Optimal Capital Income Taxation in the Borrower-Saver Model[1]

Sama Bombaywala[2]

June, 2016

Abstract

In this paper we derive the optimal level of capital taxation in the presence of agents with different discount factors. We set up a real business cycle model with patient and impatient households which borrow and lend amongst themselves, as per a borrowing constraint. Our results show that if the Ramsey planner's weights on different households are such that he is indifferent between redistribution towards patient and impatient households, the borrowing constraint is not binding. Moreover, we get the classical result of zero optimal capital taxation in the distant long run. However, if the Ramsey planner chooses the borrowing constraint to be always binding, he will favour redistribution from impatient households to patient households. As time moves forward, this ultimately leads to an increase in the optimal subsidy rate on the capital returns of patient households, a contradiction to the seminal Chamley-Judd result.

JEL Classification: E62, H21, H31

Key words: Optimal capital taxation, heterogeneous agents

[1]I am thankful to Dr. Raffaele Rossi for his constant supervision and constructive and prompt guidance at every step of the project. The paper also benefited from the discussion with Fabrizio Zilibotti and the useful comments of an anonymous, conference referee.

[2]University of Manchester, School of Social Sciences. Email: sama.bombaywala@postgrad.manchester.ac.uk

1 Introduction

In 1928, Ramsey wrote about heterogeneous degrees of patience amongst individuals and conjectured that in the long run, "the thrifty enjoy bliss and the improvident live at the subsistence level." Does Ramsey's conjecture hold in the Ramsey problem? We find that the answer is affirmative. In simple words, we derive the optimal level of capital taxation that maximizes the social welfare of a society inhabited by people of varying degrees of patience with an agent being more impatient and myopic than the other. With the aim of increasing their current utility, the impatient households borrow from the patient households, subject to a borrowing constraint. Taking a social discount factor which is a linear weight of the two different discount factors, we show that if the Ramsey planner is indifferent between redistribution between patient and impatient households, the borrowing constraint is not binding and hence the optimal taxation approaches zero in the distant long run. Till then, it is optimal to redistribute from unconstrained patient households to constrained impatient households. On the other hand, if the borrowing constraint is to be binding, the planner favours a redistribution towards patient households who will eventually account for the majority of consumption in the economy. To achieve this, the Ramsey planner will subsidize returns on the capital income at an increasing rate. Since the impatient households value future less, with passing time, the Ramsey planner will have the opportunity to tax them more and transfer the tax receipts to patient households, who value future more, in the form of subsidies.

An important result in the literature on optimal taxation is the one by Chamley (1986) and Judd (1985) who show that the optimal capital taxation should be zero in the long run. They argue that taxing capital would reduce the incentive of people to invest in capital, lowering the production activity and wages in the economy- a cost which exceeds the benefits of redistribution. However, both the papers assume perfect capital markets and agents with same time preferences. While Chamley (1986) uses an infinitely lived representative agent model to show the result, Judd (1985) includes heterogeneity in the form of capitalists and workers. The framework used by Judd (1985) resembles the

Savers-Spenders model in which one section of the society optimizes intertemporally by saving for the future while the other section survives on its wages.

The Savers-Spenders model however, lacks the borrowing side of the market. In reality, agents vary in their levels of patience and preferences towards present and future levels of consumption and savings. Models should thus have at least two different groups of agents- lenders and borrowers (Quadrini, 2011). Empirically, compared to the benchmark representative agent model, Krussel and Smith (1998) show that when heterogeneity in discount factors is taken into account, the model captures the very skewed wealth distribution observed in the data better. In this paper, accounting for this heterogeneity, we study the optimal level of capital taxation in the Borrower-Saver model which, as stated by Bilbiie, Monacelli and Perotti (2012), differs from the Savers-Spenders model in four aspects: first, there are savers and borrowers who have different time preferences, and hence, different discount factors; second, both kind of agents are intertemporal optimizers with the borrowers borrowing from the savers; third, the borrowers face a borrowing constraint which introduces financial frictions in the economy; and finally, the decision on the equilibrium level of lending and borrowing is endogenized, subject to the borrowing constraint.

In the presence of borrowing constraints, Aiyagari (1995) shows that the result of zero long run capital taxation is not welfare optimizing. Accounting for future uncertainty and the possibility of being borrowing constrained in the future, agents indulge in an over accumulation of savings in the short run. Compared to the marginal product of capital, the return on capital is lower hence making it optimal to tax capital income even in the long run. Later, Chamley (2001) generalized this result by showing that in the presence of a borrowing constraint, whenever there is positive relationship between consumption and savings, savings should be taxed. However, if there is a negative relationship between consumption and savings, savings should be subsidized.

Recently, the Borrower-Saver model has started gaining momentum in the Macroeconomic literature with Kiyotaki and Moore (1997) using an extended version of the model to study credit cycles. Some other important papers which work on or extend this model are Iacoviello and Neri (2010), Monacelli and Perotti (2011), Eggertsson and Krugman (2012), McKay and Reis (2013) and Alpanda and Zubairy (2016). With a lot of macroeconomists, policy makers and central bankers using the model to study optimal policies and various policy implications, it becomes even more important to know the very basic- the optimal level of capital taxation in a model with different degrees of patience and a borrowing constraint.

The structure of the remainder of the paper is as follows: section 2 highlights the modelled setup of our economy. Section 3 and 4 respectively discuss the equilibrium and steady state conditions of the decentralized economy. Section 5 covers the benevolent social planner's first best allocation. Section 6 is the main focus of the paper and details the Ramsey optimal taxation which covers the analysis of second best allocations. Finally, section 7 concludes. All the detailed mathematical calculations are presented in the appendix.

2 The Model

We use a real business cycle closed economy model that comprises of heterogeneous households that differ in their discount factors. The setup also comprises of firms which hire labor and rent capital to carry out their production activities. The government levies distortionary capital income taxes to finance its expenditure and runs a balanced budget constraint.

2.1 Households

Households maximize their utilities subject to their respective budget constraints and differ in two respects- the rate of time preference and labor supply. Their preferences are given by the following CES utility function:

$$\max E_t \sum_{t=0}^{\infty} \beta_s^t \left\{ \frac{c_{s,t}^{1-\sigma}}{1-\sigma} \right\},$$

where $\beta \in (0,1)$ is the discount factor, σ is the inverse of the intertemporal elasticity of substitution and s=P and I, denotes patient and impatient households respectively.

Patient households are more patient and value future more. They thus indulge in consumption smoothing by saving today. Impatient households, on the other hand, derive a higher marginal utility from consuming today and hence borrow to increase their level of current consumption. This heterogeneity is highlighted through the difference in their discount factors, β_P and β_I respectively, with $\beta_P > \beta_I$.

Further, following Campbell and Hercowitz (2005) and keeping our analysis close to Judd (1985) and Section 2 of Straub and Werning (2014), we assume that patient households do not work and accumulate enough wealth to fulfill their consumption needs. Impatient households, on the other hand, supply inelastic labour that we normalize to one.

2.1.1 Patient Households

Patient households choose the optimal level of consumption $c_{P,t}$ to maximize

$$\max E_t \sum_{t=0}^{\infty} \beta_P^t u\left(c_{P,t}\right),$$

subject to

$$c_{P,t} + d_{t+1} + k_{t+1} - (1-\delta) k_t = \left(1 + r_t^d\right) d_t + r_{k,t} k_t - \tau_{k,t}\left(r_{k,t} - \delta\right) k_t, \tag{1}$$

where k_t is the level of capital stock at the beginning of the period, d_t is the amount lent to impatient households, δ represents the rate of depreciation and $\tau_{k,t}$ is the capital income tax rate. Further r_t^d and $r_{k,t}$ represent the rate of return on private lending to impatient households and physical capital, respectively.

We derive the following first order conditions for the above problem:

$$\frac{1}{\beta_P}\left(\frac{u_{cp,t}}{u_{cp,t+1}}\right) = 1 + r_{t+1}^d, \tag{2}$$

$$r_{t+1}^{d}=\left(1-\tau_{k,t+1}\right)\left(r_{t+1}^{k}-\delta\right). \tag{3}$$

Equation 2 is the standard Euler equation which highlights the relationship between intertemporal consumption choices of the patient households and the interest rate. Equation 3 represents the no arbitrage condition between the lending rate r_t^d and the net return on capital r_t^k.

2.1.2 Impatient Households

Similar to the patient households, the impatient households maximize their utility

$$\max E_t \sum \beta_I^t u\left(c_{I,t}\right),$$

subject to

$$c_{I,t}+\left(1+r_t^d\right)d_t=w_t+d_{t+1}+T_t, \tag{4}$$

where $c_{I,t}$ is consumption, d_t is the level of borrowing from the patient households, w_t is the wage, r_t^d is the interest rate on borrowings and T_t represents lump sum transfers (when positve) or taxes (when negative) from/ to the government. Further, impatient households' level of borrowing is subject to a borrowing constraint. In the spirit of simplicity, following Monacelli and Perotti (2011), we choose the constaint to be an exogenous, fixed borrowing limit with $\bar{D}>0$:

$$d_{t+1}\leq\bar{D}. \tag{5}$$

Solving the impatient households' maximization problem subject to their budget and borrowing constraints, we derive the following first order condition:

$$\left(\frac{1-\mu_{I,t}}{\beta_I}\right)\frac{u_{ci,t}}{u_{ci,t+1}}=1+r_{t+1}^{d}. \tag{6}$$

Equation 6 gives the impatient households' Euler equation. $u_{ci,t}\mu_{I,t}$ is the Lagrange multiplier associated with the borrowing constraint and $\mu_{I,t}$ measures the value of marginally relaxing the borrowing constraint, in units of $c_{I,t}$.

The borrowing constraint alters the equilibrium by introducing credit market imperfections. In the case of perfect capital market (with no borrowing constraint), 6 becomes

$$\frac{1}{\beta_I}\left(\frac{u_{ci,t}}{u_{ci,t+1}}\right) = 1 + r^d_{t+1}. \tag{7}$$

Equating 7 and 2 in steady state, we get

$$\frac{1}{\beta_I} = \frac{1}{\beta_P},$$

which is not true with $\beta_P > \beta_I$. Thus, in a model with heterogeneous patience and no borrowing constraint, a stable steady state does not exist. This is because, without a borrowing constraint, the impatient households trade future labor with current consumption. They borrow against the future discounted value of their wages, leading their consumption to approach zero asymptotically[3]. With a borrowing constraint, however, there is a limit on the amount impatient households can borrow which results in a stable steady state and a positive level of consumption for both the households.

2.2 Firms

A perfectly competitive firm hires labor from impatient households to produce goods and pays wages w_t to its employees. The production of the firms is given by a Cobb-Douglas production function

$$y_t = F(k_t, n_{I,t}) = k_t^{\alpha} n_{I,t}^{(1-\alpha)}, \tag{8}$$

where $\alpha \in (0,1)$ to ensure monotonicity of the capital sequence[4].

Since $n_{I,t} = 1$ in equilibrium, $f(k) = F(k,1)$.

[3]Further, Le Van and Vailakis (2003) also prove that in a perfect economy with heterogeneous agents with different discount factors, even though the capital sequence approaches a particular level of capital stock, this level is not the steady state level.

[4]Refer to Becker and Foias (1987) for monotonicity of capital sequence.

The firm maximizes its profits and thus obtain the following conditions:

$$f'(k_t) = r_t^k, \tag{9}$$

$$f(k_t) - f'(k_t)\, k_t = w_t. \tag{10}$$

Equation 9 states that in equilibrium, marginal product of capital should be equal to the return on capital. Similarly, equation 10 states that in equilibrium, the marginal product of labor should be equal to the wage paid.

2.3 Government

The government runs a balanced budget constraint with no debt. The inflows from current period's tax receipts on net returns on capital are used on an exogenous level of government expenditure, g, and lump sum transfers, T_t, to impatient households

$$g + T_t = \tau_{k,t} \left(r_t^k - \delta \right) k_t. \tag{11}$$

The government thus uses capital income tax as a tool of income redistribution between households. $\tau_{k,t} > 0$ implies a tax incidence on patient households, the receipts of which are distributed to impatient households in the form of lump sum transfers. However, when $\tau_{k,t} < 0$, taxes are imposed on impatient households in lump sum and redistributed to patient households as subsidies on capital income.

2.4 Aggregate Resource Constraint

Finally, the aggregate resource constraint is as follows:

$$y_t = c_{P,t} + c_{I,t} + g + k_{t+1} - (1 - \delta)\, k_t. \tag{12}$$

3 Equilibrium (Decentralized Economy)

Definition 1 *(Competitive Equilibrium): A competitive equilibrium consists of government policies* $\{\tau_{k,t}, T_t\}_{t=0}^{\infty}$, *prices* $\{w_t, r_t^k, r_t^d, \mu_{I,t}\}_{t=0}^{\infty}$ *and private sector allocations* $\{c_{P,t}, k_{t+1}, d_{t+1}, c_{I,t}, y_t\}_{t=0}^{\infty}$, *satisfying:*

(i) private sector optimization taking government policies and prices as given, that is,

- the households' budget constraints 1 and 4, borrowing constraint 5, and optimality conditions 2, 3 and 6;
- the production function 8, and firm's optimality conditions 9 and 10;

(ii) market clearing 12 and;

(iii) the government's budget constraint 11.

4 Steady State (Decentralized Economy)

Proposition 1 *The borrowing constraint is always binding in steady state.*

Proof. Analyzing the two Euler equations, 2 and 6, in steady state

$$\begin{aligned} \left[\frac{(1-\mu_I)-\beta_I}{\beta_I}\right] &= \frac{1-\beta_P}{\beta_P}, \\ &\Rightarrow \mu_I = 1-\frac{\beta_I}{\beta_P}. \end{aligned}$$

Given our choice of β_I and β_P such that $\beta_P > \beta_I$, μ_I will be grater than zero in steady state for all τ_k, ensuring that the borrowing constraint will always be binding. ■

5 First Best Allocation

The social planner maximizes the following weighted utility function

$$\omega\beta_P^t u(c_{P,t}) + (1-\omega)\beta_I^t u(c_{I,t}), \tag{13}$$

where ω is the weight assigned by the social planner to the patient households.

Definition 2 *(First Best): The first-best equilibrium consists of allocations* $\{c_{P,t},\ c_{I,t},\ y_t\}_{t=0}^{\infty}$ *that maximize 13 subject to the production function, 8, and the market clearing condition 12.*

The social planner's problem thus becomes,

$$\max_{\{c_{I,t}, c_{P,t}, k_{t+1}\}} \sum_{t=0}^{\infty} \beta_P^t \left[\omega u\left(c_{P,t}\right) + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t u\left(c_{I,t}\right) \right],$$

subject to

$$f\left(k_t\right) - c_{P,t} - c_{I,t} - g - k_{t+1} + (1-\delta)\, k_t = 0.$$

We now need to determine the social planner's discount factor. We could either use the linear average weight of the two agents, $\omega\beta_P^t + (1-\omega)\,\beta_I^t$, or the weight of an average agent, $[\omega\beta_P + (1-\omega)\,\beta_I]^t$. Lengwiler (2005) shows that

$$\omega\beta_P^t + (1-\omega)\,\beta_I^t > [\omega\beta_P + (1-\omega)\,\beta_I]^t,$$

for $t > 1$. This implies that the discount rate (inverse of the discount factor) obtained from the average weight will be less than the discount rate obtained from the weight of an average agent. Weitzman (1998) and Gollier (2002) argue that while evaluating problems for very long horizons, the social planner should be conservative and should consider minimum growth by using the average discount factor which gives a lower discount rate. Moreover, a higher discount rate leads to a lower present value of the social welfare, implying that the derived outcome will not be the first best. We thus choose to use $\omega\beta_P^t + (1-\omega)\,\beta_I^t$ as the social planner's discount factor[5].

Analyzing the discount factor, and as mentioned in Lengwiler (2005), we notice that the impatient households weight keep falling with time and they assign almost negligible weight on distant future. In such a scenario, the average weight is thus dominated by

[5] Becker (2011) uses a similar discount factor to demonstrate the twisted turnpike property of capital in a model with heterogeneous patience.

the patient households preferences. This implies that the social discount factor is time varying and does not generate a geometric progression, as in the case of a representative agent model where the discount factor is β_P^t[6].

Proposition 2 *As $t \to \infty$, $c_{I,t} \to 0$ will maximize the social welfare.*

Proof. Let $\lambda_{F,t}$ be the Lagrange multiplier associated with the feasibility/ aggregate resource constraint. Maximizing, we get the following first order conditions:

$$\frac{\partial \mathcal{L}}{\partial c_{P,t}} : \omega c_{P,t}^{-\sigma} = \lambda_{F,t} \left[\omega + (1-\omega) \left(\frac{\beta_I}{\beta_P} \right)^t \right], \tag{14}$$

$$\frac{\partial \mathcal{L}}{\partial c_{I,t}} : (1-\omega) \left(\frac{\beta_I}{\beta_P} \right)^t c_{I,t}^{-\sigma} = \lambda_{F,t} \left[\omega + (1-\omega) \left(\frac{\beta_I}{\beta_P} \right)^t \right]. \tag{15}$$

Equating 14 and 15 we get

$$\left[\left(\frac{1-\omega}{\omega} \right) \left(\frac{\beta_I}{\beta_P} \right)^t \right]^{\frac{1}{\sigma}} c_{P,t} = c_{I,t},$$

where $t \to \infty$, $\left(\frac{\beta_I}{\beta_P} \right)^t \to 0$, implying $c_{I,t} \to 0$. ∎

This result is consistent with the early conjecture of Ramsey (1928) that the most patient consumer will be the dominant consumer while, the impatient consumers' will survive at the subsistence level. Eventually, in the long run, impatient households' consumption will reach zero. This result has also been shown in the existing literature (see for example, Becker (2011), Le Van et al (2007) and Goenka et al (2012)).

Further,

$$\frac{\partial \mathcal{L}}{\partial k_{t+1}} : \left(f'\left(k_{t+1}\right) + 1 - \delta \right) = \frac{\lambda_{F,t} \left[\omega + (1-\omega) \left(\frac{\beta_I}{\beta_P} \right)^t \right]}{\lambda_{F,t+1} \beta_P \left[\omega + (1-\omega) \left(\frac{\beta_I}{\beta_P} \right)^{t+1} \right]}.$$

Again, as $t \to \infty$ and $\left(\frac{\beta_I}{\beta_P} \right)^t \to 0$, $\left(f'\left(k_{t+1}\right) + 1 - \delta \right) \to \frac{1}{\beta_P}$. Thus in the long run, the heterogeneous agent model will collapse to the representative agent model with $c_{I,t} \to 0$

[6] Mitra (1979) discusses about a time varying discounted growth model.

and $(f'(k_{t+1}) + 1 - \delta) \rightarrow \frac{1}{\beta_P}$.

6 Ramsey Optimal Taxation

The Ramsey planner aims to maximize the social welfare, given the decentralized economy's equilibrium conditions.

Substituting 2 and 3 in 1, we derive the patient household's implementability constraint:

$$\beta_P c_{P,t}^{1-\sigma} + \beta_P c_{P,t}^{-\sigma}(d_{t+1} + k_{t+1}) - c_{P,t-1}^{-\sigma}(d_t + k_t) = 0. \tag{16}$$

Further, equating 2 and 6, we get

$$\left(1 - \mu_{I,t-1}\right)\beta_P c_{I,t-1}^{-\sigma} c_{P,t}^{-\sigma} - \beta_I c_{I,t}^{-\sigma} c_{P,t-1}^{-\sigma} = 0. \tag{17}$$

The Ramsey planner's problem could thus be defined as:

Definition 3 *(Ramsey Problem): Ramsey policy maker's problem is to maximize 13, choosing* $\{\tau_{k,t}\}_{t=0}^{\infty}$ *and allocations* $\left\{c_{P,t},\, c_{I,t},\, k_t,\, \mu_{I,t}\right\}_{t=0}^{\infty}$ *subject to 16, 17 , 12 and 5 with* $\lambda_{I,t},\ \lambda_{E,t},\ \lambda_{F,t}$ *and* $u_{cp,t}\lambda_{B,t}$ *being the Lagrange multipliers on the implementability constraint, equation equating the two Euler equations, feasibility constraint and borrowing constraint respectively.*[7]

$$\max_{\left\{c_{I,t}, c_{P,t}, k_{t+1}, d_{t+1}, \mu_{I,t}\right\}} \sum_{t=0}^{\infty} \beta_P^t \left[\omega u\left(c_{P,t}\right) + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t u\left(c_{I,t}\right)\right],$$

subject to

$$\beta_P c_{P,t}^{1-\sigma} + \beta_P c_{P,t}^{-\sigma}(d_{t+1} + k_{t+1}) - c_{P,t-1}^{-\sigma}(d_t + k_t) = 0,$$

$$\left(1 - \mu_{I,t-1}\right)\beta_P c_{I,t-1}^{-\sigma} c_{P,t}^{-\sigma} - \beta_I c_{I,t}^{-\sigma} c_{P,t-1}^{-\sigma} = 0,$$

$$f(k_t) - c_{P,t} - c_{I,t} - g - k_{t+1} + (1-\delta) k_t = 0,$$

[7] Note that from Walras' Law, if all but one markets clear in the economy, the last one will clear too. Following the law, and keeping the analysis similar to Straub and Werning (2014), we drop impatient household's budget constraint while setting the Ramsey problem.

$$d_{t+1} \leq \bar{D}.$$

As argued in the case of a social planner, we use the linear weighted average of the two discount factors, $\omega\beta_P^t + (1-\omega)\beta_I^t$, as the Ramsey planner's discount factor.

The Lagrange multiplier $u_{cp,t}\lambda_{B,t}$ helps the shadow price $\lambda_{B,t}$, measure the value, in units of $c_{P,t}$, of marginally relaxing the borrowing constraint. The borrowing constraint is binding, that is, it holds will equality whenever $\lambda_{B,t} > 0$ and is not binding when $\lambda_{B,t} = 0$.

On simplification, maximizing with respect to $\mu_{I,t}$ and as detailed in the appendix, we find that varying $\mu_{I,t}$ imposes no opportunity costs on the government. Given the Ramsey planner's optimal choice for $c_{P,t}$, $c_{I,t}$ and k_t, the value of $\mu_{I,t}$ will adjust accordingly to achieve the desired levels. The problem thus reduces to

$$\begin{aligned}
\mathcal{L} \;=\;& E_0 \sum_{t=0}^{\infty} \beta_P^t \left[\omega u(c_{P,t}) + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t u(c_{I,t}) \right] \\
&+\lambda_{I,t}\beta_P^t \left[\omega + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t\right] \left[\beta_P c_{P,t}^{1-\sigma} + \beta_P c_{P,t}^{-\sigma}(d_{t+1}+k_{t+1}) - c_{P,t-1}^{-\sigma}(d_t+k_t)\right] \\
&+\lambda_{F,t}\beta_P^t \left[\omega + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t\right] \left[f(k_t) - c_{P,t} - c_{I,t} - g - k_{t+1} + (1-\delta)k_t\right] \\
&+\lambda_{B,t}c_{P,t}^{-\sigma}\beta_P^t \left[\omega + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t\right] \left[\bar{D} - d_{t+1}\right].
\end{aligned}$$

The necessary first order conditions are:

$$\lambda_{I,0} = 0,$$

$$c_{P,t}^{-\sigma}\beta_P \left\{ \left[\omega + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t\right]\lambda_{I,t} - \left[\omega + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^{t+1}\right]\lambda_{I,t+1} \right\} = \lambda_{B,t}c_{P,t}^{-\sigma}\left[\omega + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t\right], \tag{18}$$

$$\lambda_{B,t} \geqslant 0,$$

$$(1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t c_{I,t}^{-\sigma} = \lambda_{F,t}\left[\omega + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t\right], \tag{19}$$

$$\frac{\lambda_{F,t+1}}{\lambda_{F,t}}\left(f'(k_{t+1}) + 1 - \delta\right) = \frac{\left[\omega + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t\right]}{\beta_P\left[\omega + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^{t+1}\right]} - \frac{\lambda_{B,t}c_{P,t}^{-\sigma}}{(1-\omega)\beta_P c_{I,t}^{-\sigma}}\left(\frac{\beta_P}{\beta_I}\right)^t \frac{\left[\omega + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t\right]^2}{\left[\omega + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^{t+1}\right]}, \tag{20}$$

$$\lambda_{I,t+1}\left[\omega+(1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^{t+1}\right]=\lambda_{I,t}\left[\omega+(1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^{t}\right]\left[1+\frac{(\sigma-1)\,c_{P,t}}{\sigma\,(d_{t+1}+k_{t+1})}\right]+\frac{(1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^{t}c_{I,t}^{-\sigma}-\omega c_{P,t}^{-\sigma}}{\sigma\beta_P c_{P,t}^{-\sigma-1}\,(d_{t+1}+k_{t+1})}. \tag{21}$$

Analyzing the necessary optimal conditions, we consider the three possible scenarios that the last equation generates:

- **Case 1:** Redistribution preference towards impatient households- $(1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t c_{I,t}^{-\sigma} > \omega c_{P,t}^{-\sigma}$
- **Case 2:** Indifference between redistribution- $(1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t c_{I,t}^{-\sigma} = \omega c_{P,t}^{-\sigma}$
- **Case 3:** Redistribution preference towards patient households- $(1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t c_{I,t}^{-\sigma} < \omega c_{P,t}^{-\sigma}$

and discuss each of the three possibilities in detail.

6.1 Redistribution Towards Impatient Households

Proposition 3 *With the borrowing constraint, a welfare maximizing Ramsey allocation does not exist when* $(1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t c_{I,t}^{-\sigma} > \omega c_{P,t}^{-\sigma}$.

Proof. Let $(1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t c_{I,t}^{-\sigma} > \omega c_{P,t}^{-\sigma}$. With $\lambda_{I,0}=0$, from 21, $\lambda_{I,1}>0$. Substituting $\lambda_{I,1}>0$ in 18, we get $\lambda_{B,0}<0$. This defies the Kuhn-Tucker condition according to which $\lambda_{B,t}\geqslant 0$. Hence, for a solution to exist, $(1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t c_{I,t}^{-\sigma} \leqslant \omega c_{P,t}^{-\sigma}$.[8] ■

As $t\to\infty$, $\left(\frac{\beta_I}{\beta_P}\right)^t \to 0$. For $(1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t c_{I,t}^{-\sigma} > \omega c_{P,t}^{-\sigma}$ to hold as $t\to\infty$, $c_{P,t}$ will have to be negative. This possibility could be ruled out as the patient households will not starve themselves while owning positive wealth and lending to impatient households.

Next, we go on to the case when the Ramsey planner is indifferent between redistribution between households, that is, he values the weighted discounted marginal utilities of both the households equally.

[8] Refer to the Appendix for a formal proof of the violation of the Kuhn-Tucker condition.

6.2 Indifference Between Redistribution

Proposition 4 *When* $(1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t c_{I,t}^{-\sigma} = \omega c_{P,t}^{-\sigma}$*, the borrowing constraint is not binding and the optimal capital taxation approaches zero in the distant long run.*

Proof. *Let* $(1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t c_{I,t}^{-\sigma} = \omega c_{P,t}^{-\sigma}$*. With* $\lambda_{I,0} = 0$*, from 21,* $\lambda_{I,1} = 0$*. Further,* $\lambda_{I,t} = 0$ *for all t. Substituting* $\lambda_{I,t} = 0$ *in 18, we get* $\lambda_{B,t} = 0$ *for all t and the borrowing constraint is never binding. In such a scenario, 20 becomes*

$$\left(f'\left(k_{t+1}\right)+1-\delta\right) = \frac{\lambda_{3,t}}{\lambda_{3,t+1}}\left\{\frac{\left[\omega+(1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t\right]}{\beta_P\left[\omega+(1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^{t+1}\right]}\right\},$$

and the Ramsey planner's problem reduces to the social planner's problem. As $t \to \infty$ *and* $\left(\frac{\beta_I}{\beta_P}\right)^t \to 0$*,* $f'\left(k_{t+1}\right)+1-\delta \to \frac{1}{\beta_P}$ *implying that in the long run, marginal product of capital equals the return on capital and the optimal level of capital tax rate approaches zero.* ■

Intuitively, when the borrowing constraint is not binding, the patient households are not sure how much the impatient households will borrow. To ensure smooth future level of consumption, the patient households save capital greater than the socially optimal level. In the short run, the Ramsey planner thus taxes capital. In the long run, however, as the marginal product of capital falls, the socially optimal level of capital increases. This is due to the time varying discount factor of the planner which in the long run, is dominated by the patient household's preferences. As the socially optimal level of capital and patient household's capital savings become equal, the optimal capital tax rate approaches zero.

Finally, we discuss the case of redistribution preference towards patient households.

6.3 Redistribution Towards Patient Households

Proposition 5 *When* $(1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t c_{I,t}^{-\sigma} < \omega c_{P,t}^{-\sigma}$*, the borrowing constraint is always binding and the optimal level of capital taxation is negative.*

Proof. *Let* $(1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t c_{I,t}^{-\sigma} < \omega\beta_P^t c_{P,t}^{-\sigma}$*. With* $\lambda_{I,0} = 0$*, from 21,* $\lambda_{I,1} = 0$*. Further,*

$\lambda_{1,t+1} < \lambda_{1,t}$ *for all t. Substituting* $\lambda_{1,t+1} < \lambda_{1,t}$ *in 18, we get* $\lambda_{B,t} > 0$ *for all t. Hence the borrowing constraint will always be binding. In such a scenario, from 20, we notice that as* $t \rightarrow \infty$, $\left(\frac{\beta_P}{\beta_I}\right)^t \rightarrow \infty$. *The marginal product of capital will thus keep falling with time and will approach* $-\infty$ *as* $t \rightarrow \infty$, *implying an increasing path for the optimal subsidy rate.* ■

Intuitively, when the borrowing constraint is binding, the patient household are sure that they will at least save $\bar{D}$ in the form of private lending to the impatient households and will hence get returns on these savings at the end of the period. With this certainty, the patient households are secure about their future level of returns and consumption and choose to save less in capital- less than the socially optimal level of capital. With the aim to encourage patient households to save more, the planner chooses to subsidize capital.

The results derived from the second best allocation exercise could be seen as an application of Chamley (2001) where it is shown that if consumption and savings are negatively related, in the presence of a binding borrowing constraint, it would be optimal to have the returns on capital subsidized in the long run. He does not, however, make any claim on the rate on subsidization. Similar to our impatient households who are poor, consume less in the long run and have a negative relationship between consumption and savings (or positive relationship between consumption and dissavings, here borrowing), in Chamley (2001) the "very old" are poor and consume less in distant time periods and hence generate the negative correlation between consumption and savings. Moreover, the results obtained, could also be concluded to have preserved Ramsey's conjecture according to which, the patient households are the dominant consumers and enjoy bliss in the long run.

7 Conclusion

In our paper, we consider the case of heterogeneous agents who differ in their rates of time preference- impatient households value today's consumption more than the patient households. To increase their current level of consumption, they even borrow from the patient households as per a borrowing constraint. We study the optimal level of capital income taxation in such a scenario. The results show that if the Ramsey planner is indifferent between the redistribution towards patient and impatient households, it would be optimal to initially redistribute from unconstrained patient households to constrained impatient households. In the long run, however, the optimal level of capital taxation will approach zero. Further, when the borrowing constraint is binding, the capital income tax rate in not zero in the long run. On the contrary, the Ramsey planner chooses to give subsidies to patient households and keep increasing the subsidy rate with time. These subsidies are funded by transfers from the impatient households who anyway do not favour distant future much.

Although the derived results are welfare maximizing, they are not pareto improving. The increase in the utility of patient households comes at the cost of impatient households' utility. My next exercise would thus be to derive the optimal level of pareto improving capital taxation. I am also yet to check the stability of the results in certain cases. For example, Kemp et al (1993) questions the stability of Judd (1985) results and concludes that in certain cases cycles around steady state could establish. Straub and Werning (2014) pins down the case to be when the intertemporal elasticity of substitution is greater than one and there is redistribution preference towards capitalists (here, patient households). I thus need to conduct a similar exercise and check if different discount factors and/or the borrowing constraint stop the formation of a cycle. Further, to strengthen the paper by lending a numerical perspective, I aim to derive the transition path of the optimal level of capital taxation.

For future research purposes, it would be interesting to conduct a similar study in an open economy setup and analyze the welfare optimizing allocations. With the widely debated concept of the "Global Savings Glut", it would be intriguing to theoretically capture how in the presence of patient countries (like China) and impatient ones (like US) global interest rates are affected and how international organizations could intervene to achieve an overall welfare optimization. As a starting point, Leff and Sato (1993) document the existence of international differences in savings behavior across countries.

Analyzing the same problem in the presence of other market imperfections, say for example price rigidity, and checking if our results still hold would make another interesting study. Monacelli and Perotti (2011) for example, show that in the presence of sticky prices, government spending multiplier is larger when the taxes are levied on patient households instead of favouring a transfer from impatient to patient households. But does the result hold with respect to capital income taxation as well? These and many other such issues remain the subject for related future work.

References

1. Aiyagari, S. R. (1995). Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting. Journal of Political Economy, Vol. 103, No. 6, pp. 1158-1175.

2. Alpanda, S., & Zubairy, S. (2016). Housing and Tax Policy. Journal of Money, Credit and Banking, 48(2-3).

3. Becker, R. A. (2011). Optimal Growth with Heterogeneous Agents and the Twisted Turnpike: An Example . Center for Applied Economics and Policy Research Working Paper No. 2011-008.

4. Becker, R. A., & Foias, C. (1987). A Characterization of Ramsey Equilibrium. Journal Of Economic Theory, 173-184.

5. Bilbiie, F. O., Monacelli, T., & Perotti, R. (2013). Public Debt and Redistribution with Borrowing Constraints. The Economic Journal, Volume 123, Issue 566, pages F64–F98.

6. Campbell, J., & Hercowitz, Z. (2005). The Role of Collateralized Household Debt in Macroeconomic Stabilization. NBER w.p 11330.

7. Chamley, C. (1986). Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives. Econometrica, 54 (3), pp. 607–622.

8. Chamley, C. (2001). Capital income taxation, wealth distribution and borrowing constraints. Journal of Public Economics, 55–69.

9. Eggertsson, G. B., & Krugman, P. (2012). Debt, Deleveraging, and the Liquidity Trap: A Fisher-Minsky-Koo Approach. The Quarterly Journal of Economics, Oxford University Press, vol. 127(3), pages 1469-1513.

10. Goenka, A., Cuong, L. V., & Manh-Hung, N. (2012). Existence of competitive equilibrium in an optimal growth model with heterogeneous agents and endogenous leisure. Macroeconomic Dynamics, 16, pp 33-51.

11. Gollier, C. (2002). Discounting an uncertain future. Journal of Public Economics, 149–166.

12. Iacoviello, M., & Neri, S. (2010). Housing Market Spillovers: Evidence from an Estimated DSGE Mode. American Economic Journal: Macroeconomics, pp. 125–164.

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

14. Kemp, M. C., Long, N., & Shimomura, K. (1993). Cyclical and Noncyclical Redistributive Taxation. International Economic Review, 34 (2), 415–29.

15. Kirman, A. P. (1992). Whom or What Does the Representative Individual Represent? Journal of Economic Perspectives, 117-136.

16. Kiyotaki, N., & Moore, J. (1997). Credit Cycles. Journal of Political Economy, 105(2), 211–248.

17. Krusell, P., & Smith, A. A. (1998). Income and Wealth Heterogeneity in the Macroeconomy. Journal of Political Economy, University of Chicago Press, vol. 106(5), pages 867-896.

18. Le Van, C., Nguyen, M. H., & Vailakis, Y. (2007). Equilibrium dynamics in an aggregative model of capical accumulation with heterogeneous agents and elastic labor. Journal of Mathematical Economics, 43, pp. 287-317.

19. Le Van, C., & Vailakis, Y. (2003). Existence of a competitive equilibrium in one sector growth model with heterogenous agents and irreversible investment. Economic Theory, vol. 22(4), pages 743-771.

20. Leff, N., & Sato, K. (1993). Homogeneous preferences and heterogeneous growth performance: International differences in saving and investment behavior. Kyklos, 46, 2, 203-223.

21. Lengwiler, Y. (2005). Heterogeneous Patience and the Term Structure of Real Interest Rates. The American Economic Review, Vol. 95, No. 3, pp. 890-896.

22. McKay, A., & Reis, R. (2013). The Role of Automatic Stabilizers in the U.S. Business Cycle. NBER Working Papers 19000, National Bureau of Economic Research, Inc.

23. Mitra, T. (1979). On Optimal Economic Growth with Variable Discount Rates: Existence and Stability Results. International Economic Review, Vol. 20, No. 1, pp. 133-145.

24. Monacelli, T., & Perotti, R. (2011). Redistribution and the Multiplier. Working Papers 409, IGIER (Innocenzo Gasparini Institute for Economic Research), Bocconi University.

25. Quadrini, V. (2011). Financial Frictions in Macroeconomic Fluctuations. Federal Reserve Bank of Richmond Economic Quarterly, 209–254.

26. Ramsey, F. P. (1928). A Mathematical Theory of Saving. The Economic Journal, Volume 38, Issue 152, 543-559.

27. Straub, L., & Werning, I. (2014). Positive Long Run Capital Taxation: Chamley-Judd Revisited. NBER Working Papers 20441, National Bureau of Economic Research.

28. Weitzman, M. L. (1998). Why the Far-Distant Future Should Be Discounted at Its Lowest Possible Rate. Journal od Environmental Economics and Management, 36, 201]208.

Appendix

A Ramsey planner's Problem

For any $\left(k_0, d_0, c_{P,-1}, c_{I,-1}, \mu_{I,-1}\right)$ we define the value function $V\left(k_0, d_0, c_{P,-1}, c_{I,-1}, \mu_{I,-1}\right)$ as

$$V\left(k_0, d_0, c_{P,-1}, c_{I,-1}, \mu_{I,-1}\right) = \max \sum_{t=0}^{\infty} \beta_P^t \left[\omega u\left(c_{P,t}\right) + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t u\left(c_{I,t}\right)\right] \text{ s.t.}$$

$$\beta_P c_{P,t}^{1-\sigma} + \beta_P c_{P,t}^{-\sigma}\left(d_{t+1} + k_{t+1}\right) = c_{P,t-1}^{-\sigma}\left(d_t + k_t\right)$$

$$\left(1 - \mu_{I,t-1}\right)\beta_P c_{I,t-1}^{-\sigma} c_{P,t}^{-\sigma} = \beta_I c_{I,t}^{-\sigma} c_{P,t-1}^{-\sigma}$$

$$F\left(k_t\right) + (1-\delta) k_t = c_{P,t} + c_{I,t} + g + k_{t+1}$$

Using Envelope condition:

$$V_{c_{P,-1}}\left(k_0, d_0, c_{P,-1}, c_{I,-1}, \mu_{I,-1}\right) = 0 = \frac{\partial \mathcal{L}}{\partial c_{P,-1}}$$

$$\Rightarrow \sigma \lambda_{I,0} c_{P,-1}^{-\sigma-1}\left(d_0 + k_0\right) + \sigma \lambda_{E,0} c_{P,-1}^{-\sigma-1} \beta_I c_{I,0}^{-\sigma} = 0 \tag{22}$$

Similarly,

$$V_{\mu_{I,-1}}\left(k_0, d_0, c_{P,-1}, c_{I,-1}, \mu_{I,-1}\right) = 0 = \frac{\partial \mathcal{L}}{\partial \mu_{I,-1}}$$

$$\Rightarrow \quad -\lambda_{E,0} \beta_P c_{I,-1}^{-\sigma} c_{P,0}^{-\sigma} = 0$$

$$\Rightarrow \quad \lambda_{E,0} = 0 \tag{23}$$

Substituting 23 in 22, we get $\lambda_{I,0} = \lambda_{E,0} = 0$. Further, we solve the Ramsey planner's welfare problem.

$$\frac{\partial \mathcal{L}}{\partial \mu_{I,t}} : \lambda_{E,t+1}\beta_P^{t+1}\left[\omega + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^{t+1}\right] c_{I,t}^{-\sigma} c_{P,t+1}^{-\sigma} = 0$$

With $\lambda_{E,0} = 0$ and the above FOC, $0 = \lambda_{E,1} = \lambda_{E,2} = \lambda_{E,3} = \lambda_{E,4} = ...$ This reflects the fact that varying $\mu_{I,t}$ imposes no opportunity costs on the government. Given the Ramsey planner's optimal choice for $c_{P,t}$, $c_{I,t}$ and k_t, the value of $\mu_{I,t}$ will adjust accordingly to achieve the desired levels. The problem thus reduces to

$$\begin{aligned}
\mathcal{L} &= E_0 \sum_{t=0}^{\infty} \beta_P^t \left[\omega u\left(c_{P,t}\right) + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t u\left(c_{I,t}\right)\right] \\
&\quad + \lambda_{I,t}\beta_P^t \left[\omega + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t\right]\left[\beta_P c_{P,t}^{1-\sigma} + \beta_P c_{P,t}^{-\sigma}\left(d_{t+1} + k_{t+1}\right) - c_{P,t-1}^{-\sigma}\left(d_t + k_t\right)\right] \\
&\quad + \lambda_{F,t}\beta_P^t \left[\omega + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t\right]\left[f\left(k_t\right) - c_{P,t} - c_{I,t} - g - k_{t+1} + (1-\delta) k_t\right] \\
&\quad + \lambda_{B,t} c_{P,t}^{-\sigma}\beta_P^t \left[\omega + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t\right]\left[\bar{D} - d_{t+1}\right]
\end{aligned}$$

B Derivation of 20

Maximizing with respect to capital:

$$\begin{aligned}
&\lambda_{F,t+1}\beta_P \left[\omega + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^{t+1}\right]\left(f'\left(k_{t+1}\right) + 1 - \delta\right) \\
= \;& \lambda_{F,t}\left[\omega + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^t\right] + c_{P,t}^{-\sigma}\left\{\begin{array}{c} \beta_P\left[\omega + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^{t+1}\right]\lambda_{I,t+1} \\ -\beta_P\left[\omega + (1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^{t}\right]\lambda_{I,t} \end{array}\right\}
\end{aligned}$$

From 18 and 19, this could be written as

$$\frac{\lambda_{F,t+1}}{\lambda_{F,t}}\beta_P\left[\omega+(1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^{t+1}\right]\left(f'\left(k_{t+1}\right)+1-\delta\right)$$
$$=\left[\omega+(1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^{t}\right]-\frac{\lambda_{B,t}c_{P,t}^{-\sigma}}{(1-\omega)\,c_{I,t}^{-\sigma}}\left(\frac{\beta_P}{\beta_I}\right)^{t}\left[\omega+(1-\omega)\left(\frac{\beta_I}{\beta_P}\right)^{t}\right]^{2}$$

which gives us 20.

C Kuhn-Tucker Conditions

As per the Kuhn- Tucker theorem, if the objective function and inequality constraint are at least once continuously differentiable and are concave on an open convex domain, there exists a solution to the maximization problem iff there exists $\lambda_{B,t}$ such that the following Kuhn-Tucker conditions are satisfied:

1. $\frac{\partial \mathcal{L}}{\partial x}=0$
2. $\lambda_{B,t}\geqslant 0$
3. Complementary Slackness Condition: $\lambda_{B,t}\left(\bar{D}-d_{t+1}\right)=0$
4. $\bar{D}-d_{t+1}\geqslant 0$

In our model, the Hessian matrix of the objective function, $\omega\beta_P^t u\left(c_{P,t}\right)+(1-\omega)\,\beta_I^t u\left(c_{I,t}\right)$, is given as follows:

$$H=\begin{bmatrix}-\sigma\beta_P^t\omega c_{P,t}^{-1-\sigma} & 0\\ 0 & -\sigma\beta_I^t\left(1-\omega\right)c_{P,t}^{-1-\sigma}\end{bmatrix}$$

Since $\det(H)>0$, the Hessian matrix is negative definite and is thus a strictly concave function. Moreover, the inequality constraint is linear and is concave as well. A solution to the problem will thus exist if the above mentioned Kuhn-Tucker conditions are satisfied.

In Case 1, condition number two is violated and hence, a welfare maximizing Ramsey allocation does not exist.

Further, for the complementary sackness condition to hold, either $\lambda_{B,t} = 0$ or $\bar{D} - d_{t+1} = 0$. When $\lambda_{B,t} = 0$, $\bar{D} - d_{t+1}$ might or might not be equal to zero and we say that the constraint is not binding (similar to our result from Case 2). However, if $\lambda_{B,t} > 0$, $\bar{D} - d_{t+1} = 0$ implying that the constraint is always binding (similar to our result from Case 3).