Optimal Taxation in a Limited Commitment Economy

forthcoming in the Review of Economic Studies

Yena Park*

University of Pennsylvania

JOB MARKET PAPER

Abstract

This paper studies optimal Ramsey taxation when risk sharing in private insurance markets is imperfect due to limited enforcement. In a limited commitment economy, there are externalities associated with capital and labor because individuals do not take into account that their labor and saving decisions affect aggregate labor and capital supply and wages, and thus the value of autarky. Therefore, a Ramsey government has an additional goal, which is to internalize these externalities of labor and capital to improve risk sharing, in addition to its usual goal — minimizing distortions in financing government expenditures. These two goals drive optimal capital and labor taxes in opposite directions. It is shown that the steady-state optimal capital income taxes are levied only to remove the negative externality of the capital, whereas optimal labor income taxes are set to meet the budgetary needs of the government in the long run, despite the presence of positive externalities of labor.

Keywords: Ramsey Taxation, Limited Enforcement

JEL Classification Codes: D52, E62, H21, H23

*Correspondence: Department of Economics, 3718 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104-6297. Email: parkye@sas.upenn.edu. I am especially grateful to my advisors Dirk Krueger and Harold Cole. I also would like to thank Philipp Kircher for valuable suggestions. I have benefited from constructive comments of three anonymous referees as well as discussions with seminar participants at the Federal Reserve Bank of Philadelphia, the University of Pennsylvania, the 2012 North American Econometric Society Meeting, and the 2013 SAET conference.

1 Introduction

What should optimal fiscal policy look like when private insurance against idiosyncratic income risk is not perfect? The canonical Ramsey tax literature cannot answer this question because it adopts a representative agent model framework which, in the presence of idiosyncratic risk, requires perfect insurance markets. It focuses exclusively on optimal ways to finance a given stream of government expenditures by minimizing distortions on labor supply and capital accumulation of households. Recently, however, there has been significant progress in answering this question. In the Bewley-Aiyagari class of models, where risk sharing is limited because of incomplete asset markets, several studies analyze optimal taxation considering both insurance effect and distortions of labor and capital.[1] On the other hand, not many studies investigate optimal taxation in a limited commitment model where imperfect risk sharing arises endogenously because of limited enforcement. Moreover, the existing literature focuses only on risk-sharing effects in endowment economies or a production economy with an exogenous labor supply (Krueger and Perri, 2011; Chien and Lee, 2012). No study has investigated optimal taxation in this class of imperfect risk sharing models when taxes distort both labor supply and capital accumulation. This paper fills that gap by studying optimal Ramsey taxation in a limited commitment economy, where the trade-off between efficiency and risk-sharing can be considered explicitly.

In the limited commitment economy, households that face a risk can trade a complete set of contingent insurance contracts. Private risk-sharing, however, is limited because private insurance contracts are not fully enforceable. These contracts can only be sustained through the threat of exclusion from participation in future insurance markets upon default. Thus, the trade of private insurance claims is limited by the extent to which contracts can keep households from defaulting. Since both the value of remaining in the risk-sharing contract and the value of default are affected by tax policies, the extent of risk sharing in this economy endogenously responds to the taxes set by the government. The government therefore needs to consider how the tax policies endogenously change private risk sharing as well as incentives to supply labor and accumulate capital.

In a limited commitment economy with production, in particular, the government can improve private risk-sharing using taxes because there are externalities associated with capital and labor. As Abraham and Carceles-Poveda (2006), Chien and Lee (2012), and others point

[1]For theoretical results, see Aiyagari (1995), Dávila et al. (2012); for quantitative results, see Conesa and Krueger (2006), Conesa et al. (2009).

out, an increase in aggregate capital raises the value of the default option by increasing the wage in financial autarky. This is an external cost of capital, since it has adverse effects on private risk sharing that individuals do not take into account when making their capital accumulation decisions. In the same manner, there is an external benefit of labor, because its increase lowers the wage in autarky and this effect is not internalized by households' decisions. Due to these externalities of labor and capital, a Ramsey government in a limited commitment economy has two objectives when it sets taxes on labor and capital. The first goal is standard: minimizing distortions when financing exogenous government expenditures. The second goal, which is new to Ramsey taxation literature, is to internalize the externalities of labor and capital in order to improve risk sharing in a limited commitment economy.

These two goals of the Ramsey government conflict with each other in the following sense. In the canonical Ramsey problem in which the Ramsey government does not have risk-sharing concerns, it is well known from Chamley (1986) and Judd (1985) that in order to minimize distortions, taxes on capital should be set to zero in the steady-state and positive labor income taxes are set to meet the budgetary needs of the government in the long run. On the other hand, if the government does not have to finance its expenditures using distortionary taxes and only takes into account the impact of taxes on private risk sharing, it wants to tax capital and subsidize labor in order to internalize the negative externality of capital and the positive externality of labor. These observations raise a crucial question: What is the optimal mix of labor income and capital income taxes when the Ramsey government tries to achieve these two conflicting goals simultaneously?

The main contribution of this paper to the literature is to provide an answer to this question. We show that in a limited commitment economy, long-run optimal capital income taxes are levied only to remove the externality of capital, not to finance government expenditures. On the other hand, all tax distortions due to government budgetary needs are still imposed on labor income taxes in the long run, despite the external benefit of labor.

This result implies that limited commitment provides a rationale for capital income taxes. Positive capital income taxes cause individuals to internalize the effects of capital investments on the tightness of enforcement constraints, improving private risk sharing. This positive capital income tax, in our view, does not contradict the zero capital taxation result of Chamley-Judd (Chamley, 1986; Judd, 1985), but rather can be interpreted as a generalized version of it. Even in a limited commitment economy, if the wage in financial autarky does not depend on equilibrium aggregate capital (that is, when capital and labor are not complementary in production), then there is no capital externality and thus the optimal capital income tax will be exactly zero in the steady-state. On the other hand, if the wage in financial autarky depends on the equilibrium aggregate capital, the externality of capital leads to a positive capital income tax only to internalize the externality, not because of the budgetary needs of the government.

A secondary contribution of this paper is to develop a method of characterizing a competitive equilibrium with limited commitment. *Ex-post* heterogeneity of agents due to idiosyncratic shocks makes the characterization of an equilibrium complicated in general. By adopting the methodology pioneered by Werning (2007) in a complete market economy to a limited commitment model, however, we provide a concise characterization of a competitive equilibrium with limited commitment in terms of aggregate allocations and Pareto-Negishi weights only, and show how we can use this result to employ the primal approach familiar from optimal taxation problems. With this characterization, the Ramsey problem boils down to a simple programming problem of choosing aggregate allocations, as in the canonical Ramsey tax literature but with additional constraints.

To investigate how large the optimal capital income tax is, we also provide a numerical example. The purpose of this computation is not to suggest optimal tax rates, but to examine the quantitative significance of the externality in shaping optimal taxation. The numerical example shows that capital income tax rates due to the externality can be very high, under the assumptions that cause a strong negative externality of capital such as permanent exclusion upon default and the same wage rate in and outside financial autarky.

Our paper is related to the literature on models with limited commitment. Earlier works by Kehoe and Levine (1993, 2001), Kocherlakota (1996), and Alvarez and Jermann (2000) analyze optimal risk-sharing contracts that are constrained by limited enforcement. Several studies have analyzed the effects of taxation with this limited commitment model. Krueger and Perri (2011) study the impact of the progressive income taxation on private risk sharing. They analyze taxation in an endowment economy and focus only on the risk-sharing effects of the progressivity of taxes, while our paper models a production economy and considers both risk-sharing and the distortionary effects of taxes. Chien and Lee (2012) decentralize a constrained efficient allocation using capital taxes, but their economy also abstracts from endogenous labor supply and focuses only on the risk-sharing effects, while our paper models capital accumulation and labor supply together. The richer structure of our paper enables us to take into account both risk-sharing effects and distortions of labor and capital in the design of optimal taxes. Abraham and Carceles-Poveda (2006) was the first paper to show the inefficiency of a competitive equilibrium in a limited commitment economy with capital accumulation, but they decentralize the constrained efficient allocation by putting an upper limit on capital accumulation instead of using taxes.

Our paper is also related to the large literature on optimal Ramsey taxation. One of the best known results of the Ramsey literature is the zero steady-sate capital tax result of Chamley (1986) and Judd (1985). Atkeson et al. (1999) summarize how it is robust in a wide class of models. This zero capital tax result has been obtained in a variety of other settings, including a model with human capital (Jones et al., 1997), neoclassical growth models with aggregate risk (Zhu, 1992; Chari et al., 1994), and overlapping generation models, provided that labor taxes can be conditional on age (Garriga, 2001; Erosa and Gervais, 2002). Conversely, there have been some attempts to find fundamental reasons why the capital income tax is positive in the long run in the Ramsey literature. Aiyagari (1995) argues that if idiosyncratic risk is not insurable because of incomplete insurance markets and borrowing constraints, the optimal capital income tax is positive due to precautionary saving. Recently, Acemoglu et al. (2011) show that a political economy considerations can give a rationale for optimal positive capital tax, since it relaxes the political economy constraints if politicians are more impatient than citizens. Our paper contributes to both strands of the literature. On the one hand, we provide another fundamental reason for a positive capital income tax — capital income taxes affect the value of default option in models with limited commitment, and thus the design of risk sharing in the economy. If there is no externality of capital, however, despite limited commitment, the optimal capital income tax is zero as in the celebrated Chamley-Judd result.

Finally, this paper is very closely related to the literature on the heterogeneous-agent Ramsey taxation. Methodologically, our paper is mostly related to Werning (2007), who analyzes Ramsey taxation with redistributional concerns. However, the heterogeneity Werning studies is *ex-ante* and permanent heterogeneity in productivity, while this paper introduces *ex-post* heterogeneity due to idiosyncratic shocks. Moreover, Werning focuses on the redistributional role of distortionary labor income taxes when lump-sum taxes are allowed, whereas this paper focuses on the optimal structure of capital tax and labor tax when the government has two conflicting goals. Dávila et al. (2012) analyze optimal capital taxes in the Bewley-Aiyagari class of incomplete market. Their analysis is different from Ramsey literature since there is no government expenditure and thus no need to use distortionary taxes to finance it. The constrained efficiency-concept in their paper, however, is in line with the Ramsey problem in our paper in the sense that government does not try to alter the market structure and instead only affects prices to improve welfare.

This paper is organized as follows. Section 2 presents the economy. Section 3 sets up the Ramsey problem using a primal approach, and Section 4 analyzes the properties of optimal Ramsey taxation. Section 5 provides numerical examples, and Section 6 concludes.

2 The economy

The economy is populated by a continuum of households with measure 1. Households are *ex-ante* identical, but *ex-post* heterogeneous because of idiosyncratic productivity shocks. Time is discrete. A household's productivity shock at period t, θ_t takes values in a finite set Θ with cardinality N. The productivity shock follows a Markov process with transition probability $\pi(\theta'|\theta)$ which is identical across households, and there is no *ex-ante* heterogeneity in initial productivity realizations (with same θ_0) among agents in period 0. We assume a law of large numbers, so that $\pi(\theta'|\theta)$ is the fraction of the population with productivity θ today and that will receive θ' tomorrow, and we also assume that $\pi(\theta'|\theta)$ has unique invariant measure $\Pi(\cdot)$. We denote by $\theta^t = (\theta_0, \cdots, \theta_t) \in \Theta^t$ the history of realized productivity shocks, and by $\pi_t(\theta^t) = \pi(\theta_t|\theta_{t-1}) \times \cdots \times \pi(\theta_1|\theta_0)$ the probability of history θ^t, with $\sum_{\theta^t \in \Theta^t} \pi_t(\theta^t) = 1$. We use the notation $\theta^s|\theta^t$ $(s \geq t)$ to represent θ^s, which is a possible continuation of productivity shock history θ^t and the probability of such continuation is denoted by $\pi_{s-t}(\theta^s|\theta^t) = \pi(\theta_s|\theta_{s-1}) \times \cdots \times \pi(\theta_{t+1}|\theta_t)$.

The preferences of an infinitely lived household are represented by the expected lifetime utility function,

$$\sum_{t=0}^{+\infty} \sum_{\theta^t} \beta^t \pi_t(\theta^t) \left[u(c_t(\theta^t)) - v(l_t(\theta^t)) \right],$$

where $c_t(\theta^t)$ denotes consumption and $l_t(\theta^t)$ denotes labor. The concavity of the lifetime utility is guaranteed by assuming $u' > 0, u'' < 0, v' > 0$, and $v'' > 0$. Standard Inada conditions on preferences are imposed: $\lim_{c \to 0} u'(c) = +\infty$, $\lim_{c \to +\infty} u'(c) = 0$, and $\lim_{l \to 0} v'(l) = 0$, and there exists $\bar{l} < +\infty$ such that $\lim_{l \to \bar{l}} v'(l) = +\infty$. The labor supply of the household is endogenous due to a disutility of labor. The agent who has history θ^t will supply $l_t(\theta^t)$, resulting in $\theta_t l_t(\theta^t)$ efficiency units of labor supply.

The production function of the representative firm is $F(K, L)$, where K is aggregate capital and L is aggregate labor in the economy. We assume that F is strictly increasing and concave in both of its arguments, continuously differentiable, and exhibits constant returns to scale. The firm produces output by renting labor and capital from households. To guarantee positive aggregate capital, we will assume that $F(0, L) = 0$, for all L. Also, there exists

$\bar{K} < +\infty$ such that $F(\bar{K}, \bar{L}) < \bar{K}$, where $\bar{L} = \bar{l} \sum_\theta \theta \Pi(\theta)$, ensuring that the steady-state level of output is finite. The depreciation rate for capital is denoted by δ.

By the law of large numbers, aggregate capital and labor are given as follows.

$$K_t = \sum_{\theta^{t-1}} \pi_{t-1}(\theta^{t-1}) k_t(\theta^{t-1}), \quad L_t = \sum_{\theta^t} \pi_t(\theta^t) \theta_t l_t(\theta^t)$$

As in the canonical Ramsey literature, we assume that the government can finance its exogenous stream of expenditures using debt and taxes on labor and capital income. It is also assumed that the government can fully commit to a sequence of taxes. Both capital and labor income taxes are linear, with tax rates at period t, $\tau_{k,t}$, $\tau_{l,t}$, respectively. We can relax this assumption by allowing non-linear labor income taxes and check the robustness of our results. (See the Supplementary Appendix C for non-linear tax analysis.) Lump-sum taxes are not permitted.[2]

Asset markets are assumed to be complete. Households can purchase Arrow-Debreu state-contingent consumption claims at period 0. Risk sharing in the private insurance markets is limited, however, because of limited commitment. A household has the option to renege on a risk-sharing contract at any time. Because of this option, the amount of contingent claims of a particular history that can be purchased in the private insurance market at period 0 is limited by the extent to which the contract can keep households from defaulting. Formally, households will face enforcement constraints that guarantee that a household would never be better off reverting permanently to financial autarky. After default, all assets and capital are seized and the household will be excluded from the asset and capital trading markets. Thus, after default, the household will live in financial autarky with only labor income.[3] Then the period utility of financial autarky is defined by

$$U^{aut}(\theta_t; w_t, \tau_{l,t}) = \max_{\tilde{c}_t, \tilde{l}_t} \left[u(\tilde{c}_t) - v(\tilde{l}_t) \right]$$
$$s.t. \quad \tilde{c}_t \leq (1 - \tau_{l,t}) w_t \theta_t \tilde{l}_t.$$

The wage in financial autarky is equal to the wage in equilibrium. We also assume that the government cannot discriminate between tax rates in and outside financial autarky. The enforcement constraint at θ^t is:

[2] Ruling out lump-sum taxes could be justified either by the fact that some agents cannot afford to pay lump-sum taxes or by the government's desire to redistribute income. In this paper, however, we assume that no lump-sum taxes are permitted without modeling these justifications explicitly, and we fully acknowledge that this assumption is subject to criticism by the New Dynamic Public Finance literature.

[3] This assumption is relaxed in section 4.3

$$\sum_{s=t}^{+\infty}\sum_{\theta^s|\theta^t}\beta^{s-t}\pi_{s-t}(\theta^s|\theta^t)\left[u(c_s(\theta^s))-v(l_s(\theta^s))\right]\geq\sum_{s=t}^{+\infty}\sum_{\theta^s|\theta^t}\beta^{s-t}\pi_{s-t}(\theta^s|\theta^t)U^{aut}(\theta_s;w_s,\tau_{l,s}),$$

where $c_s(\theta^s)$ specifies the consumption allocation to agents in the private insurance contract who experience history θ^s. Since there is no private information, given that the enforcement constraints are imposed, households never have an incentive to default; thus there will be no default in equilibrium.

In the Supplementary Appendix A, we present an analysis under the alternative assumption that the government cannot tax households in financial autarky. One advantage of taking this alternative assumption — no taxation in financial autarky — is that we can compare the Ramsey equilibrium allocation with the constrained efficient allocation (the solution to the constrained planner's problem who is only constrained by the enforcement constraint). This is because construction of the constrained planner's problem with a "no taxation in financial autarky" assumption is straightforward, while the construction of planner's problem with the "taxation in financial autarky" assumption (the assumption we take in this main text) is not.[4] An obvious weakness of "no taxation in financial autarky" assumption is that it raises the problem of interpretation — Since the financial autarky we are modeling is not a completely informal sector, this alternative assumption might not be realistic. Thus, in the main text, we maintain the assumption that the government cannot discriminate between tax rates in and outside financial autarky.

2.1 Competitive equilibrium with limited commitment

We define a competitive equilibrium with limited commitment which is similar to that of Kehoe and Levine (1993). We start with the household problem. We will denote by $p(\theta^t)$ the price of a contract at period 0 that specifies delivery of one unit of a consumption good at period t to a person who has experienced a history of productivity shocks θ^t. In period 0, agents are identical by assumption. Therefore, we normalize the price of the consumption good at period 0 to 1 ($p(\theta_0)=1$).

The household problem, given prices $\{w_t, r_t, p(\theta^t)\}$, linear taxes $\{\tau_{l,t}, \tau_{k,t}\}$, initial holding of one-period government bond B_0, and initial capital holding K_0 is written as:

[4] If the planner can tax labor income in financial autarky, he will tax as much as he can to make the enforcement constraints nonbinding, reverting to a complete market, which goes against the notion of constrained efficiency.

$$
\begin{aligned}
\max_{\{c_t,l_t,k_{t+1}\}} \quad & \sum_{t=0}^{+\infty}\sum_{\theta^t}\beta^t\pi_t(\theta^t)\Big[u(c_t(\theta^t))-v(l_t(\theta^t))\Big] \qquad (1)\\
\text{s.t.} \quad & \sum_{t=0}^{+\infty}\sum_{\theta^t}p(\theta^t)\Big[c_t(\theta^t)+k_{t+1}(\theta^t)-(1-\tau_{l,t})w_t\theta_t l_t(\theta^t)-(1+r_t(1-\tau_{k,t}))k_t(\theta^{t-1})\Big]\le B_0\\
& \sum_{s=t}^{+\infty}\sum_{\theta^s}\beta^{s-t}\pi_{s-t}(\theta^s|\theta^t)[u(c_s(\theta^s))-v(l_s(\theta^s))]\ge\sum_{s=t}^{+\infty}\sum_{\theta^s}\beta^{s-t}\pi_{s-t}(\theta^s|\theta^t)U^{aut}(\theta_s;w_s,\tau_{l,s})\\
& k_0=K_0.
\end{aligned}
$$

This is the standard household problem in an Arrow-Debreu equilibrium except that it has additional enforcement constraints restricting the consumption possibility set.[5] Notice that we assume that at period 0, every household has the same amount of government bond B_0 and capital K_0. Now we can define the competitive equilibrium with limited commitment.

Definition 1. *Given initial capital K_0 and government bond holding B_0, a* ***competitive equilibrium with limited commitment*** *is a sequence of allocations $\{c_t(\theta^t), l_t(\theta^t), k_{t+1}(\theta^t)\}$, prices $\{p(\theta^t), r_t, w_t\}$, and a sequence of government expenditures and taxes $\{G_t, \tau_{l,t}, \tau_{k,t}\}$ such that*

1. *given $\{p(\theta^t), r_t, w_t, \tau_{l,t}, \tau_{k,t}\}$, $\{c_t(\theta^t), l_t(\theta^t), k_{t+1}(\theta^t)\}$ solves the household problem with enforcement constraints.*
2. *$(K_t, L_t) \in \arg\max_{\tilde{K}_t,\tilde{L}_t} F(\tilde{K}_t,\tilde{L}_t) - w_t\tilde{L}_t - (r_t+\delta)\tilde{K}_t$, for all t.*
3. *government's budget constraint holds: $\sum_{t=0}^{+\infty}\sum_{\theta^t} p(\theta^t)[\tau_{l,t}w_tL_t + \tau_{k,t}r_tK_t - G_t] = B_0$.*
4. *markets clear: for all t,*
   i. *$\sum_{\theta^t}\pi_t(\theta^t)c_t(\theta^t) + K_{t+1} + G_t = F(K_t, L_t) + (1-\delta)K_t$*
   ii. *$K_t = \sum_{\theta^{t-1}}\pi_{t-1}(\theta^{t-1})k_t(\theta^{t-1})$*
   iii. *$L_t = \sum_{\theta^t}\pi_t(\theta^t)\theta_t l_t(\theta^t)$.*

3 The Ramsey problem

In this section, we formulate the Ramsey government problem. We apply the standard technique of analyzing optimal fiscal policy, the so-called "primal approach," in which the

[5] Here, we included a sequence of capital allocations in the budget constraint to explicitly show the effect of capital tax on households' decisions. For both capital and contingent claims to be traded, the prices $p(\theta^t)$ and r_t must satisfy no-arbitrage condition: $\sum_{\theta_{t+1}}\frac{p(\theta^t,\theta_{t+1})}{p(\theta^t)} = \frac{1}{1+r_{t+1}(1-\tau_{k,t+1})}$.

government chooses an allocation directly (rather than a set of tax rates), among the set of allocations that can be decentralized as a competitive equilibrium with limited commitment.

The first step of the primal approach is to characterize the set of allocations that are sustainable as a competitive equilibrium with limited commitment for some sequence of prices and taxes. In our environment with heterogeneous agents, since individual allocation is different from aggregate allocation, the simple representation of a competitive equilibrium in terms of aggregate allocation in the standard Ramsey literature does not directly apply to characterizing an equilibrium with limited commitment. The individual allocations, however, can be expressed as functions of aggregate allocations and Pareto-Negishi weights only. Thus, we can characterize a competitive equilibrium with limited commitment in terms of aggregate allocation and Pareto-Negishi weights[6] by introducing the standard implementability constraint and additional constraints on Pareto-Negishi weights. We then formulate the Ramsey problem as the problem of choosing aggregate allocation and Pareto-Negishi weights subject to a series of constraints that guarantee that the individual allocation associated with aggregate allocation and Pareto-Negishi weights can be decentralized as a competitive equilibrium with limited commitment.

The reason we can express the individual allocations as functions of aggregate allocations and Pareto-Negishi weights follows. In a competitive equilibrium with limited commitment, given an aggregate allocation, individual allocations will be assigned based on binding enforcement constraint histories, because enforcement constraints are the only source of imperfect risk sharing. Since we can find stochastic Pareto-Negishi weights that summarize all the information on binding enforcement constraints, aggregate allocation and Pareto-Negishi weights will be sufficient to characterize an equilibrium allocation. Below, we explicitly show how to construct such Pareto-Negishi weights, exploiting the cumulative Lagrange multiplier of the enforcement constraints.

This representation of equilibrium involving aggregate allocation is relevant for constructing the Ramsey problem because for the Ramsey government to achieve its goal the only thing that matters is aggregate allocation. The goal of the Ramsey government is to maximize welfare subject to an allocation being a competitive equilibrium with limited commitment. To achieve this goal, the government wants to 1. minimize distortions when financing government expenditures, and 2. internalize the externalities of capital and labor. For both objectives, only aggregate allocations matter. For the first objective, the distortions due to taxes are confined to aggregate allocations because given an aggregate allocation, individual

[6] We closely follow the treatment of Werning (2007) for this characterization.

allocations will be assigned efficiently, even though the notion of efficiency is constrained efficiency, which is efficient relative to constrained Pareto-Negishi weights. For the second objective, only aggregate labor and capital allocations matter because the only channel of the externality is autarky wage effects, which are determined by aggregate labor and capital. Thus, the characterization with aggregate allocations is relevant for the purpose of analyzing optimal Ramsey taxation in this economy.[7]

In the following section, as a stepping stone, we formulate the static problem of allocating individual consumption and labor for given aggregate consumption and labor and Pareto-Negishi weights, whose solution is individual consumption and labor which are represented as functions of aggregate allocations and Pareto-Negishi weights. We then characterize a competitive equilibrium with limited commitment using this solution.

3.1 Planner's static problem and the fictitious representative agent

With linear taxes, individual consumption and labor allocations of a competitive equilibrium with limited commitment achieve static efficiency for a given aggregate allocation. We can see this static efficiency from the equal marginal rates of substitution of consumption and labor across agents. That is, distortions due to taxation will only be confined to determination of aggregate allocations. Exploiting this static efficiency, we define the planner's static problem whose solutions are individual consumption and labor allocations of a potential competitive equilibrium with limited commitment.

For given aggregate consumption and labor (C_t, L_t) and Pareto-Negishi weights $\{M_t(\theta^t)\}$, the planner's static problem is defined as

$$U^f(C_t, L_t; M) = \max_{c_t(\theta^t), l_t(\theta^t)} \sum_{\theta^t} \pi_t(\theta^t) M_t(\theta^t) \left[u(c_t(\theta^t)) - v(l_t(\theta^t)) \right] \tag{2}$$

$$s.t. \quad \sum_{\theta^t} \pi_t(\theta^t) c_t(\theta^t) = C_t \tag{3}$$

$$\sum_{\theta^t} \pi_t(\theta^t) \theta_t l_t(\theta^t) = L_t. \tag{4}$$

We will denote the solution to the above problem by

$$h(\theta^t, C_t, L_t; M) = \left(h^c(\theta^t, C_t, L_t; M), h^l(\theta^t, C_t, L_t; M) \right). \tag{5}$$

[7] Of course, to calculate the exact level of taxes, we need to know the Pareto-Negishi weights, which will be defined and characterized below.

The planner distributes consumption and labor across households to maximize the weighted sum of expected utilites weighted by the Pareto-Negishi weights subject to aggregate feasibility. By choosing a particular Pareto-Negishi weight $\{M_t(\theta^t)\}$, the associated efficient consumption and labor can indeed be those of a competitive equilibrium with limited commitment.[8] As we will demonstrate below, setting Pareto-Negishi weight $M_t(\theta^t)$ to the sum of multipliers associated with enforcement constraints over history θ^t does exactly this, and then the individual allocations — the solution of the planner's static problem (5) — will be expressed as functions of aggregate allocations and Pareto-Negishi weights.

Following Werning (2007), we will call the value function of the planner's static problem U^f a utility of a fictitious representative agent. The superscript of U^f stands for "fictitious." The reason for this name is that we can treat this economy as the one where there is only one representative agent whose preferences are represented by the utility $U^f(C_t, L_t; M)$. Using these definitions and notations, we now can characterize a competitive equilibrium with limited commitment.

3.2 Simple characterization of a competitive equilibrium with limited commitment

The next proposition shows how we can characterize the set of aggregate allocations and Pareto-Negishi weights that can be supported as a competitive equilibrium with limited commitment. That is, we derive conditions that the aggregate allocations and Pareto-Negishi weights have to satisfy so that individual allocations associated with these aggregate allocations and Pareto-Negishi weights can be decentralized as a competitive equilibrium with limited commitment. Imposing these constraints on the Ramsey government's problem then ensures that any aggregate allocation and Pareto-Negishi weights chosen by the government can be supported as a competitive equilibrium with limited commitment. For this proposition, we need an additional assumption on utility function and allocations.

Assumption 2. *There exist constants $\zeta_1, \zeta_2 < +\infty$ such that for all t and θ^t,*

$$\left|u(c_t(\theta^t))\right| \leq \zeta_1 u'(c_t(\theta^t))c_t(\theta^t), \quad \left|v(l_t(\theta^t))\right| \leq \zeta_2 v'(l_t(\theta^t))\theta_t l_t(\theta^t).$$

[8]This planner's static problem is exactly the same as that of Werning (2007) except that the planner here faces stochastic Pareto-Negishi weights while the planner in Werning (2007) faces fixed weights. Werning used the term, 'market weights' instead of Pareto-Negishi weights. The weights in that paper are nonstochastic because they depend only on an innate heterogeneity in productivity that is invariant over time.

Assumption 2 is technical assumption we need when we construct a competitive equilibrium from the aggregate allocation and Pareto-Negishi weights. We will construct individual allocations, prices and Lagrange multipliers that satisfy all the first order conditions of both the household problem and the firm's problem as well as market clearing conditions. In the presence of an infinite sequence of enforcement constraints, however, first order conditions of the Lagrangian might not be sufficient conditions of a household's optimality even if the objective function is concave and a constraint set is convex, because the infinite sum in the Lagrangian might not converge. Assumption 2 guarantees convergence of the infinite sum of Lagrange multipliers, validating the Lagrangian method.[9] Notice that Assumption 2 is a joint requirement on the allocation and the utility functions. If both consumption and labor are uniformly bounded away from zero, this assumption is satisfied automatically.[10]

Proposition 3. *Suppose Assumption 2 holds. Given initial government bond holding B_0, initial capital holding K_0, and initial capital tax rate $\tau_{k,0}$, an aggregate allocation $\{C_t, K_t, L_t\}$ can be supported as a competitive equilibrium with limited commitment if and only if there exist Pareto-Negishi weights $\{M_t(\theta^t)\}$ so that the following conditions (i) - (v) hold.*

(*i*) ***Resource constraint:*** $C_t + K_{t+1} + G_t = F(K_t, L_t) + (1-\delta)K_t \quad \text{for all } t.$ (6)

(*ii*) ***Implementability constraint:***

$$\sum_t \sum_{\theta^t} \beta^t \pi_t(\theta^t) \left[U_c^f(C_t, L_t; M) h^c(\theta^t, C_t, L_t; M) + U_L^f(C_t, L_t; M)\theta_t h^l(\theta^t, C_t, L_t; M)\right]$$
$$= U_c^f(C_0, L_0; M)\left\{\left[1 + (F_K(K_0, L_0) - \delta)(1-\tau_{k,0})\right] K_0 + B_0\right\}, \tag{7}$$

where $h^c(\cdot)$, $h^l(\cdot)$ are the solution of planner's static problem (2).

(*iii*) ***Enforcement constraint:*** *for all θ^t,*

$$\sum_{s=t}^{+\infty} \sum_{\theta^s} \beta^{s-t} \pi_{s-t}(\theta^s|\theta^t) \left[u(h^c(\theta^s, C_s, L_s; M)) - v(h^l(\theta^s, C_s, L_s; M))\right]$$
$$\geq \sum_{s=t}^{+\infty} \sum_{\theta^s} \beta^{s-t} \pi_{s-t}(\theta^s|\theta^t) U^{aut}(\theta_s, L_s, K_s; \tau_{l,s}) \tag{8}$$

$$\text{where} \quad U^{aut}(\theta_s, L_s, K_s; \tau_{l,s}) = \max_{\tilde{c}_s, \tilde{l}_s} \left[u(\tilde{c}_s) - v(\tilde{l}_s)\right] \quad s.t. \quad \tilde{c}_s \leq F_L(K_s, L_s)(1-\tau_{l,s})\theta_s \tilde{l}_s$$

$$\text{for labor income tax rates } \{\tau_{l,t}\} \text{ s.t. } (1-\tau_{l,t}) = -\frac{U_L^f(C_t, L_t; M)}{F_L(K_t, L_t) U_c^f(C_t, L_t; M)}.$$

(*iv*) ***Monotonicity:*** $M_{t+1}(\theta^t, \theta_{t+1}) \geq M_t(\theta^t) \geq 1 \quad \text{for all } \theta^t, \theta_{t+1}, \text{ and } M_0(\theta_0) = 1.$ (9)

[9] Assumption 2 is a modified version of condition (c) of Proposition 4.6. of Alvarez and Jermann (2000). Alvarez and Jermann (2000) made the same assumption to guarantee that each term of summation associated with enforcement constraints converges to a finite limit in the Lagrangian. This assumption is needed for the "if" part of Proposition 3. The "only if" part of this proposition still holds without Assumption 2.

[10] For a further sufficient condition, see Alvarez and Jermann (2000)

(*v*) ***High MRS:*** *If the enforcement constraint at* (θ^t, θ_{t+1}) *is not binding, then* $M_t(\theta^t) = M_{t+1}(\theta^t, \theta_{t+1})$.

Given an aggregate allocation, individual allocations can then be computed using equation (5).[11]

In the next subsection, we will prove the "only if" part of the proposition. See the Appendix A for the proof of the "if" part of the proposition.

Proposition 3 says that the aggregate allocation and Pareto-Negishi weights $\{C_t, L_t, K_t, \{M_t(\theta^t)\}\}$ that satisfy *(i)-(v)* can be implemented as a competitive equilibrium.[12] That is, there will be some prices and taxes that support the individual allocations associated with the aggregate allocation $\{C_t, L_t, K_t\}$ and Pareto-Negishi weights $\{M_t(\theta^t)\}$ — the solution of planner's static problem defined in the previous subsection — as a competitive equilibrium. Notice that conditions *(iii)*,*(iv)*, and *(v)* exactly capture the risk-sharing rule of a competitive equilibrium with limited commitment. The relative size of $\{M_t(\theta^t)\}$ across history $\{\theta^t\}$ at period t will determine the consumption share of the agent with θ^t. At time 0, the initial Pareto-Negishi weight is equal to 1 across all agents because there is no *ex-ante* heterogeneity. If it were a standard Arrow-Debreu model, then this weight would be fixed. In the presence of enforcement constraints, however, this weight increases over time whenever the enforcement constraint is binding. The consumption share of agents with non-binding constraints will then drift downward because these agents' Pareto-Negishi weights stay constant, while those of others increase. On the other hand, the consumption share of agents whose constraints are binding will jump up to guarantee that these agents do not renege on their contracts.

3.3 Proof of the "only if" part of Proposition 3

In this subsection, we prove the "only if" part of Proposition 3. Deriving conditions in Proposition 3 is important in formulating the Ramsey problem because they characterize a

[11]We are borrowing the name of the last condition, "High MRS" from Alvarez and Jermann (2000). The name comes from the fact that the MRS between today's consumption and tomorrow's consumption of the non-binding agent is the highest. Holding Pareto-Negishi weights constant guarantees the highest MRS for the agents with non-binding constraints.

[12]The enforcement constraint is specified for the specific tax rate in autarky, which is designated in the proposition because of the assumption that the Ramsey government cannot discriminate between tax rates in and outside autarky. Since the labor income tax rate in autarky is implied by aggregate allocations and Pareto-Negishi weights, the characterization of a competitive equilibrium with limited commitment is given only with aggregate allocations and Pareto-Negishi weights.

competitive equilibrium with limited commitment using a system of usable equalities and inequalities that can be directly incorporated into the government maximization problem.

The key step in this proof is deriving the implementability constraint. As in the canonical Ramsey problem, an implementability constraint is a household budget constraint whose prices and taxes are substituted out by the first order conditions of the household and firm. One difference is that the implementability constraint of a competitive equilibrium with limited commitment is expressed in terms of an aggregate allocation and Pareto-Negishi weights, while that in a canonical Ramsey problem is expressed in terms of the representative agent's allocation. To derive the implementability constraint, we closely follow the treatment of Werning (2007), with modification of stochastic Pareto-Negishi weights. Agents in this economy are *ex-post* heterogeneous because of idiosyncratic shocks that are drawn every period, while the agents in Werning (2007) are *ex-ante* heterogeneous in their innate abilities that are invariant over time. Thus, the stochastic Pareto-Negishi weights of a competitive equilibrium with limited commitment are more involved than fixed weights in Werning (2007). We now give the proof.

First, we define the Lagrangian for the household problem (1) by attaching Lagrange multipliers λ^{KL}, $\{\beta^t \pi_t(\theta^t)\mu^{KL}(\theta^t)\}$ to the budget constraint and enforcement constraints, respectively. Then, following Marcet and Marimon (2011), we construct cumulative Lagrange multipliers for this problem, which is denoted by $M_t(\theta^t) \equiv 1 + \sum_{\theta^s \preceq \theta^t} \mu^{KL}(\theta^s),$ for all t and θ^t.

Now we re-express the Lagrangian of the household problem using these cumulative Lagrange multipliers:

$$\begin{aligned} L = & \max_{\{c_t, l_t, k_{t+1}\}} \sum_{t=0}^{+\infty} \sum_{\theta^t} \beta^t \pi_t(\theta^t) \left[\begin{array}{l} M_t(\theta^t)\left[u(c_t(\theta^t)) - v(l_t(\theta^t))\right] \\ \qquad - \{M_t(\theta^t) - 1\}\, U^{aut}(\theta_t; w_t, \tau_{l,t}) \end{array} \right] \\ & + \lambda^{KL} \left\{ \sum_t \sum_{\theta^t} p(\theta^t) \left[\begin{array}{l} w_t(1-\tau_{l,t})\theta_t l_t(\theta^t) + (1 + r_t(1-\tau_{k,t}))k_t(\theta^{t-1}) \\ -c_t(\theta^t) - k_{t+1}(\theta^t) \end{array} \right] - B_0 \right\}. \end{aligned} \tag{10}$$

Let $\{C_t, L_t, K_{t+1}\}$ be the equilibrium aggregate allocations given tax rates $\{\tau_{l,t}, \tau_{k,t}\}$ and let $\{M_t(\theta^t) | M_t(\theta^t) = 1 + \sum_{\theta^s \preceq \theta^t} \mu^{KL}(\theta^t), \theta^t \in \Theta^t\}$ be the equilibrium Pareto-Negishi weights. Then, individual allocations of a competitive equilibrium with limited commitment solve the planner's static problem (2). That is, the solution of (2) is equated with the individual allocation of an equilibrium, $\left(h^c(\theta^t, C_t, L_t; M), h^l(\theta^t, C_t, L_t; M)\right) = (c_t(\theta^t), l_t(\theta^t)).$[13]

[13]Since only the relative size of the weights $\{M_t(\theta^t)\}$ across $\theta^t \in \Theta^t$ matters for the planner's static problem, any constant times $\{M_t(\theta^t)\}_{\theta^t \in \Theta^t}$ derives the same individual allocations for a fixed aggregate allocation. However, by imposing condition (*iv*) and (*v*) of Proposition 3, we pin down Pareto-Negishi weights

Note that by the envelope theorem and the first order conditions of the planner's static problem,

$$U_c^f(C_t, L_t; M) = \lambda_t^f = M_t(\theta^t)u'(c_t(\theta^t)) \tag{11}$$

$$U_L^f(C_t, L_t; M) = \mu_t^f = -\frac{1}{\theta_t}M_t(\theta^t)v'(l_t(\theta^t)), \tag{12}$$

where λ_t^f and μ_t^f are Lagrange multipliers associated with (3) and (4), respectively.

Then, using (11), (12), and the first order conditions of Lagrangian of the household problem (10), equilibrium prices can be expressed in terms of marginal rates of substitution of a fictitious representative agent whose utility function is $U^f(C, L; M)$: for all t and θ^t,

$$p(\theta^t) = \beta^t \pi_t(\theta^t)\frac{M_t(\theta^t)u'(c_t(\theta^t))}{u'(c_0(\theta_0))} = \beta^t \pi_t(\theta^t)\frac{U_c^f(C_t, L_t; M)}{U_c^f(C_0, L_0; M)}, \tag{13}$$

$$w_t(1-\tau_{l,t}) = \frac{v'(l_t(\theta^t))}{\theta_t u'(c_t(\theta^t))} = -\frac{U_L^f(C_t, L_t; M)}{U_c^f(C_t, L_t; M)}. \tag{14}$$

Interest rates and wages are computed from the firm's problem in an equilibrium: for all t,

$$r_t = F_K(K_t, L_t) - \delta, \tag{15}$$

$$w_t = F_L(K_t, L_t). \tag{16}$$

By substituting out all prices, taxes, and individual allocations in the budget constraint of the household with the functions of aggregate allocations we derived, we can obtain the implementability constraint:

$$\sum_t \sum_{\theta^t} \beta^t \pi_t(\theta^t)\left[U_c^f(C_t, L_t; M)h^c(\theta^t, C_t, L_t; M) + U_L^f(C_t, L_t; M)\theta_t h^l(\theta^t, C_t, L_t; M)\right]$$
$$= U_c^f(C_0, L_0; M)\left\{\left[1 + (F_K(K_0, L_0) - \delta)(1-\tau_{k,0})\right]K_0 + B_0\right\}.$$

The remaining part of the proof of the "only if " part of Proposition 3 is obvious. The resource constraint and enforcement constraints come from the definition of the equilibrium as stated. Monotonicity conditions and high MRS conditions are obviously satisfied since $\{M_t(\theta^t)\}$ is the sum of nonnegative Lagrange multipliers across histories, which completes the proof.

This characterization of a competitive equilibrium with limited commitment is similar to that of the canonical Ramsey taxation literature except that we are using a fictitious

$\{M_t(\theta^t)\}$ so that $\{M_t(\theta^t)\}$ are equal to the cumulative multipliers of the enforcement constraints. This characterization of $\{M_t(\theta^t)\}$ makes the construction of a competitive equilibrium with limited commitment the most convenient in the proof of the "if" part of the proposition.

representative agent's utility and we need three additional conditions on $\{M_t(\theta^t)\}$ so that $\{M_t(\theta^t)\}$ exactly reflects the degree of risk sharing constrained by limited enforcement. As we show in the proof of the "if" part of Proposition 3 (in the Appendix A), $\{M_t(\theta^t)\}$ that satisfies these three additional conditions (*iii*), (*iv*), and (*v*) of Proposition 3 is equal to the cumulative Lagrange multipliers of the enforcement constraints of the household problem in an equilibrium.

In summary, Proposition 3 shows how we can characterize the set of aggregate allocations that can be implemented as a competitive equilibrium with limited commitment using five conditions. The Ramsey government then maximizes social welfare over the set of aggregate allocations restricted by these conditions. Finally, the equilibrium individual allocation that is associated with the aggregate allocation and Pareto-Negishi weights chosen by the government can be derived by solving the planner's static problem (2) in subsection 3.1, and equilibrium prices and taxes can be constructed using the first order conditions of households and firms.[14]

3.4 Solving the Ramsey problem

Proposition 3 enables us to formulate the Ramsey problem as the problem of choosing aggregate allocations and Pareto-Negishi weights among the implementable set:

Ramsey Problem (RP)

$$\max_{\{C_t,L_t,K_t,\{M_t(\theta^t)\}\}\in E^{KL}} \quad \sum_t \sum_{\theta^t} \beta^t \pi_t(\theta^t) \left[u(h^c(\theta^t, C_t, L_t; M)) - v(h^l(\theta^t, C_t, L_t; M))\right],$$

$$where \quad E^{KL} = \left\{ \left\{C_t, K_{t+1}, L_t, \{M_t(\theta^t)\}\right\} \middle| \begin{array}{l} \text{satisfy conditions } (i), (ii), (iii), \\ (iv), \text{ and } (v) \text{ of Proposition 3} \end{array} \right\},$$

and the constraint set of the Ramsey problem, E^{KL} is assumed to be nonempty.[15]

[14] As in the canonical Ramsey literature, we assume that if there are multiple competitive equilibria associated with given tax rates $\{\tau_{k,t}, \tau_{l,t}\}$, the government can choose the equilibrium that yields the highest utility.

[15] We acknowledge that the constraint set of the Ramsey problem, E^{KL} could be empty in general. For example, if the exogenous government expenditure is too big, then there is no equilibrium that supports such a government expenditure, even without limited commitment. The constraint set of the Ramsey problem will be nonempty if there exists a competitive equilibrium with limited commitment that can be implemented by some tax policy for given government expenditures and initial government debt. In this paper, we only consider exogenous government expenditures $\{G_t\}$ and initial government debt B_0 such that the constraint set of the Ramsey problem is nonempty, without characterizing the condition for nonemptiness — the

We want to reformulate this Ramsey problem as a simple programming problem, but it is not easy to directly incorporate the last condition — (v) the high MRS condition — into such a programming problem because it is a constraint that is imposed only when the enforcement constraint is not binding. We can drop this condition, however, because it turns out that the high MRS condition is not a binding constraint in the maximization problem. The intuition of this result is as follows. To maximize social welfare, the intertemporal marginal rates of substitution should be equalized across agents whenever the enforcement constraint allows it. Otherwise, there will be a Pareto improving allocation assignment without violating any constraint. Formally, we can show this by the following. First, we construct a relaxed Ramsey problem (RRP) where condition (v) — high MRS condition — is dropped. Then, we will show that the optimal solution of the RRP satisfies condition (v). It then follows that the solution of the RRP solves the original Ramsey problem (RP).

Relaxed Ramsey Problem (RRP)

[1] Given $\{\tau_{l,t}\}$,

$$\max_{\{C_t,L_t,K_{t+1},M_t(\theta^t)\}} \quad \sum_t \sum_{\theta^t} \beta^t \pi_t(\theta^t) \left[u(h^c(\theta^t, C_t, L_t; M)) - v(h^l(\theta^t, C_t, L_t; M))\right]$$

$$s.t. \quad C_t + K_{t+1} + G_t \leq F(K_t, L_t) + (1-\delta)K_t \tag{17}$$

$$\sum_t \sum_{\theta^t} \beta^t \pi_t(\theta^t) \left\{ \begin{array}{l} U_c^f(C_t, L_t; M) h^c(\theta^t, C_t, L_t; M) \\ +U_L^f(C_t, L_t; M)\theta_t h^l(\theta^t, C_t, L_t; M) \end{array} \right\}$$
$$= U_c^f(C_0, L_0; M) \left\{[1 + (F_K(K_0, L_0) - \delta)(1 - \tau_{k,0})]K_0 + B_0\right\} \tag{18}$$

$$\sum_{s=t}^{+\infty} \sum_{\theta^s} \beta^{s-t} \pi_{s-t}(\theta^s|\theta^t) \left\{u(h^c(\theta^s, C_s, L_s; M)) - v(h^l(\theta^s, C_s, L_s; M))\right\}$$
$$\geq \sum_{s=t}^{+\infty} \sum_{\theta^s} \beta^{s-t} \pi_{s-t}(\theta^s|\theta^t) U^{aut}(\theta_s, L_s, K_s; \tau_{l,s}) \tag{19}$$

$$M_t(\theta^t) \geq M_{t-1}(\theta^{t-1}) \tag{20}$$

[2] $1 - \tau_{l,t} = \frac{v'(h^l(\theta^t))}{u'(h^c(\theta^t))F_L(K_t,L_t)\theta_t}$, for all t and θ^t

The implementability constraint (18) in the RRP does not contain prices and taxes because we already substituted them out using optimality conditions of the household. On the other hand, the enforcement constraints still depend on the labor income taxes $\{\tau_{l,t}\}$ in autarky. However, the Ramsey problem does not maximize over this labor income tax

condition for existence. However, whenever the constraint set is nonempty, such as in the setting we used for numerical example in section 5, our analysis applies.

rates of autarky because, by assumption, the government cannot discriminate between tax rates in and outside autarky. The Ramsey government chooses an aggregate allocation and Pareto-Negishi weights to maximize the expected lifetime utility given labor income taxes $\{\tau_{l,t}\}$ in autarky; then, labor income tax rates of autarky will be pinned down by condition [2] of the RRP, which is exactly coming from the no discrimination assumption on tax rates. That is, we substitute out tax rates in autarky after maximizing over allocations, not before maximization. The reason why we do not substitute out tax rates in autarky before maximization is because of the externality. If the Ramsey government solves the maximization problem after substituting out the tax rates in autarky by condition [2], then it is solving a problem in which households internalize the autarky effects, which contradicts the assumption that households behave competitively by taking prices and the value of financial autarky as given when they decide on allocations. Technically, when we solve for allocations in the RRP, we will derive first order conditions of the maximization problem given $\{\tau_{l,t}\}$ first. Only then, will we substitute out the labor income tax rates in the first order conditions using the condition [2] of the RRP. Finally, we can solve for allocations and Pareto-Negishi weights using the first order conditions.

The following proposition shows that the solution of the RRP solves the original Ramsey problem (RP).

Proposition 4. *Suppose that an aggregate allocation and Pareto-Negishi weights* $\left\{C_t^R, L_t^R, K_t^R, \{M_t^R(\theta^t)\}\right\}$ *solve the RRP. Then* $\left\{C_t^R, L_t^R, K_t^R, \{M_t^R(\theta^t)\}\right\}$ *satisfy constraints* (v) *of Proposition 3.*

Proof See the Appendix A. ■

Thus, we can analyze the optimal taxation by solving the RRP. We start by composing a Lagrangian for the RRP. We attach Lagrange multipliers γ_t, λ, $\{\beta^t \pi_t(\theta^t)\mu(\theta^t)\}$, and $\{\phi(\theta^t)\}$ to the constraints (17), (18), (19), and (20), respectively. By collecting terms on $\Big(u(h^c(\theta^t)) - v(h^l(\theta^t))\Big)$, we cumulate the Lagrange multipliers of enforcement constraints over history. By defining a cumulative multiplier as

$$\xi_t(\theta^t) \equiv 1 + \sum_{\theta^s \preceq \theta^t} \mu_s(\theta^s),$$

we can rewrite the Lagrangian of the RRP:

$$
\begin{aligned}
\max_{C_t,L_t,K_{t+1},M_t(\theta^t)} \quad & \sum_t \beta^t W(C_t,L_t;M,\xi,\lambda) - \sum_t\sum_{\theta^t}\beta^t\pi_t(\theta^t)\{\xi_t(\theta^t)-1\}U^{aut}(\theta_t,L_t,K_t;\tau_{l,t}) \\
& +\sum_t \gamma_t\left[F(K_t,L_t)+(1-\delta)K_t-C_t-K_{t+1}-G_t\right] \\
& -\lambda U_c^f(C_0,L_0;M)\left\{\left[1+(F_K(K_0,L_0)-\delta)(1-\tau_{k,0})\right]K_0+B_0\right\} \\
& +\sum_t\sum_{\theta^t}\phi(\theta^t)[M_t(\theta^t)-M_{t-1}(\theta^{t-1})],
\end{aligned}
$$

where $M = \{M_t(\theta^t)\}$, $\xi = \{\xi_t(\theta^t)\}$, and the pseudo-utility function $W(C_t,L_t;M,\xi,\lambda)$ is defined by

$$
W(C_t,L_t;M,\xi,\lambda) = \sum_{\theta^t}\pi_t(\theta^t)\left[\begin{array}{l} \xi_t(\theta^t)\left[u(h^c(\theta^t,C_t,L_t;M))-v(h^l(\theta^t,C_t,L_t;M))\right] \\ +\lambda U_c^f(C_t,L_t;M)h^c(\theta^t,C_t,L_t;M) \\ +\lambda U_L^f(C_t,L_t;M)\theta_t h^l(\theta^t,C_t,L_t;M) \end{array}\right].
$$

Throughout the paper, the optimal taxation analysis exploits only the first order necessary conditions of the relaxed Ramsey problem.[16] Notice that the first order conditions of the Ramsey problem involve the derivative of the pseudo-utility function which is the combination of the period utility function, the implementability constraint, and enforcement constraints. Thus, it is useful to know how these marginal pseudo-utilities are related to the marginal utility of the fictitious representative agent.

Lemma 5. *The marginal pseudo utility with respect to C_t and L_t can be expressed as:*

$$
\begin{aligned}
W_c(C_t,L_t;M,\xi,\lambda) &= U_c^f(t)\left[1+\lambda\left(1+\frac{U_{cc}^f(t)}{U_c^f(t)}C_t\right)\right] \\
W_L(C_t,L_t;M,\xi,\lambda) &= U_L^f(t)\left[1+\lambda\left(1+\frac{U_{LL}^f(t)}{U_L^f(t)}L_t\right)\right].
\end{aligned}
$$

Proof See the Appendix A. ■

The marginal pseudo-utilities in Lemma 5 are represented in familiar forms as in the canonical Ramsey literature. The only difference is that here we are using the utility of the fictitious representative agent and general equilibrium elasticities, $\frac{U_{cc}^f(t)}{U_c^f(t)}C_t$, $\frac{U_{LL}^f(t)}{U_L^f(t)}L_t$, which are also expressed in terms of the utility of a fictitious representative agent.

[16]It is well known that the set of allocations that satisfies the implementability constraint is not necessarily convex. Thus, first order necessary conditions of the Ramsey problem might not be sufficient for the optimality, as in the canonical Ramsey literature. Moreover, by adding limited commitment, enforcement constraints might make the constraint set "even more non-convex." For the theoretical results we derive in this paper, however, the necessity of first order conditions is sufficient. Since the first order conditions are necessary for the optimality, the properties of taxes we derive using the first order conditions are satisfied by the optimal tax schedule.

4 Properties of optimal taxation

4.1 Optimal capital income taxes and labor income taxes

The goal of the Ramsey government is to maximize welfare subject to an allocation being a competitive equilibrium with limited commitment. In the canonical Ramsey problem without limited commitment, the goal of the government is essentially to minimize distortions while financing government expenditures. In a limited commitment economy, however, the government has an additional issue to consider: externalities of capital and labor. We can see these externalities from the first order conditions of the relaxed Ramsey problem:

$$\cdot W_c(C_t, L_t; M, \xi, \lambda) F_L(K_t, L_t) - \sum_{\theta^t} \pi_t(\theta^t)\{\xi_t(\theta^t) - 1\} \frac{\partial U^{aut}(\theta_t, L_t, K_t; \tau_{l,t})}{\partial L_t}$$
$$= -W_L(C_t, L_t; M, \xi, \lambda) \tag{21}$$
$$\cdot W_c(C_t, L_t; M, \xi, \lambda) + \sum_{\theta^{t+1}} \beta \pi_{t+1}(\theta^{t+1}) \left\{\xi_{t+1}(\theta^{t+1}) - 1\right\} \frac{\partial U^{aut}(\theta_{t+1}, L_{t+1}, K_{t+1}; \tau_{l,t+1})}{\partial K_{t+1}}$$
$$= \beta W_c(C_{t+1}, L_{t+1}; M, \xi, \lambda)[F_K(K_{t+1}, L_{t+1}) + 1 - \delta]. \tag{22}$$

The first order conditions of the RRP have additional terms associated with the effects of aggregate labor and capital on financial autarky, which do not appear in the canonical Ramsey problem. Those new terms appear because the value of financial autarky endogenously responds to the wage. Recall that the wage in financial autarky is equal to the equilibrium wage. For a production function that has the property $F_{LK} > 0$ and $F_{LL} < 0$ (for example, a Cobb-Douglas function), an increase in aggregate labor decreases the wage in financial autarky and an increase in aggregate capital increases the wage in financial autarky. Thus, the intratemporal condition (21) has an additional marginal benefit term, $\sum_{\theta^t} \pi_t(\theta^t)\{\xi_t(\theta^t) - 1\} \frac{\partial U^{aut}(\theta_t, L_t, K_t; \tau_{l,t})}{\partial L_t}$, which captures the positive externality of labor, in the sense that a one-unit increase in labor lowers the wage in financial autarky, which relaxes enforcement constraints. On the other hand, the intertemporal condition (22) has an additional marginal cost term $\sum_{\theta^{t+1}} \beta \pi_{t+1}(\theta^{t+1})\{\xi_{t+1}(\theta^{t+1}) - 1\} \frac{\partial U^{aut}(\theta_{t+1}, L_{t+1}, K_{t+1}; \tau_{k,t+1})}{K_{t+1}}$, which captures the negative externality of capital. A one-unit increase in capital raises the wage in financial autarky, which tightens enforcement constraints. These financial autarky wage effects are called externalities because households do not internalize these autarky effects when they decide on labor and capital.[17]

[17] We used the terms "negative" and "positive" to refer the externalities of capital and labor, respectively. These terms are based on whether the externalities improve risk sharing. We say capital has negative

Thus, the Ramsey government in a limited commitment economy essentially has two objectives: 1. minimizing distortions while financing the government expenditures, and 2. internalizing the externalities of capital and labor to improve risk sharing. It is well known that the standard Ramsey government, which only considers the first goal, sets capital taxes to zero in the steady-state and imposes all distortions due to the government budgetary needs on the labor income taxes in the long run. The next proposition characterizes optimal taxation of the government which considers both goals in a limited commitment economy, which is distinctive from the conventional Ramsey taxation. For this proposition, we make an additional assumption on an initial tax rate, as in the canonical Ramsey taxation literature.

Assumption 6. *$\tau_{k,0} \leq \bar{\tau}_k$, and $\bar{\tau}_k$ is sufficiently small that this constraint is binding.*

This assumption is needed because if there is no restriction on initial capital taxation, the government will levy taxes on initial capital so high that it can finance government expenditures without additional distortionary taxation.

Proposition 7. *If $F_{LK} > 0$ and $F_{LL} < 0$, then the optimal tax system in the steady-state satisfies*

$$\begin{aligned} \tau_k &= \frac{\bar{\chi}}{W_c \cdot [F_K - \delta]} > 0, \quad \textit{where} \quad \bar{\chi} = \sum_{\theta^t} \pi_t(\theta^t) \left\{ \xi_t(\theta^t) - 1 \right\} \frac{\partial U^{aut}(\theta_t; w_t, \tau_{l,t})}{\partial K} > 0 \\ 1 - \tau_l &= \frac{1 + \lambda \left(1 + \frac{U^f_{cc}}{U^f_c} C \right)}{1 + \lambda \left(1 + \frac{U^f_{LL}}{U^f_L} L \right) - \frac{\bar{\Delta}}{U^f_L}}, \quad \textit{where} \quad \frac{\bar{\Delta}}{U^f_L} = - \sum_{\theta^t} \pi_t(\theta^t) \frac{\theta_t}{v'(l(\theta^t))} \frac{\partial U^{aut}(\theta_t; w_t, \tau_{l,t})}{\partial L} > 0. \end{aligned}$$

Proof See the Appendix A. ■

The condition of the this proposition — $F_{LK} > 0$, $F_{LL} < 0$ — holds for the standard Cobb-Douglas production function. This proposition shows that optimal capital taxes are not zero, even in the steady-state. It is precisely the external cost of capital normalized by $W_c \cdot (F_k - \delta)$ in the steady-state that leads to positive capital income taxes.

This positive capital income tax could be interpreted in two ways. One interpretation is that in a limited commitment model, there is a fundamental reason for taxing capital income — the externality of capital, which is distinguished from the canonical Ramsey literature. Another interpretation is that this positive capital income tax result can be considered a

---

externalities since its externalities harm risk sharing, and we say labor has positive externalities since its externalities improve risk sharing.

generalized version of the Chamley-Judd result. This is because capital income taxes are levied only to remove the negative externality of capital, but not to finance government expenditures.

Notice that by normalizing with W_c, the capital income tax depends on the general equilibrium elasticity $\frac{U^f_{cc}}{U^f_c}C$ and distortion cost λ, because $W_c = U^f_c\left[1+\lambda\left(1+\frac{U^f_{cc}}{U^f_c}C\right)\right]$ from Lemma 5. Since these terms are relevant to raising government revenues, one might think that positive capital taxes are levied not only to internalize the externality of capital but also to share the cost of distortions due to the revenue burden with labor income taxes. The dependence on an elasticity and a shadow cost, however, arises only because they affect the opportunity cost of capital disinvestment, which is the proper normalization for the Ramsey government. From the perspective of the Ramsey government, the opportunity cost of capital disinvestment is not only forgone utility but also forgone government surplus. Thus, the normalization term includes an elasticity and a shadow cost. Still, the only reason for levying capital income taxes in this economy is to internalize the externaltiy.

Even though the purpose of the capital income tax is to internalize the externality, the revenue from the capital income taxes will be used for government expenditures because we do not allow government transfers and we assume that any tax revenue is used for such expenditures, as is commonly assumed in the canonical Ramsey literature. If government expenditures are big enough that the revenue from the capital income tax is not enough to meet the budgetary needs of the government, the remaining budgetary needs will be financed by labor income taxes and accumulated government assets in the long run. Formally, the level of the steady-state labor income tax will be determined by λ, which implies how binding the government budget constraint (or equivalently, the implementability constraint) is. Thus, the labor income tax is responsible for the remaining government expenditures, despite the presence of external benefit of labor.

It is important to notice that it is the existence of externality, not limited commitment itself, that distinguishes the structure of capital and labor income taxes in our model from that of canonical Ramsey problem. The following corollary shows that even in a limited commitment economy, if there is no externality of capital, then the optimal capital tax should be zero, reverting back to the classic Chamley-Judd result.

Corollary 8. *(i) If $F_{LK} = 0$ and $F_{LL} = 0$, then the optimal tax system in the steady-state satisfies*

$$\tau_k = 0, \quad \tau_l = 1 - \frac{1+\lambda\left[1+\frac{U^f_{cc}}{U^f_c}C\right]}{1+\lambda\left[1+\frac{U^f_{LL}}{U^f_L}L\right]} \in (0,1).$$

(ii) If $F_{LK} = 0$ and $F_{LL} < 0$, then the optimal tax system in the steady-state satisfies

$$\tau_k = 0, \quad \tau_l = 1 - \frac{1 + \lambda \left(1 + \frac{U^f_{cc}}{U^f_c} C\right)}{1 + \lambda \left(1 + \frac{U^f_{LL}}{U^f_L} L\right) - (1 - \tau_l) \frac{\Delta}{U^f_L}},$$

$$\textit{where} \quad \frac{\Delta}{U^f_L} = - \sum_{\theta^t} \pi_t(\theta^t) \frac{\theta_t}{v'(l(\theta^t))} \frac{\partial U^{aut}(\theta_t; w_t, \tau_{l,t})}{\partial w(1 - \tau_l)} \frac{\partial w}{\partial L} > 0.$$

In case (i) of Corollary 8, the wage in financial autarky does not depend on aggregate labor and capital. For example, the production function with functional form $F(K, L) = \alpha L + f(K)$ has this property. In this case, since there is no externality in the economy, the government has only one goal: minimizing distortions when financing government expenditures. Thus, the steady-state capital income taxes will be zero in the long run, and labor income taxes are levied to fund remaining government expenditures in the long run, which is exactly the Chamley-Judd result.

In case (ii) of Corollary 8, the wage in autarky depends only on aggregate labor. For example, the production function with functional form $F(K, L) = g(L) + f(K)$, with $g''(L) < 0$ has this property. Since there is no externality of capital, the steady-state optimal capital taxes are zero. On the other hand, even though there is an external benefit of labor, all distortions due to the budgetary needs of the government are still imposed on labor income taxes in the long run, as long as the implementability constraint is binding. Comparison of labor income taxes of case (i) and case (ii) is not obvious because shadow cost λ and general equilibrium elasticities take different values. See the Supplementary Appendix B for more discussion on this.

Corollary 8 shows that limited commitment itself does not change the main result of the Ramsey taxation literature, which is zero capital tax in the steady-state. What matters for the structure of the steady-state labor and capital taxes is whether there is an externality in a limited commitment economy.

4.2 Understanding optimal taxation: balancing conflicting objectives

In this subsection, we try to explain the optimal taxation properties of Proposition 7 as the result of balancing two conflicting goals of the government. As we discussed above, the Ramsey problem is indeed a problem of finding tax policies that achieve two objectives together — minimizing distortions and internalizing externalities. We will show that these two objectives of the government push capital and labor income taxes in opposite directions by comparing optimal taxes with only one of the two objectives respectively.

The optimal taxes of the government that only cares about minimizing distortion was already analyzed in Corollary 8. To minimize distortions, capital income taxes should be set to zero in the steady-state and labor taxes are positive to fund the remaining government expenditures in the long run.

We can see that the second goal of the government (internalizing externalities) is driving capital income taxes and labor income taxes in opposite directions by analyzing optimal tax in the presence of lump-sum taxes. When lump-sum taxes are allowed, the government does not have to use distortionary taxes to finance its expenditures, because the government can always choose lump-sum taxes such that the implementability constraint is no longer binding. Thus, the government will have only one goal, which is to internalize externalities.

The properties of optimal taxation when lump-sum taxes are allowed are summarized in the following proposition. Notice that we only allow the government to use lump-sum taxes to finance government expenditures. In other words, lump-sum taxes cannot be used to relax the enforcement constraints. Technically, this restriction can be satisfied by assuming that the government does not levy lump-sum taxes in financial autarky. This assumption might seem unnatural, but our purpose in analyzing this case is simply to determine which direction the second goal of the government will drive the tax rates. Thus, we assume the absence of lump-sum taxes in financial autarky only to secure this end. Still, the government cannot discriminate between labor income taxes in and outside autarky.

Proposition 9. *With lump-sum taxes, the optimal tax system satisfies: for all t,*

$$\tau_{l,t} < 0,$$

$$\tau_{k,t+1} = \frac{1}{U_c^f(t+1)[F_K(t+1)-\delta]} \sum_{\theta^{t+1}} \pi_{t+1}(\theta^{t+1})\{\xi_{t+1}(\theta^{t+1})-1\} \frac{\partial U^{aut}(\theta_{t+1}; w_{t+1}, \tau_{l,t+1})}{\partial K_{t+1}} > 0.$$

Proof See the Appendix A. ■

This proposition provides the sign of capital and labor income taxes when the government only cares about risk sharing. Since there is a positive externality of aggregate labor, the government will subsidize labor ($\tau_{l,t} < 0$), and since there is a negative externality of aggregate capital, the government will tax capital ($\tau_{k,t} > 0$). Thus, the two goals of the government drive taxes in opposite directions, and the properties of optimal taxation in Proposition 7 are the result of balancing these conflicting objectives.

4.3 Extension: relaxed assumptions on financial autarky

We have analyzed optimal taxation under the assumption that if an agent defaults, then he is excluded from the financial market permanently. Although this assumption is commonly adopted in the limited commitment literature, it is harsher than the real-world punishment on default. Can we apply the analysis of an economy with permanent exclusion punishment for default to an economy which has less harsh punishments? The answer depends on the form of punishment on default. We consider two different alternatives.

First we consider the relaxed financial autarky where the agent is excluded from contingent claim trading, but is allowed to save at an exogenously given state-uncontingent interest rate r_d, as in Krueger and Perri (2006). With this relaxed assumption on financial autarky, the analysis of optimal taxation does not change qualitatively. In the steady-state, capital taxes are levied to internalize externality of capital, and labor income taxes are levied to finance remaining government expenditures. The sign of capital income tax is positive as in the permanent exclusion, because the increase of capital raises the wage in this relaxed financial autarky, increasing the value of default. Since the interest rate in autarky r_d is exogenously given, there are no additional autarky price effects.

Second, we consider the temporary exclusion from financial market after default. Following Krebs et al. (2011) and Azariadis and Kaas (2013), we assume that if an agent defaults, he cannot trade in the financial market during the default period, and that in any subsequent period the agent regains full access to the financial market with probability $1-\alpha$, where α is exogenously given. We show that the fundamental reason for taxing capital income and labor income in the steady-state still applies with this temporary exclusion, but the sign of optimal taxation could be different from that of permanent exclusion.

We denote the expected continuation value of an agent who regains access to financial markets at θ_s by $V^d(\omega = 0, \theta_s; \{w_t, r_t, \tau_{l,t}, \tau_{k,t}\}_{t=s}^{\infty})$, where $\omega = 0$ implies zero asset when the agent re-enters the financial market. Each household takes the value function V^d as given and solves its utility maximization problem subject to budget constraint and modified enforcement constraints which are written as: for all t and θ^t,

$$\begin{aligned}
&\sum_{s=t}^{+\infty}\sum_{\theta^s}\beta^{s-t}\pi(\theta^s|\theta^t)[u(c_s(\theta^s)) - v(l_s(\theta^s))] \\
&\qquad \geq \sum_{s=t}^{+\infty}\sum_{\theta^s}\alpha^{s-t}\beta^{s-t}\pi(\theta^s|\theta^t)U^{aut}(\theta_s; w_s, \tau_{l,s}) \\
&\qquad\quad +(1-\alpha)\sum_{s=t+1}^{+\infty}\sum_{\theta^s}\alpha^{s-t-1}\beta^{s-t}\pi(\theta^s|\theta^t)V^d(0, \theta_s; \{w_j, r_j, \tau_{l,j}, \tau_{k,j}\}_{j=s}^{\infty}).
\end{aligned}$$

The value function $V^d : R \times \Theta \to R$ is endogenously determined, since we require the equilibrium condition, $V^d(\cdot,\cdot;\{w_{s+j}, r_{s+j}, \tau_{l,s+j}, \tau_{k,s+j}\}_{j=0}^{\infty}) = V(\cdot,\cdot;\{w_t, r_t, \tau_t, \tau_{k,t}\}_{t=0}^{\infty})$, where V is the equilibrium value function associated with the household problem[18] and $\{w_{s+j}, r_{s+j}, \tau_{l,s+j}, \tau_{k,s+j}\}_{j=0}^{\infty} = \{w_t, r_t, \tau_t, \tau_{k,t}\}_{t=0}^{\infty}$.

We can easily show that optimal tax formulas which take temporary exclusion into account have exactly the same form as in Proposition 7 with modified externality terms $(\bar{\chi}, \bar{\Delta})$. That is, the fundamental reason for taxing capital and labor does not change; in the steady-state, capital income tax is levied only to internalize externality and labor income tax is levied to finance the remaining government expenditures. However, the externalities have somewhat different characteristics, because the value of default depends not only on equilibrium wages but also equilibrium interest rates, as we can see from the following decomposition of the externality of capital $\bar{\chi}_{t+1}$:

$$
\begin{aligned}
&\bar{\chi}_{t+1} = \bar{\chi}_{t+1,1} + \bar{\chi}_{t+1,2}, \\
\textit{where}\quad &\bar{\chi}_{t+1,1} = \sum_{\theta^{t+1}} \pi_{t+1}(\theta^{t+1})\{\sum_{\theta^s \preceq \theta^{t+1}} \alpha^{t+1-s}\mu(\theta^s)\} \frac{\partial U^{aut}(\theta_{t+1}; w_{t+1}, \tau_{l,t+1})}{\partial w_{t+1}} \frac{\partial w_{t+1}}{\partial K_{t+1}} > 0 \\
&\bar{\chi}_{t+1,2} = (1-\alpha) \sum_{\theta^{t+1}} \sum_{\theta^s \preceq \theta^{t+1}} \beta^s \pi_s(\theta^s)\{\sum_{\theta^l \preceq \theta^{s-1}} \alpha^{s-l-1}\mu(\theta^l)\} \left[\begin{array}{l} \frac{\partial V^d(0,\theta_s;\{w_{t'}, r_{t'}, \tau_{l,t'}, \tau_{k,t'}\}_{t'=s}^{\infty})}{\partial w_{t+1}} \frac{\partial w_{t+1}}{\partial K_{t+1}} \\ + \frac{\partial V^d(0,\theta_s;\{w_{t'}, r_{t'}, \tau_{l,t'}, \tau_{k,t'}\}_{t'=s}^{\infty})}{\partial r_{t+1}} \frac{\partial r_{t+1}}{\partial K_{t+1}} \end{array} \right].
\end{aligned}
$$

The externality of capital during the financial autarky $\bar{\chi}_{t+1,1}$ is positive because of the autarky wage effect, but the externality of capital after regaining access to financial markets $\bar{\chi}_{t+1,2}$ has an ambiguous sign because autarky interest rate effects $\frac{\partial V^d(0,\theta_s;\{w_{t'}, r_{t'}, \tau_{l,t'}, \tau_{k,t'}\}_{t'=s}^{\infty})}{\partial r_{t+1}} \frac{\partial r_{t+1}}{\partial K_{t+1}}$ and autarky wage effects $\frac{\partial V^d(0,\theta_s;\{w_{t'}, r_{t'}, \tau_{l,t'}, \tau_{k,t'}\}_{t'=s}^{\infty})}{\partial w_{t+1}} \frac{\partial w_{t+1}}{\partial K_{t+1}}$ have the opposite signs. In the limit, when $\alpha = 1$, the result goes back to the permanent exclusion case, but when α is small, the sign of capital tax could be reversed.[19] Further analysis on the sign of externalities is limited, because the value of default (V^d) and Lagrange multipliers endogenously respond to α. Regardless of the sign of taxes, however, the fundamental reason for taxing capital income and labor income remains true even with this relaxed assumption.

4.4 Outside the steady-state: special forms of preferences

So far, we have characterized optimal Ramsey taxation only in the steady-state. With special forms of preferences, we can extend the analysis even outside the steady-state. For

[18] The first argument of V is the initial wealth $(1 + r_0(1 - \tau_{k,0}))K_0 + B_0$, and the second argument of V is θ_0. This equilibrium condition implies that we assume rational expectation.

[19] With similar arguments, we can see the sign of labor externalities is also ambiguous.

this, we denote the time-additively separable utility function of the fictitious representative agent as

$$U(C, L; M) = \sum_{t=0}^{+\infty} \beta^t U^f(C_t, L_t; M),$$

where $C = (C_0, C_1, \cdots)$ and $L = (L_0, L_1, \cdots)$ are sequences of aggregate consumption and aggregate labor, respectively. The next proposition provides the conditions under which the steady-state optimal structure of taxes we derived in Proposition 7 can be applied even outside the steady-state.

Proposition 10. *If the utility function of the fictitious representative agent satisfies that either (i) $U(C, L; M)$ is homothetic in (C, L), or (ii) $U(C, L; M) = H(G(C), L; M)$, where G is homothetic in C, then for all $t \geq 1$, the optimal capital income tax system has the following property:*

$$\tau_{k,t+1} = \frac{1}{W_c(t+1)[F_K(t+1) - \delta]} \sum_{\theta^{t+1}} \pi_{t+1}(\theta^{t+1})\{\xi_{t+1}(\theta^{t+1}) - 1\} \frac{\partial U^{aut}(\theta_{t+1}; w_{t+1}, \tau_{k,t+1})}{\partial K_{t+1}}.$$

Proof See the Appendix A. ■

Notice that the conditions in Proposition 10 are exactly the conditions that we require on the representative agent's utility function to make zero capital taxes optimal even outside the steady-state in a canonical Ramsey tax analysis. These conditions are indeed the conditions under which the optimal uniform commodity taxation principle applies. In a limited commitment economy, the same conditions are imposed on the utility of a fictitious representative agent to make outside steady-state optimal taxes equal to the normalized externality of capital. The following two examples of individual utility functions guarantee that the utility of the fictitious representative agent satisfies the conditions in Proposition 10.

$$u(c) = \frac{c^{1-\sigma}}{1-\sigma}, \quad v(l) = \alpha \frac{l^\gamma}{\gamma}, \quad \sigma > 0, \quad \gamma > 1 \tag{23}$$

$$u(c, l) = \frac{(cl^{-\gamma})^{1-\sigma}}{1-\sigma}, \quad \sigma \leq 1, \quad 0 < \gamma < 1 \tag{24}$$

It is worth mentioning that the outside steady-state analysis does not apply to another form of Cobb-Douglas utility which is used often in the macro and public finance literature,

$$u(c, l) = \frac{(c^\gamma (1-l)^{1-\gamma})^{1-\sigma}}{1-\sigma}, \tag{25}$$

which represents the utility of leisure instead of the disutility of labor. With this form of utility, the optimal capital income taxes outside the steady-state are not equivalent to normalized externalities of capital. We can understand the different implications of utility (24) and utility (25) in relation to the requirement for uniform commodity taxation in the classic Ramsey literature. The utility of the form (24) is homothetic in (c, l), thus it satisfies the condition for uniform commodity taxation. The utility of the form (25), however, is homothetic in $(c, 1-l)$, not in (c, l). Since leisure $(1-l)$ cannot be taxed directly, this utility is against the condition for the uniform commodity taxation. Thus, with utility (25), the optimal capital taxation property in Proposition 10 does not apply outside the steady-state.

4.5 Comparison to the incomplete markets case

In this subsection, we relate the results in this paper to the results of studies on optimal taxation in the Bewley-Aiyagari class of incomplete market models. Private risk sharing is not perfect in either the limited commitment model or the incomplete market model. The reasons for this imperfect risk sharing, however, are different, and this difference leads to diverging optimal tax results between the two classes of models.

Dávila et al. (2012) study constrained efficient allocation and its implementation in the incomplete market model with uninsurable idiosyncratic shocks and borrowing constraints. The concept of constrained efficiency in their paper is in line with our paper in the sense that both papers look at efficiencies that do not alter the market structure and do not force any transfers between consumers.[20] They show that there is a pecuniary externality of capital in incomplete markets. More specifically, there are two effects of capital accumulation that are not considered when agents make their saving decisions. First, an increase in capital leads to a higher wage, which increases the amount of risk the agent is exposed to by scaling up the share of labor income that is stochastic. Thus, there can be negative externality of capital. The second effect of increasing capital is that the lower interest rate helps the poor (households with low wealth) and hurts the rich (households with high wealth), which is welfare-improving in the view of the utilitarian planner. This is a positive externality of capital. Depending on calibration, one of these two externalities will dominate the other. Thus, in the incomplete market, the optimal capital income tax can be either positive or negative, as the government tries to internalize the externalities of capital.[21]

[20]However, the government in Dávila et al. (2012) does not have to finance government expenditures.

[21]Aiyagari (1995) argues that in the presence of idiosyncratic risk, an incomplete market with borrowing constraint gives the rationale for the positive capital income tax, which is different from Dávila et al. (2012).

In a limited commitment model, however, there are different types of externalities, which lead to positive capital income taxes. This externality in a limited commitment economy works through autarky prices, whereas the pecuniary externality in the incomplete markets works through equilibrium prices. The pecuniary externality through equilibrium prices does not cause a welfare loss in a limited commitment model because of the complete asset market structure, but it impacts welfare in an economy with incomplete markets because the change in relative prices induces a change in the feasible consumption set. Thus, the externality of capital in a limited commitment model always leads to positive capital income taxes, which is different from the incomplete market case.

5 Numerical example

In this section, we provide a quantitative investigation of the theoretical results we have obtained. The purpose of this computation is not to suggest the optimal tax rate based on a full calibration, but to examine a quantitative significance of the externality in optimal taxation.

To find a numerical solution to the Ramsey problem, we exploit the first order necessary conditions of the Ramsey problem. That is, we numerically solve for an aggregate allocation and Pareto-Negishi weights that satisfy the first order conditions (21), (22), and all the constraints of the RRP. [22]

This result, however, is crucially dependent on the fact that the government optimally chooses the level of government expenditure that enters the household's utility function. Since the Euler equation must hold for government expenditures, the government tries to make the pre-tax return to capital equal to the time discount rate. In an equilibrium with incomplete market, however, the after-tax interest rate is always less than the time discount rate because there is capital overaccumulation due to the precautionary motive. This shows that optimal capital income tax in Aiyagari (1995)'s economy is positive even in the long run.

[22] Necessity of first order conditions is enough for theoretical results, because the properties of taxation we derive using first order conditions should be satisfied by the optimal taxation. The use of the first order necessary conditions for a numerical solution, however, is justified only when they are also sufficient for the optimality. As we discussed above, the constraint set of the Ramsey problem is not convex in general, and deriving sufficient conditions for convexity is challenging in this problem, because convexity depends not only on the curvature of utility and the production function, but also on the law of motion of capital. Thus, we prove the validity of using first order conditions in a rather crude way: We check whether the aggregate allocation that satisfies the first order conditions is unique by varying an initial guess. In our numerical examples, we could find unique solution regardless of the initial guess. Since first order conditions are necessary, if the numerical solution for the first order condition is unique, then it is optimal.

5.1 Computational issues and the partial solution method

Solving Ramsey problems numerically is not straightforward because of some issues associated with limited commitment. First, we will briefly explain the computation procedure of the canonical representative agent Ramsey problem used in the literature and why it is difficult to apply that method to our problem. The canonical Ramsey problem, which is defined as a maximization of expected lifetime utility of an agent subject to resource constraints and implementability constraint, can be solved numerically using a two-step procedure. First, for a fixed Lagrange multiplier of an implementability constraint λ, optimal policy functions of consumption, labor, and capital can be computed using the first order conditions of the Ramsey problem. Second, we can evaluate the implementability constraint by plugging in the policy functions we derived in the first step, and then we can adjust λ until the implementability constraint is satisfied for a fixed initial debt of the government. Notice that to solve for λ, we need to solve for the policy functions of both the transition path and the steady-state because the implementability constraint requires allocations of the entire path of infinite horizon. Computing transitional dynamics in a limited commitment model, however, is a challenging task. The main difficulty comes from the fact that policy functions in a limited commitment model depend on the entire history of productivity shocks. In the case of a steady-state, we can compute stationary policy functions using a recursive Lagrangian approach, but due to difficulty of computing transition policy functions, it is not easy to solve the Ramsey problem fully in our model.[23]

In this paper, we avoid this issue by inferring λ for each steady-state labor income tax rate and computing the steady-state optimal capital income tax rates which are matched to each labor income tax rate. In this way, we can solve for the set of optimal *mixes* of long-run labor income tax and capital income tax, where each mix is associated with different level of λ — proxy for initial government debt. However, we cannot pin down the exact optimal pair for given initial debt. The exact optimal pair can only be determined by solving for λ for given initial debt using transition policy functions. In this paper, we only provide the set of long-run optimal mixes of labor income tax and capital income tax. Each labor income tax rate is associated with different level of initial debt, since labor income is increasing in λ and λ is increasing in initial government debt, even though the exact initial debt associated with each λ cannot be inferred. Monotone relationship between λ and government debt is

[23] We might think of a brute force approach of computing optimal taxation; that is, computing a competitive equilibrium for every given sequence of tax rates directly and finding the taxes that give the highest utility. Then, however, solving for the entire path of the taxes becomes a truly challenging task.

Table 1: Calibrated Parameters

Category		Symbol	Value
Preference	utility[a]	$u(c,l) = \log c - \alpha \frac{l^\gamma}{\gamma}$	$\alpha = 1$, $\gamma = 3$
	discout factor	β	$\beta = 0.85$
Technology	production function	$F(K,L) = K^\nu L^{1-\nu}$	$\nu = 0.25$
	depreciation rate	δ	$\delta = 0.1$
Productivity Shock	shock process[b]	$\log \theta_t = \rho \log \theta_{t-1} + \epsilon_t$	$\rho = 0.92$
		$\epsilon_t \sim iidN(0, \sigma_\epsilon^2)$	$\sigma_\epsilon^2 = 0.05$
Government	government expenditure[c]	G	$G = 0.115$

[a]The Frisch elasticity of labor supply is $\frac{1}{\gamma-1} = 0.5$, which is in the middle of estimates of prime-age males and married women in the literature. MaCurdy (1981) estimates the elasticity of prime-age males to be in the range 0.1 – 0.45. Blundell et al. (1993) estimate elasticities of married women in the U.K. to be in the 0.5 – 1.0 range.

[b]Both ρ and σ_ϵ^2 are set in the middle of estimates in the literature. Heathcote (2005) reports that in the literature, the estimate of ρ is in the range of 0.88 – 0.96, and σ_ϵ^2 is in the range of 0.014 – 0.063.

[c]The government expenditure is set so that the average ratio of government expenditure to output is 0.20.

clear from the fact that the Lagrange multiplier λ is the shadow price of the government debt — If the economy begins with a large initial government debt, then the multiplier λ will be high since the implementability constraint is more binding. See the Supplementary Appendix D for the detailed computation procedure and algorithm.

5.2 Calibration and results

Calibrated parameters are summarized in Table 1. We set the discount factor β to 0.85, which is slightly below the standard values from the dynamic general equilibrium literature. In the limited commitment literature, the discount factor tends to be set low to avoid full risk-sharing.[24] Since our model assumes a very harsh punishment on default, not only seizing all capital and assets but also imposing permanent exclusion from any intertemporal trade, the degree of risk-sharing in our economy is very high. Thus, a low discount factor β is needed to avoid full risk-sharing. For the same reason, the capital income share ν is set to 0.25, which is also slightly lower than the standard value. See the the Supplementary Appendix E for sensitivity analysis with respect to β and ν.

Figure 1 shows the inferred Lagrange multiplier of the implementability constraint, λ,

[24]Alvarez and Jermann (2001) used 0.65 and Chien and Lee (2012) used 0.75.

Figure 1: λ, inferred multiplier

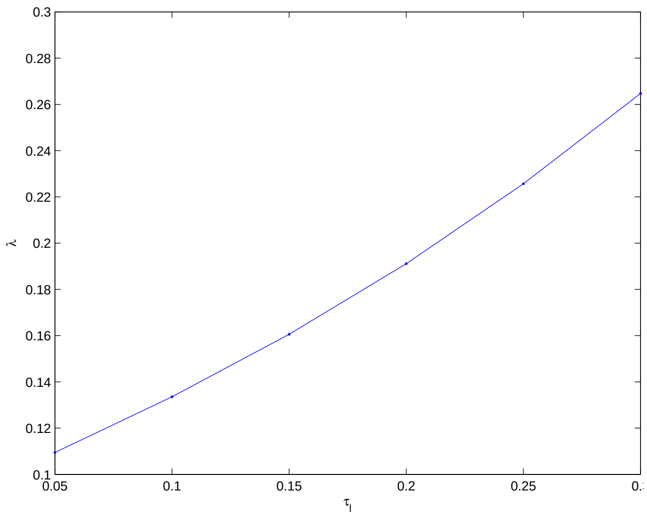
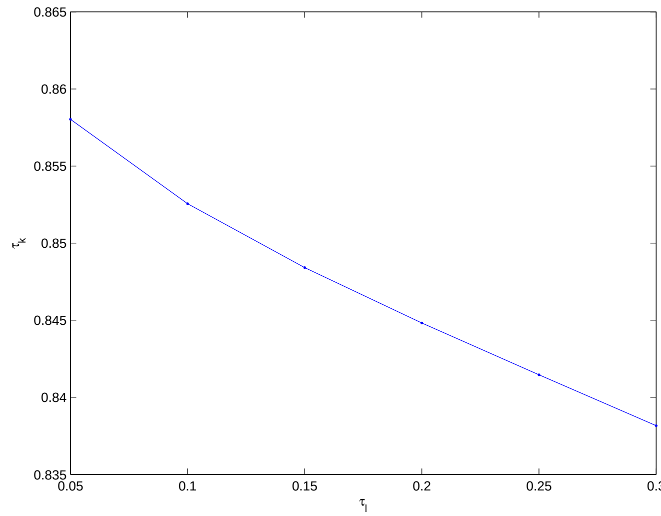


Figure 2: τ_k, optimal s.s. capital tax



for each steady-state labor income tax rate.[25] A labor income tax is an increasing function of λ because labor income taxes are responsible for covering the government's budgetary needs, as we discussed in section 4. Then, since λ is increasing in initial government debt, each labor income tax rate is associated with different level of initial government debt.

Figure 2 shows the optimal mixes of labor income tax and capital income tax in the steady-state. Recall that the level of the optimal capital tax rate depends on both the externality cost of capital ($\sum_{\theta^t} \pi\{M_t(\theta^t) - 1\} \frac{\partial U^{aut}(\theta_t)}{\partial K}$) and the normalization ($W_c \cdot (F_K - \delta)$). We discuss how these terms respond to the changes in labor income taxes. For log utility, there is no direct impact of labor income tax rates on the externality, since the increase in value of financial autarky for one unit of wage increase ($\frac{\partial U^{aut}}{\partial w}$) is independent of the labor income tax rate ($\frac{\partial U^{aut}}{\partial K} = \frac{\partial U^{aut}}{\partial w} \frac{\partial w}{\partial K} = \frac{1}{w} F_{LK} = \frac{\nu}{K}$). However, there will be indirect effects of labor income taxes through the changes in aggregate capital and labor. Since a higher labor income tax decreases the value of financial autarky more than the value of staying in the contract,[26] it will relax the enforcement constraint, resulting in more risk-sharing. Through this channel, higher labor income tax rates decrease the marginal benefit of capital because the government does not have to increase future consumption for the purpose of decreasing a household's incentive to default. Hence, aggregate capital will decrease, and this will increase both the externality term and normalization term. Figure 2 shows that the effect of normalization

[25] We can compute λ for each $\hat{\tau}$ by iterating on λ, until $\hat{\tau}_l = \tau_l(\lambda) = 1 + \frac{U_L^f(C,L)}{U_c^f(C,L) F_L(K,L)}$ is satisfied.

[26] This is because the labor income is the only source of income in financial autarky.

dominates, and thus that capital tax rate decreases as the labor income tax increases.

From this numerical example, we can see that optimal capital tax rates are very high, for the functional forms and the parameters that we chose. As we mentioned above, the purpose of these examples is not to suggest optimal tax rates based on a full calibration, but instead to illustrate that the externality of capital might be quantitatively significant as a rationale for positive capital taxes.

6 Conclusion

In this paper, we have studied optimal Ramsey taxation in a limited commitment model. The goal of the Ramsey government in this economy is to maximize welfare subject to an allocation being a competitive equilibrium with limited commitment. To achieve this goal, the government faces two conflicting objectives: 1. minimizing distortions when financing government expenditures, and 2. internalizing the externality of labor and capital to improve risk-sharing. When balancing these two objectives, the steady-state optimal capital taxes are levied only to remove the external cost of capital. All the remaining budgetary needs of the government will be financed using labor income taxes in the long run, even though labor has a positive externality.

Therefore, our analysis shows that there is a good reason to tax capital income in a limited commitment economy — the externality of capital. Our result, however, can also be interpreted as a version of the famous zero capital tax result of Chamley-Judd. If there is no capital externality in a limited commitment economy, then the result reverts to Chamley-Judd's result. If there is an external cost of capital, then the result is a generalization of the Chamley-Judd's result because capital taxes are levied only to remove the externality.

This result implies that when thinking about the real world implications of optimal Ramsey tax analysis, the market structure matters crucially for the optimal level of capital income taxes and labor income taxes. Thus, any policy prescriptions should be based on a thorough investigation of the relevant financial market structure and its associated frictions.

We see this paper as the first attempt to consider the risk-sharing effects endogenous to tax codes in the Ramsey literature. If private insurance markets are not perfect, a tax system implemented by a government will affect private risk sharing. Thus, that government must consider how the tax rates endogenously change risk sharing in private markets when it optimally sets its tax system to finance government expenditures. This paper provides an analysis of optimal Ramsey taxation focusing on one source of imperfect private risk sharing, limited commitment. Analyzing optimal Ramsey taxation with other sources of imperfect risk sharing would be an important and complementary future work.

References

Abraham, A. J. and E. Carceles-Poveda (2006). Endogenous incomplete markets, enforcement constraints, and intermediation. *Theoretical Economics 1*, 439–459.

Acemoglu, D., M. Golosov, and A. Tsyvinski (2011, August). Political economy of ramsey taxation. *Journal of Public Economics 95*, 467–475.

Aiyagari, S. R. (1995, December). Optimal capital income taxation with incomplete markets, borrowing constraints, and constant discounting. *Journal of Political Economy 103*(6), 1158–1175.

Alvarez, F. and U. J. Jermann (2000). Efficiency, equilibrium, and asset pricing with risk of default. *Econometrica 68*(4), 775–797.

Alvarez, F. and U. J. Jermann (2001). Quantitative asset pricing implications of endogenous solvency constraints. *Review of Financial Studies 14*(4), 1117–1151.

Atkeson, A., V. Chari, and P. Kehoe (1999). Taxing capital income: a bad idea. *Quarterly Review Sum*, 3–17.

Azariadis, C. and L. Kaas (2013, March). Endogenous credit limits with small default costs. *Journal of Economic Theory 148*(2), 806–824.

Blundell, R., C. Meghir, and P. Neves (1993, September). Labour supply and intertemporal substitution. *Journal of Econometrics 59*(1âĂŞ2), 137–160.

Chamley, C. (1986, May). Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica 54*(3), 607–622.

Chari, V. V., L. J. Christiano, and P. J. Kehoe (1994, August). Optimal fiscal policy in a business cycle model. *Journal of Political Economy 102*(4), 617–652.

Chien, Y. and J. Lee (2012). Externality cost of capital investment in limited commitment. *KDI Journal of Economic Policy II*, 17–40.

Conesa, J. C., S. Kitao, and D. Krueger (2009, March). Taxing capital? not a bad idea after all! *American Economic Review 99*(1), 25–48.

Conesa, J. C. and D. Krueger (2006, October). On the optimal progressivity of the income tax code. *Journal of Monetary Economics 53*(7), 1425–1450.

Dávila, J., J. H. Hong, P. Krusell, and J.-V. Ríos-Rull (2012). Constrained efficiency in the neoclassical growth model with uninsurable idiosyncratic shocks. *Econometrica*.

Erosa, A. and M. Gervais (2002, August). Optimal taxation in life-cycle economies. *Journal of Economic Theory 105*(2), 338–369.

Garriga, C. (2001). Optimal fiscal policy in overlapping generations models. *Working Paper*.

Heathcote, J. (2005, January). Fiscal policy with heterogeneous agents and incomplete markets. *Review of Economic Studies 72*(1), 161–188.

Jones, L. E., R. E. Manuelli, and P. E. Rossi (1997, March). On the optimal taxation of capital income. *Journal of Economic Theory 73*(1), 93–117.

Judd, K. L. (1985, October). Redistributive taxation in a simple perfect foresight model. *Journal of Public Economics 28*(1), 59–83.

Kehoe, T. J. and D. K. Levine (1993). Debt-constrained asset markets. *Review of Economic Studies 60*(4), 865–888.

Kehoe, T. J. and D. K. Levine (2001). Liquidity constrained markets versus debt constrained markets. *Econometrica 69*(3), 575–598.

Kocherlakota, N. R. (1996). Implications of efficient risk sharing without commitment. *Review of Economic Studies 63*(4), 595–609.

Krebs, T., M. Kuhn, and M. Wright (2011). Human capital risk, contract enforcement, and the macroeconomy. *Working Paper*.

Krueger, D. and F. Perri (2006, January). Does income inequality lead to consumption inequality? evidence and theory. *Review of Economic Studies 73*(1), 163–193.

Krueger, D. and F. Perri (2011, May). Public versus private risk sharing. *Journal of Economic Theory 146*(3), 920–956.

MaCurdy, T. E. (1981, December). An empirical model of labor supply in a life-cycle setting. *Journal of Political Economy 89*(6), 1059–1085.

Marcet, A. and R. Marimon (2011). Recursive contracts. *Working Paper*.

Werning, I. (2007). Optimal fiscal policy with redistribution. *Quarterly Journal of Economics 122*(3), 925–967.

Zhu, X. (1992, December). Optimal fiscal policy in a stochastic growth model. *Journal of Economic Theory 58*(2), 250–289.

A Appendix: Proofs

Proof of "if" part of Proposition 3 We prove by construction.
Suppose $\{C_t, L_t, K_t, \{M_t(\theta^t)\}\}$ satisfies $(i), (ii), (iii), (iv)$ and (v). Recall that fictitious representative agent utility was defined as follows: for all t,

$$\begin{aligned} U^f(C_t, L_t; M) \quad &= \quad \max_{c_t(\theta^t), l_t(\theta^t)} \sum_{\theta^t} \pi_t(\theta^t) M_t(\theta^t) \left[u(c_t(\theta^t)) - v(l_t(\theta^t)) \right] \\ s.t. \quad (\lambda_t^f) \quad & \sum_{\theta^t} \pi_t(\theta^t) c_t(\theta^t) = C_t \qquad (26) \\ (\mu_t^f) \quad & \sum_{\theta^t} \pi_t(\theta^t) \theta_t l_t(\theta^t) = L_t. \qquad (27) \end{aligned}$$

We denote the solution of this static problem as $\Big(h^c(\theta^t, C_t, L_t; M), h^l(\theta^t, C_t, L_t; M)\Big)$. Then, individual allocations of competitive equilibrium with limited commitment are constructed as follows:

$$c_t(\theta^t) = h^c(\theta^t, C_t, L_t; M), \quad l_t(\theta^t) = h^l(\theta^t, C_t, L_t; M), \quad \text{for all } t.$$

With given allocations and Pareto-Negishi weights, we define prices and taxes as follows.

$$\begin{aligned} r_t &= F_K(K_t, L_t) - \delta \\ w_t &= F_L(K_t, L_t) \\ p(\theta^t) &= \frac{\beta \pi_t(\theta^t) U_c^f(C_t, L_t; M)}{U_c^f(C_0, L_0; M)} = \frac{\beta^t \pi_t(\theta^t) M_t(\theta^t) u'(h^c(\theta^t, C_t, L_t; M))}{M_0(\theta_0) u'(h^c(\theta_0, C_0, L_0; M))} \\ (1 - \tau_{k,t+1}) &= \frac{\frac{U_c^f(C_t, L_t; M)}{\beta U_c^f(C_{t+1}, L_{t+1}; M)} - 1}{F_K(K_{t+1}, L_{t+1}) - \delta} = \frac{\frac{M_t(\theta^t) u'(h^c(\theta^t, C_t, L_t; M))}{\beta M_{t+1}(\theta^{t+1}) u'(h^c(\theta^{t+1}, C_{t+1}, L_{t+1}; M))} - 1}{F_K(K_{t+1}, L_{t+1} - \delta)} \\ (1 - \tau_{l,t}) &= -\frac{U_L^f(C_t, L_t; M)}{F_L(K_t, L_t) U_c^f(C_t, L_t; M))} = -\frac{M_t(\theta^t) \frac{1}{\theta_t} v'(h^l(\theta^t, C_t, L_t; M))}{F_L(K_t, L_t) M_t(\theta^t) u'(h^c(\theta^t, C_t, L_t; M))} \end{aligned}$$

Then, by the definition of $\{r_t, w_t\}$, the firm's optimality is satisfied.

Resource constraints hold by assumption (i), and by constraints of the static problem (26) and (27), the goods market and labor market clear.

After dividing left hand side (LHS) and right hand side (RHS) of the implementability constraint (ii) by $U_c^f(C_0, L_0; M)$, we substitute out following terms in the LHS:

$$\frac{\beta^t \pi_t(\theta^t) U_c^f(C_t, L_t; M)}{U_c^f(C_0, L_0; M)} = p(\theta^t)$$
$$\frac{\beta^t \pi_t(\theta^t) U_L^f(C_t, L_t; M)}{U_c^f(C_0, L_0; M)} = -\frac{\beta^t \pi_t(\theta^t) U_c^f(C_t, L_t; M)(1 - \tau_{l,t} w_t)}{U_c^f(C_0, L_0; M)} = -p(\theta^t)(1 - \tau_{l,t}) w_t,$$

and substitute out $F_K(K_0, L_0) - \delta$ in the RHS by r_0. Then, we get a budget constraint (BC) of households. Enforcement constraints (EC) of the household hold by assumption (iii).

The only issue that remains to be verified is that $\{c_t(\theta^t), l_t(\theta^t)\}$ is optimal for the household, given $(w, r, p, \tau_k, \tau_l)$. It will suffice to find the nonnegative multipliers associated with the BC and EC and verify that they are a saddle. First, we construct the nonnegative Lagrange multipliers associated with BC and EC.

The Lagrange multiplier of BC, λ^{KL} is defined by

$$\lambda^{KL} = M_0(\theta_0) u'(h^c(\theta_0, C_0, L_0; M)) = U_c^f(C_0, L_0; M).$$

The Lagrange multiplier of EC at time 0 is set to zero, because enforcement constraints at period 0 are not binding for all agents due to the absence of *ex-ante* uncertainty:

$$\mu^{KL}(\theta_0) = 0.$$

For $t > 0$, the multiplier of EC at $\theta^t \in \Theta^t$, $\mu^{KL}(\theta^t)$ is defined recursively by

$$\mu^{KL}(\theta^{t-1}, \theta_t) = M_t(\theta^{t-1}, \theta_t) - M_{t-1}(\theta^{t-1}). \tag{28}$$

Then, we get the following equivalence:

$$\left\{ 1 + \sum_{\theta^s \preceq \theta^t} \mu_s^{KL}(\theta^s) \right\} = M_t(\theta^t) \quad \forall t, \forall \theta^t. \tag{29}$$

Finally, we need to verify that these multipliers together with allocations are indeed a saddle. First, by the assumption (iv), $\mu_t^{KL}(\theta^t) \geq 0$ is satisfied for all θ^t; and by the assumption (v), $\mu^{KL}(\theta^{t+1}) = 0$ if the enforcement constraint for θ^{t+1} is not binding. By construction of the multipliers, it is clear that multipliers minimize the Lagrangian problem of the household.

Second, we will show that $\{c_t(\theta^t), l_t(\theta^t)\}$ maximizes $L(\cdot, \lambda^{KL}, \{\mu^{KL}(\theta^t)\})$. We first check the convergence of each component of sums in the Lagrangian, as a technical requirement. We can check that by the following.

$$\begin{aligned}
& \left|\sum_t \sum_{\theta^t} \pi_t(\theta^t) u(c_t(\theta^t)) + \sum_t \sum_{\theta^t} \beta^t \pi_t(\theta^t) \mu^{KL}(\theta^t) \sum_{s=t}^{+\infty} \sum_{\theta^s} \beta^{s-t} \pi(\theta^s|\theta^t) u(c_s(\theta^s))\right| \\
= & \left|\sum_t \sum_{\theta^t} \beta^t \pi_t(\theta^t) \left\{1 + \sum_{\theta^s \preceq \theta^t} \mu^{KL}(\theta^t)\right\} u(c(\theta^t))\right| \\
\leq & \sum_t \sum_{\theta^t} \beta^t \pi_t(\theta^t) \left\{1 + \sum_{\theta^s \preceq \theta^t} \mu^{KL}(\theta^t)\right\} \zeta_1 u'(c(\theta^t)) c(\theta^t) \\
< & +\infty,
\end{aligned}$$

where the second inequality comes from Assumption 2 and the third inequality comes from the implementability constraint.

Similarly,

$$\begin{aligned}
& \left|\sum_t \sum_{\theta^t} \pi_t(\theta^t) v(l_t(\theta^t)) + \sum_t \sum_{\theta^t} \beta^t \pi_t(\theta^t) \mu^{KL}(\theta^t) \sum_{s=t}^{+\infty} \sum_{\theta^s} \beta^{s-t} \pi(\theta^s|\theta^t) v(l_s(\theta^s))\right| \\
= & \left|\sum_t \sum_{\theta^t} \beta^t \pi_t(\theta^t) \left\{1 + \sum_{\theta^s \preceq \theta^t} \mu^{KL}(\theta^t)\right\} v(l(\theta^t))\right| \\
\leq & \sum_t \sum_{\theta^t} \beta^t \pi_t(\theta^t) \left\{1 + \sum_{\theta^s \preceq \theta^t} \mu^{KL}(\theta^t)\right\} \zeta_2 v'(l(\theta^t)) \theta_t l(\theta^t) \\
< & +\infty.
\end{aligned}$$

By the argument about the convergence of each component of sums, showing optimality of $\{c_t(\theta^t), l_t(\theta^t), k_{t+1}(\theta^t))\}$ given multipliers is equivalent to verify that [27]

$$\begin{aligned}
& \sum_t \sum_{\theta^t} \beta^t \pi_t(\theta^t) \Big[u(c_t(\theta^t)) - v(l_t(\theta^t))\Big] + \sum_t \sum_{\theta^t} \beta^t \pi_t(\theta^t) \mu^{KL}(\theta^t) \sum_{s=t}^{+\infty} \sum_{\theta^s} \beta^{s-t} \pi_{s-t}(\theta^s|\theta^t) \left[u(c_s(\theta^s)) - v(l_s(\theta^s))\right] \\
& + \lambda^{KL} \sum_t \sum_{\theta^t} p(\theta^t) \Big[w_t(1-\tau_{l,t}) \theta_t l_t(\theta^t) - c_t(\theta^t)\Big] \\
\geq & \sum_t \sum_{\theta^t} \beta^t \pi_t(\theta^t) \Big[u(\hat{c}_t(\theta^t)) - v(\hat{l}_t(\theta^t))\Big] + \sum_t \sum_{\theta^t} \beta^t \pi_t(\theta^t) \mu^{KL}(\theta^t) \sum_{s=t}^{+\infty} \sum_{\theta^s} \beta^{s-t} \pi_{s-t}(\theta^s|\theta^t) \Big[u(\hat{c}_s(\theta^s)) - v(\hat{l}_s(\theta^s))\Big] \\
& + \lambda^{KL} \sum_t \sum_{\theta^t} p(\theta^t) \Big[w_t(1-\tau_{l,t}) \theta_t \hat{l}_t(\theta^t) - \hat{c}_t(\theta^t)\Big], \\
& \text{for all } \left\{\hat{c}_t(\theta^t), \hat{l}_t(\theta^t), \hat{k}_t(\theta^t)\right\}.
\end{aligned}$$

[27]Notice that the lifetime budget constraint of a household is written without sequences of capital allocation, which is obtained by manipulating the budget constraint with no arbitrage condition: $\sum_{\theta^{t+1}} p(\theta^{t+1})[1 + r_{t+1}(1-\tau_{k,t+1})] = p(\theta^t)$.

Equivalently, we want to show the following.

$$\begin{aligned}
& \sum_t \sum_{\theta^t} \beta^t \pi_t(\theta^t) \left[\left\{ 1 + \sum_{\theta^s \preceq \theta^t} \mu^{KL}(\theta^s) \right\} \Big(u(c_t(\theta^t)) - v(l_t(\theta^t)) \Big) \right] + \lambda \sum_t \sum_{\theta^t} p(\theta^t) \left[w_t(1-\tau_{l,t})\theta_t l_t(\theta^t) - c_t(\theta^t) \right] \\
\geq \; & \sum_t \sum_{\theta^t} \beta^t \pi_t(\theta^t) \left[\left\{ 1 + \sum_{\theta^s \preceq \theta^t} \mu^{KL}(\theta^s) \right\} \Big(u(\hat{c}_t(\theta^t)) - v(\hat{l}_t(\theta^t)) \Big) \right] + \lambda \sum_t \sum_{\theta^t} p(\theta^t) \left[w_t(1-\tau_{l,t})\theta_t \hat{l}_t(\theta^t) - \hat{c}_t(\theta^t) \right].
\end{aligned}$$

By concavity of u and convexity of v, we have

$$u(\hat{c}_t(\theta^t)) \leq u(c_t(\theta^t)) + u'(c_t(\theta^t)) \left[\hat{c}_t(\theta^t) - c_t(\theta^t) \right] \tag{30}$$

$$-v(\hat{l}(\theta^t)) \geq -v(l_t(\theta^t)) - v'(l_t(\theta^t)) \left[\hat{l}_t(\theta^t) - l_t(\theta^t) \right]. \tag{31}$$

Also, by the definition of $p(\theta^t)$, λ^{KL}, and $(1-\tau_{l,t})$ we get

$$\begin{aligned}
\beta^t \pi_t(\theta^t) M_t(\theta^t) u'(h^c(\theta^t, C_t, L_t; M)) &= p(\theta^t) M_0(\theta_0) u'(h^c(\theta_0, C_0, L_0; M)) \\
&= \lambda^{KL} p(\theta^t) \\
\beta^t \pi_t(\theta^t) M_t(\theta^t) \frac{1}{\theta_t} v'(h^l(\theta^t, C_t, L_t; M)) &= w_t(1-\tau_{l,t}) \beta^t \pi_t(\theta^t) M_t(\theta^t) u'(h^c(\theta^t, C_t, L_t; M)) \\
&= \lambda^{KL} p(\theta^t) w_t(1-\tau_{l,t}).
\end{aligned}$$

And by (28), we get first order conditions of the household problem.

$$\left\{ 1 + \sum_{\theta^s \preceq \theta^t} \mu_s^{KL}(\theta^s) \right\} \beta^t \pi_t(\theta^t) u'(h^c(\theta^t, C_t, L_t; M)) = \lambda^{KL} p(\theta^t) \tag{32}$$

$$\left\{ 1 + \sum_{\theta^s \preceq \theta^t} \mu_s^{KL}(\theta^s) \right\} \beta^t \pi_t(\theta^t) \frac{1}{\theta_t} v'(h^l(\theta^t, C_t, L_t; M)) = \lambda^{KL} p(\theta^t) w_t(1-\tau_{l,t}). \tag{33}$$

Then, using (30), (31), (32), and (33), we obtain the desired inequality:

$$\begin{aligned}
& \sum_t \sum_{\theta^t} \beta^t \pi_t(\theta^t) \left[\left\{ 1 + \sum_{\theta^s \preceq \theta^t} \mu^{KL}(\theta^s) \right\} \Big(u(\hat{c}_t(\theta^t)) - v(\hat{l}_t(\theta^t)) \Big) \right] + \lambda \sum_t \sum_{\theta^t} p(\theta^t) \Big[w_t(1-\tau_{l,t})\theta_t \hat{l}_t(\theta^t) - \hat{c}_t(\theta^t) \Big] \\
\leq \; & \sum_t \sum_{\theta^t} \beta^t \pi_t(\theta^t) \left[\left\{ 1 + \sum_{\theta^s \preceq \theta^t} \mu^{KL}(\theta^s) \right\} \Big(u(c_t(\theta^t)) - v(l_t(\theta^t)) \Big) \right] \\
& + \sum_t \sum_{\theta^t} \beta^t \pi_t(\theta^t) \left[\left\{ 1 + \sum_{\theta^s \preceq \theta^t} \mu^{KL}(\theta^s) \right\} \Big\{ u'(c_t(\theta^t)) \left[\hat{c}_t(\theta^t) - c_t(\theta^t) \right] - v'(l_t(\theta^t)) \left[\hat{l}_t(\theta^t) - l_t(\theta^t) \right] \Big\} \right] \\
& + \lambda \sum_t \sum_{\theta^t} p(\theta^t) \Big[w_t(1-\tau_{l,t})\theta_t \hat{l}_t(\theta^t) - \hat{c}_t(\theta^t) \Big] \\
= \; & \sum_t \sum_{\theta^t} \beta^t \pi_t(\theta^t) \left[\left\{ 1 + \sum_{\theta^s \preceq \theta^t} \mu^{KL}(\theta^s) \right\} \Big(u(c_t(\theta^t)) - v(l_t(\theta^t)) \Big) \right] + \lambda \sum_t \sum_{\theta^t} p(\theta^t) \Big[w_t(1-\tau_{l,t})\theta_t l_t(\theta^t) - c_t(\theta^t) \Big].
\end{aligned}$$

$\blacksquare$

Proof of Proposition 4 Suppose not. Then there exists $(\hat{\theta}^t, \hat{\theta}_{t+1})$ at which the enforcement constraint is not binding and

$$M^R_{t+1}(\hat{\theta}^t, \hat{\theta}_{t+1}) > M^R_t(\hat{\theta}^t). \tag{34}$$

From the first order conditions of the planner's static problem and the envelope theorem at t and $t+1$, we get

$$\frac{\beta\pi_{t+1}(\hat{\theta}^t, \hat{\theta}_{t+1}) M^R_{t+1}(\hat{\theta}^t, \hat{\theta}_{t+1}) u'(h^c(\hat{\theta}^t, \hat{\theta}_{t+1}))}{\pi_t(\hat{\theta}^t) M^R_t(\hat{\theta}^t) u'(h^c(\hat{\theta}^t))} = \frac{\beta\pi_t(\hat{\theta}^t, \hat{\theta}_{t+1})\lambda^f_{t+1}}{\pi_t(\hat{\theta}^t)\lambda^f_t} = \frac{\beta\pi_{t+1}(\hat{\theta}^t, \hat{\theta}_{t+1}) U^f_c(C^R_{t+1}, L^R_{t+1}; M^R)}{\pi_t(\hat{\theta}^t) U^f_c(C^R_t, L^R_t; M^R)},$$
$$\frac{\beta\pi_{t+1}(\hat{\theta}^t, \hat{\theta}_{t+1}) M^R_{t+1}(\hat{\theta}^t, \hat{\theta}_{t+1}) \frac{1}{\hat{\theta}^{t+1}} v'(h^l(\hat{\theta}^t, \hat{\theta}_{t+1}))}{\pi_t(\hat{\theta}^t) M^R_t(\hat{\theta}^t) \frac{1}{\hat{\theta}^t} v'(h^l(\hat{\theta}^t))} = \frac{\beta\pi_{t+1}(\hat{\theta}^t, \hat{\theta}_{t+1})\mu^f_{t+1}}{\pi_t(\hat{\theta}^t)\mu^f_t} = \frac{\beta\pi_{t+1}(\hat{\theta}^t, \hat{\theta}_{t+1}) U^f_L(C^R_{t+1}, L^R_{t+1}; M^R)}{\pi_t(\hat{\theta}^t) U^f_L(C^R_t, L^R_t; M^R)}.$$

Let's denote

$$\begin{aligned} q^c(\hat{\theta}^t, \hat{\theta}_{t+1}) &= \frac{\beta\pi_t(\hat{\theta}^t,\hat{\theta}_{t+1}) U^f_c(C^R_{t+1}, L^R_{t+1}; M^R)}{\pi_t(\hat{\theta}^t) U^f_c(C^R_t, L^R_t; M^R)} \\ q^l(\hat{\theta}^t, \hat{\theta}_{t+1}) &= \frac{\beta\pi_{t+1}(\hat{\theta}^t,\hat{\theta}_{t+1}) U^f_L(C^R_{t+1}, L^R_{t+1}; M^R)}{\pi_t(\hat{\theta}^t) U^f_L(C^R_t, L^R_t; M^R)}. \end{aligned}$$

Notice that $q^c(\hat{\theta}^t, \hat{\theta}_{t+1})$ ($q^l(\hat{\theta}^t, \hat{\theta}_{t+1})$) is the ratio of Ramsey planner's shadow price of consumption (labor) between $\hat{\theta}^t$ and $\hat{\theta}_{t+1}$. Using (34), we get

$$u'(h^c(\hat{\theta}^t)) > \frac{1}{q^c(\hat{\theta}^t, \hat{\theta}_{t+1})} \beta\pi(\hat{\theta}^{t+1}|\hat{\theta}^t) u'(h^c(\hat{\theta}^t, \hat{\theta}_{t+1})) \tag{35}$$
$$\frac{1}{\hat{\theta}_t} v'(h^l(\hat{\theta}^t)) > \frac{1}{q^l(\hat{\theta}^t, \hat{\theta}_{t+1})} \beta\pi(\hat{\theta}^{t+1}|\hat{\theta}^t) \frac{1}{\hat{\theta}_{t+1}} v'(h^l(\hat{\theta}^t, \hat{\theta}_{t+1})). \tag{36}$$

Now we will construct a (ϵ_1, ϵ_2)-variation ($\epsilon_1, \epsilon_2 > 0$) of the Pareto-Negishi weights, $\{\tilde{M}_t(\theta^t)\}$ and that of the aggregate allocation, $\{\tilde{C}_t, \tilde{L}_t, \tilde{K}_t\}$ as follows.

$$\tilde{M}_s(\theta^s) = \begin{cases} M^R_s(\theta^s) - \epsilon_1 & \text{if } \theta^s = (\hat{\theta}^t, \hat{\theta}_{t+1}) \\ M^R_s(\theta^s) + \epsilon_2 & \text{if } \theta^s = \hat{\theta}^t \\ M^R_s(\theta^s) & o.w. \end{cases}$$

$$\begin{cases} \tilde{C}_t = C^R_t + \pi_t(\hat{\theta}^t)\left(h^c(\hat{\theta}^t, \tilde{C}_t, \tilde{L}_t; \tilde{M}) - h^c(\hat{\theta}^t, C^R_t, L^R_t; M^R)\right) \\ \tilde{C}_{t+1} = C^R_{t+1} + \pi_{t+1}(\hat{\theta}^{t+1})\left(h^c(\hat{\theta}^{t+1}, \tilde{C}_{t+1}, \tilde{L}_{t+1}; \tilde{M}) - h^c(\hat{\theta}^{t+1}, C^R_{t+1}, L^R_{t+1}; M^R)\right) \\ \tilde{C}_s = C^R_s \qquad \text{if } s \notin \{t, t+1\} \end{cases}$$
$$\begin{cases} \tilde{L}_t = L^R_t + \pi_t(\hat{\theta}^t)\hat{\theta}_t\left(h^l(\hat{\theta}^t, \tilde{C}_t, \tilde{L}_t; \tilde{M}) - h^l(\hat{\theta}^t, C^R_t, L^R_t; M^R)\right) \\ \tilde{L}_{t+1} = L^R_{t+1} + \pi_{t+1}(\hat{\theta}^{t+1})\hat{\theta}_{t+1}\left(h^l(\hat{\theta}^{t+1}, \tilde{C}_{t+1}, \tilde{L}_{t+1}; \tilde{M}) - h^l(\hat{\theta}^{t+1}, C^R_{t+1}, L^R_{t+1}; M^R)\right) \\ \tilde{L}_s = L^R_s \qquad \text{if } s \notin \{t, t+1\} \end{cases}$$

We construct $\{\tilde{K}_{s+1}\}$ so that it satisfies resource constraints for all s: $\tilde{K}_{s+1} = F(\tilde{K}_s, \tilde{L}_s) + (1-\delta)\tilde{K}_s - \tilde{C}_s$.

Notice that

$$\left(h^c(\theta^s, \tilde{C}_s, \tilde{L}_s; \tilde{M}), h^l(\theta^s, \tilde{C}_s, \tilde{L}_s; \tilde{M})\right) = \left(h^c(\theta^s, C_s^R, L_s^R; M^R), h^l(\theta^s, C_s^R, L_s^R; M^R)\right), \quad \forall \theta^s \notin \{\hat{\theta}^t, (\hat{\theta}^t, \hat{\theta}_{t+1})\}.$$

That is, we changed the Pareto-Negishi weights and aggregate allocation so that they only change individual allocations for history $\hat{\theta}^t$ and $(\hat{\theta}^t, \hat{\theta}_{t+1})$. Also notice that $\{U_c^f(s), U_L^f(s)\}$ are kept constant for all s after variation. $\Big($This is because $U_c^f(s) = \tilde{M}(\tilde{\theta}^s)u'(h^c(\tilde{\theta}^s)), \quad U_L^f(s) = \tilde{M}(\tilde{\theta}^s)\frac{1}{\tilde{\theta}_s}v'(h^l(\tilde{\theta}^s))$ for all $\tilde{\theta}^s$ and we know that Pareto-Negishi weights and individual allocations are not changed for all $\tilde{\theta}^s \notin \{\hat{\theta}^t, \hat{\theta}^{t+1}\}$.$\Big)$

Since $(\epsilon_1, \epsilon_2) > 0$, and using $\tilde{M}_s(\theta^s)u'(h^c(\theta^s)) = \lambda_s^m \quad and \quad \frac{1}{\theta_s}\tilde{M}_s(\theta^s)v'(h^l(\theta^s)) = \mu_s^m, \quad \forall \theta^s \in \Theta^s$, we know that

$$\begin{aligned}
h^c(\hat{\theta}^t, \tilde{C}_t, \tilde{L}_t; \tilde{M}) &> h^c(\hat{\theta}^t, C_t^R, L_t^R; M^R) && (37)\\
h^l(\hat{\theta}^t, \tilde{C}_t, \tilde{L}_t; \tilde{M}) &< h^l(\hat{\theta}^t, C_t^R, L_t^R; M^R) && (38)\\
h^c(\hat{\theta}^{t+1}, \tilde{C}_{t+1}, \tilde{L}_{t+1}; \tilde{M}) &< h^c(\hat{\theta}^{t+1}, C_{t+1}^R, L_{t+1}^R; M^R) && (39)\\
h^l(\hat{\theta}^{t+1}, \tilde{C}_{t+1}, \tilde{L}_{t+1}; \tilde{M}) &> h^l(\hat{\theta}^{t+1}, C_{t+1}^R, L_{t+1}^R; M^R) && (40)
\end{aligned}$$

Then, t-period utility of $\hat{\theta}^t$ and $t+1$-period utility of history $\hat{\theta}^{t+1}$ will decrease. Then, by (35) and (36), the continuation value at $\hat{\theta}^t$ will increase.

By picking up small $(\epsilon_1, \epsilon_2) > 0$, the enforcement constraints and the monotonicity will be satisfied for all histories.

In addition, for a given ϵ_1, we can always find ϵ_2 so that the implementability constraint holds. The reason we can find such ϵ_2 value is the following. For a given $\epsilon_1 > 0$, we will find $\epsilon_2 > 0$ that satisfies following equation:

$$\underbrace{\beta^t \pi_t(\hat{\theta}^t) \begin{bmatrix} U_c^f(C_t^R, L_t^R; M^R)\left(h^c(\hat{\theta}^t, \tilde{C}_t, \tilde{L}_t; \tilde{M}) - h^c(\hat{\theta}^t, C_t^R, L_t^R; M^R)\right) \\ +U_L^f(C_t^R, L_t^R; M^R)\hat{\theta}_t\left(h^l(\hat{\theta}^t, \tilde{C}_t, \tilde{L}_t; \tilde{M}) - h^l(\hat{\theta}^t, C_t^R, L_t^R; M^R)\right) \end{bmatrix}}_{\equiv \Delta_t(\epsilon_2) > 0}$$
$$+ \underbrace{\beta^{t+1}\pi_{t+1}(\hat{\theta}^{t+1}) \begin{bmatrix} U_c^f(C_{t+1}^R, L_{t+1}^R; M^R)\left(h^c(\hat{\theta}^{t+1}, \tilde{C}_{t+1}, \tilde{L}_{t+1}; \tilde{M}) - h^c(\hat{\theta}^{t+1}, C_{t+1}^R, L_{t+1}^R; M^R)\right) \\ +U_L^f(C_{t+1}^R, L_{t+1}^R; M^R)\hat{\theta}_t\left(h^l(\hat{\theta}^{t+1}, \tilde{C}_{t+1}, \tilde{L}_{t+1}; \tilde{M}) - h^l(\hat{\theta}^{t+1}, C_{t+1}^R, L_{t+1}^R; M^R)\right) \end{bmatrix}}_{\equiv \Delta_{t+1}(\epsilon_1) < 0} = 0. \tag{41}$$

(41) guarantees that the implementability constraint is satisfied because our $(\epsilon_1, \epsilon_2)-$variation did not change U_c^f, U_L^f and all other individual allocations for $\theta^s \notin \{\hat{\theta}^t, \hat{\theta}^{t+1}\}$. There always exists

a $\epsilon_2 > 0$ that satisfies (41) for a given $\epsilon_1 > 0$, because given $\epsilon_1 > 0$, $\Delta_{t+1}(\epsilon_1) < 0$ and $\Delta_t(\epsilon_2)$ has property: $\Delta_t'(\cdot) > 0$ and $\Delta_t(0) = 0$. Then we can see that $\left\{\tilde{C}_t, \tilde{L}_t, \tilde{K}_t, \{\tilde{M}_t(\theta^t)\}\right\}$ satisfies all constraints of the RRP but increases expected lifetime utility (the objective function of the RRP) because of (35), (36), and (37)–(40), which contradicts optimality of $\left\{C_t^R, L_t^R, K_t^R, \{M_t^R(\theta^t)\}\right\}$. ■

To prove Lemma 5, we first state following lemma, which shows that the Pareto weight of the Ramsey government $\xi_t(\theta^t)$ and the Pareto-Negishi weight $M_t(\theta^t)$ are equivalent in the optimal solution.

Lemma 11. *Let $\{M_t^R(\theta^t)\}$ be the Ramsey equilibrium Pareto-Negishi weights that solve the RRP and let $\{\xi_t(\theta^t)\}$ be the cumulative multipliers of the enforcement constraints in the RRP, which are defined as $\xi_t(\theta^t) = 1 + \sum_{\theta^s \preceq \theta^t} \mu_s(\theta^s)$, $\forall \theta^t$. Then, for all θ^t, $M_t^R(\theta^t) = \xi_t(\theta^t)$.*

Proof Since both the relaxed Ramsey problem and the household problem of an equilibrium have the same objective function and face exactly the same enforcement constraints, the Lagrange multipliers of enforcement constraints are the same for both problems. Since $\xi_t(\theta^t)$ and $M_t(\theta^t)$ are the cumulative multipliers of each problem, respectively, they are equal to each other. ■

Proof of Lemma 5

$$
\begin{aligned}
W_c(C_t, L_t; M, \xi, \lambda) &= \sum_{\theta^t} \pi_t(\theta^t) \left[\xi_t(\theta^t) u'(h^c(\theta^t)) \frac{\partial h^c(\theta^t)}{\partial C_t} + \lambda \left\{ U_{cc}^f(t) h^c(\theta^t) + U_c^f(t) \frac{\partial h^c(\theta^t)}{\partial C_t} \right\}\right] \\
&= \sum_{\theta^t} \pi_t(\theta^t) \left[U_c^f(t) \frac{\partial h^c(\theta^t)}{\partial C_t} + \lambda \left\{ U_{cc}^f(t) h^c(\theta^t) + U_c^f(t) \frac{\partial h^c(\theta^t)}{\partial C_t} \right\}\right] \\
&= U_c^f(t) \left[1 + \lambda \left(1 + \frac{U_{cc}^f(t)}{U_c^f(t)} C_t\right)\right]
\end{aligned}
$$

The second equality comes from Lemma 11 and (11), and the third equality is because the sum of changes of individual allocations is equal to the change of aggregate allocation.

$$
\begin{aligned}
W_L(C_t, L_t; M, \xi, \lambda) &= \sum_{\theta^t} \pi_t(\theta^t) \left[-\xi_t(\theta^t) v'(h^l(\theta^t)) \frac{\partial h^l(\theta^t)}{\partial L_t} + \lambda \left\{ U_{LL}^f(t) \theta_t h^l(\theta^t) + U_L^f(t) \theta_t \frac{\partial h^l(\theta^t)}{\partial L_t} \right\}\right] \\
&= \sum_{\theta^t} \pi_t(\theta^t) \left[U_L^f(t) \theta_t \frac{\partial h^l(\theta^t)}{\partial L_t} + \lambda \left\{ U_{LL}^f(t) \theta_t h^l(\theta^t) + U_L^f(t) \theta_t \frac{\partial h^l(\theta^t)}{\partial L_t} \right\}\right] \\
&= U_L^f(t) \left[1 + \lambda \left(1 + \frac{U_{LL}^f(t)}{U_L^f(t)} L_t\right)\right]
\end{aligned}
$$

Again, second equality is comes from Lemma 11 and (12), and the third equality is because the sum of changes of individual allocations is equal to the change of aggregate allocation. ■

Proof of Proposition 7 First order conditions of the relaxed Ramsey problem are:

$$\cdot \quad W_L(C_t, L_t; M, \xi, \lambda) - \bar{\Delta}_t = -W_c(C_t, L_t; M, \xi, \lambda) F_L(K_t, L_t), \tag{42}$$
$$where \quad \bar{\Delta}_t = \sum_{\theta^t} \pi_t(\theta^t) \left\{ \xi_t(\theta^t) - 1 \right\} \frac{\partial U^{aut}(\theta_t; w_t, \tau_{l,t})}{\partial L_t}$$

$$\cdot \quad W_c(C_t, L_t; M, \xi, \lambda) = \beta W_c(C_{t+1}, L_{t+1}; M, \xi, \lambda) \{F_K(K_{t+1}, L_{t+1}) + 1 - \delta\} - \beta \bar{\chi}_{t+1}, \tag{43}$$
$$where \quad \bar{\chi}_{t+1} = \sum_{\theta^{t+1}} \pi_{t+1}(\theta^{t+1}) \left\{ \xi_{t+1}(\theta^{t+1}) - 1 \right\} \frac{\partial U^{aut}(\theta_{t+1}; w_{t+1}, \tau_{l,t+1})}{\partial K_{t+1}}.$$

First, we derive optimal capital income taxes. By equating the intertemporal condition of a competitive equilibrium with limited commitment and that of the Ramsey government (43), we get

$$\begin{aligned} & \frac{\beta W_c(t+1)}{W_c(t)} [F_K(t+1) + 1 - \delta] - \frac{\beta \bar{\chi}_{t+1}}{W_c(t)} \\ = \; & \frac{\beta U_c^f(t+1)}{U_c^f(t)} [F_K(t+1) + 1 - \delta] - \tau_{k,t+1}(F_k(t+1) - \delta) \frac{\beta U_c^f(t+1)}{U_c^f(t)}. \end{aligned}$$

We need to show that $\frac{W_c(t)}{U_c^f(t)}$ is constant in the steady-state. By Lemma 5, it is sufficient to show that $\frac{U_{cc}^f(t)}{U_c^f(t)} C_t$ is constant in the steady-state. Since $\{C_t, L_t\}$ are constant and distributions of $\{c_t(\theta^t), l_t(\theta^t)\}$ are invariant in the steady-state, there exists time-invariant Pareto-Negishi weights $\{\tilde{M}_t(\theta^t)\}$ whose distribution is equivalent to $\{\alpha_t M_t(\theta^t)\}$ for some constant, $\alpha_t > 0$, and whose mean is normalized to 1 ($\sum_{\theta^t} \pi(\theta^t) \tilde{M}_t(\theta^t) = 1$). Then, for every t, $U^f(C_t, L_t; M) = \alpha_t \cdot U^f(C_t, L_t; \tilde{M})$, for some $\alpha_t > 0$, which implies $\frac{U_{cc}^f(t)}{U_c^f(t)} C_t$ is constant in the steady-state. Thus, we get

$$\tau_{k,t+1} \;=\; \frac{1}{W_c(t+1)[F_K(t+1) - \delta]} \sum_{\theta^{t+1}} \pi_{t+1}(\theta^{t+1}) \left\{ \xi_{t+1}(\theta^{t+1}) - 1 \right\} \frac{\partial U^{aut}(\theta_{t+1}; w_{t+1}, \tau_{l,t+1})}{\partial K_{t+1}}.$$

Notice that $\frac{\partial U^{aut}(\theta_{t+1})}{\partial K_{t+1}} = \frac{\partial U^{aut}(\theta_{t+1})}{\partial w_{t+1}} \frac{\partial w_{t+1}}{\partial K_{t+1}} > 0$. Also, $\{\xi_{t+1}(\theta^{t+1}) - 1\} > 0$ in the long run, since $\xi_{t+1}(\theta^{t+1})$ is the cumulative Lagrange multiplier of the enforcement constraints and $\xi_0(\theta_0) = 1$. Also, the normalization term, $W_c(t+1)$ is always positive at the optimum. We can see that from the first order condition of the Ramsey problem with respect to C_t,

$$\beta^t W_c(t) = \gamma_t,$$

where γ_t is the Lagrange multiplier for resource constraint. Since $\gamma_t > 0$ for all t, we get $W_c(t) > 0$, for all t. Thus, the optimal capital income taxes are positive in the long run.

Now we derive optimal labor income taxes. By equating the intratemporal condition of a competitive equilibrium with limited commitment and that of the Ramsey government, we can derive the following expression for the labor income taxes.

$$
\begin{aligned}
1-\tau_{l,t} &= \frac{W_c(C_t,L_t;M,\xi,\lambda)}{U_c^f(C_t,L_t;M)} \frac{U_L^f(C_t,L_t;M))}{W_L(C_t,L_t;M,\xi,\lambda)-(1-\tau_{l,t})\Delta_t}, \\
& \quad \textit{where} \quad (1-\tau_{l,t})\Delta_t = (1-\tau_{l,t}) \underbrace{\sum_{\theta^t} \pi(\theta^t)\left\{\xi_t(\theta^t)-1\right\} \frac{\partial U^{aut}(\theta_t;w_t,\tau_{l,t})}{\partial w^t(1-\tau_{l,t})} \frac{\partial w_t}{\partial L_t}}_{=\Delta_t} < 0 \\
&= \frac{1+\lambda\left(1+\frac{U_{cc}^f(t)}{U_c^f(t)}C_t\right)}{1+\lambda\left(1+\frac{U_{LL}^f(t)}{U_L^f(t)}L_t\right)-(1-\tau_{l,t})\frac{\Delta_t}{U_L^f(t)}}, \\
& \quad \textit{where} \\
& \quad \lim_{t\to+\infty} \frac{\Delta_t}{U_L^f(t)} = \lim_{t\to+\infty} -\sum_{\theta^t} \pi(\theta^t) \frac{\theta_t}{v'(l(\theta^t))}\left\{\frac{\xi_t(\theta^t)}{M_t(\theta^t)} - \frac{1}{M_t(\theta^t)}\right\} \frac{\partial U^{aut}(\theta_t;w_t,\tau_{l,t})}{\partial w_t(1-\tau_{l,t})} \frac{\partial w_t}{\partial L_t} \\
& \quad\quad = -\sum_{\theta^t} \pi(\theta^t) \frac{\theta_t}{v'(l(\theta^t))} \frac{\partial U^{aut}(\theta_t;w_t,\tau_{l,t})}{\partial w(1-\tau_l)} \frac{\partial w}{\partial L}
\end{aligned}
$$

The last equality holds because $\xi_t(\theta^t) = M_t(\theta^t)$ by Lemma 11, and $\lim_{t\to+\infty} M_t(\theta^t) = +\infty$ ■

Proof of Proposition 9 When allowing lump-sum taxes, the implementability constraint is expressed as follows.

$$
\begin{aligned}
& \sum_t \sum_{\theta^t} \beta^t \pi_t(\theta^t) \left[U_c^f(C_t,L_t;M) h^c(\theta^t,C_t,L_t;M) + U_L^f(C_t,L_t;M)\theta_t h^l(\theta^t,C_t,L_t;M)\right] \\
& = U_c^f(C_0,L_0;M)\left\{\left[1+(F_K(K_0,L_0)-\delta)(1-\tau_{k,0})\right]K_0 + B_0 - T\right\}, \\
& \textit{where} \quad T = \sum_t \sum_{\theta^t} p(\theta^t) T_t.
\end{aligned}
$$

Among the five conditions characterizing a competitive equilibrium with limited commitment, only the implementability constraint will be changed by the introduction of lump-sum taxes.

In the Ramsey problem, the Ramsey government will also maximize over T. The first order condition with respect to T is $U_c^m(0)\lambda = 0$, implying $\lambda = 0$. This is because the government can always choose T such that an implementability constraint is no longer binding. Then, Lemma 5 implies that

$$
W_c(C_t,L_t;M,\xi,\lambda) = U_c^f(C_t,L_t;M), \quad W_L(C_t,L_t;M,\xi,\lambda) = U_L^f(C_t,L_t;M).
$$

Using these equalities, the first order conditions of the Ramsey problem, (21), (22), and the first order conditions of a competitive equilibrium with limited commitment, (13) and (14), we can derive following labor income taxes and capital income taxes. For all t,

$$1-\tau_{l,t} = \frac{1}{1-(1-\tau_{l,t})\frac{\Delta_t}{U_L^f(t)}}, \tag{44}$$

$$where \quad \Delta_t = \sum_{\theta^t} \pi_t(\theta^t)\{\xi_t(\theta^t)-1\}\frac{\partial U^{aut}(\theta_t; w_t, \tau_{l,t})}{\partial(1-\tau_{l,t})w_t}\frac{\partial w_t}{\partial L_t} < 0,$$

$$\tau_{k,t+1} = \frac{1}{U_c^f(t+1)[F_K(t+1)-\delta]}\sum_{\theta^{t+1}} \pi_{t+1}(\theta^{t+1})\{\xi_{t+1}(\theta^{t+1})-1\}\frac{\partial U^{aut}(\theta_{t+1}; w_{t+1}, \tau_{l,t+1})}{\partial K_{t+1}} > 0.$$

We know $\tau_{l,t} < 1$ for all t, since otherwise $L_t = 0$ for all t. We can also prove $\tau_{l,t} < 0$, for all t, by contradiction. Suppose $\tau_{l,t} \geq 0$. Then, the left hand side of (44) is $0 \leq 1-\tau_{l,t} \leq 1$. Then, because $\Delta_t < 0$, $U_L^f(t) < 0$ and $0 \leq (1-\tau_{l,t}) \leq 1$, the right hand side of (44) is either greater than 1 or negative, which leads to a contradiction. ■

Proof of Proposition 10 When the fictitious representative agent's utility is time-additively separable and satisfies the condition (i) or (ii) of Proposition 10, we can easily show that $\frac{U_{cc}^f(t)C_t + U_{LC}^f(t)L_t}{U_c^f(t)}$ is constant over t. Then, $\frac{W_c(t)}{U_c^f(t)} = 1+\lambda+\lambda(\frac{U_{cc}^f(t)C_t + U_{LC}^f(t)L_t}{U_c^f(t)})$ is constant over time, which implies

$$\tau_{k,t+1} \quad = \quad \frac{1}{W_c(t+1)[F_K(t+1)-\delta]}\sum_{\theta^{t+1}} \pi_{t+1}(\theta^{t+1})\left\{\xi_{t+1}(\theta^{t+1})-1\right\}\frac{\partial U^{aut}(\theta_{t+1}; w_{t+1}, \tau_{l,t})}{\partial K_{t+1}}.$$

■