political and commitment structure. I compare three economies over the transitional path: an economy with the optimal tax with commitment, an economy with the optimal tax without commitment, and a political economy with sequential voting. Using the generalized Euler equation, I characterize the Markov perfect equilibria of the cases without commitment. Additionally, through quantitative exercises, I obtain two main findings. First, a lack of commitment hinders the government from managing the evolution of inequalities in the long run but instead makes it pursue more income with a lower tax rate in the short run. This incapability results in substantially lower welfare in the case without commitment. Second, given a lack of commitment, the economy with sequential voting yields significantly different macroeconomic and distributional implications from the economy with the optimal policy. In the political economy, the government considers only the interests of the median voter, who is middle class and reluctant to bear larger distortions from a higher tax rate. Therefore, the political economy becomes more efficient but less equal, leading to a worse welfare outcome with the utilitarian criterion.

JEL classification: E61, H11, P16.

Keywords: Income Tax, Commitment, Time-Consistent Policy, Political Economy, Voting

---

*I would like to thank José-Víctor Ríos-Rull for his extensive comments and suggestions that have improved the paper substantially. I also have benefited from helpful comments by Adrien Auclert, Juan Carlos Conesa, Tatyana Koreshkova, Dirk Krueger, Virgiliu Midrigan, Makoto Nakajima, Svetlana Pashchenko, Ponpoje Porapakkarm, Gianluca Violante, Minchul Yum, seminar participants at the University of Melbourne, and session participants at the Jinan IESR Quantitative Macro Workshop. Discussions with Ji-Woong Moon have been constructive at all stages of this project. I also thank Yunho Cho, Pantelis Kazakis, Qian Li, and Hoonsuk Park for their generous help. A previous version of the paper circulated with the title "Democracy or Optimal Policy: Income Tax Decisions without Commitment". All remaining errors are mine.

†Institute for Advanced Research, Shanghai University of Finance and Economics, No.111 Wuchuan Road, Shanghai, China 200433. E-mail: jangys724@gmail.com

1 Introduction

Public and fiscal policies are essentially subject to a lack of government commitment because political procedures sequentially determine the policy executor. Previous studies have found that a lack of commitment can yield substantial differences in the implications of designing and implementing policies (Kydland and Prescott, 1977; Calvo, 1978; Barro and Gordon, 1983; Lucas and Stokey, 1983; Klein and Ríos-Rull, 2003; Klein, Krusell and Ríos-Rull, 2008). However, relatively few studies have considered how the government's political structure affects the design of public and fiscal policies, given the difficulty in devising a proper framework. Investigating this issue requires models that incorporate heterogeneous agents because political decisions—from selecting policymakers to implementing policies—widely interact at the individual level. In addition to heterogeneous agents, because a lack of commitment leads successive governments to make strategic choices, solving a dynamic game of consecutive governments is essential.

In this direction of research, there are two issues. The first is that not much is known about the theoretical features of the equilibrium of this dynamic game of successive governments with heterogeneous agents. Although a few studies—e.g., Klein, Krusell and Ríos-Rull (2008)— have succeeded in analyzing the strategic behavior of consecutive governments in Markov perfect equilibria (MPE), they do not examine how the government's strategic choice interacts with individual heterogeneity. A theoretical characterization for the MPE with heterogeneous agents is required to understand the underlying economic forces behind the interactions of successive, strategic governments with individual heterogeneity.

The second issue is that solving this dynamic game entails a substantial computational burden. For example, political-economy models, originally developed by Krusell, Quadrini and Ríos-Rull (1996); Krusell and Ríos-Rull (1999), have three types of equilibrium objects—individual decisions, the aggregate law of motion for the distribution of households, and the endogenous government policy function—that have to be consistent with one another in equilibrium. One might consider using Krusell and Smith's (1998) method to achieve their consistency; however, this approach is ineffective for this class of models. First, more than one aggregate law of motion increases the computational burden exponentially in this simulation-based method. The existence of the government policy function leads to adding another outer loop to the outer loop in their method. Second, the government policy function is severely nonlinear because political decisions, which shape the government policy function, are sensitive to the distribution of individuals. This nonlinearity is not well-captured by the parameterized law of motion in Krusell and Smith (1998).

In this paper, I (i) characterize the Markov perfect equilibria of the dynamic game of consecutive governments in a heterogeneous agent model with uninsurable idiosyncratic income risk; (ii) develop a numerical computational method for solving this game. For the characterization of the MPE, I have employed the generalized Euler equation (GEE) approach, proposed initially by Klein et al. (2008). Obtained by computing the first-order condition of the government's choice, the GEE provides insights into the underlying economic forces the government considers in making policy decisions. Regarding my numerical solution method, to handle the aforementioned computational issues, I take ideas from the backward induction method of Reiter (2010). Because the backward induction method is not designed for economies in the MPE but for those with aggregate uncertainty, I make variations to this method to address the characteristics of the MPE—e.g., the existence of off the equilibrium—while preserving its computational benefits. My method is a non-simulation-based approach as in Reiter (2010), which substantially improves computational efficiency. Furthermore, my solution method approximates the aggregate laws of motion, including the government policy function, through a non-parametric approach as in Reiter (2010), thereby enabling me to capture the nonlinearity.

I characterize and solve for the MPE of this game in the canonical model of Aiyagari (1994) with wealth effects of labor supply, in which the government's tax/transfer system is endogenously determined according to its political and commitment structure. I assume a simple government financing rule to better understand the fundamental roles of the political and commitment structure: the government levies a flat tax from labor and capital income and redistributes its revenue to households through lump-sum transfers after covering a given size of government spending.

Specifically, I compare three economies: an economy with the time-inconsistent optimal income tax with the Ramsey planner (with commitment), an economy with the time-consistent optimal tax (without commitment), and a political economy with sequential voting (without commitment). In the economy with the time-inconsistent optimal policy, because the government can commit to all future tax policies, it chooses a sequence of income taxes that maximizes the utilitarian welfare function over the transitional path. By contrast, in the time-consistent optimal case, the government can only decide a tax rate for the next period and cannot commit to it after that. Thus, the government sequentially chooses a tax policy maximizing the utilitarian welfare function under this commitment constraint, and this action continues perpetually. Finally, following a seminal study by Krusell and Ríos-Rull (1999), the political economy with voting has two political parties whose unique goal is to win election in each period, meaning a lack of commitment, through the majority's support. The two parties propose tax rates on which households vote. Because the policy dimension is one in my policy exercise, the dominant strategy of the two parties is to offer the most preferred tax rate of the median voter. To prevent multiple equilibria, I assume that one party always wins when votes are tied.

In characterizing the MPE, I find that when deciding on taxes, the government strikes a balance between two economic forces: the pecuniary externality—coming from income changes caused by variations in the factor composition of individual income—and income redistribution—caused by altered transfers following a tax rate change. Note that in response to a change in income tax rate, these two forces oppositely affect individual welfare. An increase in income tax rate, for example, decreases overall individual savings, thereby reducing the aggregate capital. This reduction in the aggregate capital increases the equilibrium interest rate but reduces the equilibrium market wage. These factor price changes have different welfare implications across individuals. The consumption-poor (-rich), whose income composition tends to be toward labor (capital) income, stand to lose (gain) welfare. This pecuniary externality is considered by the government.

Meanwhile, an increased tax rate raises the level of lump-sum transfers, which increases (reduces) after-tax incomes for the consumption-poor (-rich). Because these two forces play an opposite role for each individual, the government weighs the two forces, the extent to which is hugely reliant on the political structure. Under the utilitarian government, the optimal policy tends to be more in favor of the consumption-poor, who prefers larger redistribution through more substantial transfers. But with sequential voting, income taxes are determined solely by the median voter, who is reluctant to bear larger distortions from a higher tax rate.

Another notable feature revealed by the GEE is that when deciding on taxes, the government in the MPE does not consider distortions—caused by decisions on consumption and labor—at the individual level. This result contrasts with findings in the optimal tax literature. This stark difference occurs because the economies in this model are not centralized with a social planner but decentralized with a government making endogenous decisions. Because individuals optimally make decisions on consumption, savings, and labor in competitive equilibrium, the government has no room for improvement regarding individual allocations. Therefore, the government focuses on balancing the pecuniary externality via variations in the factor composition of income and income redistribution through transfers.

For quantitative exercises, I solve for and compare the aforementioned three economies. To solve the Ramsey problem, I have employed the numerical approach in Dyrda and Pedroni (2022). I solve for the economy with the time-consistent optimal income tax and the political economy with sequential voting by using my numerical solution method. Then, I compare their equilibrium results over the transitional path.

In the first set of these exercises, to understand the effects of the commitment structure, I compare the economy with the time-inconsistent optimal income tax—Ramsey solution—to the economy with the time-consistent optimal policy. I find that the commitment instrument makes substantial differences in the aggregate economy, inequality, and welfare. The economy with the Ramsey planner is less efficient but more equal by levying larger income taxes. Moreover, welfare gain, measured by the consumption equivalent variation, in the economy with the time-inconsistent optimal income tax (+2.19 percent) is much greater than that with the time-consistent case (+0.59 percent). These findings imply that the Ramsey planner achieves a better welfare outcome by putting less weight on efficiency and managing the evolution of inequalities in consumption and leisure, through the controlling of after-tax incomes with a sequence of transfers. This management for a sequence of transfers is possible thanks to the commitment instrument, which is unavailable to the government without commitment. A lack of commitment leads the government with the time-consistent optimal income tax to focus on sparing more income in the short run, putting less weight on the management of inequalities in the long run.

In the other exercise, I examine the effects of the political structure by comparing the economy with the time-consistent optimal policy and the political economy with sequential voting. I find that given a lack of commitment, the political structure brings about non-negligible differences in the macroeconomy, inequality, and welfare. The political economy shows more frugal income taxes than the optimal policy economy, thus leading to a more efficient but less equal economy. In my model, because the dimension for tax is single, the majority voting rule boils down to the median voting rule. As a result, the equilibrium income tax in the political economy is equivalent to income tax most preferred by the median voter. Because this median voter is an individual, it considers only his interest. When the median voter is substantially different from the consumption-poor, which is observed in my model result, he is reluctant to bear substantial distortions. Instead, he prefers a lower tax rate while not considering a desirable level of inequality. As a result, welfare, measured by the utilitarian criterion, is much lower in the political economy (-2.13 percent) than in the economy with the time-consistent optimal income tax (+0.57 percent).

These findings imply that commitment and political structure are hard to ignore because they bring different policy designs that cause disparities in the aggregate economy, inequality, and welfare.

This paper belongs to the stream of political macroeconomic literature that examines the implications of governments' political and commitment structure in designing public policies. Motivated by the seminal studies of Aiyagari and Peled (1995); Krusell, Quadrini and Ríos-Rull (1996); Krusell and Ríos-Rull (1999), several papers have investigated the effects of the political procedure on policy decisions from a macroeconomic perspective. Corbae, D'Erasmo and Kuruscu (2009) studies how political governments make decisions on income taxation in response to the increased inequality in wages in the U.S. They find that the increased inequality in wages raises the equilibrium income tax rate without commitment. The study of Corbae et al. (2009) is similar to my work in the sense that both studies compare a series of economies with heterogeneous agents according to the political and commitment structure of the government. However, in contrast to my model, Corbae et al. (2009) employs the preference of Greenwood, Hercowitz and Huffman (1988), which lacks wealth effects of labor supply and focuses on analyzing equilibrium results in the long-run.[1]

[1] The authors mention that this choice is made to mitigate the computational burden. Note that the wealth effects of labor supply are crucial for the macroeconomy and welfare because changes in transfers affect efficiency through this channel.

Song, Storesletten and Zilibotti (2012) is another study using a political economy. Their goal is to understand intergenerational conflict through public policy instruments. The different objective leads to a different model selection. While they consider an overlapping generations model in partial equilibrium, this paper uses an infinite-horizon model in general equilibrium. Farhi, Sleet, Werning and Yeltekin (2012) is also related to my work because they address the choice of income tax without commitment. However, their approach is different from mine. Whereas Farhi et al. (2012) solves the planner's centralized problem in a dynamic Mirrleesian model, I solve households' decentralized problems in an incomplete markets model with uninsurable idiosyncratic risk.

The characterization of the MPE in this paper relies hugely on two studies: Klein, Krusell and Ríos-Rull (2008) and Davila, Hong, Krusell and Ríos-Rull (2012). However, the results in this paper show some stark differences from these two papers. Klein et al. (2008) introduces the generalized Euler equation approach in a model without commitment. However, because their model lacks individual heterogeneity, the underlying economic forces are substantially different. In their model, the government's policy choice is closely related to distortions related to decisions on consumption and labor supply. In my model, such distortions are borne by individuals, and the government does not care about them. Instead, the government internalizes two economic forces driven by individual heterogeneity—the pecuniary externality related to the income composition and income redistribution with altered transfers. Davila et al. (2012) investigates the features of constrained efficient allocation in the standard incomplete markets model. They find that constrained efficiency can be achieved when the social planner takes into account both distortions related to dynamic consumption allocation at the individual level and the pecuniary externality driven by changes in the factor composition of income at the aggregate level. In my model, the government considers not distortions related to individual dynamic consumption allocation but the pecuniary externality. This difference occurs because, in my model, the government is endogenous but not a social planner, and individuals optimally make decisions on consumption in competitive equilibrium.

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 Markov-perfect equilibria. The first is Klein, Krusell and Ríos-Rull's (2008) approach, which is a local solution method using the generalized Euler equation. This method is accurate and efficient but not general enough to handle the class of heterogeneous agent models. My method in this paper is a global solution method applicable to heterogeneous agent models. The other approach is 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 with 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, the endogenous policy function could show a severe non-linearity that is not well-captured by the parameterized law of motion in Krusell and Smith's (1998) method. My method is an efficient non-simulation-based solution approach that captures the non-linearity through 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 using the generalized Euler equation. 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 function, heterogeneous households make decisions on consumption, savings and labor supply at the intensive margin, as in standard incomplete markets models. A notable difference from the standard models is the setting of its tax policy. The tax policy is endogenously determined, according to the political and commitment structure of government. In equilibrium, the tax policy, individual decisions, and the evolution of the distribution are consistent with one another under the political and commitment structure.

2.1 Environment

The model economy is populated by a continuum of infinitely lived households. Their preference follows

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

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.

It is worth spending more time on the above preference. Note that the preference here captures wealth effects of labor supply. By contrast, Corbae, D'Erasmo and Kuruscu (2009) employed the preference in Greenwood, Hercowitz and Huffman (1988) that lacks wealth effects of labor supply, to mitigate the computational burden. Such wealth effects are crucial for welfare analysis, closely related to efficiency loss. An increase in transfers, for example, decreases overall labor supply, shrinking the size of the aggregate economy and playing a role in reducing welfare.

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

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

where 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.

The government obtains its tax revenue by levying taxes on household capital and labor income at proportional flat tax rate, τ_t.[2] Given a tax revenue, the government covers government spending G, and the rest is used for lump-sum transfers/tax T_t. The government runs a balanced budget each period:

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

Note that T_t can be either transfers (positive) or a tax (negative).

2.2 Competitive Equilibrium, Exogenous Policy

In this section, I define competitive equilibrium, given an exogenous tax policy first. Let me 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.

[2] In a later section, I am relaxing this assumption.

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 solves

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

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 income taxes $\{\tau_t\}_{t=0}^{\infty}$, a sequential competitive equilibrium (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 $\{v_t(\cdot)\}_{t=0}^{\infty}$, a sequence of distributions over the state space $\{\mu_t(\cdot)\}_{t=1}^{\infty}$, such that for all t

(i) Given $\{\tau_t\}_{t=0}^{\infty}$ 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 fucntion.

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

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

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

(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

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

On the other hand, 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. 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 $\mu' = \Gamma(\mu,\tau,\tau')$ 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 $\Psi(\mu,\tau)$, a recursive competitive equilibrium (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:

$$w(\mu) = (1-\theta)\left(\frac{K}{N}\right)^{\theta} \tag{13}$$
$$r(\mu) = \theta\left(\frac{K}{N}\right)^{\theta-1} - \delta. \tag{14}$$

(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 income tax is endogenously determined. I model the tax choice in three ways: the time-inconsistent optimal income tax with commitment (Ramsey problem); the time-consistent optimal income tax without commitment; and income tax determined by majority voting. Let me begin with the Ramsey problem.

Definition 2.3.1. ***The Ramsey Problem:***
A SEC with the Time-inconsistent Optimal Income Tax with Commitment
Given μ_0, the government chooses $\{\tau_t\}_{t=0}^{\infty}$ such that

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

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 that 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-inconsistent issue.

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

Definition 2.3.2. ***A 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

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

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 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 $\Psi(\mu, \tau) = \tau^{WO}(\mu, \tau)$.

In the economy with the optimal income tax without commitment, the government implements the 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. Note that the government cannot commit to the future tax rate from the period after the next period. Thus, once a chosen tax rate $\tilde{\tau}'$ deviates from the equilibrium tax policy function $\Psi(\cdot)$, tax rates thereafter follow the equilibrium tax policy function $\Psi(\cdot)$ because the government cannot commit to the future tax policy after one period. (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, which is shown in (19). In equilibrium, for each aggregate state (μ, τ), the government's choice $\tau^{WO}(\mu, \tau)$ should be equal to the equilibrium tax function $\psi(\mu, \tau)$, which is presented in (iv).

Definition 2.3.3. ***A Political RCE with Sequential Voting.***

(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 $(a, \epsilon; \mu, \tau)$, households choose $\psi(a, \epsilon; \mu, \tau)$ such that

$$\psi(a, \epsilon; \mu, \tau) = \underset{\tilde{\tau}'}{\text{argmax}} \int \hat{V}(a, \epsilon; \mu, \tau, \tilde{\tau}')\mu(\mathbf{d}(a \times \epsilon)) \tag{20}$$

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

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

$$\tau' = \tilde{\tau}', \text{ and thereafter } \tau'' = \Psi(\mu', \tau' = \tilde{\tau}') \tag{21}$$

$$\mu' = \Gamma(\mu, \tau, \tilde{\tau}), \text{ and thereafter } \mu'' = \Gamma(\mu', \tilde{\tau}, \tau'' = \Psi(\mu', \tau' = \tilde{\tau}')). \tag{22}$$

(iii) For each (μ, τ), the median voting outcome $\tau^M(\mu, \tau)$ satisfies

$$\int_{\{\psi(a,\epsilon;\mu,\tau) \leq \tau^M(\mu,\tau)\}} \mu(\mathbf{d}(a \times \epsilon)) \geq \frac{1}{2} \tag{23}$$

$$\int_{\{\psi(a,\epsilon;\mu,\tau) \geq \tau^M(\mu,\tau)\}} \mu(\mathbf{d}(a \times \epsilon)) \geq \frac{1}{2}. \tag{24}$$

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

The political economy with sequential voting follows a dynamic game between two political parties, as in Krusell and Ríos-Rull (1999). These parties compete with one another to take power, and the winner is determined bymajority voting by households on income taxes that the two parties proposed for each period—a lack of commitment. One-dimensional voting with a single-peaked preference leads the most preferred policy of the median voters to be supported by the majority. As a result, the dominant strategy of these two parties is a policy preferred by the median voter. To avoid multiple equilibria, I assume that one party always wins when the votes are tied.

Condition (ii) implies that each household solves the one-time deviation problem in (20), resulting in $\psi(\cdot)$, the most preferred tax of households associated with a state of $(a, \epsilon; \mu, \tau)$. As in the case with the optimal policy without commitment, a lack of government commitment makes households believe that future tax rates after one period will follow a sequence of income taxes induced by the equilibrium tax policy function $\Psi(\cdot)$ as shown in (21). The law of motion for the distribution of households has to capture all changes in the evolution of these distributions caused by the one-time deviation problem of households, which is presented in (22).

Following Corbae et al. (2009), I use condition (iii) to define the median voter. I sort the agents by the most preferred tax rate of households $\psi(\cdot)$ and find $\tau^M(\cdot)$ for each (μ, τ). Condition (iv) implies that in the political equilibrium, the median voting outcome $\tau^M(\mu, \tau)$ should be equal to the equilibrium tax function $\Psi(\mu, \tau)$ for each (μ, τ).

3 Characterization of the Markov Perfect Equilibria

Despite the above definitions clearly showing how and when the government or the median voter makes a decision on income tax, underlying economic trade-offs are hard to observe that the policymaker takes into account. In this section, I will explain what kinds of underlying economic trade-offs exist and how the policymaker weighs these trade-offs in determining tax policy, using the generalized-Euler equation approach, which is proposed by Klein, Krusell and Ríos-Rull (2008).

The generalized-Euler equation of the policymaker reveals economic forces behind the policymaker's decision through its first-order condition. The first-order condition can be derived by using the Benveniste-Scheinkman condition—the envelope condition—to eliminate terms related to the partial derivative of the value function. Here, I begin with this analysis for the case of the time-consistent optimal income tax without commitment. For ease of computation, I omit the endogenous labor supply decision, the general Euler equation with which is shown in Appendix. To obtain the first-order condition of the government, I take the partial derivative of the value of the government $\hat{V}$ in income tax for the next period $\tilde{\tau}'$ near its equilibrium value τ':

$$\begin{aligned}
0 =& \frac{d}{d\tilde{\tau}'}\Big|_{\tilde{\tau}'=\tau'} \int \hat{V}(a, \epsilon; \mu, \tau, \tilde{\tau}')\mu(d(a \times \epsilon)) \\
=& \int \frac{d}{d\tilde{\tau}'}\Big|_{\tilde{\tau}'=\tau'} \Big[u((1-\tau)w(\mu)\epsilon + (1 + r(\mu)(1-\tau))a + T - \tilde{g}^a(a, \epsilon; \mu, \tau, \tilde{\tau}')) \\
&+ \beta \sum_{\epsilon'|\epsilon} \pi_{\epsilon'|\epsilon} v(\tilde{g}^a(a, \epsilon; \mu, \tau, \tilde{\tau}'), \epsilon'; \mu', \tilde{\tau}') \Big] \mu(\mathbf{d}(a \times \epsilon))
\end{aligned} \tag{25}$$

Note that the tilde over g^a means that the deviation of $\tilde{\tau}'$ from its equilibrium value τ' makes the decision rule for assets $\tilde{g}^a$ different from that on the equilibrium path g^a.

An obscure part in computing the FOC (25) is the derivative of v in μ. Let m_q denote the $q-th$ moment of μ. I assume that there exists $Q \in \mathbb{N}$ such that $\{m_q\}_{q=1}^Q$ is a sufficient statistics of μ. This assumption allows me to replace μ with $\{m_q\}_{q=1}^Q$ in the value function. Then, the FOC (25) is given by:

$$
\begin{aligned}
0 = \int \Bigg[& - u'(c) \cdot \frac{\partial \tilde{g}^a(a, \epsilon; \{m_q\}_{q=1}^Q, \tau, \tau')}{\partial \tau'} \\
& + \beta \sum_{\epsilon'} \pi_{\epsilon,\epsilon'} \Bigg\{ \frac{\partial v(a', \epsilon', \{m'_q\}_{q=1}^Q, \tau')}{\partial a'} \cdot \frac{\partial \tilde{g}^a(a, \epsilon; \{m_q\}_{q=1}^Q, \tau, \tau')}{\partial \tau'} \\
& + \sum_{q=1}^{Q} \Big\langle \frac{\partial v(a', \epsilon', \{m'_q\}_{q=1}^Q, \tau')}{\partial m'_q} \cdot \frac{dm'_q}{d\tau'} \Big\rangle + \frac{\partial v(a', \epsilon', \{m'_q\}_{q=1}^Q, \tau')}{\partial \tau'} \Bigg\} \Bigg] \mu(\mathbf{d}(a \times \epsilon)). \qquad (26)
\end{aligned}
$$

I will substitute out the derivative terms of the value $\frac{\partial v}{\partial a'}$, $\frac{\partial v}{\partial m'_q}$, and $\frac{\partial v}{\partial \tau'}$ using the Benveniste-Scheinkman condition. First, $\frac{\partial v}{\partial a'}$ is given by:

$$
\frac{\partial v(a, \epsilon; \{m_q\}_{q=1}^Q, \tau)}{\partial a} = u'(c)(1 + r(K)(1 - \tau)) \qquad (27)
$$

where $K = m_1$ is the first moment of μ—the mean—, which is equivalent to the aggregate capital. Let me define ω as a wedge for the consumption Euler Equation:

$$
\begin{aligned}
\omega(a, \epsilon; \{m_q\}_{q=1}^Q, \tau) &= -u'(c) + \beta(1 + r(K')(1 - \tau')) \sum_{\epsilon'} \pi_{\epsilon,\epsilon'} u'(c') \\
&= -u'(c) + \beta(1 + r(K')(1 - \Psi(\mu, \tau))) \sum_{\epsilon'} \pi_{\epsilon,\epsilon'} u'(c'). \qquad (28)
\end{aligned}
$$

Substituting (28) and (27) into the FOC (26) gives:

$$
\begin{aligned}
0 = \int \Bigg[& \omega(a, \epsilon; \{m_q\}_{q=1}^Q, \tau) \cdot \frac{\partial \tilde{g}^a(a, \epsilon; \{m_q\}_{q=1}^Q, \tau, \tau')}{\partial \tau'} \\
& + \beta \sum_{\epsilon'} \pi_{\epsilon,\epsilon'} \Bigg\{ \sum_{q=1}^{Q} \Big\langle \frac{\partial v(a', \epsilon', \{m'_q\}_{q=1}^Q, \tau')}{\partial m'_q} \cdot \frac{dm'_q}{d\tau'} \Big\rangle + \frac{\partial v(a', \epsilon', \{m'_q\}_{q=1}^Q, \tau')}{\partial \tau'} \Bigg\} \Bigg] \mu(\mathbf{d}(a \times \epsilon)). \\
& \qquad (29)
\end{aligned}
$$

Now, I will substitute out $\frac{\partial v}{\partial \tau}$ using the Benveniste-Scheinkman condition. $\frac{\partial v}{\partial \tau}$ is given by:

$$\begin{aligned}\frac{\partial v(a,\epsilon;\{m_q\}_{q=1}^Q,\tau)}{\partial \tau} =& u'(c)\Big(-w(m_1)\epsilon - r(m_1)a + \frac{\partial T}{\partial \tau} - \frac{\partial g^a(a,\epsilon;\{m_q\}_{q=1}^Q,\tau)}{\partial \tau}\Big) \\ &+ \beta \sum_{\epsilon'} \pi_{\epsilon,\epsilon'} \Big\{ \frac{\partial v(a',\epsilon';\{m'_q\}_{q=1}^Q,\tau')}{\partial a'} \cdot \frac{\partial g^a(a,\epsilon;\{m_q\}_{q=1}^Q,\tau)}{\partial \tau} \\ &+ \sum_{q=1}^Q \Big\langle \frac{\partial v(a',\epsilon';\{m'_q\}_{q=1}^Q,\tau')}{\partial m'_q} \cdot \frac{dm'_q}{d\tau} \Big\rangle + \frac{\partial v(a',\epsilon';\{m'_q\}_{q=1}^Q,\tau')}{\partial \tau'} \cdot \frac{\partial \Psi(\mu,\tau)}{\partial \tau} \Big\}. \end{aligned} \tag{30}$$

With $\frac{\partial T}{\partial \tau} = rk + wN$, (30) is rearranged to:

$$\begin{aligned}\frac{\partial v(a,\epsilon;\tau,\{m_q\}_{q=1}^Q)}{\partial \tau} =& u'(c)\Big(w(m_1)(N-\epsilon) + r(m_1)(K-a)\Big) \\ &+ \omega(a,\epsilon,\tau,\{m_q\}_{q=1}^Q) \cdot \frac{\partial g^a(a,\epsilon;\tau,\{m_q\}_{q=1}^Q)}{\partial \tau} \\ &+ \beta \sum_{\epsilon'} \pi_{\epsilon'|\epsilon} \Big\{ \sum_{q=1}^Q \Big\langle \frac{\partial v(a',\epsilon';\tau',\{m'_q\}_{q=1}^Q)}{\partial m'_q} \cdot \frac{dm'_q}{d\tau} \Big\rangle \\ &+ \frac{\partial v(a',\epsilon';\tau',\{m'_q\}_{q=1}^Q)}{\partial \tau'} \cdot \frac{\partial \Psi(\tau,\mu)}{\partial \tau} \Big\}. \end{aligned} \tag{31}$$

Note that $u'(c)\big(w(m_1)(N-\epsilon) + r(m_1)(K-a)\big)$ means a welfare change from income changes via redistribution with transfers T following an increase in the current tax τ. This effect is in favor of lower-labor income and -wealth, implying lower ϵ and a relative to N and K, respectively, as these households benefits from greater transfers following an increasing in income taxes. I define this redistribution effect with transfers T as χ to be used later:

$$\chi(a,\epsilon;\tau,\{m_q\}_{q=1}^Q) = u'(c)\big(w(m_1)(N-\epsilon) + r(m_1)(K-a)\big) \tag{32}$$

The remaining is to substitute out $\frac{\partial v(a',\epsilon';\tau',\{m'_q\}_{q=1}^Q)}{\partial m'_q}$. It is hard to know how large Q is required for sufficient statistics for μ. Here, I assume $Q = 1$, implying $m_1 = K$ is sufficient to capture the evolution of distributions as in Krusell and Smith (1998). An alternative interpretation of this assumption is that the government considers only changes in the future prices, not higher moments for the future distribution in determining income taxes. This assumption enables me to characterize the MPE further.

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

$$
\begin{aligned}
\frac{\partial v(a,\epsilon;\tau,K)}{\partial K} =& u'(c)\Big((1-\tau)(f_{NK}(K)\epsilon + f_{KK}(K)a) + \frac{\partial T}{\partial K}\Big) \\
& - u'(c)\frac{g^a(a,\epsilon;\tau,K)}{\partial K} + \beta\sum_{\epsilon'}\pi_{\epsilon'|\epsilon}\Big\{\frac{\partial v(a',\epsilon';\tau',K')}{\partial a'}\cdot\frac{g^a(a,\epsilon;\tau,K)}{\partial K} \\
& + \frac{\partial v(a',\epsilon_j';\tau',K')}{\partial K'}\cdot\frac{\partial\Gamma(K,\tau,\tau')}{\partial K} + \frac{\partial v(a',\epsilon';\tau',K')}{\partial \tau'}\cdot\frac{\partial\Psi(\tau,K)}{\partial K}\Big\}.
\end{aligned}
\tag{33}
$$

$u'(c)\big((1-\tau)(f_{NK}(K)\epsilon + f_{KK}(K)a) + \frac{\partial T}{\partial K}\big)$ 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, Hong, Krusell and Ríos-Rull (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 proceeds with further steps following Davila et al. (2012). Because f is homogeneous of degree 1, $f_{KK}(K,N)K + fKN(K,N)N = 0$. On top of that, 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:

$$
\begin{aligned}
\frac{\partial v(a,\epsilon;\tau,K)}{\partial K} =& u'(c)\big((1-\tau)(-\frac{\epsilon}{N} + \frac{a}{K})f_{KK}(K)K + f_K(K)\tau\big) + \omega(a,\epsilon;\tau,K)\cdot\frac{g^a(a,\epsilon;\tau,K)}{\partial K} \\
& + \beta\sum_{\epsilon'}\pi_{\epsilon'|\epsilon}\Big\{\frac{\partial v(a',\epsilon_j';\tau',K')}{\partial K'}\cdot\frac{\partial\Gamma(K,\tau,\tau')}{\partial K} + \frac{\partial v(a',\epsilon';\tau',K')}{\partial \tau'}\cdot\frac{\partial\Psi(\tau,K)}{\partial K}\Big\}
\end{aligned}
\tag{34}
$$

Note that $u'(c)\big((1-\tau)(-\frac{\epsilon}{N} + \frac{a}{K})f_{KK}(K)K + f_K(K)\tau\big)$ indicates a welfare change following a shift of the factor composition of income that is driven by general equilibrium effects after an increase in K. This channel has been thoroughly investigated by Davila, Hong, Krusell and Ríos-Rull (2012). Note that because $f_{KK}(K)K < 0$, $f_K(K)\tau > 0$, the sign of this term depends hugely on the size of $(-\frac{\epsilon}{L} + \frac{a}{K})$. For example, let me assume that there is an increase in K. In regard to this change, if $\frac{\epsilon}{L}$ is substantially larger than $\frac{a}{K}$—the factor income is biased to labor (the factor composition of income of the consumption-poor)—$\frac{\partial v}{\partial K}$ is positive because w increases while r is reduced in general equilibrium. On the other hand, the consumption-rich, whose factor income is biased more toward capital, are more likely to experience a loss in welfare because a decline in r plays a role in reducing their income. I will define this pecuniary externality through changes in the composition of individual income as ϕ to be used later:

$$\phi(a, \epsilon; \tau, K) = u'(c)\big((1-\tau)(-\frac{\epsilon}{N} + \frac{a}{K})f_{KK}(K)K + f_K(K)\tau\big). \tag{35}$$

I will ease the notations by employing sequential ones. I refers to v_{i,X_t} as the partial derivative of v in X_t for individual i in period t. Then, $\frac{\partial v}{\partial K}$ and $\frac{\partial v}{\partial \tau}$ are rearranged to:

$$v_{i,K_t} = \phi_{i,t} + \omega_{i,t} \cdot g^a_{i,K_t} + \beta E_{i,t}[v_{K_{t+1}}\Gamma_{K_t} + v_{\tau_{t+1}}\Psi_{K_t}] \tag{36}$$

$$v_{i,\tau_t} = \chi_{i,t} + \omega_{i,t} \cdot g^a_{i,\tau_t} + \beta E_{i,t}[v_{K_{t+1}}\Gamma_{i,\tau_t} + v_{\tau_{t+1}}\Psi_{\tau_t}] \tag{37}$$

where $g^a_{i,K_t} = \frac{\partial g^a}{\partial K}$, $g^a_{i,\tau_t} = \frac{\partial g^a}{\partial \tau}$, $\Gamma_{K_t} = \frac{\partial \Gamma(K,\tau,\tau')}{\partial K}$, $\Gamma_{\tau_t} = \frac{\partial \Gamma(K,\tau,\tau')}{\partial \tau}$,$\Psi_{K_t} = \frac{\partial \Psi(K,\tau)}{\partial K}$, and $\Psi_{\tau_t} = \frac{\partial \Psi(K,\tau)}{\partial \tau}$. $E_{i,t}$ means the conditional expectation of individual i in period t.

I will eliminate the partial derivatives of the future values by successively substituting themselves out. Then, $v_{i,K}$ is given by:

$$\begin{aligned} v_{i,K_t} = & \underbrace{\omega_{i,t} \cdot g^a_{i,K_t} + \sum_{s=1}^{\infty} \beta^s E_{i,t}\left[\left(\frac{\Xi\Gamma_{t+s-1}}{\Xi K_t} \cdot g^a_{i,K_{t+s}} + \frac{\Xi\Psi_{t+s-1}}{\Xi K_t} \cdot g^a_{i,\tau_{t+s}}\right) \cdot \omega_{t+s}\right]}_{\text{PDV of lifetime wedges over Euler Equation following } K_t \uparrow} \\ & + \underbrace{\phi_{i,t} + \sum_{s=1}^{\infty} \beta^s E_{i,t}\left[\frac{\Xi\Gamma_{t+s-1}}{\Xi K_t} \cdot \phi_{t+s}\right]}_{\substack{\text{PDV of changes in lifetime income via} \\ \text{variations in factor composition of income} \\ \text{following } K_t \uparrow}} + \underbrace{\sum_{s=1}^{\infty} \beta^s E_{i,t}\left[\frac{\Xi\Psi_{t+s-1}}{\Xi K_t} \cdot \chi_{t+s}\right]}_{\substack{\text{PDV of changes in lifetime income} \\ \text{via redistribution with transfers} \\ \text{following } K_t \uparrow}}. \end{aligned} \tag{38}$$

where $\frac{\Xi\Gamma_{t+s-1}}{\Xi K_t}$ is the sum of all the variation in $\Gamma_{t+s-1} = K_{t+s}$ caused by a change in K_t and $\frac{\Xi\Psi_{t+s-1}}{\Xi K_t}$ is the sum of all the variation in $\Psi_{t+s-1} = \tau_{t+s}$ caused by a change in K_t.[3] These operators are required because one-time change in capital stock K_t causes itself to adjust gradually over time. Note that $\frac{\Xi\Gamma_{t+s-1}}{\Xi K_t}$ is positive because of the mean-reverting dynamics of capital. Likewise $\frac{\Xi\Psi_{t+s-1}}{\Xi K_t}$ is positive because the government tend to impose a grater tax on an economy with a good deal of the aggregate capital.

Equation (38) implies that a change in K_t leads individuals to face three different effects over time. First, a change in K_t affects individual consumption allocations over time, which is presented in the sum of the first two terms. This sum indicates the present discount value of the lifetime wedges over the consumption Euler equation. Second, a variation in K_t changes individuals‘ lifetime income through variations in the factor composition of income over time. This force is observed in the sum of the third and fourth terms. Finally, a change in K_t affects individuals' lifetime income via redistribution caused by altered transfers over time, which is shown in the fifth

[3]Appendix B shows their formal definition with recursive operators.

term.

It is worth examining several features for the effects above. The first effect–related to individual dynamic allocations—has no impact on individual values. Note that the first two terms comprise the product the wedge from the consumption Euler equation (ω) and the partial derivative of the saving decision rule in K and in τ ($g^a_{i,K}$ and $g^a_{i,\tau}$) for each period. This product is always zero for each period because when an individual, for example, does not hit the borrowing constraint, ω is zero, and when it does so, both $g^a_{i,K}$ and $g^a_{i,\tau}$ are zero. Such no impact takes place because this economy is not centralized by a social planner but, with a government endogenously making decisions, a decentralized economy where individuals optimally and efficiently choose consumption and savings in competitive equilibrium.

As previously mentioned, the second effect—relevant to income changes caused by variations in the factor composition of individual income—can differently affect individuals' value according to their income factor composition. For instance, when K_t increases, the sign of ϕ is reliant on the composition of income. Note that in general equilibrium, this capital increase gives rise to a reduction in r but an increase in w. Equation (35) indicates that in response to these price changes, those whose income tends to be toward labor income have a benefit. Considering the mean-reverting dynamics of K_t, $\frac{\Xi\Gamma_{t+s-1}}{\Xi K_t}$ is positive over time. Therefore, for labor income-biased individuals—the consumption-poor, an increase in K_t has a positive impact on value through changes in the factor composition of individual income. On the other hand, for capital income-biased individuals, the non-consumption-poor, increased K reduces their value.

The final effect—associated with changes in income via redistribution caused by altered transfers—differs across individuals. Equation (32) suggests that the sign of the final effect depends on individuals' relative labor productivity and asset holdings to the aggregate factors. Note that $\frac{\Xi\Psi_{t+s-1}}{\Xi K_t}$ is positive because the government tends to levy a more significant tax on an economy with a good deal of capital. Therefore, the final effect benefits the poor while negatively influencing the rich.

Likewise, v_{i,τ_t} is given by:

$$v_{i,\tau_t} = \underbrace{\omega_{i,t} \cdot g^a_{i,\tau_t} + \sum_{s=1}^{\infty} \beta^s E_{i,t}\left[\left(\frac{\Xi\Gamma_{t+s-1}}{\Xi\tau_t} \cdot g^a_{i,K_{t+s}} + \frac{\Xi\Psi_{t+s-1}}{\Xi\tau_t} \cdot g^a_{i,\tau_{t+s}}\right) \cdot \omega_{t+s}\right]}_{\text{PDV of lifetime wedges over Euler Equation following } \tau_t \uparrow}$$
$$+ \underbrace{\sum_{s=1}^{\infty} \beta^s E_{i,t}\left[\frac{\Xi\Gamma_{t+s-1}}{\Xi\tau_t} \cdot \phi_{t+s}\right]}_{\substack{\text{PDV of changes in lifetime income via} \\ \text{variations in factor composition of income} \\ \text{following } \tau_t \uparrow}} + \underbrace{\chi_{i,t} + \sum_{s=1}^{\infty} \beta^s E_{i,t}\left[\frac{\Xi\Psi_{t+s-1}}{\Xi\tau_t} \cdot \chi_{t+s}\right]}_{\substack{\text{PDV of changes in lifetime income} \\ \text{via redistribution with transfers} \\ \text{following } \tau_t \uparrow}}. \quad (39)$$

where $\frac{\Xi\Gamma_{t+s-1}}{\Xi\tau_t}$ is the sum of all the variation in $\Gamma_{t+s-1} = K_{t+s}$ caused by a change in τ_t; $\frac{\Xi\Psi_{t+s-1}}{\Xi\tau_t}$ is the sum of all the variation in $\Psi_{t+s-1} = \tau_{t+s}$ caused by a change in τ_t. The sign of $\frac{\Xi\Gamma_{t+s-1}}{\Xi\tau_t}$ is negative because a greater tax reduce overall savings, and that of $\frac{\Xi\Psi_{t+s-1}}{\Xi\tau_t}$ is positive given its mean-revert dynamics.

Equation (39) shows that one-time change in τ_t causes individuals to confront three different effects over time, as in the case with a change in K_t. First, a change in τ_t affects individual consumption allocations over time, which is presented in the sum of the first two terms, of which value is zero as previously mentioned. Second, a change in τ_t changes individuals' lifetime income through variations in the factor composition of income over time. This force is observed in the sum of the third term. Finally, a change in τ_t affects individuals' lifetime income via redistribution caused by altered transfers over time, which is shown in the fourth and the fifth terms.

Substituting (38) and (39) into the FOC (29), the FOC is given by:

$$\begin{aligned}
0 = \int \Bigg[E_{i,t}\Bigg[&\omega_{i,t} \cdot g^a_{i,\tau_{t+1}} + \sum_{s=1}^{\infty} \beta^s \Bigg\{ \omega_{t+s} \cdot \Bigg(g^a_{i,K_{t+s}} \cdot \left(\frac{\partial K_{t+1}}{\partial \tau_{t+1}} \cdot \frac{\Xi\Gamma_{t+s-1}}{\Xi K_{t+1}} + \frac{\Xi\Gamma_{t+s-1}}{\Xi\tau_{t+1}} \right) \\
&+ g^a_{i,\tau_{t+s}} \cdot \left(\frac{\partial K_{t+1}}{\partial \tau_{t+1}} \cdot \frac{\Xi\Psi_{t+s-1}}{\Xi K_{t+1}} + \frac{\Xi\Psi_{t+s-1}}{\Xi\tau_{t+1}} \right) \Bigg) \Bigg\} \\
&+ \sum_{s=1}^{\infty} \beta^s \Bigg\{ \phi_{t+s} \cdot \left(\frac{\partial K_{t+1}}{\partial \tau_{t+1}} \cdot \frac{\Xi\Gamma_{t+s-1}}{\Xi K_{t+1}} + \frac{\Xi\Gamma_{t+s-1}}{\Xi\tau_{t+1}} \right) \Bigg\} \\
&+ \sum_{s=1}^{\infty} \beta^s \Bigg\{ \chi_{t+s} \cdot \left(\frac{\partial K_{t+1}}{\partial \tau_{t+1}} \cdot \frac{\Xi\Psi_{t+s-1}}{\Xi K_{t+1}} + \frac{\Xi\Psi_{t+s-1}}{\Xi\tau_{t+1}} \right) \Bigg\} \Bigg] \Bigg] \mu(\mathbf{d}(a \times \epsilon)).
\end{aligned} \tag{40}$$

As mentioned before, the first three terms are zero because individuals' optimal decisions on consumption and savings. Thus, the above equation is rearranged to:

$$\begin{aligned}
&\underbrace{-\int \Bigg[E_{i,t}\Bigg[\Bigg\{ \sum_{s=1}^{\infty} \beta^s \phi_{t+s} \cdot \left(\frac{\partial K_{t+1}}{\partial \tau_{t+1}} \cdot \frac{\Xi\Gamma_{t+s-1}}{\Xi K_{t+1}} + \frac{\Xi\Gamma_{t+s-1}}{\Xi\tau_{t+1}} \right) \Bigg\} \Bigg] \Bigg] \mu(\mathbf{d}(a \times \epsilon))}_{\substack{\text{Sum of PDV of changes in lifetime income} \\ \text{via variations in factor composition of income following } \tau_{t+1} \uparrow \text{ over individual)}}} \\
&= \underbrace{\int \Bigg[E_{i,t}\Bigg[\Bigg\{ \Bigg\{ \sum_{s=1}^{\infty} \beta^s \Bigg\{ \chi_{t+s} \cdot \left(\frac{\partial K_{t+1}}{\partial \tau_{t+1}} \cdot \frac{\Xi\Psi_{t+s-1}}{\Xi K_{t+1}} + \frac{\Xi\Psi_{t+s-1}}{\Xi\tau_{t+1}} \right) \Bigg\} \Bigg] \Bigg] \mu(\mathbf{d}(a \times \epsilon))}_{\substack{\text{Sum of PDV of changes in lifetime income} \\ \text{via redistribution with transfer following } \tau_{t+1} \uparrow \text{ over individuals}}}
\end{aligned} \tag{41}$$

where $\left(\frac{\partial K_{t+1}}{\partial \tau_{t+1}} \cdot \frac{\Xi\Gamma_{t+s-1}}{\Xi K_{t+1}} + \frac{\Xi\Gamma_{t+s-1}}{\Xi\tau_{t+1}} \right)$ is negative, and $\left(\frac{\partial K_{t+1}}{\partial \tau_{t+1}} \cdot \frac{\Xi\Psi_{t+s-1}}{\Xi K_{t+1}} + \frac{\Xi\Psi_{t+s-1}}{\Xi\tau_{t+1}} \right)$ is positive.[4]

Equation (41) implies that when determining taxes, the government takes into account two

[4]Recall that $\frac{\partial K_{t+1}}{\partial \tau_{t+1}} < 0$, $\frac{\Xi\Gamma_{t+s-1}}{\Xi K_{t+1}} > 0$, $\frac{\Xi\Gamma_{t+s-1}}{\Xi\tau_{t+1}} < 0$, $\frac{\Xi\Psi_{t+s-1}}{\Xi K_{t+1}} < 0$, and $\frac{\Xi\Psi_{t+s-1}}{\Xi\tau_{t+1}} > 0$.

underlying forces: the pecuniary externality relevant to changes in the income factor composition and income redistribution across individuals with transfers. Note that the government does not consider distortions observed in the consumption Euler equation; because individuals optimally choose consumption and saving in competitive equilibrium, there is no room for improvement in terms of the government. As a result, when making tax decisions, the government internalizes the above two forces—the pecuniary externality and income redistribution with altered transfers.

Equation (41) can be interpreted from the view of the government's cost-benefit analysis. Because the government takes the utilitarian's view, and each force is weighted with the individual marginal utility of consumption, the government favors the consumption-poor. Let us assume an increase in τ' causing an increase in transfers and a gradual reduction in the aggregate capital over time. Because this change brings about an increase in r and a decrease in w, the government will regard the force related to income changes through variations in individuals' income composition as a cost. But the government will view the channel associated with redistribution via altered transfers as a benefit because the consumption-poor benefits from these increased. In response to a reduction in τ', of course, the government will take the opposite view regarding cost-benefit.

Equation (41) also shows how individual heterogeneity plays a role in the determination of taxes. Without individual heterogeneity, the cross-sectional distribution is de-generated, and the force related to income redistribution via altered transfers—$\chi =$ becomes zero. Thus Equation (41) is rearranged to:

$$-E_t\Big[\sum_{s=1}^{\infty}\beta^s\phi_{t+s}\cdot\Big(\frac{\partial K_{t+1}}{\partial\tau_{t+1}}\cdot\frac{\Xi\Gamma_{t+s-1}}{\Xi K_{t+1}}+\frac{\Xi\Gamma_{t+s-1}}{\Xi\tau_{t+1}}\Big)\Big]=0 \tag{42}$$

where

$$\phi_t=u'(c_{i,t})\big((1-\tau_t)(-\frac{\epsilon_{i,t}}{L_t}+\frac{a_{i,t}}{K_t})f_{KK}(K_t)K_t+f_{K_t}(K_t)\tau_t\big)=u'(c_{i,t})\cdot f_{K_t}(K_t)\cdot\tau_t$$

The above equation indicates that to satisfy the above condition, the force associated with the pecuniary externality has to be null, consistent with previous literature findings: the optimal tax in the representative agent model is zero. This finding means that individual heterogeneity shaped by the interaction between their decisions and uninsurable idiosyncratic income risk brings about the pecuniary externality and insurance channels via transfers, thereby making the optimal tax rate positive. Another interesting investigation for the optimal condition (41) is to compare this to the planner's optimal condition in Davila et al. (2012):

$$\omega_{i,t} + \beta \int E_{i,t}\big[\phi_{t+1}\big]\mu(\mathbf{d}(a \times \epsilon)) = 0 \tag{43}$$

where

$$\omega_{i,t} = -u'(c_{i,t}) + \beta(1 + r(K_{t+1})) \sum_{\epsilon'} \pi_{\epsilon'|\epsilon} u'(c_{t+1}) \tag{44}$$

$$\phi_{i,t} = u'(c_{i,t})\big((-\frac{\epsilon_{i,t}}{L_t} + \frac{a_{i,t}}{K_t}) f_{KK}(K_t) K_t\big) \tag{45}$$

In contrast with the government's optimal condition (41) 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 as well as distortions embedded in the consumption Euler equation. This distinction takes place because of different assumptions between the two economies. Davila, Hong, Krusell and Ríos-Rull's (2012) economy is centralized: the social planner can manipulate individual saving decisions while preserving constraints caused by incomplete markets and uninsurable idiosyncratic income risk. This centralized economy assumption makes the consumption Euler equation non-zero. But 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.

Finally, the government's optimal condition (41) allows a theoretical investigation for the difference between the case with time-consistent optimal policy and with sequential voting. As mentioned previously, the majority voting game boils down to the median voter's choice. Therefore, in this political economy, the optimal condition is given by:

$$\begin{aligned} 0 =& \frac{d}{d\tilde{\tau}'}\Big|_{\tilde{\tau}'=\tau'} \hat{V}_m(a, \epsilon; \tau, \mu, \tilde{\tau}') \\ =& \frac{d}{d\tilde{\tau}'}\Big|_{\tilde{\tau}'=\tau'} \Big[u((1-\tau)w(\mu)\epsilon + (1 + r(\mu)(1-\tau))a + T - \tilde{g}_{a,m}(a, \epsilon; \tau, \mu : \tilde{\tau}')) \\ &+ \beta \sum_{\epsilon'} \pi_{\epsilon'|\epsilon} v_m(\tilde{g}_a(a, \epsilon; \tau, \mu : \tilde{\tau}'), \epsilon', \mu', \tilde{\tau}') \Big] \end{aligned} \tag{46}$$

where m subscript refers to the median voter. Using the Benveniste-Scheinkman condition, I rearrange this condition with χ_i, η_i, and ϕ_i as in the optimal time-consistent case. In the economy with sequential voting, the government offers voters τ_{t+1}^m satisfying

$$\underbrace{-E_{m,t}\left[\left\{\sum_{s=1}^{\infty}\beta^s\phi_{m,t+s}\cdot\left(\frac{\partial K_{t+1}}{\partial\tau_{t+1}}\cdot\frac{\Xi\Gamma_{t+s-1}}{\Xi K_{t+1}}+\frac{\Xi\Gamma_{t+s-1}}{\Xi\tau_{t+1}}\right)\right\}\right]}_{\substack{\text{PDV of changes in lifetime income}\\ \text{via variations in factor composition of income following } \tau_{t+1}\uparrow \text{ for the median voter}}}$$
$$=\underbrace{E_{m,t}\left[\left\{\left\{\sum_{s=1}^{\infty}\beta^s\left\{\chi_{m,t+s}\cdot\left(\frac{\partial K_{t+1}}{\partial\tau_{t+1}}\cdot\frac{\Xi\Psi_{t+s-1}}{\Xi K_{t+1}}+\frac{\Xi\Psi_{t+s-1}}{\Xi\tau_{t+1}}\right)\right\}\right]}_{\substack{\text{PDV of changes in lifetime income}\\ \text{via redistribution with transfer following } \tau_{t+1}\uparrow \text{ for the median voter}}}. \tag{47}$$

No integral appears over individual because the government considers only what the median voter prefers. Therefore, what matters is where the median voter is. When his position is far from the consumption-poor's, The median voter's income composition is more likely to be biased toward capital than the consumption-poor, thereby leading the cost from variations in the factor composition of income to be smaller in response to an increase in τ_{t+1}. Therefore, this smaller cost requires a less generous transfer system. Because the median voter's position is a quantitative issue, I will address its implications in a later section.

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 in the previous section helps us better understand the government's decisions on policies, it is not that 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 should be determined according to the political and commitment structure while consistent with individual decisions and the law of motion for the distribution of households. In other words, three equilibrium objects—individual decisions, 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 a Markov-perfect equilibrium.

I address the above computational issues by taking ideas from the backward induction method of Reiter (2010). The author introduced a non-simulation-based solution method to solve an incomplete markets economy with aggregate uncertainty. As in Krusell and Smith's (1998), 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 it is defined across the aggregate finite grid points. However, the way of finding $\Gamma(\cdot)$ is 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., 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 much less required in Reiter (2010) than in Krusell and Smith (1998), which improves computational efficiency for 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 $\Psi(\cdot)$ consistent with the political and commitment structure. This is feasible because, with the value function, these endogenous tax functions can be directly obtained by solving (17) and (20).

However, I wish to clarify that I cannot directly apply the Reiter's (2010) method to the model in this paper because of the existence of off-equilibrium paths. In the incomplete markets economy with aggregate uncertainty, for which Reiter's (2010) method is originally designed, the distribution of aggregate shocks (TFP) Z is stationary. 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, an economy in Markov-perfect equilibrium does not have this property. For example, in the political economy with sequential voting, the vote on policies are obtained by comparing one-time deviation policies. Some tax paths would not be reached on the equilibrium path.

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, 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

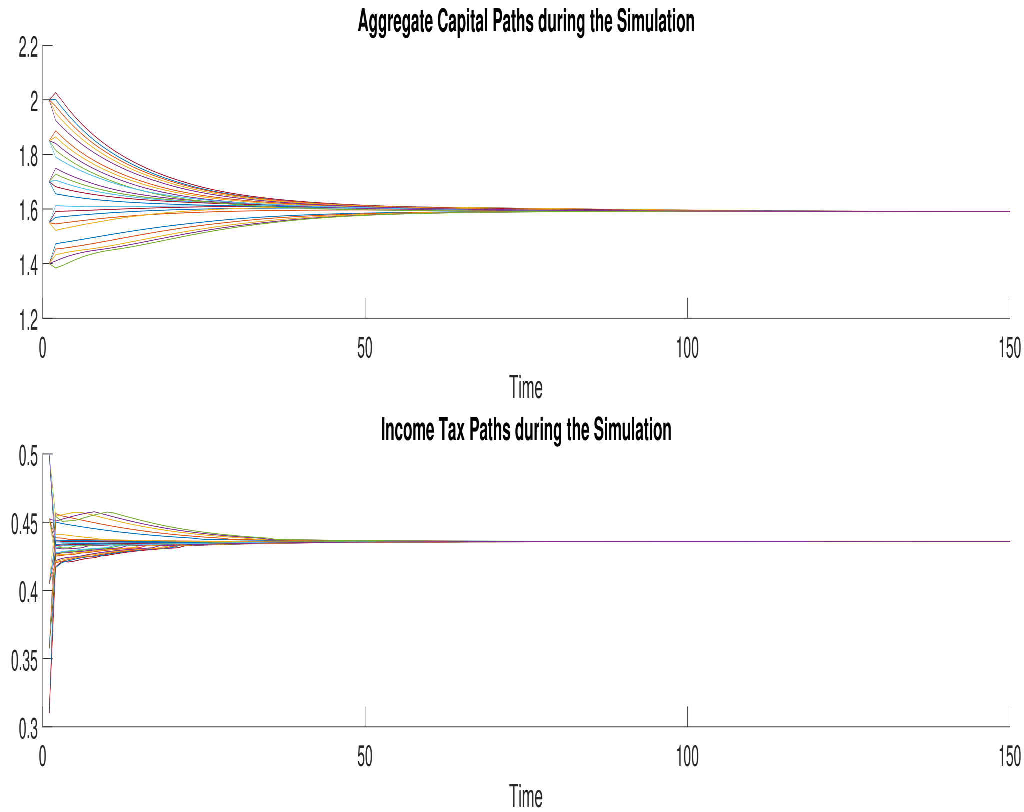

Figure 1: Transitions from off the Equilibrium to the Equilibrium

equilibrium to the steady-state equilibrium in the political economy with voting. Finally, I modify the way of constructing reference distributions, which is required to update the proxy distributions in Reiter (2010), by reflecting the features of the Markov-perfect equilibrium, 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, with its performances in efficiency and accuracy.

Because of these somewhat complex variations in Reiter's (2010) method, one might consider simply using Krusell and Smith's (1998) method to solve this model. However, their approach would not be efficient in addressing this class of models in Markov-perfect equilibrium. 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 Krusell and Smith's (1998) original algorithm, thereby exponentially increasing the computational burden. Second, the parametric assumption of Krusell and Smith's (1998) approach would act as a barrier because the equilibrium tax function $\Psi(\cdot)$ could be severely nonlinear in the aggregate state. The parametric

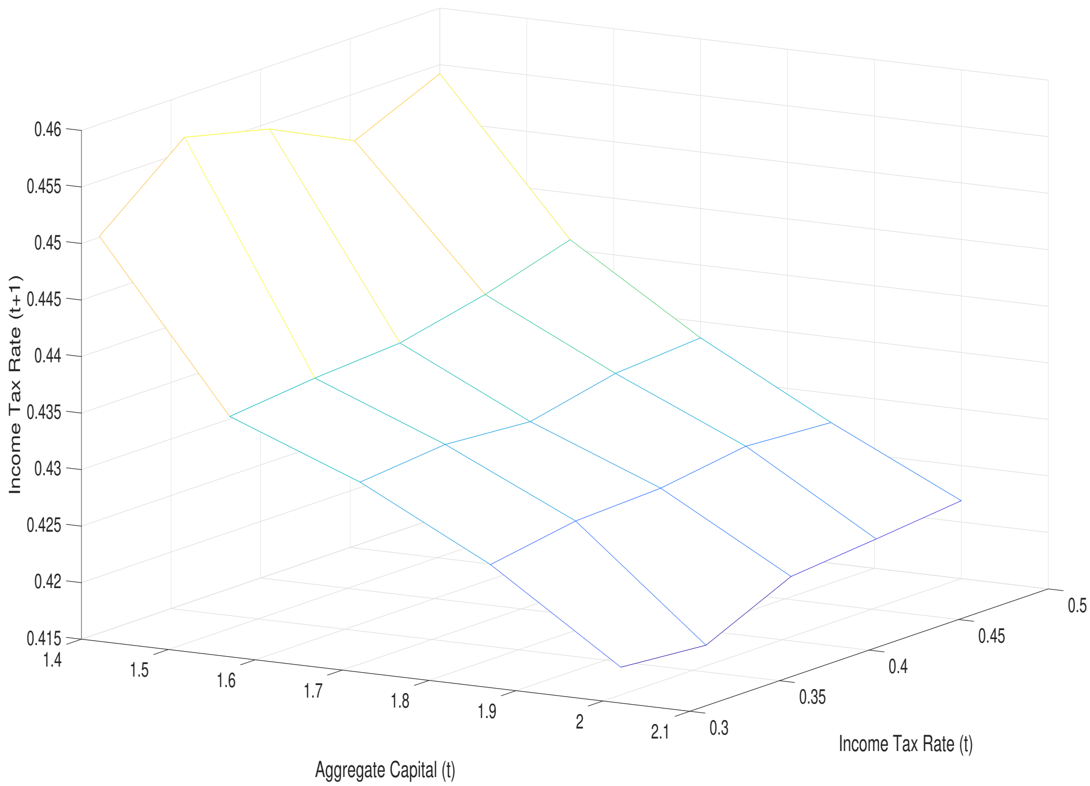

Figure 2: Income Tax Function $\Psi(m_i, \tau_k)$

assumption works well when the law of motion for household distributions $\Gamma(\cdot)$ is close to linear. I find that although this linearity still appears in $\Gamma(\cdot)$, $\Psi(\cdot)$ shaped by the median voter's choice is severely nonlinear, as shown in Figure 2.[5]

When solving the Ramsey problem, I have employed the approach in Dyrda and Pedroni (2022), parameterizing the transitional path of income taxes as follows:

$$\tau_t = \left(\sum_{i=0}^{m_{x0}} \alpha_i^x P_i(t) \right) \exp(\lambda t) + (1 - \exp(-\lambda t)) \left(\sum_{j=0}^{m_{xF}} \beta_j^x P_j(t) \right), \quad t \leq t_F \tag{48}$$

where $\{P_i(t)\}_{i=1}^{m_{x0}}$ and $\{P_j(t)\}_{j=0}^{m_{xF}}$ are families of Chebyshev polynomial; m_{x0} and m_{xF} are orders of the polynomial apporoximation for the short-run and long-run dynamics; $\{\alpha_i^x\}_{i=0}^{m_{x0}}$ 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 $\{\alpha_0^x, \alpha_1^x, \alpha_2^x, \beta_0^x, \beta_1^x, \beta_2^x, \lambda\}$

[5] Corbae et al. (2009) employed Krusell and Smith's (1998) method to solve a similar economy to mine but without wealth effects of labor supply. Such difficulties might lead them to omit wealth effects of labor supply although adding more states to the forecasting rules.

that maximizes the welfare function of the utilitarian government at time 0.[6]

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 the parameters requires solving the stationary distribution of the model to match 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.

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
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 calibrated 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 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 $\underline{a} = 0$. 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 up 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 that is closer to its empirical counterpart, 0.044.[7]

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

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

6 Results

In this section, I quantitatively explore how differences in the commitment and political structure affect the aggregate economy, inequality, and welfare. For this, I conduct two sets of counterfactual exercises. First, to examine the effects of the commitment structure, I compare the economy with the time-consistent optimal policy to the economy with the time-inconsistent optimal policy with the Ramsey planner. Second, to figure out the impacts of the political structure, I contrast the economy with the time-consistent optimal policy with the economy with sequential voting. I assume that the initial economy begins at the calibrated steady-state through all these exercises, and I compare their equilibrium results over the transitional path.

6.1 Time-consistent versus -inconsistent Optimal Tax

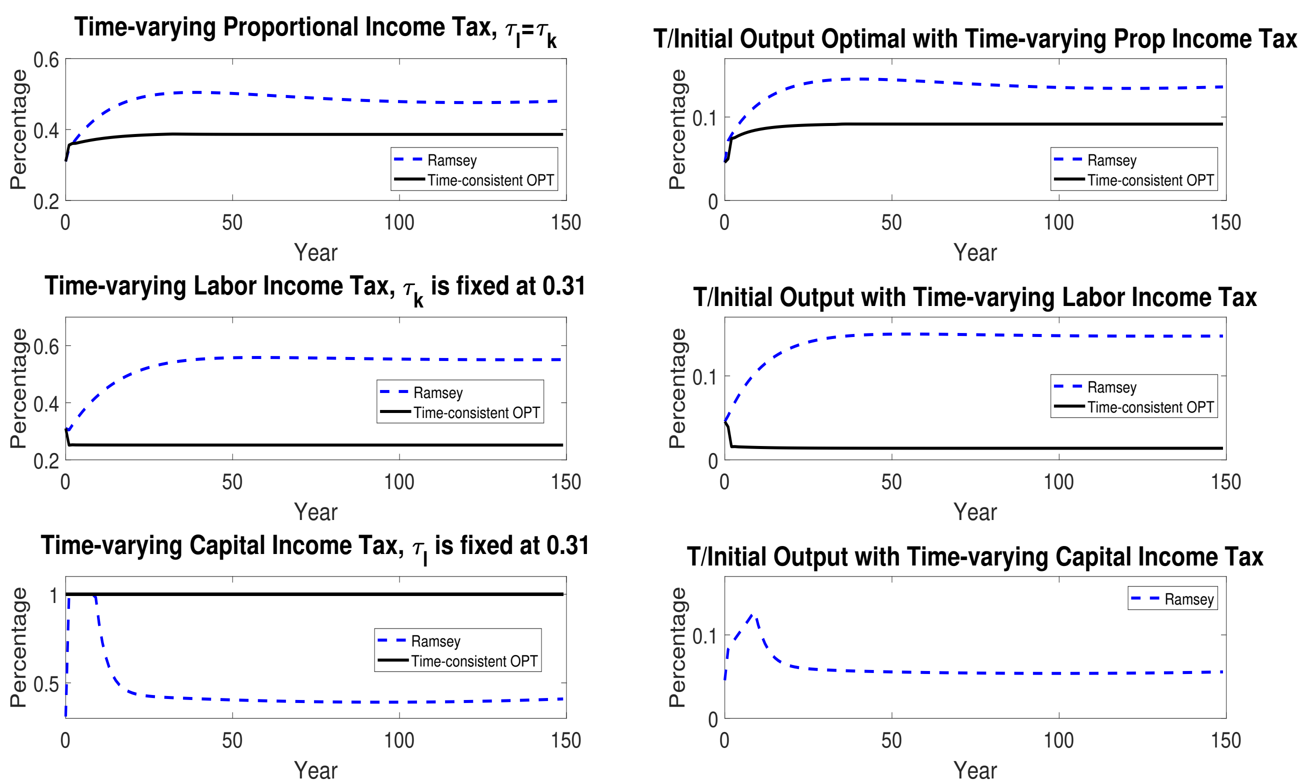

Figure 3: Time-consistent and -inconsistent Optimal Policies: Tax/Transfers Transition Paths

Figure 3 shows the time-consistent and inconsistent optimal taxes and the implied ratio of transfers to the initial output. The first raw indicates economies where income tax is allowed to be varying over time; the second raw dose economies where only labor income tax is allowed to be time-varying while capital income tax is fixed at 0.31; and the last raw does economies where only capital income tax is allowed to be time-varying while labor income tax is fixed at 0.31. Note that the time-consistent capital income tax is not well-established in my model. With a lack of commitment, the government levies capital income tax at a confiscatory rate, leading the economy to shrink. In my model, because the government purchase is constant, households have to pay income taxes or a lump-sum tax to finance it. Under this shrinking economy, because their income is also reduced, levying a lump-sum tax is required. In the case with capital income tax without commitment, this lump-sum tax level is so high that households with low productivity cannot have a positive consumption under their budget set, preventing the sustainability of this economy. Therefore, I focus on analyzing results with the cases with proportional income and labor income taxes going forward.[8]

Let me begin with the cases with proportional income tax. The top-left panel of Figure 3 implies that the Ramsey planner with the time-inconsistent optimal policy chooses more substantial income taxes than the government with the time-consistent optimal policy over the whole transitional path. The Ramsey planner gradually increases income taxes by 16 percentage points. But with a lack of commitment, the optimal income tax rate raises taxes by 2 percentage points. This gap in the 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 that in the case without commitment by 4.7 percentage points.

Table 2: Welfare Outcomes According to Commitment Structure

Welfare (CEV)	Time-inconsistency	Time-consistency
OPT INC TAX	+2.19%	+0.57%
OPT Labor INC TAX ($\tau_k = 0.31$)	+2.20%	-1.24%
OPT Capital INC TAX ($\tau_l = 0.31$)	+2.85%	-

This distinction for income taxes brings about 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 one does so 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 commitment structure. Note that welfare increases when the overall level of consumption and leisure increases and their inequality is reduced.

[8] I report results for the case with the time-inconsistent optimal capital income tax in Appendix.

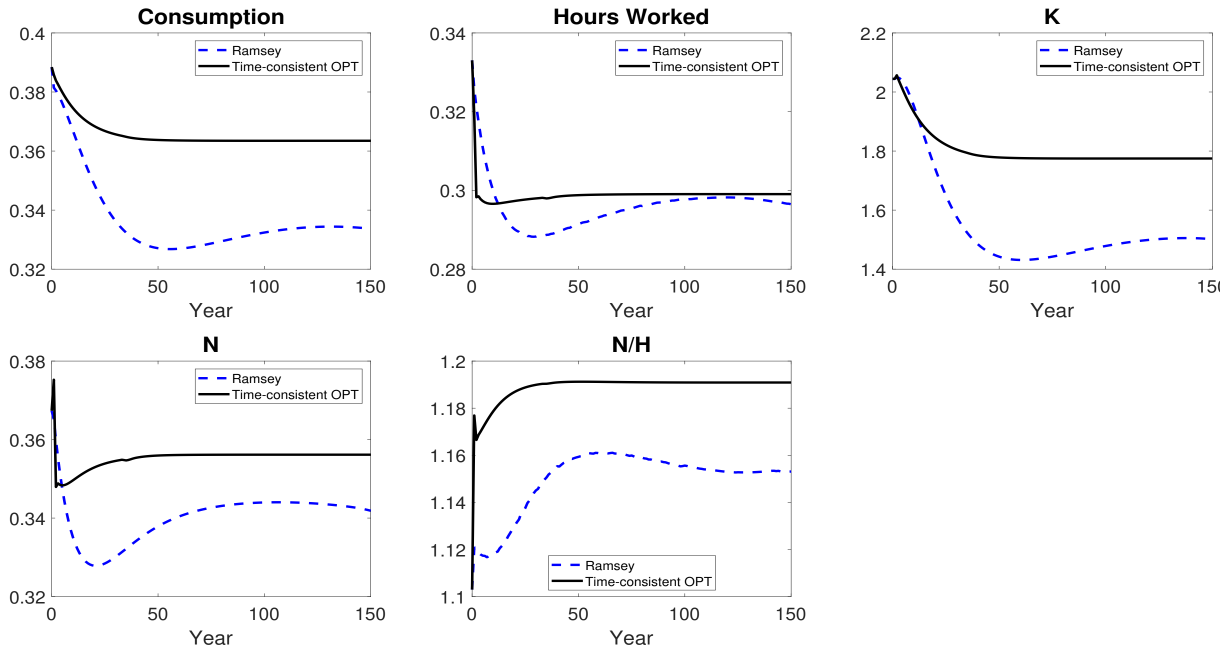

Figure 4: Time-consistent and -inconsistent Optimal Income Tax: Aggregate Outcomes

Figure 4 depicts changes in the levels of the aggregate variables. Figure 4 suggests that the case without commitment generates more efficient outcomes. All the aggregate variables in the economy with the time-consistent optimal policy are larger than that with time-inconsistent optimal policy. This result may be obvious because lower taxes in the time-consistent case bring about fewer distortions. However, this finding seems obscure to understand the welfare consequences because consumption is substantially larger in the case with the time-consistent policy, and the gaps in hours worked are not significant.

Figure 5 shows inequalities in consumption, hours worked, wealth and after-tax income. Figure 5 suggests that more substantial 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 around 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 Ramsey planner achieves better welfare outcomes by better balancing efficiency and redistribution through the controlling of a sequence of transfers over the transitional path, leading to more equal after-tax incomes. This management is allowed thanks to the commitment instrument. By contrast, the government with the time-consistent optimal policy is less competent in managing the trade-off between efficiency and redistribution over the transitional path because it is under a lack of commitment. As a result,

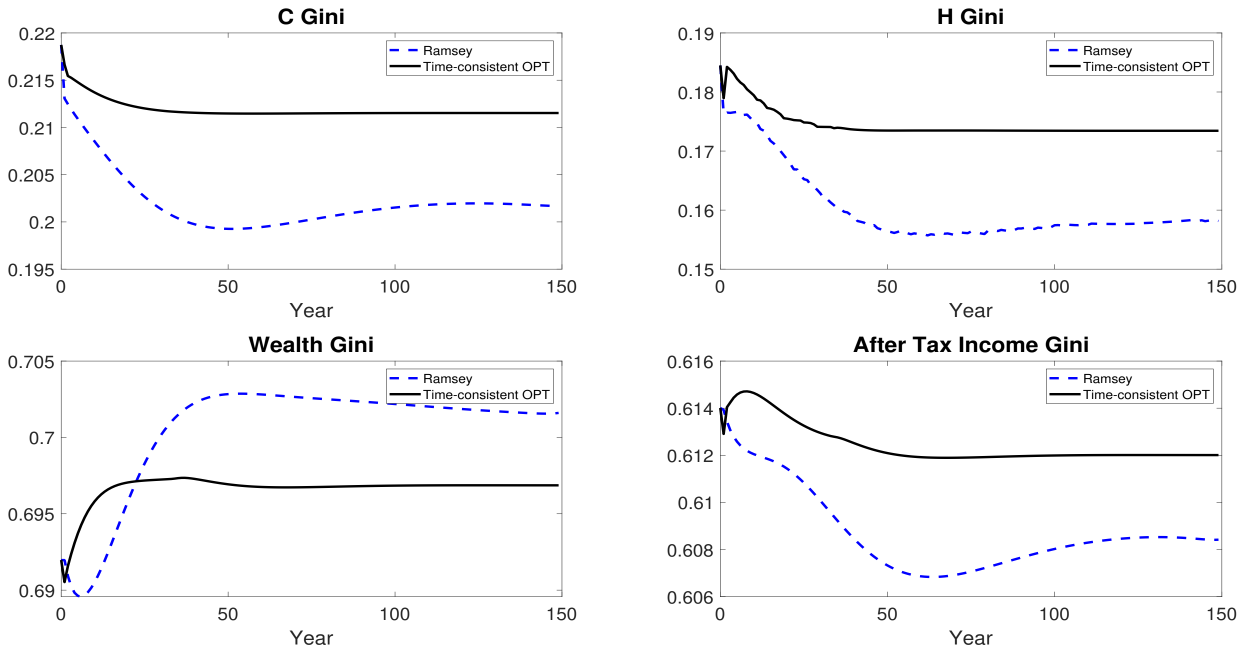

Figure 5: Time-consistent and -inconsistent Optimal Income Taxes: Distributional Outcomes

the government without commitment implements less tax to spare more incomes in the short run while forsaking welfare gains from reduced inequalities in the long run.

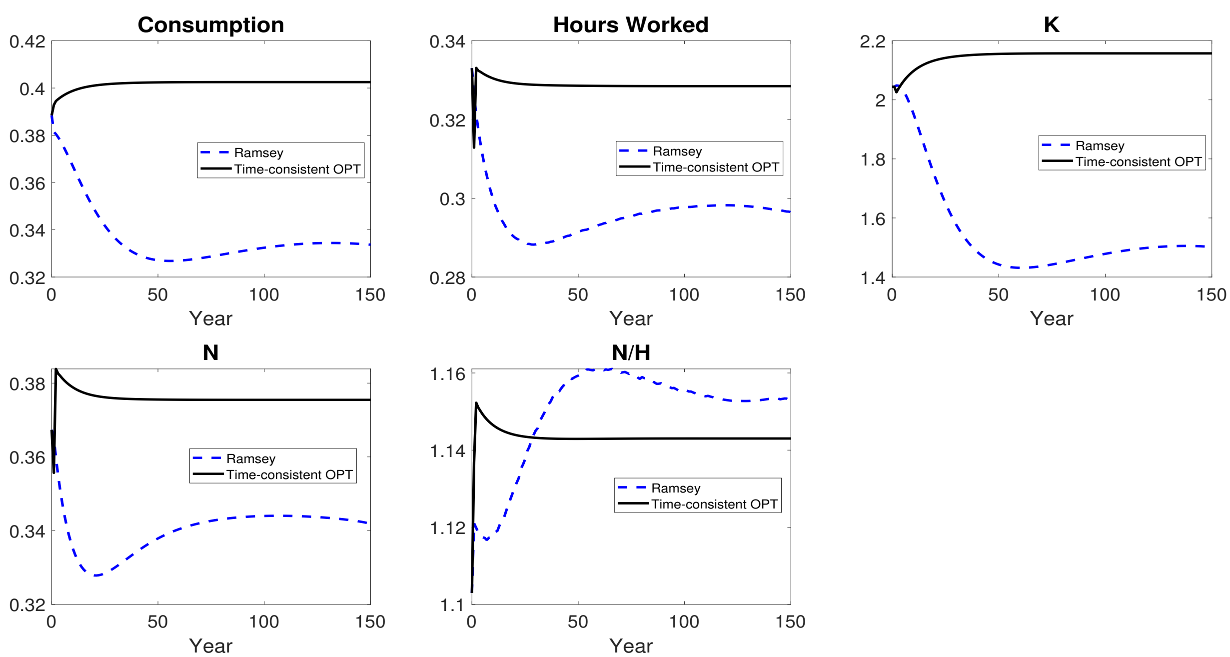

Figure 6: Time-consistent and -inconsistent Optimal Labor Taxes: Aggregate Outcomes

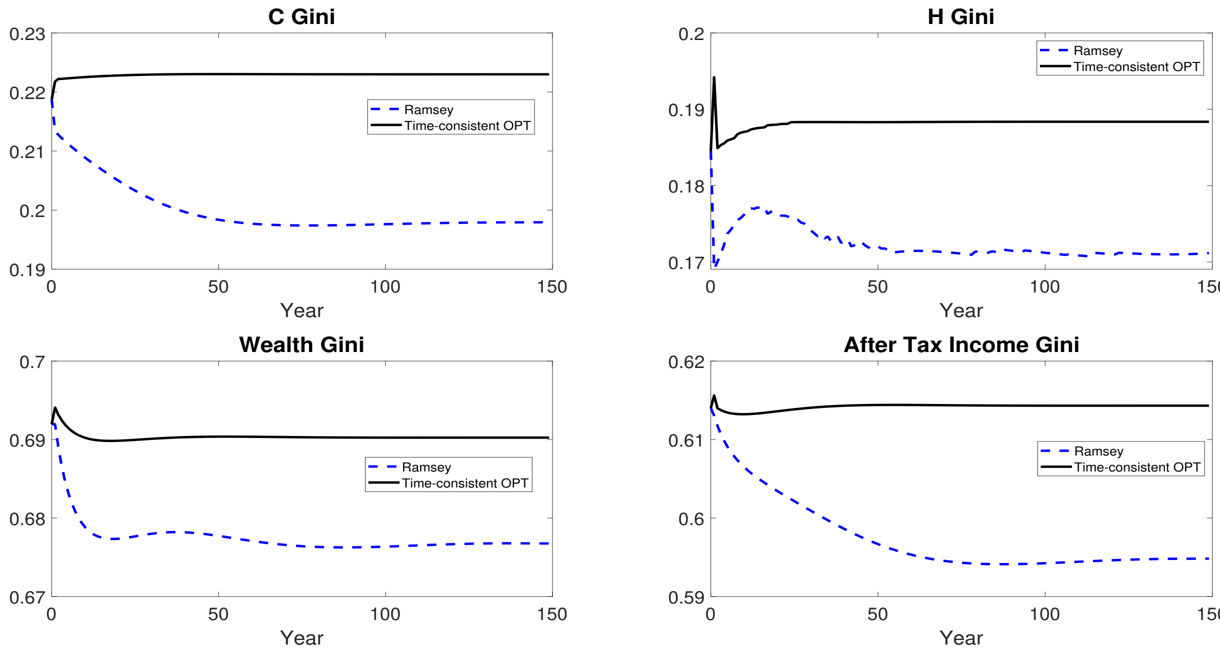

Figure 7: Time-consistent and -inconsistent Optimal Labor Taxes: Distributional Outcomes

The optimal labor income tax cases show similar implications for the macroeconomy, inequality, and welfare. As can be seen in Table 2, an evident difference is that the gap in the welfare outcomes is more substantial according to the commitment structure. This quantitative difference is driven by the longer working hours in the case with the time-consistent optimal policy, as shown in Figure 6; however, as Figure 7 shows, the evolution of distributions is still crucial in understanding their welfare consequences.

6.2 Time-consistent Optimal Tax versus Sequential Voting

In this section, given a lack of commitment, I examine how the political structure affects equilibrium outcomes. Figure 8 shows the tax/transfers paths according to the political structure. Figure 8 implies that the economy with sequential voting has more frugal income taxes both in proportional income tax and in labor income tax. In the proportional income tax system, the political economy shows lower tax rates than the initial steady state's. This reduction in tax rates converts transfers to a lump-sum tax in the economy with sequential voting. In the labor income case, the economy with sequential voting generates tax rates lower than 5 percent. This lower tax rate brings about a lump-sum tax of 0.1 percent of the initial GDP.

As mentioned in the previous section, because the dimension of taxes is one in my model, the majority voting rule boils down to the median voting rule. Thus, the tax outcomes in the political economies are equivalent to tax rates most preferred by the median voter. Figure 8 implies that the

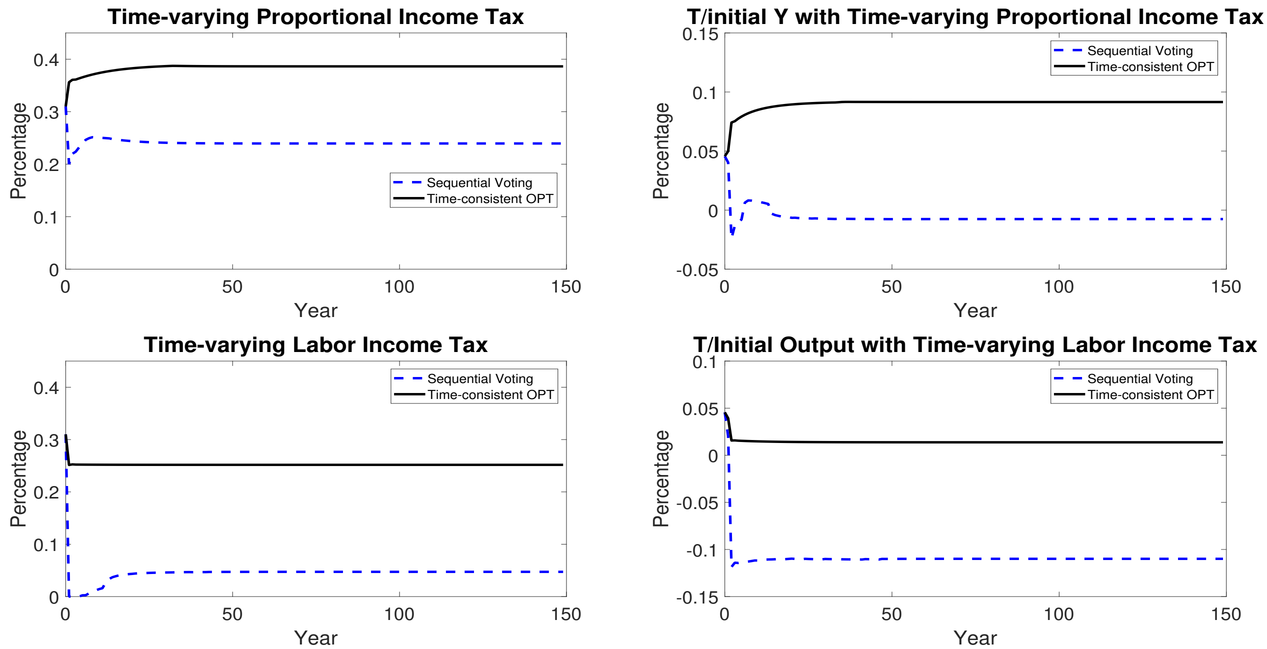

Figure 8: Time-consistent Optimal Policy vs. Sequential Voting: Tax Transition Paths

median voter is substantially different from the consumption-poor because this voter is not willing to bear large distortions, preferring a lower tax rate with fewer transfers.

This median voter's decision considering only his own efficiency affects aggregate outcomes. Figure 9 implies that the economy with sequential voting shows larger aggregate variables. While the aggregate consumption declines with the time-consistent optimal policy, it increases with sequential voting. And the overall level of hours worked, the aggregate capital, and the efficiency unit of the aggregate labor are also higher in the economy with sequential voting. These results are driven by fewer distortions—driven by lower tax rates—that the median voter prefers and chooses.

Figure 10 shows the distributional outcomes according to the political structure. Figure 10 implies that the political economy does not bring a reduction in inequality. Because he takes only into account his own efficiency, the median voter has no interest in overall inequality. Of course, the median voter considers how much transfers he would get in voting. However, this consideration does not mean that the median voter thinks of a desirable level of inequality. Therefore, the level of transfers relies on the median voter's position. In my quantitative model, the median voter is substantially different from the consumption-poor, thereby preferring lower tax rates with few transfers. As a result, the economy's overall inequalities with sequential voting are greater than with the time-consistent optimal income tax.

This median voter's attitude for efficiency and redistribution brings stark differences to wel-

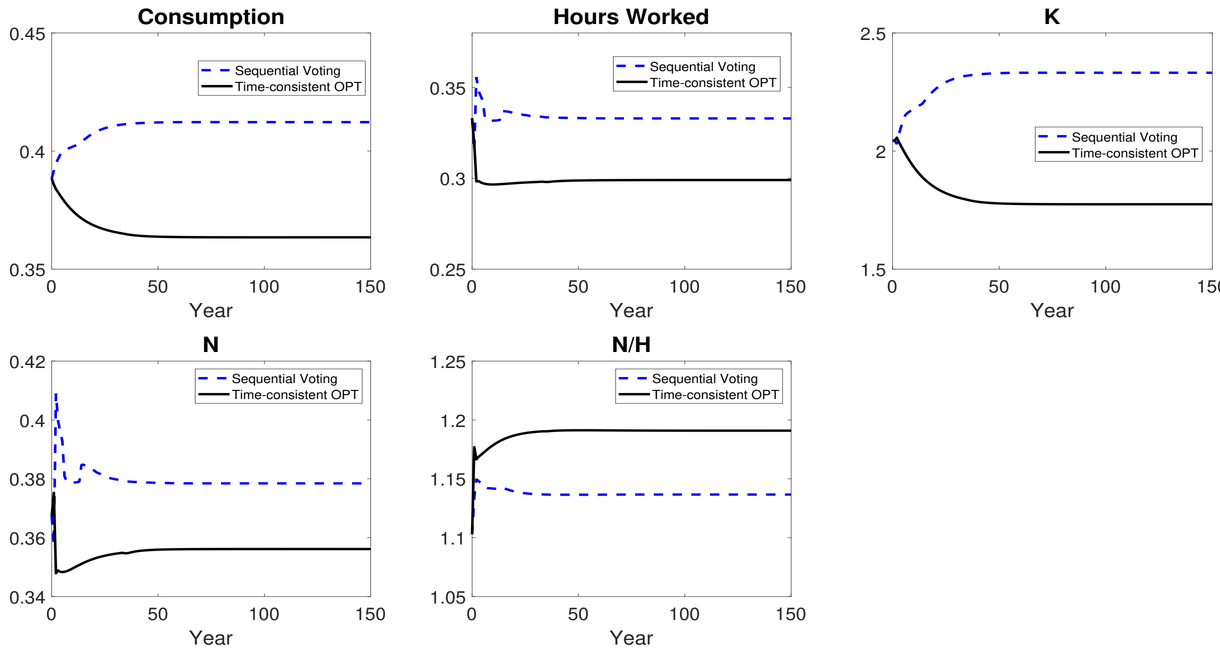

Figure 9: Optimal Policy and Sequential Voting for Income Tax: Aggregate Outcomes

Table 3: Welfare Outcomes According to Political Structure

Welfare (CEV)	Optimal Policy	Sequential Voting
INC TAX	+0.57%	-2.13%
Labor INC TAX ($\tau_k = 0.31$)	-1.24%	-7.8%

fare outcomes between the case with sequential voting and that with the time-consistent optimal policy. Table 3 implies that welfare losses with sequential voting are substantially associated with its more significant inequalities. Although the increased consumption plays a role in improving welfare, longer hours worked and more significant inequalities in consumption and hours worked overwhelm the buoyant force, leading to welfare's aggravation.

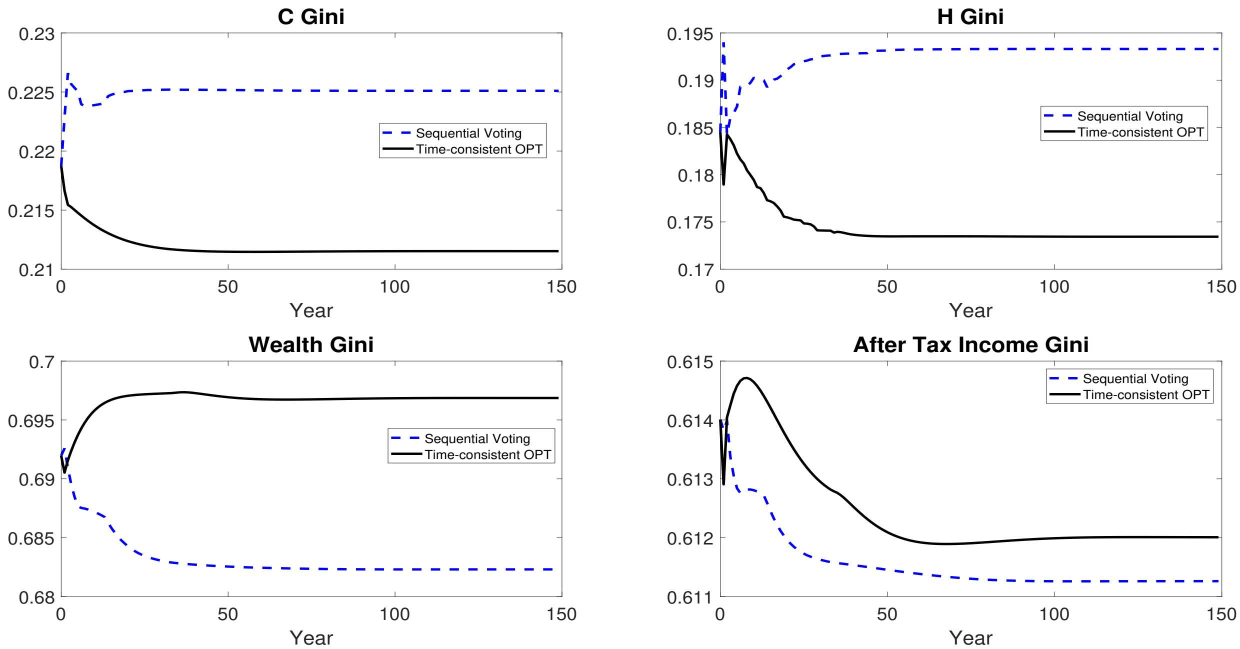

Figure 10: Optimal Policy and Sequential Voting for Income Tax: Distributional Outcomes

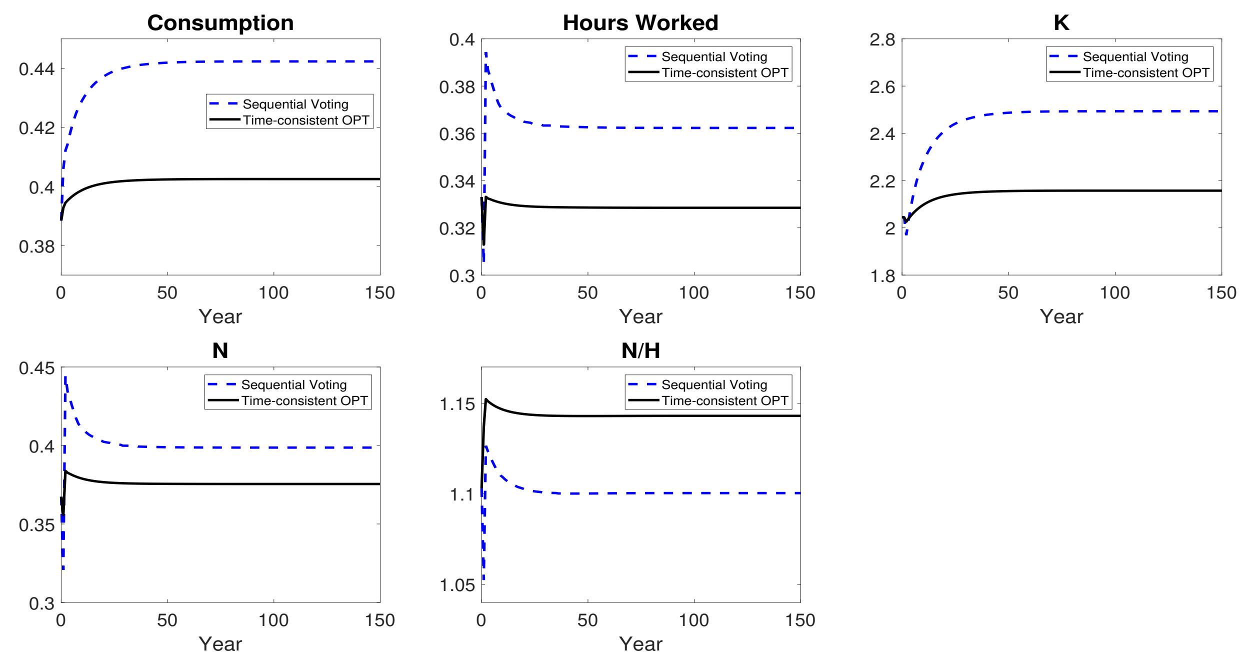

Figure 11: Optimal Policy and Sequential Voting for Labor Income Tax: Aggregate Outcomes

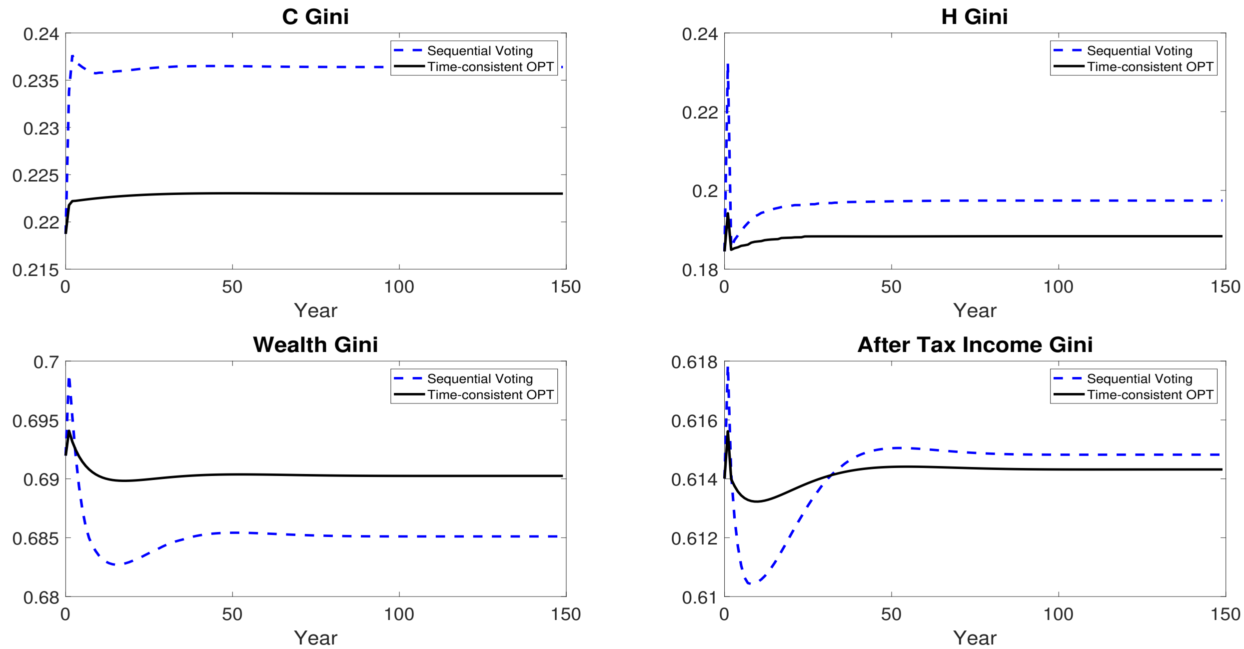

Figure 12: Optimal Policy and Sequential Voting for Labor Income Tax: Distributional Outcomes

The cases with labor income tax shows similar implications for the macroeconomy, inequality, and welfare. As can be seen in Table 3, an evident difference is that the gap in the welfare outcomes is greater according to the political structure. This quantitative difference is driven by larger gaps in labor taxes between these two economies, as can be seen in the bottom-left panel of Figure 8. In particular, the median voter in the political economy pursues very low tax rates, causing the substantial lump-sum tax. Although these quantitative difference appears, the macroeconomic and distributional implications and welfare consequences are very similar to the case with proportional income tax: Although the increased consumption plays a role in improving welfare, longer hours worked and more significant inequalities in consumption and hours worked overwhelm the buoyant force, leading to welfare's deterioration.

7 Conclusion

This paper examines how differences in the government's political and commitment structure affect the macroeconomy, inequality, and welfare. I characterize the MPE with heterogeneous agents using the generalized Euler equation; develop a numerical solution method for this game; and apply this method to the standard incomplete markets model with uninsurable idiosyncratic income risk, in which its tax/transfer system is endogenously determined according to its government's political and commitment structure.

I find that both—commitment and political structure—are significantly crucial for the aggregate economy, inequality, and welfare. Commitment and political system endogenously shape the government's behaviors, resulting in different dynamics of taxes. As a result, these different policies give rise to disparities in individual decisions, causing the aggregate economy, inequality, and welfare to differ.

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 would make 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. Kim's (2021) method would make this direction reachable. She extends Reiter's (2010) backward induction method to solve an overlapping generations model with aggregate uncertainty. 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.

_ **and Dan Peled**, "Social insurance and taxation under sequential majority voting and utilitarian regimes," *Journal of Economic Dynamics and Control*, 1995, *19* (8), 1511–1528.

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.

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.

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

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.

Dyrda, Sebastian and Marcelo Pedroni, "Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Income Risk," *Review of Economic Studies, Forthcoming*, 2022.

Farhi, Emmanuel, Christopher Sleet, Ivan Werning, and Sevin Yeltekin, "Non-linear capital taxation without commitment," *Review of Economic Studies*, 2012, *79* (4), 1469–1493.

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.

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.

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

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.

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 voting outcome or optimal policy without government 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 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 political economy with sequential voting and its stationary distribution. In this political equilibrium, the voted policies are 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 way of constructing the reference distributions in Reiter (2002, 2010), reflecting the features of economies in the MPE wherein how many times a policy off the equilibrium takes place is unknown before simulations.

Here, I show how to apply the algorithm to the political economy with sequential voting, which is the most complicated and informative in the three economies. 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 voting outcome, 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 (voting). 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 voting outcome. Explicitly,

$$m' = G(m_{i_m}, \tau_{i_\tau}, \tau') \tag{49}$$

$$\tau' = P(m_{i_m}, \tau_{i_\tau}) \tag{50}$$

$$\tau' = \tau^m(m_{i_m}, \tau_{i_\tau}) \tag{51}$$

$$w = W(m_{i_m}, \tau_{i_\tau}) \tag{52}$$

$$T = TR(m_{i_m}, \tau_{i_\tau}) \tag{53}$$

(49) is to approximate Γ, (50) is to do Ψ, (51) is for the voting outcome, (52) is the mapping for the market wage, and (53) is the mapping for trasnfers.

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

$$v^{n+1}(a, \epsilon; m_{i_m}, \tau_{i_\tau}) = \max_{c,a',n} u(c, 1-n) + \beta \sum_{\epsilon'} v^n(a', \epsilon_j, m', \tau') \tag{54}$$

such that

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

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

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 $\tilde{T}$.

$$\tilde{m}' = \int g_a^{n+1}(a, \epsilon; m_{i_m}, \tau_{i_\tau}) \mu(da, \epsilon; m_{i_m}, \tau_{i_\tau}) \tag{55}$$

$$\tilde{w} = (1-\theta)\left(\frac{m_i}{N}\right)^{\theta} \tag{56}$$

$$\tilde{T} = \tau_{i_\tau}(r(m_{i_m}, \tau_{i_\tau}) m_i + w(m_{i_m}, \tau_{i_\tau}) N) \tag{57}$$

where

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

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 (54) 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

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

such that

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

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.

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

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

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

where

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

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 $\psi^q(a, \epsilon, m, \tau)$ as follows:

$$\psi^q(a, \epsilon; m_{i_m}, \tau_{i_\tau}) = \operatorname*{argmax}_{\tilde{\tau}'} \hat{v}(a, \epsilon; m_{i_m}, \tau_{i_\tau}, \tilde{\tau}') \tag{58}$$

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

b) For each aggregate grid (m_{i_m}, τ_{i_τ}), using the proxy distribution $\mu(a, \epsilon; m_{i_m}, \tau_{i_\tau})$, compute the policy outcome $\tau^{m,q}(m_{i_m}, \tau_{i_\tau})$ that satisfies

$$\int_{\{\psi^q(a,\epsilon;\mu,\tau) \leq \tau^{m,q}(m_{i_m}, \tau_{i_\tau})\}} \mu(da, \epsilon; m_{i_m}, \tau_{i_\tau}) \geq \frac{1}{2} \tag{59}$$

$$\int_{\{\psi^q(a,\epsilon;\mu,\tau) \geq \tau^{m,q}(m_{i_m}, \tau_{i_\tau})\}} \mu(da, \epsilon; m_{i_m}, \tau_{i_\tau}) \geq \frac{1}{2} \tag{60}$$

$$\tag{61}$$

c) For each aggregate grid (m_{i_m}, τ_{i_τ}), if $P^q(m_{i_m}, \tau_{i_\tau}) = \tau^{m,q}(m_{i_m}, \tau_{i_\tau})$, 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 $P^{q+1} = \omega \cdot P^q + (1 - \omega) \cdot \tau^{m,q}$, 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 a political economy with sequential voting 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 median voting rule 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

$$d((m_0, \tau_0), (m_1, \tau_1)) = (m_0 - m_1)^{-4} + (\tau_0 - \tau_1)^{-4} \tag{62}$$

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 $(m_{\tilde{t}}, \tau_{\tilde{t}}) \in \Big([m_{i_m}, m_{i_m+1}), [\tau_{i_\tau}, \tau_{i_\tau+1})\Big)$,

$$\mu^r(a, \epsilon; m_{i_m}, \tau_{i_\tau}) = \sum_{\tilde{t}=1}^{T \times N_m \times N_\tau} \frac{d((m_{i_m}, \tau_{i_\tau}), (m_{\tilde{t}}, \tau_{\tilde{t}}))}{d_1(m_{i_m}, \tau_{i_\tau})} \mu_{\tilde{t}}(a, \epsilon). \tag{63}$$

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.

$$\min_{\{\mu_{i_a,i_\epsilon}\}_{i_a=1,i_\epsilon=1}^{N_a,N_\epsilon}} \sum_{i_a=1}^{N_a}\sum_{i_\epsilon=1}^{N_\epsilon}\Big(\mu_{i_a,i_\epsilon} - \mu^r_{i_a,i_\epsilon}(i_m,i_\tau)\Big)^2 \tag{64}$$

$$\sum_{i_a=1}^{N_a} \mu_{i_a,1_\epsilon} = \sum_{i_a=1}^{N_a} \mu^r_{i_a,i_\epsilon}(i_m,i_\tau) \text{ for } i = 1,\cdots,N_\epsilon \tag{65}$$

$$\sum_{i_\epsilon=1}^{N_\epsilon}\sum_{i_a=1}^{N_a} \mu_{i_a,i_\epsilon}\cdot a_{i_a} = m_{i_m} \tag{66}$$

$$\mu_{i_a,i_\epsilon} \geq 0 \tag{67}$$

The first-order condition for μ_{i_a,i_ϵ} with λ_i as the LaGrange multiplier for (65) and ω the multiplier (66) is

$$2(\mu_{i_a,i_\epsilon} - \mu^r_{i_a,i_\epsilon}(i_m,i_\tau)) - \lambda_i - \omega\cdot a_{i_a} = 0 \tag{68}$$

If I ignore the non-negative constraints for probabilities in (67), I have N_ϵ constraint in (65). 1 constraint in (66) and $N_a \times N_\epsilon$ first-order conditions in (67). These are a system of $N_a\times N_\epsilon + N_\epsilon + 1$ linear equations in $\Big(\{\mu_{i_a,i_\epsilon}\}_{i_a=1,i_\epsilon=1}^{N_a,N_\epsilon}, \{\lambda_{i_\epsilon}\}_{i_\epsilon}^{N_\epsilon}, \omega\Big)$.

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 13. 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

$$\mathbf{A}((i_\epsilon - 1)\times N_a + i_a, (i_\epsilon - 1)\times N_a + i_a) = 2 \tag{69}$$

$$\mathbf{A}(N_\epsilon \times N_a + i_\epsilon, (i_\epsilon - 1)\times N_a + i_a) = 1 \text{ for } i_\epsilon = 1,\cdots,N_\epsilon \tag{70}$$

$$\mathbf{A}(N_\epsilon \times N_a + N_\epsilon + 1, (i_\epsilon - 1)\times N_a + i_a)) = a_{i_a}. \tag{71}$$

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

$$\mathbf{A}((i_\epsilon - 1)\times N_a + i_a, N_\epsilon \times N_a + i_\epsilon) = -1 \tag{72}$$

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

$$\mathbf{A}((i_\epsilon - 1) \times N_a + i_a, N_\epsilon \times N_a + N_\epsilon + 1) = -a_{i_a}. \tag{73}$$

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

$$\mathbf{b}((i_\epsilon - 1) \times N_a + i_a) = 2\mu^r_{i_a,i_\epsilon}(i_m, i_\tau) \tag{74}$$

$$\mathbf{b}(N_\epsilon \times N_a + i_\epsilon) = \sum_{i_a=1}^{N_a} \mu^r_{i_a,i_\epsilon}(i_m, i_\tau) \tag{75}$$

$$\mathbf{b}(N_\epsilon \times N_a + N_\epsilon + 1) = m_{i_m}. \tag{76}$$

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

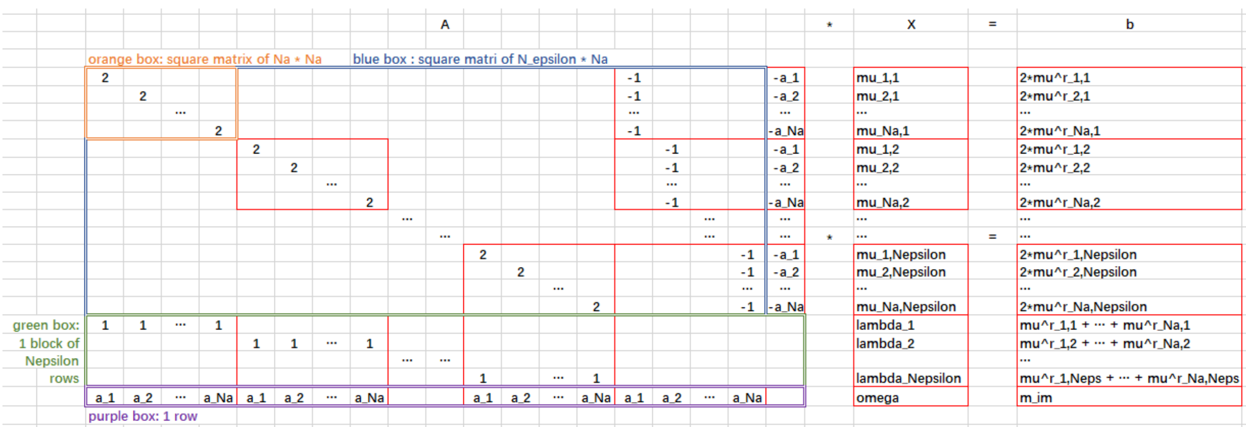

Figure 13: $\mathbf{A} \times \mathbf{x} = \mathbf{b}$

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

$$J_0 = \{j | 1 \leq j \leq N_a \times N_\epsilon \text{ and } \mathbf{x}(j) < 0\}. \tag{77}$$

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 (77).

Table 4 shows the setting of the grids in this paper. With this setting, I continue to repeat the

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)

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 efficient in a usual personal computer.

Table 5: Accuracy and Efficiency of the Solution Method

	OPT w/o Commitment	Voting
Run time	11.1 min	15.8 min
DH of m at EQ	0.394%	0.539%
DH of w at EQ	0.048%	0.046%
DH of τ at EQ	0.153%	0.263%
AVG(DH) of m	0.668%	0.577%
AVG(DH) of w	0.251%	0.202%
AVG(DH) of τ	0.129%	0.244%
MAX(DH) of m	2.133%	2.44%
MAX(DH) of w	0.949%	0.935%
MAX(DH) of τ	0.415%	1.4%

$AVG(\cdot)$ 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 Definition of $\frac{\Xi\Gamma_{t+s-1}}{\Xi K_t}$, $\frac{\Xi\Psi_{t+s-1}}{\Xi K_t}$, $\frac{\Xi\Gamma_{t+s-1}}{\Xi\tau_t}$, and $\frac{\Xi\Psi_{t+s-1}}{\Xi\tau_t}$

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

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

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

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