

# Risk aversion, intertemporal substitution and optimal fiscal policy over the business cycle

Preliminary and incomplete

Anastasios G. Karantounias \*

March 16, 2011

## Abstract

When the intertemporal elasticity of substitution is disentangled from risk aversion as in the recursive preferences of Epstein and Zin (1989) and Weil (1990) (EZW henceforth), news about future consumption and leisure affect the equilibrium value of the government's portfolio of securities and therefore the extent to which the government has to resort to distortionary taxation in order to finance government expenditures. This paper studies the implications of this channel for the optimal capital and labor income taxation over the business cycle and the implied asset positions of the government. In contrast to the case of time-additive expected utility, the excess burden of taxation becomes time-varying, causing *taxation* of the intertemporal margin and *variation* of taxation at the intratemporal margin. Optimal policy becomes a mechanism of endogenous persistence.

JEL CLASSIFICATION: D80; E62; H21; H63.

KEY WORDS: Ramsey plan, Epstein-Zin, recursive utility, labor tax, capital tax, martingale.

---

\*Karantounias : Federal Reserve Bank of Atlanta (e-mail: anastasios.karantounias@atl.frb.org). I am thankful to Demian Pouzo and Vasco Carvalho and to seminar participants at the University of California at Davis, Pompeu Fabra, the European University Institute and LUISS Guido Carli University. The views expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System.

# 1 Introduction

This paper analyzes optimal fiscal policy in a business cycle model with the recursive preferences of Epstein and Zin (1989) and Weil (1990) (EZW henceforth). I consider the stochastic growth model with technology and government expenditure shocks and complete markets. Lump-sum taxes are not available. Instead, there are proportional taxes on capital and labor income. Distortionary taxes give rise to an optimal policy problem, in which a benevolent planner is choosing taxes and debt in order to maximize the utility of the representative household. My analysis is normative – no attempts to explain actual time series of fiscal variables will be made.

The questions are classic: how should the policy maker use taxes and debt during booms and recessions and in response to high and low fiscal shocks so that welfare is maximized? Should there be taxation of capital or labor income? How should the return on government debt vary? The formulation of these classic questions makes it evident that the policy problem entails a dynamic decision that is intertwined with a decision under risk. These two dimensions are encoded in *asset prices*, which are central to the policy maker since they inform him about the desirability of debt and therefore the extent to which he has to resort to distortionary taxation. In particular, optimal policy prescriptions revolve around the theme of fiscal *hedging*, i.e. the usage of state-contingent securities as an insurance device against shocks in the government budget constraint, prescribing a reduction in returns on government debt in the case of high fiscal or low technology shocks and an increase in returns on government debt in the case of low fiscal or high technology shocks.<sup>1</sup>

As is well known though from Epstein and Zin (1989) and Weil (1990), standard expected utility, by forcing the coefficient of risk aversion to equate the inverse of intertemporal elasticity of substitution (IES), fails to make a distinction between the dynamic dimension and the risk dimension and imposes indifference to the temporal resolution of uncertainty. Besides being theoretically unappealing, this feature may be a drawback also on the grounds of keeping the market price of risk artificially low, giving rise to the various empirical asset pricing puzzles and creating concern about the proper use of asset prices in minimizing tax distortions. As a response, there is by now a voluminous theoretical and empirical literature in asset pricing that tries to increase the market price of risk by making a distinction between these two concepts, and a growing macro-finance literature, that tries to retain the basic structure of the real business cycle model and augment it with richer pricing kernels in order to match well both quantities and asset prices.<sup>2</sup> What I am interested in are the

---

<sup>1</sup>See Berndt et al. (2010) for an empirical investigation of fiscal hedging in U.S. data.

<sup>2</sup>See among others Hansen et al. (2006), Bansal and Yaron (2004) and Thomas D. Tallarini (2000) and

implications of this new model of the pricing kernel for optimal allocation of tax distortions over the business cycle. Therefore, I am going to deviate from the stochastic growth model with distortionary taxes only in the preferences dimension. The model I build would reduce to the standard business cycle model that Chari et al. (1994) analyzed, if I equate the risk aversion parameter with the inverse of the IES parameter and to Chamley (1986) and Judd (1985), if I shut off uncertainty.

The crucial element that enters the picture with EZW preferences is the temporal composition of risk. In particular, the stochastic discount factor depends on a *short run* component of consumption growth and leisure, and on a novel *long run* component of continuation utility, i.e. the present value of future streams of consumption and leisure. Innovations in continuation utility or else *news* about the present value of future consumption and leisure affect therefore asset prices. When there is preference for early resolution of uncertainty,<sup>3</sup> good news about future utility reduce the stochastic discount factor and therefore increase the return that the investor requires to hold an Arrow claim to consumption at that contingency. As a result, the policymaker has to take into account how his tax-debt plan is going to affect the future profile of consumption and leisure since this is going to affect the market value of government debt.

Before I proceed to my results it is worth summarizing the main results about optimal taxation over the business cycle coming from the work of Chari et al. (1994) with standard expected utility: essentially *constant* labor taxes and essentially *zero* ex-ante taxation of capital income. A further result is that labor taxes inherit the stochastic properties of exogenous shocks, thus optimal tax rates cannot constitute a distinct source of persistence in the economy. The driving force behind these results is the cost of distortionary taxation, a multiplier that captures the welfare loss stemming from the absence of lump-sum taxes. The optimal policy prescription is to make this cost *constant* over dates and states, which agrees with the conventional intuition of “tax smoothing”.

In contrast to this prescription, I show that the continuation value channel in asset prices leads to a *time-varying* cost of distortionary taxation. This leads to *variation* in taxation of the intratemporal margin and to *taxation* of the intertemporal margin. I show explicitly the dependence of the optimal labor tax on the curvature properties of the period utility of consumption and leisure, which captures the short-run component of the pricing kernel, and on the time-varying cost of distortionary taxation, which captures the long-run component of the pricing kernel. For example, when there is indifference to the temporal resolution of

---

the respective surveys by Campbell and Cochrane.

<sup>3</sup>The typical calibration of EZW preferences features a coefficient of risk aversion that is larger than the inverse of IES, implying a preference for early resolution of uncertainty.

uncertainty, for standard period utility aggregators of consumption and leisure, the labor tax varies only as long as the Frisch elasticity of labor supply varies, taxing higher events with low Frisch elasticity and lower events with high Frisch elasticity. In my model though, even if the Frisch elasticity is constant, the labor tax varies due to variation in the cost of distortionary taxation, leading to a higher labor tax when the cost of taxation increases and a lower when it decreases.

From the work of Chari et al. (1994) and Zhu (1992) we know that there is a multiplicity of capital taxes and debt that can implement the optimal allocation and therefore only the ex-ante tax rate is uniquely determined. With standard expected utility, I show explicitly the dependence of the ex-ante taxation of capital income on the elasticity of marginal period utility of consumption with respect to consumption and labor. Variation in the sum of these two elasticities is a necessary condition for taxation of capital income.<sup>4</sup> With EZW preferences though, there can be ex-ante taxation of capital income even in the case of a *constant* sum of elasticities, due to the variation in the cost of distortionary taxation.

The reason for the time-varying cost of distortionary taxation will become clear if we think of the dynamic tradeoffs that the planner is facing. Facing a fiscal shock, the planner has to decide whether to tax (labor or capital income) or to postpone taxation to the future by issuing one period state-contingent debt. The crucial variable in his problem is the market value of his debt or asset portfolio. An increased amount of debt (or reduced amount of assets) for a particular contingency next period entails on the one hand a marginal *cost*, since increased debt in the future will have to be paid back with future distortionary taxes and a marginal *benefit*, since debt relaxes the current financing constraint of the government and allows less taxes today. These are the only two tradeoffs with time-additive expected utility. Equating the marginal benefit with the marginal cost leads to a *constant* marginal cost of distortionary taxation, which consists the source of the tax-smoothing results in the literature. With EZW preferences though, the planner has to consider the effect of increased debt on continuation utility and therefore on the market value of his portfolio. Increased debt leads to a negative innovation in utility which –when there is early preference for resolution of uncertainty– increases asset prices and therefore the value of the government’s portfolio. This is beneficial for the government if it borrows on average from the private sector, since by selling claims to consumption next period at a high price allows a lower tax rate in the current period, and harmful if the government lends on average to the private sector, since the government buys claims to consumption at a high price, forcing it to tax higher the current period.

---

<sup>4</sup>This condition would be sufficient in the deterministic model of Chamley (1986) and Judd (1985). Therefore, the zero-tax result emerges at the steady state, since elasticities would be constant.

As a result, with a preference for early resolution of uncertainty, the cost of distortionary taxation *increases* when there is a positive innovation in the household's wealth (government securities and capital), leading to an incentive to increase the labor tax rate and *decreases* when there is a negative innovation in the household's wealth, leading an incentive to decrease the labor tax. Thus, high fiscal shocks or low technological shocks, that are insured against with assets command a lower tax, whereas low fiscal shocks or high technology shocks, for which the planner is ex-ante issuing debt, lead to a higher tax. As a result, the tax rate on labor income becomes *countercyclical*.

Furthermore, I show that the cost of distortionary taxation exhibits a martingale-like behavior, inducing persistence to the policy variables and to the optimal allocation. This is in contrast to the results of Lucas and Stokey (1983) and Chari et al. (1994), who show –theoretically in the case of Lucas and Stokey (1983) in an economy without capital and quantitatively in the capital case of Chari et al. (1994)– that the labor tax rate inherits the stochastic properties of the exogenous shocks. With EZW preferences though, optimal tax rates become persistent in order to properly affect the forward-looking continuation utilities and therefore asset prices. The persistence in the Ramsey allocation is captured by an additional state variable, which is the household's wealth in period marginal utility units. I show that the planner needs to keep track of this state variable in order to be able to determine the cost of distortionary taxation. When there is no distinction between risk aversion and IES, the cost of distortionary taxation is constant, which makes tracking wealth for the optimal policy problem *redundant*. Therefore, government debt, by affecting the household's wealth, is assuming a novel role besides the traditional fiscal absorber one.

The quantitative importance of the EZW pricing kernel for the optimal taxation plan is pending. The sensitivity of the response of the cost of distortionary taxation towards innovations in the household's wealth depends on the size of the difference between the risk aversion coefficient and the inverse of IES. However, the theoretical martingale-like results for the cost of taxation and the labor tax suggest that even if the conditional volatility of the tax rate is small due to a small difference between the two relevant preference parameters, the unconditional volatility of the labor tax will be high due to the optimal persistence that the planner induces.

...

## 1.1 Related literature

[To be written]

## 2 Economy

Time is discrete and the horizon is infinite. As Chari et al. (1994) and Zhu (1992), I use the canonical stochastic growth model with distortionary labor and capital income taxation, with the exception of the recursive preferences. The economy is populated by a representative household that consumes, works, accumulates capital and trades in complete asset markets. There is uncertainty in the economy stemming from exogenous technology and government expenditure shocks and it is captured by the variable  $s$ . Let  $s^t = (s_0, s_1, \dots, s_t)$  denote the partial history of shocks till time  $t$ . For simplicity, I am going to assume that the exogenous shocks take finite or countable values. Let  $\pi_t(s^t)$  denote the probability of the partial history  $s^t$  and let  $s_0$  be given, so that  $\pi_0(s_0) = 1$ .

### 2.1 Preferences

The representative household derives utility from random sequences of consumption  $\{c\} \equiv \{c_t(s^t)\}_{t \geq 0, s^t}$  and leisure  $\{l\} \equiv \{l_t(s^t)\}_{t \geq 0, s^t}$ . The notation makes explicit that consumption and leisure at time  $t$  are measurable functions of history  $s^t$  (or adapted to information at time  $t$ ). There is one unit of time to allocate between labor and leisure, thus labor is  $h_t(s^t) = 1 - l_t(s^t)$ . The household ranks consumption and leisure plans following a recursive utility criterion of Kreps and Porteus (1978). In particular, let  $V_t$  denote the household's utility at time  $t$  (as adapted to information at  $t$ ).  $V_t$  follows the recursion

$$V_t = W(u(c_t, 1 - h_t), \mu_t(V_{t+1})). \quad (1)$$

The household derives utility from a composite good that consists of consumption and leisure,  $u(c_t, 1 - h_t)$ , and from the certainty equivalent of continuation utility  $\mu_t \equiv \phi^{-1}(E_t \phi(V_{t+1}))$ , where  $E_t$  denotes the conditional expectation operator given information at  $t$  with respect to the measure  $\pi_t$ , and  $\phi(\cdot)$  is a concave, increasing function that is capturing atemporal risk aversion. The time preference of the household between the composite good today and the certainty equivalent of continuation utility is captured by the time aggregator  $W$ .

I will focus my analysis on the case of isoelastic preferences of Epstein and Zin (1989) and Weil (1990) (EZW henceforth), and use a CES time aggregator and a power utility certainty equivalent. In particular, the EZW preferences take the form

$$V_t = [(1 - \beta)u(c_t, 1 - h_t)^{1-\rho} + \beta(E_t V_{t+1}^{1-\gamma})^{\frac{1-\rho}{1-\gamma}}]^{\frac{1}{1-\rho}}. \quad (2)$$

The parameter  $1/\rho$  captures the intertemporal elasticity of substitution, whereas the

parameter  $\gamma$  represents risk aversion in the calculation of the certainty equivalent. These preferences reduce to standard time-additive expected utility when  $\rho = \gamma$ . Furthermore, they take into account the temporal resolution of uncertainty and can exhibit preference for early ( $\rho < \gamma$ ) or late ( $\rho > \gamma$ ) resolution of uncertainty, whereas with expected utility there is indifference about the temporal resolution of uncertainty ( $\rho = \gamma$ ).

It is useful for later purposes to bear in mind the monotonic transformation  $v_t \equiv \frac{V_t^{1-\rho}-1}{(1-\beta)(1-\rho)}$ , which will be called the  $\rho$ -transformation.<sup>5</sup> The utility recursion (2) becomes in this case

$$v_t = U_t + \beta \frac{\left[ E_t [1 + (1-\beta)(1-\rho)v_{t+1}]^{\frac{1-\gamma}{1-\rho}} \right]^{\frac{1-\rho}{1-\gamma}} - 1}{(1-\beta)(1-\rho)}, \quad (3)$$

where  $U(c_t, 1-h_t) \equiv \frac{u_t^{1-\rho}-1}{1-\rho}$ , with respective derivatives  $U_i = u^{-\rho}u_i, i = c, l$ . For the rest of the paper, I will refer to  $U$  and  $U_i, i = c, l$  as *period* utility and *period* marginal utility respectively.

Of particular interest is the case when the intertemporal elasticity of substitution becomes unity,  $\rho = 1$ . Then (2) becomes  $V_t = u_t^{1-\beta} \mu_t^\beta$ , and applying the  $\rho$ -transformation for  $\rho = 1$ ,  $v_t = \frac{\ln V_t}{1-\beta}$ , we get the risk-sensitive recursion<sup>6 7</sup>

$$v_t = \ln u_t + \frac{\beta}{(1-\beta)(1-\gamma)} \ln E_t \exp [(1-\beta)(1-\gamma)v_{t+1}]. \quad (5)$$

It will be useful to define

$$m_{t+1} \equiv \frac{V_{t+1}^{1-\gamma}}{E_t V_{t+1}^{1-\gamma}}, t \geq 0, \quad (6)$$

---

<sup>5</sup>Applying the respective  $\gamma$ -transformation  $f(V) \equiv \frac{V^{1-\gamma}-1}{(1-\beta)(1-\gamma)}$  on (2) delivers the representation used in Weil (1990).

<sup>6</sup>The risk-sensitive recursion in (5) has an intimate link with the multiplier preferences of Hansen and Sargent (2001), that capture the decision maker's fear of misspecification of the probability model  $\pi_t$  that are described by the recursion

$$v_t^{\text{HS}} = U_t + \beta \min_{m_{t+1} \geq 0, E_t m_{t+1} = 1} [E_t m_{t+1} v_{t+1}^{\text{HS}} + \theta E_t m_{t+1} \ln m_{t+1}], \quad (4)$$

for a positive penalty parameter  $\theta > 0$  and some period return function  $U$ . The multiplier preferences (4) reduce to expected utility for  $\theta = \infty$ . Assume that  $\gamma > 1$ , and let the period utility in (4) be logarithmic. Then the risk-sensitive recursion and the multiplier preferences recursion are equivalent if we set  $\theta^{-1} \equiv -(1-\beta)(1-\gamma)$ .

<sup>7</sup>For a decision-theoretic axiomatization of the multiplier preferences see Strzalecki (2010a). For a deeper analysis of the relationship between the temporal resolution of uncertainty and ambiguity aversion see Strzalecki (2010b).

with  $m_0 \equiv 1$ . Note that  $m_{t+1}$  is positive since  $V_{t+1}$  is positive and that  $E_t m_{t+1} = 1$ . So  $m_{t+1}$  can be interpreted as a change of measure of the conditional probability density  $\pi_{t+1}(s_{t+1}|s^t)$ , or, in other words, a conditional likelihood ratio. Similarly, define the product of the conditional likelihood ratios as

$$M_t(s^t) \equiv \prod_{i=1}^t m_i(s^i), \quad (7)$$

with the normalization  $M_0 \equiv 1$ . This object has the interpretation of an unconditional likelihood ratio and is a martingale with respect to measure  $\pi$ .

Let the technology in the economy be captured by a constant returns to scale production function  $F$ , and let  $k_{t+1}(s^t)$  denote capital for the beginning of period  $t + 1$ , as function of information at  $t$ . The resource constraint in the economy takes the form

$$c_t(s^t) + k_{t+1}(s^t) - (1 - \delta)k_t(s^{t-1}) + g_t(s^t) = F(s_t, k_t(s^{t-1}), h_t(s^t)). \quad (8)$$

**Household's problem.** The representative household consumes  $c_t(s^t)$ , works  $h_t(s^t)$  at wage rate  $w_t(s^t)$ , accumulates capital  $k_{t+1}(s^t)$  that depreciates at the rate  $\delta$  and can be rented at rental rate  $r_t(s^t)$ , pays proportional labor income taxes with rate  $\tau_t^H(s^t)$  and capital income taxes with rate  $\tau_t^K(s^t)$  and trades in complete markets. The household's problem is  
Choose  $\{c_t(s^t), h_t(s^t), k_{t+1}(s^t), b_{t+1}(s^{t+1})\}_{t \geq 0, s^t}$  to maximize

$$V_0(\{c\}, \{h\})$$

subject to

$$c_t(s^t) + k_{t+1}(s^t) + \sum_{s_{t+1}} p_t(s_{t+1}, s^t) b_{t+1}(s^{t+1}) \leq (1 - \tau_t^H(s^t)) w_t(s^t) h_t(s^t) + [(1 - \tau_t^K(s^t)) r_t(s^t) + (1 - \delta)] k_t(s^{t-1}) + b_t(s^t),$$

the non-negativity constraints for consumption and capital  $c_t(s^t), k_{t+1}(s^t) \geq 0$  and the feasibility constraint for labor  $h_t(s^t) \in [0, 1]$ , where  $k_0$  and  $b_0$  given. The term  $b_{t+1}(s^{t+1})$  stands for the holdings at history  $s^t$  of an Arrow security that delivers next period one unit of consumption if the state is  $s_{t+1}$  and zero units otherwise, and trades at the price  $p_t(s_{t+1}, s^t)$  in terms of the history-contingent consumption  $c_t(s^t)$ .

The household is also facing a no-Ponzi game condition that takes the form

$$\lim_{t \rightarrow \infty} \sum_{s^t} q_t(s^t) [k_{t+1}(s^t) + \sum_{s^{t+1}} p_t(s_{t+1}, s^t) b_{t+1}(s^{t+1})] \geq 0, \quad (9)$$

where  $q_t(s^t) \equiv \prod_{i=0}^{t-1} p_i(s_{i+1}, s^i)$ ,  $q_0 \equiv 1$ . In other words,  $q_t$  stands for the price of an Arrow-Debreu contract at  $t = 0$ .

Define the *after-tax* gross return on capital

$$R_{t+1}^K(s^{t+1}) \equiv (1 - \tau_{t+1}^K(s^{t+1}))r_{t+1}(s^{t+1}) + 1 - \delta. \quad (10)$$

By a no-arbitrage argument we can show that the price of this return has to be unity

$$\sum_{s_{t+1}} p_t(s_{t+1}, s^t) R_{t+1}^K(s^{t+1}) = 1. \quad (11)$$

This relationship will be also derived as an optimality condition of the household.

The no-arbitrage condition (11) together with the no-Ponzi game condition (9) allows us to derive the intertemporal budget constraint of the household at  $t = 0$

$$\sum_{t=0}^{\infty} \sum_{s^t} q_t(s^t) [c_t(s^t) - (1 - \tau_t^H(s^t))w_t(s^t)h_t(s^t)] \leq R_0^K k_0 + b_0. \quad (12)$$

**Firms.** A price-taking firm operates the constant returns to scale technology,  $y_t = F(s_t, K_t, H_t)$ . The firms rents capital and labor services and maximizes profits. Factor markets are competitive and therefore profit maximization leads to  $w_t = F_H(s^t)$  and  $r_t = F_K(s^t)$ .

**Government.** The government taxes labor and capital income and issues state-contingent debt in order to finance the exogenous government expenditures. The dynamic government constraint takes the form

$$b_t(s^t) + g_t(s^t) \leq T_t(s^t) + \sum_{s^{t+1}} p_t(s_{t+1}, s^t) b_{t+1}(s^{t+1}), \quad (13)$$

where

$$T_t(s^t) \equiv \tau_t^H(s^t)w_t(s^t)h_t(s^t) + \tau_t^K(s^t)r_t(s^t)k_t(s^{t-1}),$$

the total tax revenues of the government. When  $b_t > 0$ , the government borrows from the household and when  $b_t < 0$ , the government lends to the household. The respective no-Ponzi game condition for the government is

$$\lim_{t \rightarrow \infty} \sum_{s^{t+1}} q_{t+1}(s^{t+1}) b_{t+1}(s^{t+1}) \leq 0. \quad (14)$$

This condition has the interpretation that the government should not have debt at infinity.

**Definition 1.** *A competitive equilibrium with taxes is a stochastic process for prices  $\{p, w, r\}$ , an allocation  $\{c, h, k, b\}$  and a government policy  $\{g, \tau^H, \tau^K, b\}$  such that:*

- *Given prices  $\{p, w, r\}$  and taxes  $\{\tau^H, \tau^K\}$ , the allocation  $\{c, h, k, b\}$  solves the households's problem.*
- *Given  $\{w, r\}$ , firms maximize profits and therefore  $w_t(s^t) = F_H(s^t)$ ,  $r_t(s^t) = F_K(s^t)$ .*
- *Markets clear*

$$c_t(s^t) + k_{t+1}(s^t) - (1 - \delta)k_t(s^{t-1}) + g_t(s^t) = F(s_t, k_t(s^{t-1}), h_t(s^t)).$$

- *The government budget constraint holds.*

$$b_t(s^t) = T_t(s^t) - g_t(s^t) + \sum_{s^{t+1}} p_t(s_{t+1}, s^t) b_{t+1}(s^{t+1}).$$

## 2.2 Competitive equilibrium characterization

I am now in the position to characterize the competitive equilibrium. It is instructive to depict first the household's optimality conditions for a general utility index  $V_0$  and then proceed to the EZW case. The first-order conditions for an interior solution with respect to Arrow securities and capital are <sup>8</sup>

$$p_t(s_{t+1}, s^t) = \frac{\partial V_0 / \partial c_{t+1}(s^{t+1})}{\partial V_0 / \partial c_t(s^t)}$$

and

---

<sup>8</sup>These two conditions deliver the no-arbitrage condition (11).

$$\frac{\partial V_0}{\partial c_t(s^t)} = \sum_{s^{t+1}} \frac{\partial V_0}{\partial c_{t+1}(s^{t+1})} R_{t+1}^K(s^{t+1}).$$

The labor supply decision is governed by

$$\frac{-\partial V_0 / \partial h_t(s^t)}{\partial V_0 / \partial c_t(s^t)} = (1 - \tau_t^H(s^t)) w_t(s^t),$$

which equates the marginal rate of substitution between consumption and leisure with the after tax wage. For the EZW preferences, the derivatives with respect to consumption and labor take the form <sup>9</sup>

$$\begin{aligned} \frac{\partial V_0}{\partial c_t(s^t)} &= (1 - \beta) V_0^\rho \beta^t \pi_t(s^t) M_t(s^t)^{\frac{\rho-\gamma}{1-\gamma}} U_c(s^t), \\ \frac{\partial V_0}{\partial h_t(s^t)} &= -(1 - \beta) V_0^\rho \beta^t \pi_t(s^t) M_t(s^t)^{\frac{\rho-\gamma}{1-\gamma}} U_l(s^t), \end{aligned}$$

where  $M_t$  defined in (7).

Thus the labor supply condition becomes

$$\frac{U_l(s^t)}{U_c(s^t)} = \frac{u_l(s^t)}{u_c(s^t)} = (1 - \tau_t^H(s^t)) w_t(s^t). \quad (15)$$

The equilibrium condition that equates the price of an Arrow security with the marginal rate of substitution of the household is therefore

$$p_t(s_{t+1}, s^t) = \beta \pi_{t+1}(s_{t+1} | s^t) \left( \frac{V_{t+1}(s^{t+1})}{\mu_t(V_{t+1})} \right)^{\rho-\gamma} \frac{U_c(s^{t+1})}{U_c(s^t)}, \quad (16)$$

or

$$p_t(s_{t+1}, s^t) = \beta \pi_{t+1}(s_{t+1} | s^t) m_{t+1}(s^{t+1})^{\frac{\rho-\gamma}{1-\gamma}} \frac{U_c(s^{t+1})}{U_c(s^t)}, \quad (17)$$

---

<sup>9</sup>The derivative with respect to  $c_t$  can be calculated recursively from the relationship

$$\frac{\partial V_t}{\partial c_{t+i}} = \frac{\partial V_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial V_{t+1}} \frac{\partial V_{t+1}}{\partial c_{t+i}}.$$

Similarly for  $h_t$ .

in terms of the conditional likelihood ratio (6). The change of measure  $M_t$  allows also a concise expression for the price of an Arrow-Debreu contract at  $t = 0$

$$q_t(s^t) = \beta^t \pi_t(s^t) M_t(s^t)^{\frac{\rho-\gamma}{1-\gamma}} \frac{U_c(s^t)}{U_c(s_0)}. \quad (18)$$

Furthermore, at the optimum the two asymptotic conditions (9) and (14) have to hold with equality, which lead to two transversality conditions with respect to capital and Arrow securities

$$\lim_{t \rightarrow \infty} \sum_{s^t} \beta^t \pi_t(s^t) M_t(s^t)^{\frac{\rho-\gamma}{1-\gamma}} U_c(s^t) k_{t+1}(s^t) = 0 \quad (19)$$

$$\lim_{t \rightarrow \infty} \sum_{s^{t+1}} \beta^{t+1} \pi_{t+1}(s^{t+1}) M_{t+1}(s^{t+1})^{\frac{\rho-\gamma}{1-\gamma}} U_c(s^{t+1}) b_{t+1}(s^{t+1}) = 0 \quad (20)$$

The stochastic discount factor  $S_{t+1}$  with EZW utility is

$$S_{t+1} \equiv \beta \left( \frac{V_{t+1}}{\mu_t} \right)^{\rho-\gamma} \frac{U_{c,t+1}}{U_{ct}} = \beta m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} \frac{U_{c,t+1}}{U_{ct}} \quad (21)$$

The disentanglement of risk aversion and intertemporal elasticity of substitution ( $\rho \neq \gamma$ ) introduces *continuation values* scaled by their certainty equivalent  $\mu_t$ . In log-linear setups this translates into the innovation in the log of continuation utility or else “news” about the present value of future consumption and leisure streams. The temporal composition of risk therefore matters for the stochastic discount factor and asset pricing and is the channel that allows to increase the market price of risk in Bansal and Yaron (2004) and Hansen et al. (2006).

**Add a more elaborate analysis of the “news” channel.**

**Remark 1.** *The presence of continuation values in the stochastic discount factor is changing non-trivially the implications for optimal policy, by changing the planner’s means of affecting equilibrium prices. The forward-looking nature of continuation values will contribute to richer dynamics of the Ramsey plan.*

### 3 Ramsey problem

I am going to solve the Ramsey problem under commitment. The problem of the planner is to choose at period zero the competitive equilibrium that maximizes the utility of the

representative household. I follow the primal approach of Lucas and Stokey (1983) and eliminate taxes and equilibrium prices from the competitive equilibrium conditions. As a result, the problem of the planner reduces to a problem of choosing allocations that satisfy the resource constraint (8) and *implementability* constraints, i.e. constraints that allow the optimal allocation to be implemented as a competitive equilibrium.

### 3.1 Implementability constraints

I am going to work with the household's *wealth*  $W_t$ , which comprises of the holdings of Arrow securities  $b_t$  and the capital wealth  $R_t^K k_t$ ,

$$W_t(s^t) \equiv b_t(s^t) + R_t^K(s^t)k_t(s^{t-1}). \quad (22)$$

Given the above definition and the no-arbitrage condition (11) we can recast the household's problem in terms of wealth. In particular, note that

$$\begin{aligned} \sum_{s_{t+1}} p_t(s_{t+1}, s^t) W_{t+1}(s^{t+1}) &= \sum_{s_{t+1}} p_t(s_{t+1}, s^t) [b_{t+1}(s^{t+1}) + R_{t+1}^K(s^{t+1})k_{t+1}(s^t)] \\ &= \sum_{s_{t+1}} p_t(s_{t+1}, s^t) b_{t+1}(s^{t+1}) + k_{t+1}(s^t). \end{aligned}$$

Therefore, the household's dynamic budget constraint becomes

$$c_t(s^t) + \sum_{s_{t+1}} p_t(s_{t+1}, s^t) W_{t+1}(s^{t+1}) = (1 - \tau_t^H(s^t))w_t(s^t)h_t(s^t) + W_t(s^t).$$

Using now (15) and (16) to eliminate labor taxes and equilibrium prices leads to

$$U_{ct}W_t = \Omega_t + \beta E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} U_{c,t+1}W_{t+1}, \quad (23)$$

where

$$\Omega_t \equiv U_{ct}[c_t - (1 - \tau_t^H)w_t h_t] = U_{ct}c_t - U_{lt}h_t. \quad (24)$$

So  $\Omega_t$  stands for consumption minus after tax labor income, in *period* marginal utility of consumption units. Note that  $\Omega_t$  is a function of consumption and labor only,  $\Omega_t = \Omega(c_t, h_t)$ .

We can summarize all this discussion in terms of a proposition

**Proposition 1.** *The Ramsey planner faces the following implementability constraints:*

$$\begin{aligned} U_{c_t}W_t &= \Omega_t + \beta E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} U_{c,t+1}W_{t+1}, t \geq 1 \\ U_{c_0}W_0 &= \Omega_0 + \beta E_0 m_1^{\frac{\rho-\gamma}{1-\gamma}} U_{c,1}W_1 \end{aligned}$$

where  $W_0 \equiv [(1 - \tau_0^K)F_K(s_0, k_0, h_0) + 1 - \delta]k_0 + b_0$ ,  $c_t, k_{t+1} \geq 0$ ,  $h_t \in [0, 1]$  and  $(k_0, b_0, \tau_0^K, s_0)$  given. Furthermore, the two transversality conditions (19) and (20) have to be satisfied. The conditional likelihood ratio  $m_{t+1}, t \geq 0$ , defined in (6), is endogenously determined by continuation values that follow the recursion (2).

**Remark 2.** *Complete markets allow the collapse of the household's dynamic budget constraint to a unique intertemporal budget constraint. However, maintaining the dynamic budget constraint of the household is more convenient for a recursive formulation. Also I am abstracting from the taxation of initial capital and take  $\tau_0^K$  as given.*

A consumption-labor-capital allocation  $(c, h, k)$  that satisfies the resource constraint (8) and the constraints of proposition 1 can be implemented as a competitive equilibrium by recovering prices  $\{p, w, r\}$ , tax rates  $\{\tau^H, \tau^K\}$  and government debt policies  $\{b\}$  from the household's optimality conditions and budget constraints. Note that, as Zhu (1992) and Chari et al. (1994) have shown, we can recover in a unique way only the labor tax  $\tau^H$ , whereas there is a *multiplicity* of capital tax and debt policies  $\{\tau^K, b\}$  that can implement the same allocation as a competitive equilibrium with prices  $\{p, w, r\}$ . The reason behind this result is that an implementable allocation  $\{c, h, k\}$  uniquely determines only the household's *wealth*  $W_t$ , which can be generated by a multiplicity of capital tax and debt policies through (22). However, it is well known that we can uniquely determine the *ex-ante* tax rate on capital income  $\bar{\tau}_{t+1}^K(s^t)$ , which is restricted to be function of history  $s^t$ ,<sup>10</sup>

$$\bar{\tau}_{t+1}^K(s^t) \equiv \frac{\sum_{s_{t+1}} p_t(s_{t+1}, s^t) (r_{t+1}(s^{t+1}) + 1 - \delta) - 1}{\sum_{s_{t+1}} p_t(s_{t+1}, s^t) r_{t+1}(s^{t+1})} \quad (25)$$

**Definition 2.** *The Ramsey problem is to maximize at  $t = 0$  the utility of the representative household subject to the implementability constraints of proposition 1 and the resource*

<sup>10</sup>The ex-ante tax rate is associated with the ex-post tax rate from the equation

$$\bar{\tau}_{t+1}^K(s^t) = \frac{\sum_{s_{t+1}} p_t(s_{t+1}, s^t) r_{t+1}(s^{t+1}) \tau_{t+1}^K(s^{t+1})}{\sum_{s_{t+1}} p_t(s_{t+1}, s^t) r_{t+1}(s^{t+1})}.$$

constraint (8).

### 3.2 Recursive formulation

I will break the Ramsey problem into two subproblems and form the problem recursively from period one onward. For that purpose, let  $z_t$  denote the household's wealth in period marginal utility terms,  $z_t \equiv U_{ct}W_t$ . Given this definition, rewrite the dynamic implementability constraint (23) as

$$z_t = \Omega_t + \beta E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} z_{t+1}, t \geq 0$$

The variable  $z_t$  will serve as a state variable in a recursive formulation of the Ramsey problem. Note though that from period one onward  $z_t$  is forward-looking variable that is not inherited from the past. This creates the necessity to determine the space in which it lives. Let

$$\begin{aligned} Z(s_1, k_1) \equiv & \left\{ z \mid \exists \{c_t, h_t\}_{t \geq 1}, \{k_{t+1}, z_{t+1}, V_{t+1}\}_{t \geq 1}, \text{ with } c_t, k_{t+1} \geq 0 \text{ and } h_t \in [0, 1] \right. \\ & \text{such that:} \\ & z_t = \Omega(c_t, h_t) + \beta E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} z_{t+1}, t \geq 1 \\ & c_t + k_{t+1} - (1 - \delta)k_t + g_t = F(s_t, k_t, h_t), t \geq 1 \\ & \text{where } m_{t+1} \text{ defined as in (6) with } V_{t+1}, t \geq 1 \text{ following recursion (2)} \\ & \left. \text{and } \lim_{t \rightarrow \infty} E_1 \beta^t \left( \frac{M_{t+1}}{M_1} \right)^{\frac{\rho-\gamma}{1-\gamma}} z_{t+1} = 0 \right\} \end{aligned} \quad (26)$$

$Z(s_1, k_1)$  denotes the set of values of the variable  $z_t$  at  $t = 1$ , that can be *generated* by a competitive equilibrium with taxes, when the initial shock is  $s_1$  and the capital inherited from period zero  $k_1$ .

I am going to represent the problem under commitment recursively by keeping track of the natural state variables  $(k_t, s_t)$  and the variable  $z_t$ , that captures the commitment of the planner to his past promises, so my state vector will be  $X_t \equiv (z_t, k_t, s_t)$ . Note that due to the forward-looking nature of  $z_t$ , the set  $Z(s_1, k_1)$  has to be determined. A (computationally intensive) way to proceed would be the recursive calculation of this set along the lines of Kydland and Prescott (1980).

In particular, let  $z_1 \in Z(s_1, k_1)$  and break the problem into two subproblems: the problem

starting from period one onward and the problem at the initial period  $t = 0$ . Let  $H(z_1, k_1, s_1)$  denote the value function of the subproblem from period one onward, i.e.

**Problem 1.** *The problem from period one onward is*

$$H(z_1, k_1, s_1) \equiv \max V_1(\{c, h\}_{t \geq 1})$$

subject to

$$\begin{aligned} z_t &= \Omega(c_t, h_t) + \beta E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} z_{t+1}, t \geq 1 \\ c_t + k_{t+1} - (1 - \delta)k_t + g_t &= F(s_t, k_t, h_t), t \geq 1, \end{aligned}$$

the non-negativity constraints for consumption, capital, labor, the TVC, where  $m_{t+1}$  denotes the conditional likelihood ratio that is generated by the utility recursion (2).

**Bellman equation.** Let the exogenous shocks follow a Markov process with transition probabilities  $\pi(s'|s)$ . The functional equation that determines the value function  $H(\cdot)$  takes the form

$$H(z, k, s) = \max_{c, h, k', z'_s} \left[ (1 - \beta)u(c, 1 - h)^{1-\rho} + \beta \left[ \sum_{s'} \pi(s'|s) H(z'_s, k', s')^{1-\gamma} \right]^{\frac{1-\rho}{1-\gamma}} \right]^{\frac{1}{1-\rho}}$$

subject to

$$z = \Omega(c, h) + \beta \sum_{s'} \pi(s'|s) \frac{H(z'_s, k', s')^{\rho-\gamma}}{\left[ \sum_{s'} \pi(s'|s) H(z'_s, k', s')^{1-\gamma} \right]^{\frac{\rho-\gamma}{1-\gamma}}} z'_s \quad (27)$$

$$c + k' - (1 - \delta)k + g_s = F(s, k, h) \quad (28)$$

$$c, k' \geq 0, h \in [0, 1] \quad (29)$$

$$z'_s \in Z(s', k') \quad (30)$$

$k'$  and  $z'_s$  stand for capital and period marginal utility-weighted assets next period. The notation captures the fact that the planner is choosing  $z'_s$  next period at the particular contingency  $s'$ .

**Remark 3.** *The value function shows up in the dynamic implementability constraint that the planner is facing. The reason is the presence of continuation value in the stochastic*

discount factor (21) due to the fact that the temporal resolution of uncertainty matters.

**Policy functions.** The policy functions from period one onward are functions of the state  $(z, k, s)$

$$\begin{aligned} c &= c(z, k, s) \\ h &= h(z, k, s) \\ k' &= k'(z, k, s) \\ z'_{s'} &= z'_{s'}(z, k, s) \end{aligned}$$

### 3.3 Initial period

A full description of the Ramsey planner's problem requires the description of the problem at the initial period:

**Problem 2.** *The problem at  $t = 0$  is*

$$\bar{H}_0(b_0, k_0, s_0, \tau_0^K) \equiv \max_{c_0, h_0, k_1, z_{1,s_1}} \left[ (1 - \beta)u(c_0, 1 - h_0)^{1-\rho} + \beta \left[ \sum_{s_1} \pi_1(s_1|s_0) H(z_{1,s_1}, k_1, s_1)^{1-\gamma} \right]^{\frac{1-\rho}{1-\gamma}} \right]^{\frac{1}{1-\rho}}$$

subject to

$$U_{c_0} \left[ ((1 - \tau_0^K)F_K(s_0, k_0, h_0) + 1 - \delta)k_0 + b_0 \right] = \Omega(c_0, h_0) + \beta \sum_{s_1} \pi_1(s_1|s_0) \frac{H(z_{1,s_1}, k_1, s_1)^{\rho-\gamma}}{\left[ \sum_{s_1} \pi_1(s_1|s_0) H(z_{1,s_1}, k_1, s_1)^{1-\gamma} \right]^{\frac{\rho-\gamma}{1-\gamma}}} z_{1,s_1} \quad (31)$$

$$c_0 + k_1 - (1 - \delta)k_0 + g_0 = F(s_0, k_0, h_0) \quad (32)$$

$$c_0, k_1 \geq 0, h_0 \in [0, 1], \quad (33)$$

$$z_{1,s_1} \in Z(s_1, k_1) \quad (34)$$

where  $(k_0, b_0, s_0, \tau_0^K)$  given.

**Discussion.** The problem at the initial period is different from period one onward due to the presence of the initial debt and capital  $(b_0, k_0)$  and to the exogenously fixed initial tax rate on capital income  $\tau_0^K$ , that functions as a lump-sum tax. As a result, the overall value

of the Ramsey problem  $\bar{H}_0$  and the policy functions of problem 2  $(c_0, h_0, k_1, z_{1,s_1})$  will depend now on  $(b_0, k_0, s_0, \tau_0^K)$ . Note that the initial value of the forward-looking variable  $z_1$  that was taken as *given* in the subproblem 1 is chosen *optimally* in order to solve the problem at  $t = 0$ . In this sense the variable  $z$  is a pseudo-state variable, i.e. a jump variable that is treated as fixed in order to capture the commitment of the planner to his past promises and help formulate the problem recursively, as Kydland and Prescott (1980) have demonstrated.

## 4 Optimality conditions

I am going to present the first-order conditions in a compact way by utilizing a “tilde” notation, that denotes a convenient scaling with the value function  $H$  and the discount factor  $\beta$ ,

$$\tilde{Y}_t \equiv H_t^{-\rho}(1 - \beta)^{-1}Y_t \quad (35)$$

for a generic random variable  $Y_t = Y(z_t, k_t, s_t)$ . Furthermore, define

$$\omega \equiv \sum_{s'} \pi(s'|s) m_{s'}^{\frac{\rho-\gamma}{1-\gamma}} z'_{s'}, \quad (36)$$

where  $m'_{s'}$  stands for the conditional likelihood ratio,  $m'_{s'} \equiv \frac{H(z'_{s'}, k', s')^{1-\gamma}}{\sum_{s'} \pi(s'|s) H(z'_{s'}, k', s')^{1-\gamma}}$ . The term  $\omega$  appears in the left hand side of (27) and is instrumental to the interpretation of the Ramsey plan. It can be roughly thought of as the *market value of the household's portfolio*.

Let  $\Phi$  and  $\lambda$  be the multipliers on the dynamic implementability constraint (27) and the resource constraint (28) respectively. Note that at the optimal solution the multipliers will also be functions of the state  $\Phi = \Phi(z, k, s)$  and  $\lambda = \lambda(z, k, s)$ . Then the first-order conditions for an interior solution at points of differentiability of the value function are <sup>11</sup>

---

<sup>11</sup>The “tilde” notation employed is directly related to the  $\rho$ -transformation in (3): Let  $v(z, k, s) \equiv \frac{H(z, k, s)^{1-\rho}-1}{(1-\beta)(1-\rho)}$ . Then  $v_i(z, k, s) = (1 - \beta)^{-1} H(z, k, s)^{-\rho} H_i(z, k, s) \equiv \tilde{H}_i(z, k, s), i = z, k$ . The multipliers  $\tilde{\Phi}, \tilde{\lambda}$  correspond to the multipliers of the respective functional equation for the transformed value function  $v(z, k, s)$ .

$$c : U_c + \tilde{\Phi}\Omega_c = \tilde{\lambda} \quad (37)$$

$$h : -U_l + \tilde{\Phi}\Omega_h = -\tilde{\lambda}F_H \quad (38)$$

$$k' : \tilde{\lambda} = \beta \sum_{s'} \pi(s'|s) m_{s'}^{\prime \frac{\rho-\gamma}{1-\gamma}} \tilde{H}_k(z'_{s'}, k', s') + \beta \tilde{\Phi} \frac{\partial \omega}{\partial k'} \quad (39)$$

$$z'_{s'} : \pi(s'|s) m_{s'}^{\prime \frac{\rho-\gamma}{1-\gamma}} \tilde{H}_z(z'_{s'}, k', s') + \tilde{\Phi} \frac{\partial \omega}{\partial z'_{s'}} = 0 \quad (40)$$

$\Omega_i, i = c, h$  stand for the partial derivative of  $\Omega$  with respect to consumption and labor. Furthermore, the derivatives of the market value of the household's portfolio with respect to capital and period marginal utility-weighted assets are

$$\frac{\partial \omega}{\partial k'} = (1 - \beta)(\rho - \gamma) \sum_{s'} \pi(s'|s) m_{s'}^{\prime \frac{\rho-\gamma}{1-\gamma}} \tilde{H}_k(z'_{s'}, k', s') H(z'_{s'}, k', s')^{-(1-\rho)} \eta'_{s'} \quad (41)$$

$$\frac{\partial \omega}{\partial z'_{s'}} = \pi(s'|s) m_{s'}^{\prime \frac{\rho-\gamma}{1-\gamma}} [(1 - \beta)(\rho - \gamma) \tilde{H}_z(z'_{s'}, k', s') H(z'_{s'}, k', s')^{-(1-\rho)} \eta'_{s'} + 1], \quad (42)$$

where

$$\eta'_{s'} \equiv z'_{s'} - m_{s'}^{\prime \frac{1-\rho}{1-\gamma}} \sum_{s'} \pi(s'|s) m_{s'}^{\prime \frac{\rho-\gamma}{1-\gamma}} z'_{s'}. \quad (43)$$

Using (41) and (42), the first-order conditions with respect to  $k'$  and  $z'_{s'}$  are finally becoming

$$k' : \tilde{\lambda} = \beta \sum_{s'} \pi(s'|s) m_{s'}^{\prime \frac{\rho-\gamma}{1-\gamma}} \tilde{H}_k(z'_{s'}, k', s') [1 + (1 - \beta)(\rho - \gamma) H(z'_{s'}, k', s')^{-(1-\rho)} \eta'_{s'} \tilde{\Phi}] \quad (44)$$

$$z'_{s'} : \tilde{H}_z(z'_{s'}, k', s') + \tilde{\Phi} [1 + (1 - \beta)(\rho - \gamma) \tilde{H}_z(z'_{s'}, k', s') H(z'_{s'}, k', s')^{-(1-\rho)} \eta'_{s'}] = 0 \quad (45)$$

Similarly, let  $\Phi_0$  and  $\lambda_0$  denote the multipliers on the initial period constraints (31) and (32) respectively and let the “tilde” notation (35) refer now to the scaling with the value function  $\tilde{H}_0$  at  $t = 0$ , to get

$$c_0 : \quad U_{c0} + \tilde{\Phi}_0 \left[ \Omega_{c0} - U_{cc,0} W_0 \right] = \tilde{\lambda}_0 \quad (46)$$

$$h_0 : \quad -U_{l0} + \tilde{\Phi}_0 \left[ \Omega_{h0} + U_{cl,0} W_0 - U_{c0} (1 - \tau_0^K) F_{KH,0} k_0 \right] = -\tilde{\lambda}_0 F_{H0}$$

$$k_1 : \quad \tilde{\lambda}_0 = \beta \sum_{s_1} \pi(s_1 | s_0) m_{1,s_1}^{\frac{\rho-\gamma}{1-\gamma}} \tilde{H}_k(z_{1,s_1}, k_1, s_1) \cdot \left[ 1 + (1 - \beta)(\rho - \gamma) H(z_{1,s_1}, k_1, s_1)^{-(1-\rho)} \eta_{1,s_1} \tilde{\Phi}_0 \right] \quad (47)$$

$$z_{1,s_1} : \quad \tilde{H}_z(z_{1,s_1}, k_1, s_1) + \tilde{\Phi}_0 \left[ 1 + (1 - \beta)(\rho - \gamma) \tilde{H}_z(z_{1,s_1}, k_1, s_1) H(z_{1,s_1}, k_1, s_1)^{-(1-\rho)} \eta_{1,s_1} \right] = 0, \quad (48)$$

where  $W_0 = [(1 - \tau_0^K) F_K(s_0, k_0, h_0) + (1 - \delta)] k_0 + b_0$ , the household's initial wealth.

**Envelope conditions:** The envelope conditions for the endogenous state variables are

$$H_z(z, k, s) = -\Phi \quad (49)$$

$$H_k(z, k, s) = \lambda(1 - \delta + F_k). \quad (50)$$

Note that the binding directions of (27) and (28) lead to  $\Phi > 0$  and  $\lambda > 0$ , so that  $H_z(z, k, s) < 0$  and  $H_k(z, k, s) > 0$ . As a consequence, the respective ‘tilde’ derivatives become  $\tilde{H}_z = -\tilde{\Phi} < 0$  and  $\tilde{H}_k = \tilde{\lambda}(1 - \delta + F_k) > 0$ .

## 5 Analysis

### 5.1 Cost of distortionary taxation

Our main object of the analysis will be the multiplier  $\tilde{\Phi}$ , which captures the shadow value of the constraints that the competitive equilibrium imposes in the second-best world. In first-best world with lump-sum taxes available,  $\tilde{\Phi}$  would be zero. For that reason it has been dubbed in the literature as the *excess burden* of distortionary taxation.

The richer structure of the stochastic discount factor manifests itself with the presence of the value function in the dynamic implementability constraint (27). Its importance is rooted in the fact that changes in the period marginal-utility weighted assets  $z'_s$ , and in capital  $k'$  affect through the continuation value the market value of the household's asset portfolio  $\omega$ .

Consider the first-order condition (45), which captures the dynamic tradeoffs that the planner is facing in increasing  $z'_{s'}$ , and decompose it to three terms:

$$\underbrace{\tilde{H}_z(z'_{s'}, k', s')}_{\text{MC of increasing } z'_{s'}} + \underbrace{\tilde{\Phi}}_{\text{MB of relaxing IC}} + \underbrace{\tilde{\Phi}(1 - \beta)(\rho - \gamma)\tilde{H}_z(z'_{s'}, k', s')H(z'_{s'}, k', s')^{-(1-\rho)}\eta'_{s'}}_{\text{price effect of } z'_{s'} \text{ (MB/MC)}} = 0. \quad (51)$$

The first-order condition equates the marginal costs and benefits of increasing  $z'_{s'}$ . An increase in  $z'_{s'}$  (e.g. by issuing debt) at contingency  $s'$  has a marginal cost since it is associated with more taxation in the future ( $H_z(z'_{s'}, k', s') < 0$ ) but it entails also a marginal benefit by relaxing the implementability constraint and allowing therefore less taxation today. These are the first and second terms respectively in (51) and the only two terms relevant if we were in the expected utility world of Chari et al. (1994).

**Price effect of continuation value.** However, an increase in  $z'_{s'}$  has a third effect which I will term the *price effect*, since it affects through the value function the equilibrium *price* and therefore the market value of the household's portfolio  $\omega$  (36). This effect is captured by the third term in (51) and depends on the preference for early (late) resolution of uncertainty ( $\rho < (>)\gamma$ ) and on the term  $\eta'_{s'}$  that is defined in (43). At first, note that  $\eta'_{s'}$  has the following property:

**Lemma 1.** (*"Innovation" property*)

$$\sum_{s'} \pi(s'|s) m_{s'}'^{\frac{\rho-\gamma}{1-\gamma}} \eta'_{s'} = 0 \quad (52)$$

*Proof.* Use the definition of (43) to get

$$\begin{aligned} \sum_{s'} \pi(s'|s) m_{s'}'^{\frac{\rho-\gamma}{1-\gamma}} \eta'_{s'} &= \sum_{s'} \pi(s'|s) m_{s'}'^{\frac{\rho-\gamma}{1-\gamma}} z'_{s'} - \sum_{s'} \pi(s'|s) m_{s'}'^{\frac{1-\rho}{1-\gamma}} m_{s'}'^{\frac{\rho-\gamma}{1-\gamma}} \sum_{s'} \pi(s'|s) m_{s'}'^{\frac{\rho-\gamma}{1-\gamma}} z'_{s'} \\ &= \sum_{s'} \pi(s'|s) m_{s'}'^{\frac{\rho-\gamma}{1-\gamma}} z'_{s'} - \left( \sum_{s'} \pi(s'|s) m_{s'}' \right) \sum_{s'} \pi(s'|s) m_{s'}'^{\frac{\rho-\gamma}{1-\gamma}} z'_{s'} = 0, \end{aligned}$$

since  $\sum_{s'} \pi(s'|s) m_{s'}' = 1$  by construction.  $\square$

The above lemma instructs us that  $\eta'_{s'}$  can take both positive and negative values. Thus, if there is a preference for early resolution of uncertainty ( $\rho < \gamma$ ) and

- $\eta'_{s'} > 0$ , then there is a marginal *benefit* in increasing  $z'_{s'}$  (*positive price effect*).

- $\eta'_{s'} < 0$ , then there a marginal *cost* (*negative price effect*).

The direction of the inequalities is reversed for the opposite case of  $\rho > \gamma$ . The variable  $\eta'_{s'}$  is by construction a measure of how much the household's assets weighted by the period marginal utility of consumption  $z'_{s'}$  are above or below a multiple of  $\omega$ . In the case of  $\rho = 1$ , there is a sharper interpretation of  $\eta'_{s'}$  as the conditional forecast error (or innovation) of the assets  $z'_{s'}$  under the measure  $\pi_t \cdot M_t$ . I will refer to  $\eta'_{s'}$  as the *innovation* even in the case of  $\rho \neq 1$ , which explains the naming of lemma 1.

Collect now the terms that involve the derivative of the value function  $H_z(z'_{s'}, k', s')$  in (45), use the envelope condition for next period to get  $\tilde{H}_z(z'_{s'}, k', s') = -\tilde{\Phi}'_{s'}$ , and rewrite (45) in terms of  $\tilde{\Phi}^{-1}$

$$\frac{1}{\tilde{\Phi}'_{s'}} = \frac{1}{\tilde{\Phi}} + (1 - \beta)(\rho - \gamma)H(z'_{s'}, k', s')^{-(1-\rho)}\eta'_{s'}. \quad (53)$$

Follow the same steps for the optimality condition at  $t = 0$ , (48) and use sequence notation to finally get

$$\frac{1}{\tilde{\Phi}_{t+1}} = \frac{1}{\tilde{\Phi}_t} + (1 - \beta)(\rho - \gamma)H_{t+1}^{-(1-\rho)}\eta_{t+1}, t \geq 0. \quad (54)$$

The above equation determines the law of motion of the shadow cost of distortinary taxation  $\tilde{\Phi}_t$  or else the shadow cost of household's assets weighted by the period marginal utility of consumption  $z_t$ , and plays a crucial role in the dynamics of the solution.

**Add comment for deterministic** ( $\eta_{t+1} \equiv 0$ ).

Note that for the case of time-additive expected utility ( $\rho = \gamma$ ), we would have

$$\frac{1}{\tilde{\Phi}_t} = \frac{1}{\tilde{\Phi}_{t-1}} = \dots = \frac{1}{\tilde{\Phi}_1} = \frac{1}{\tilde{\Phi}_0},$$

so the cost of distortinary taxation would stay *constant*.

**Remark 4.** *In the case of time-additive expected utility ( $\rho = \gamma$ ), the planner chooses assets or debt  $z_t$  in such a way so that the cost of distortinary taxation remains constant, whereas when  $\rho \neq \gamma$ , there is time variation in  $\tilde{\Phi}_t$  that reflects the price effect of  $z_t$ .*

The time-varying nature of  $\tilde{\Phi}$  is also the deeper reason for keeping  $z$  as a state variable:

**Proposition 2.** (“Debt matters”) *EZW preferences contribute an additional state variable  $z_t$  to the optimal taxation problem. With expected utility  $\rho = \gamma$ , it would be sufficient to keep track only of  $(k, s)$ , because the cost of distortionary taxation  $\tilde{\Phi}$  is not time-varying.*

*Proof.* [To be shortened] For  $\rho = \gamma$ , the Bellman equation (3.2) becomes

$$H(z, k, s) = \max_{c, h, k', z'_{s'}} \left[ (1 - \beta)u(c, 1 - h)^{1-\rho} + \beta \sum_{s'} \pi(s'|s) H(z'_{s'}, k', s')^{1-\rho} \right]^{\frac{1}{1-\rho}}$$

subject to

$$z = \Omega(c, h) + \beta \sum_{s'} \pi(s'|s) z'_{s'} \quad (55)$$

and the resource constraint (28). Note at first that the value function is not appearing anymore in the dynamic implementability constraint, since with expected utility it does not affect the stochastic discount factor. Furthermore, the first-order conditions become

$$\begin{aligned} c : & \quad (1 - \beta)H(z, k, s)^\rho u^{-\rho} u_c + \Phi \Omega_c = \lambda \\ h : & \quad -(1 - \beta)H(z, k, s)^\rho u^{-\rho} u_l + \Phi \Omega_h = -\lambda F_h(s, k, h) \\ k' : & \quad H(z, k, s)^{-\rho} \lambda = \beta \sum_{s'} \pi(s'|s) H(z'_{s'}, k', s')^{-\rho} H_k(z'_{s'}, k', s') \\ z'_{s'} : & \quad -H(z'_{s'}, k', s')^{-\rho} H_z(z'_{s'}, k', s') = H(z, k, s)^{-\rho} \Phi \end{aligned}$$

Use now the envelope condition (49) and utilize the tilde notation to get

$$\tilde{\Phi}(z'_{s'}, k', s') = \tilde{\Phi}(z, k, s) \quad (56)$$

Thus the multiplier  $\tilde{\Phi}$  remains *constant*  $\tilde{\Phi}(z, k, s) = \bar{\Phi}, \quad \forall (z, k, s)$ . This can be easily seen also from the law of motion (54) for  $\rho = \gamma$ . Utilize also the envelope condition with respect to capital (50) and the “tilde notation” to get finally the following first-order conditions:

$$\begin{aligned}
u^{-\rho}u_c + \bar{\Phi}\Omega_c &= \tilde{\lambda} \\
-u^{-\rho}u_l + \bar{\Phi}\Omega_h &= -\tilde{\lambda}F_h \\
\tilde{\lambda} &= \beta \sum_{s'} \pi(s'|s)\tilde{\lambda}'_{s'}(1 - \delta + F_{k,s'})
\end{aligned}$$

Note that if we knew  $\bar{\Phi}$ , then the optimality conditions above together with the resource constraint (28) would allow us to derive the policy functions for  $(c, h, k')$  as functions of  $(k, s)$  and the fixed value  $\bar{\Phi}$ . In other words, the dynamic implementability constraint (55) is necessary for the determination of the solution only as long as it affects the shadow cost of distortionary taxation  $\tilde{\Phi}$ . Since for the case of expected utility this cost is *not* time-varying, but constant to some value  $\bar{\Phi}$  independent of  $(z, k, s)$ , the planner does not have to keep track of the period marginal utility-weighted total assets  $z$ , but only of the state variable  $(k, s)$ . Note that if we had an economy without capital, we would need to keep track only of  $s$ , and this would be a manifestation of the *history-independence* result of Lucas and Stokey (1983).

So the problem for  $\rho = \gamma$  reduces to the Chari et al. (1994) case with a constant  $\bar{\Phi}$ . In order to see that clearly, transform the utility recursion (2) by using the monotonic transformation  $V_t^{\text{CCK}} \equiv \frac{V_t^{1-\rho}-1}{(1-\rho)(1-\beta)}$ . The Bellman equation for the Chari et al. problem is

$$V^{\text{CCK}}(k, s; \bar{\Phi}) = \max_{c, h, k'} [U^{\text{CCK}}(c, h; \bar{\Phi}) + \beta \sum_{s'} \pi(s'|s)V^{\text{CCK}}(k', s'; \bar{\Phi})]$$

subject to the resource constraint (28), where  $c, k' \geq 0, h \in [0, 1]$  and

$$U^{\text{CCK}}(c, h; \bar{\Phi}) \equiv \frac{u(c, 1-h)^{1-\rho} - 1}{1-\rho} + \bar{\Phi}\Omega(c, h). \quad (57)$$

Note that the period return function is augmented so that it takes into account the relevant part of the implementability constraint  $\Omega$  and the fixed value  $\bar{\Phi}$ . The solution to the above problem delivers policy functions for  $(c, h, k')$  as functions of  $(k, s)$  and  $\bar{\Phi}$  from period one onward.<sup>12</sup>  $\square$

<sup>12</sup>A full characterization requires solving in the initial period the problem

$$\max_{\{c_0, h_0, k_1\}} [U^{\text{CCK}}(c_0, h_0; \bar{\Phi}) - \bar{\Phi}u_0^{-\rho}u_{c_0}a_0W_0 + \beta \sum_{s_1} \pi(s_1|s_0)V^{\text{CCK}}(k_1, s_1; \bar{\Phi})] \quad (58)$$

subject to the initial resource constraint. Remember that  $W_0$  stands for initial wealth. Note that the initial

The time-varying  $\tilde{\Phi}_t$  has the following properties:

**Proposition 3.** (*Properties of the cost of distortionary taxation*)

1. (Monotonicity) *Let the household have a preference for early resolution of uncertainty ( $\gamma > \rho$ ). Then*

- *if  $\eta_{t+1} > 0$ , then  $\tilde{\Phi}_{t+1} > \tilde{\Phi}_t$*
- *if  $\eta_{t+1} < 0$ , then  $\tilde{\Phi}_{t+1} < \tilde{\Phi}_t$*

*The inequalities are reversed in the case of a preference for late resolution of uncertainty ( $\gamma < \rho$ ).*

2. (Martingale Characterization) *Then inverse of  $\tilde{\Phi}_t$  is a martingale with respect to the beliefs  $\pi_t \cdot M_t$  and as a result,  $\tilde{\Phi}_t$  is a submartingale with respect to the beliefs  $\pi_t \cdot M_t$ .*

*Proof.* Part (1) is obvious from the law of motion of  $1/\tilde{\Phi}$  (54). For the second part, take conditional expectations with respect to the measure  $\pi_t \cdot M_t$  in (54) to get

$$\begin{aligned} E_t m_{t+1} \frac{1}{\tilde{\Phi}_{t+1}} &= \frac{1}{\tilde{\Phi}_t} E_t m_{t+1} + (1 - \beta)(\rho - \gamma) E_t m_{t+1} H_{t+1}^{-(1-\rho)} \eta_{t+1} \\ &= \frac{1}{\tilde{\Phi}_t} + (1 - \beta)(\rho - \gamma) E_t m_{t+1} H_{t+1}^{-(1-\rho)} \eta_{t+1}, \end{aligned}$$

since  $E_t m_{t+1} = 1$ . Recalling the definition of  $m_{t+1}$  allows rewriting the expectation in the second term as

$$\begin{aligned} E_t m_{t+1} H_{t+1}^{-(1-\rho)} \eta_{t+1} &= E_t \frac{H_{t+1}^{1-\gamma}}{\mu_t^{1-\gamma}} H_{t+1}^{-(1-\rho)} \eta_{t+1} = E_t \frac{\mu_t^{\rho-\gamma}}{\mu_t^{1-\gamma}} \left( \frac{H_{t+1}}{\mu_t} \right)^{\rho-\gamma} \eta_{t+1} \\ &= \mu_t^{\rho-1} E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} \eta_{t+1} = 0, \end{aligned}$$

since  $E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} \eta_{t+1} = 0$ , according to lemma 1. Thus  $1/\tilde{\Phi}_t$  is a martingale according to the measure  $\pi_t \cdot M_t$ . Furthermore, since the function  $f(x) = 1/x$  is convex for  $x > 0$ , an application of the conditional Jensen's inequality leads to  $E_t m_{t+1} \frac{1}{X_{t+1}} \geq \frac{1}{E_t m_{t+1} X_{t+1}}$ . Set now  $X_t = 1/\tilde{\Phi}_t$  and use the martingale result to finally get  $E_t m_{t+1} \tilde{\Phi}_{t+1} \geq \tilde{\Phi}_t$ .  $\square$

---

policy functions are functions of  $\bar{\Phi}$  and the initial state  $(b_0, k_0, s_0)$ . The final step is to adjust appropriately  $\bar{\Phi}$  so that the intertemporal implementability constraint is satisfied.

Note that since  $\tilde{\Phi}$  is a positive submartingale, it converges almost surely according to the measure  $\pi \cdot M$  to a random variable. A natural question that proposition 3 leaves unanswered is about the stochastic properties of  $\tilde{\Phi}$  under the original probability measure that describes uncertainty  $\pi$ .

**Proposition 4.** *Stochastic properties of  $\tilde{\Phi}$  according to  $\pi$ . To be completed.*

## 6 Capital and labor income taxation

The time-varying nature of the cost of distortionary taxation  $\tilde{\Phi}$  is changing drastically the implications for the taxation of the intertemporal and the intratemporal margin.

### 6.1 Capital tax

The optimal choice of capital  $k'$  is affecting also the market value of the household's wealth  $\omega$ . This can be seen by the optimality condition (39) which equates the marginal cost of increasing  $k'$  by one unit the current period with the marginal benefit of having one additional unit of capital in the beginning of next period and the marginal benefit *or* cost that an additional unit of  $k'$  has on  $\omega$ .

Turning now into sequence notation and using (54) in order to rewrite (44), allows us to see more explicitly the dependence of the optimal capital decision on the time-varying  $\tilde{\Phi}_t$ ,

$$\tilde{\lambda}_t = \beta E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} \tilde{H}_k(z_{t+1}, k_{t+1}, s_{t+1}) \frac{\tilde{\Phi}_t}{\tilde{\Phi}_{t+1}}. \quad (59)$$

Furthermore, using the envelope condition with respect to capital  $\tilde{H}_k = \tilde{\lambda}(1 - \delta + F_K)$ , and updating one period delivers

$$E_t S_{t+1}^* (1 - \delta + F_{K,t+1}) = 1, \quad (60)$$

where

$$S_{t+1}^* \equiv \beta m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} \frac{\tilde{\lambda}_{t+1}/\tilde{\Phi}_{t+1}}{\tilde{\lambda}_t/\tilde{\Phi}_t}. \quad (61)$$

I will call  $S_{t+1}^*$  the *social* stochastic discount factor.<sup>13</sup>  $S_{t+1}^*$  is an object that captures how the planner discounts the *pre-tax* return on capital at the *second-best* allocation.  $S_{t+1}^*$  contrasts to the market stochastic discount factor  $S_{t+1} \equiv \beta m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} U_{c,t+1}/U_{c,t}$ . It is easy to see that in a first-best world with lump-sum taxes available, we would have  $S_{t+1}^* = S_{t+1}$ . The social discount factor  $S_{t+1}^*$  can differ though from  $S_{t+1}$  in the second-best world, and is useful in summarizing the optimal wedge at the *intertemporal* margin. More specifically, recall the definition of the ex-ante tax rate on capital income (25) and use (60) to get

$$\bar{\tau}_{t+1}^K = \frac{E_t[S_{t+1} - S_{t+1}^*](1 - \delta + F_{K,t+1})}{E_t S_{t+1} F_{K,t+1}}. \quad (62)$$

Thus, there is a positive (negative) tax rate on capital income if the numerator of (62) is positive (negative). Another way to think about the sign of the numerator is in terms of the *size* of the (non-centered) *covariances* of the social and the market stochastic discount factors with the pre-tax capital return  $1 - \delta + F_{K,t+1}$ . A (non-centered) covariance of the market stochastic discount factor with pre-tax return on capital that is larger (smaller) than the respective non-centered covariance of the social stochastic discount factor, leads to a positive (negative) ex-ante tax rate  $\bar{\tau}_{t+1}^K > 0$  ( $\bar{\tau}_{t+1}^K < 0$ ).

The social stochastic discount factor  $S_{t+1}^*$  is associated with the market stochastic discount factor  $S_{t+1}$  as follows

$$\begin{aligned} S_{t+1}^* &= S_{t+1} \frac{\tilde{\lambda}_{t+1}/(\tilde{\Phi}_{t+1} U_{c,t+1})}{\tilde{\lambda}_t/(\tilde{\Phi}_t U_{c,t})} \\ &= S_{t+1} \frac{\frac{1}{\tilde{\Phi}_{t+1}} + \frac{\Omega_{c,t+1}}{U_{c,t+1}}}{\frac{1}{\tilde{\Phi}_t} + \frac{\Omega_{c,t}}{U_{c,t}}}, \end{aligned} \quad (63)$$

where I used the optimality condition with respect to consumption (37) in the second line. Thus, the difference in the two discount factors can be expressed as

---

<sup>13</sup> $S_{t+1}^*$  depends on the state  $X = (z, k, s)$  and the shock  $s'$  as follows:

$$S_{t+1}^* = S^*(s', X) = \beta \frac{H(z'_{s'}(X), k'(X), s')^{\rho-\gamma}}{\sum_{s'} [\pi(s'|s) H(z'_{s'}(X), k'(X), s')^{1-\gamma}]^{\frac{\rho-\gamma}{1-\gamma}}} \frac{\tilde{\lambda}(z'_{s'}(X), k'(X), s')}{\tilde{\lambda}(X)} \frac{\tilde{\Phi}(X)}{\tilde{\Phi}(z'_{s'}(X), k'(X), s')}.$$

$$\frac{S_{t+1} - S_{t+1}^*}{S_{t+1}} = \frac{\frac{1}{\tilde{\Phi}_t} - \frac{1}{\tilde{\Phi}_{t+1}} + \frac{\Omega_{c,t}}{U_{c,t}} - \frac{\Omega_{c,t+1}}{U_{c,t+1}}}{\frac{1}{\tilde{\Phi}_t} + \frac{\Omega_{c,t}}{U_{c,t}}}. \quad (64)$$

The difference in the two discount factors and, consequently, a wedge at the intertemporal margin, can be attributed to *time variation* in the marginal cost of distortionary taxation  $\tilde{\Phi}$  and to *time variation* in  $\Omega_c/U_c$ . Remember that  $\Omega$  stands for consumption minus after-tax labor income, weighted by the period marginal utility of consumption. Calculating  $\Omega_c/U_c$  delivers

$$\frac{\Omega_c}{U_c} = 1 - \epsilon_{cc} - \epsilon_{ch}, \quad (65)$$

where  $\epsilon_{cc} \equiv -U_{cc}c/U_c$  and  $\epsilon_{ch} \equiv U_{cl}h/U_c$ , i.e. the own and cross elasticity of the period marginal utility of consumption with respect to consumption and labor.<sup>14</sup>

**Proposition 5.** (*Sources of capital taxation*) *The ex-ante tax rate on capital income  $\bar{\tau}_{t+1}^K$  is positive (negative) iff*

$$E_t S_{t+1} (1 - \delta + F_{K,t+1}) \left[ \underbrace{\left( \frac{1}{\tilde{\Phi}_t} - \frac{1}{\tilde{\Phi}_{t+1}} \right)}_{\text{variation in } \tilde{\Phi}_t} + \underbrace{(\epsilon_{cc,t+1} + \epsilon_{ch,t+1} - \epsilon_{cc,t} - \epsilon_{ch,t})}_{\text{variation in period elasticities}} \right] > (<) 0$$

*If  $\epsilon_{cc} + \epsilon_{ch}$  is constant, then the only reason for taxing the intertemporal margin comes from variation in the cost of distortionary taxation  $\tilde{\Phi}_t$ .*

Let's interpret what proposition 5 implies about capital income taxation. Assume that we are in a deterministic economy. Then  $\tilde{\Phi}_t$  is *constant* and capital income is taxed (subsidized) if the sum of the own and cross elasticities is increasing (decreasing), as can be seen from the formula in proposition 5. A necessary and sufficient condition for a zero capital tax at *every* period in the deterministic economy is that the sum of the elasticities of the period marginal utility of consumption is *constant* (which implies that  $S_{t+1}^* = S_{t+1}$ ). The sum of the elasticities is constant at the deterministic steady state, which delivers the zero tax result of Chamley (1986) and Judd (1985). Turning now to the stochastic case of Chari et al. (1994) and Zhu (1992) with time-additive expected utility ( $\rho = \gamma$ ), variation in  $\tilde{\Phi}_t$  is again absent. The analysis remains essentially the same as in the deterministic case, with the nuance of integrating properly over next period's shocks the product of the market discount

<sup>14</sup>Note that I have multiplied each elasticity with minus unity.

factor, the pre-tax gross return on capital and the change in the sum of period elasticities. Variation in the sum of period elasticities is a *necessary* condition for an intertemporal wedge, since a constant sum of period elasticities would imply  $S_{t+1}^* = S_{t+1}$ . For the case of EZW preferences, even in the case of constant period elasticities, the variation in  $\tilde{\Phi}_t$  consists a novel source of taxation at the intertemporal margin, reflecting the presence of continuation values in equilibrium prices.

## 6.2 Labor tax

We can eliminate  $\tilde{\lambda}$  and combine the first-order conditions with respect to consumption and labor (37)-(38) in order to derive the optimal wedge in labor supply as

$$\tau^H = \frac{U_c F_H - U_l}{U_c F_H} = -\tilde{\Phi} \left[ \frac{\Omega_c}{U_c} + \frac{\Omega_h}{U_c F_H} \right],$$

which, by using  $\frac{\Omega_h}{U_c F_H} = \frac{\Omega_h}{U_l} \frac{U_l}{U_c F_H} = \frac{\Omega_h}{U_l} (1 - \tau^H)$ , becomes

$$\tau^H(z, k, s) = \frac{-\tilde{\Phi} \left[ \frac{\Omega_c}{U_c} + \frac{\Omega_h}{U_l} \right]}{1 - \tilde{\Phi} \frac{\Omega_h}{U_l}}. \quad (66)$$

Note that

$$\frac{\Omega_h}{U_l} = -1 - \epsilon_{hh} - \epsilon_{hc}, \quad (67)$$

where  $\epsilon_{hh} \equiv -U_{ll}h/U_l$  and  $\epsilon_{hc} \equiv U_{lc}c/U_l$ , the own and cross elasticity of the period marginal disutility of labor with respect to labor and consumption. So, by using (65) and (67), the optimal labor tax becomes<sup>15</sup>

$$\tau_t^H = \tilde{\Phi}_t \frac{\epsilon_{cc,t} + \epsilon_{ch,t} + \epsilon_{hh,t} + \epsilon_{hc,t}}{1 + \tilde{\Phi}_t (1 + \epsilon_{hh,t} + \epsilon_{hc,t})}, \quad t \geq 1. \quad (69)$$

---

<sup>15</sup>Similarly, the labor tax at  $t = 0$  is

$$\tau_0^H = \tilde{\Phi}_0 \frac{\epsilon_{cc} + \epsilon_{ch} + \epsilon_{hh} + \epsilon_{hc} + (1 - \tau_0^K) \frac{F_{KH}}{F_H} k_0 - (\epsilon_{cc} + \epsilon_{hc}) c_0^{-1} W_0}{1 + \tilde{\Phi}_0 (1 + \epsilon_{hh} + \epsilon_{hc} - \epsilon_{hc} c_0^{-1} W_0)}, \quad (68)$$

where the respective elasticities and marginal products are evaluated at the initial allocation.

The optimal labor tax varies either due to variation in the elasticities of the period marginal utility of consumption and the period marginal disutility of labor *or* due to variation in  $\tilde{\Phi}_t$ . Expression (69) illustrates the two separate sources of taxation also at the intratemporal margin, the period utility channel (the period elasticities in (69)) and the continuation value channel ( $\tilde{\Phi}_t$ ).

In the next subsection I am going to obtain a sharper characterization of the optimal labor and capital taxes by assuming particular functional forms for the aggregator  $u$ , that will deliver explicit formulas for the period elasticities  $\epsilon_{ij}$ ,  $i, j = c, h$ .

### 6.3 Aggregators $u$

Consider two specifications of  $u$ :

#### 6.3.1 Separability of consumption and leisure (zero capital tax for expected utility)

Consider preferences that are separable between consumption and leisure and isoelastic in consumption with parameter  $\rho$ ,<sup>16</sup>

$$u(c, 1 - h) = [c^{1-\rho} - (1 - \rho)v(h)]^{\frac{1}{1-\rho}}. \quad (70)$$

For specification (70) we get the own elasticities  $\epsilon_{cc} = \rho$  and  $\epsilon_{hh} = v''(h)h/v'(h)$  and *zero* cross elasticities  $\epsilon_{ch} = \epsilon_{hc} = 0$ . Note that for this aggregator, the inverse of the elasticity of the marginal disutility of labor represents the *Frisch* elasticity of labor supply  $\varepsilon_F(h) = 1/\epsilon_{hh} = 1/\frac{v''(h)h}{v'(h)}$ .

Chari et al. (1994) have shown that these preferences imply a *zero* ex-ante tax rate on capital income. This is easily seen from the formula in proposition 5. The elasticity of the period marginal utility of consumption is *constant* and therefore, for the case of expected utility ( $\rho = \gamma$ , which implies a constant  $\tilde{\Phi}$ ), we have the equality of the social and the market discount factor  $S_{t+1}^* = S_{t+1}$ , leading to a zero ex-ante tax on capital income.<sup>17</sup>

Thus, for these preferences, when  $\rho \neq \gamma$ , the variation in the cost of distortionary taxation creates *taxation* of the intertemporal margin. In particular, use the law of motion (54) and

---

<sup>16</sup>We can see in a clearer way the separability between consumption and leisure by applying the  $\rho$ -transformation to get the period utility function  $U(c, 1 - h) = \frac{c^{1-\rho} - 1}{1-\rho} - v(h)$ . For  $\rho = 1$ , the utility recursion becomes  $V_t = \exp[(1 - \beta)(\ln c - v(h)) + \beta \ln \mu_t]$ .

<sup>17</sup>The zero tax result would obviously hold also in a deterministic economy for the aggregator (70).

proposition 5, to find that

$$\bar{\tau}_{t+1}^K > 0 (< 0) \text{ iff } E_t S_{t+1} (1 - \delta + F_{K,t+1}) [-(1 - \beta)(\rho - \gamma) H_{t+1}^{-(1-\rho)} \eta_{t+1}] > 0 (< 0). \quad (71)$$

Since  $\eta_{t+1} \equiv z_{t+1} - m_{t+1}^{\frac{1-\rho}{1-\gamma}} E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} z_{t+1}$ , takes *both* positive and negative values, the sign of the above expression is not obvious.

The optimal labor tax (69) simplifies as follows:

$$\tau_t^H = \tilde{\Phi}_t \frac{\rho + \frac{1}{\varepsilon_F(h_t)}}{1 + \tilde{\Phi}_t (1 + \frac{1}{\varepsilon_F(h_t)})}, t \geq 1, \quad (72)$$

illustrating that variability in the optimal labor tax is due either to variability in the Frisch elasticity  $\varepsilon_F$  or to the cost of distortionary taxation  $\tilde{\Phi}_t$ .<sup>18</sup>

**Constant Frisch elasticity.** Let  $v(h) = a_h \frac{h^{1+\phi_h}}{1+\phi_h}$ , which implies a *constant* Frisch elasticity  $\varepsilon_F = 1/\phi_h$ . Expression (72) shows that the labor tax is *constant* in the case of expected utility ( $\tilde{\Phi}_t = \bar{\Phi}$ ), leading to *perfect* tax smoothing. For  $\rho \neq \gamma$ , this is not true anymore, due to the variation of  $\tilde{\Phi}$ . In particular,

**Proposition 6.** (*Labor tax with constant Frisch elasticity*)

1. *The labor tax takes the form*

$$\frac{1}{\tau_{t+1}^H} = \frac{1}{\tau_t^H} + \frac{(1 - \beta)(\rho - \gamma)}{\rho + \phi_h} H_{t+1}^{-(1-\rho)} \eta_{t+1}, t \geq 1. \quad (73)$$

2. (*Monotonicity*) *Let the household have a preference for early resolution of uncertainty  $\rho < \gamma$  (the direction of the inequalities is reversed for the case of  $\rho > \gamma$ ). Then*

- *if  $\eta_{t+1} > 0$ , then  $\tau_{t+1}^H > \tau_t^H$  (because  $\tilde{\Phi}_{t+1}$  increases)*
- *if  $\eta_{t+1} < 0$ , then  $\tau_{t+1}^H < \tau_t^H$  (because  $\tilde{\Phi}_{t+1}$  decreases)*

3. (*Martingale Characterization*) *The inverse of the labor tax  $1/\tau_t^H$  is a martingale with respect to the beliefs  $\pi_t \cdot M_t$  and therefore,  $\tau_t^H$  is a submartingale with respect to  $\pi_t \cdot M_t$ .*

<sup>18</sup>The labor tax at  $t = 0$  would differ due to the initial wealth and capital,

$$\tau_0^H = \tilde{\Phi}_0 \frac{\rho + \frac{1}{\varepsilon_F(h_0)} + (1 - \tau_0^K) \frac{F_{KH}}{F_H} k_0 - \rho c_0^{-1} W_0}{1 + \tilde{\Phi}_0 (1 + \frac{1}{\varepsilon_F(h_0)})}.$$

*Proof.* Write the inverse of the labor tax as

$$\frac{1}{\tau_t^H} = \frac{1 + \phi_h}{\rho + \phi_h} + \frac{1}{\rho + \phi_h} \frac{1}{\tilde{\Phi}_t},$$

and note that  $1/\tau_t^H$  is an *affine* function of  $1/\tilde{\Phi}_t$ . Given the law of motion of  $\tilde{\Phi}$  (54), this allows us to write the law of motion of the labor tax as in (73). Notice the close resemblance of the law of motion of the labor tax (73) to the law of motion of the cost of distortionary taxation (54), a fact that leads to similar conclusions about monotonicity and martingale-like properties as in proposition 3.  $\square$

### 6.3.2 Balanced-growth consistent preferences of Chari et al. (1994)

Preferences (70) do not allow for balanced growth in a deterministic version of our economy, unless  $\rho = 1$ . Consider a Cobb-Douglas aggregator

$$u(c, 1 - h) = c^\theta (1 - h)^{1-\theta}, \quad (74)$$

where  $\theta \in (0, 1)$ . These preferences feature non-separabilities between consumption and leisure and have been a workhorse for the RBC literature. They consist the basis of the quantitative experiments of Chari et al. (1994). According to (74), the elasticities of the period marginal utility of consumption are  $\epsilon_{cc} = 1 - \theta(1 - \rho)$  and  $\epsilon_{ch} = (1 - \theta)(1 - \rho)h/(1 - h)$ , and the elasticities of the period marginal disutility of labor are  $\epsilon_{hh} = (\theta + (1 - \theta)\rho)h/(1 - h)$  and  $\epsilon_{hc} = \theta(1 - \rho)$ .<sup>19</sup> Note that, in contrast to the separable preferences, there is variation in the sum of the elasticities of the marginal utility of consumption (since  $\epsilon_{ch}$  depends on labor  $h$ ), leading to the taxation of capital even in the expected utility case ( $\rho = \gamma$ ), constituting the base of the quantitative experiments of capital taxation of Chari et al. (1994). The variation due to the marginal utility channel would be absent if  $\rho = 1$  ( $\epsilon_{ch} = 0$ , since then (74) becomes a subcase of the separable preferences (70)) and we would have only the novel variation in  $\tilde{\Phi}_t$  as a channel for the taxation of capital. The labor tax for  $t \geq 1$  becomes<sup>20</sup>

<sup>19</sup>The Frisch elasticity is  $\varepsilon_F(h) = \frac{1-h}{h} \frac{1-\theta(1-\rho)}{\rho}$ .

<sup>20</sup>At  $t = 0$  we have

$$\tau_0^H = \tilde{\Phi}_0 \frac{1 + \frac{h_0}{1-h_0} + (1 - \tau_0^K) \frac{F_{KH}}{F_H} k_0 - c_0^{-1} W_0}{1 + \tilde{\Phi}_0 \left[ 1 + \theta(1 - \rho) + (\theta + (1 - \theta)\rho) \frac{h_0}{1-h_0} - \theta(1 - \rho) c_0^{-1} W_0 \right]}.$$

$$\tau_t^H = \tilde{\Phi}_t \frac{1 + \frac{h_t}{1-h_t}}{1 + \tilde{\Phi}_t \left[ 1 + \theta(1-\rho) + (\theta + (1-\theta)\rho) \frac{h_t}{1-h_t} \right]}, \quad (75)$$

which for  $\rho = 1$  simplifies further to

$$\tau_t^H = \tilde{\Phi}_t \frac{1 + \frac{1}{\varepsilon_F(h_t)}}{1 + \tilde{\Phi}_t \left[ 1 + \frac{1}{\varepsilon_F(h_t)} \right]}, \quad (76)$$

with  $\varepsilon_F$  standing for the Frisch elasticity,  $\varepsilon_F(h) = \frac{1-h}{h}$ . This formula could have been seen also from (71), since  $\rho = 1$  restores separability.

## 7 Numerical results

[Incomplete]

Numerical computation. The endogenous grid method (see Barillas and Fernandez-Villaverde (2007)) is promising for the solution of the Chari et al. (1994) and may be also useful for the current problem.

Production function:

$$y_t = k_t^\alpha (A_t h_t)^{1-\alpha} \quad (77)$$

Main numerical examples:

1. Preferences of Chari et al. (1994) with  $\rho = 1$ . For  $\rho = \gamma$ , this leads to a zero ex-ante capital tax rate and variation in labor taxes due to variation in the Frisch elasticity only.
2. Explore also the case of *perfect tax smoothing* with constant Frisch elasticity.
3. Growth
  - deterministic trend as in Chari et al. (1994) and Benigno and Woodford (2006)
  - stochastic trend as in Thomas D. Tallarini (2000) and others.

[To be completed]

## 8 Concluding remarks

[To be written]

## A Optimality conditions

Define the Lagrangian of the Ramsey problem

$$\begin{aligned}
L = & \left[ (1 - \beta)u(c, 1 - h)^{1-\rho} + \beta \left[ \sum_{s'} \pi(s'|s) H(z'_{s'}, k', s')^{1-\gamma} \right]^{\frac{1-\rho}{1-\gamma}} \right]^{\frac{1}{1-\rho}} \\
& - \Phi \left[ z - \Omega(c, h) - \beta \sum_{s'} \pi(s'|s) \frac{H(z'_{s'}, k', s')^{\rho-\gamma}}{\left[ \sum_{s'} \pi_{s'|s} H(z'_{s'}, k', s')^{1-\gamma} \right]^{\frac{\rho-\gamma}{1-\gamma}}} z'_{s'} \right] \\
& - \lambda [c + k' - (1 - \delta)k + g_s - F(s, k, h)].
\end{aligned}$$

The first-order necessary conditions for an interior solution are

$$\begin{aligned}
c : & \quad [(1 - \beta)u^{1-\rho} + \beta\mu^{1-\rho}]^{\frac{\rho}{1-\rho}} (1 - \beta)u^{-\rho} u_c + \Phi \Omega_c(c, h) = \lambda \\
h : & \quad -[(1 - \beta)u^{1-\rho} + \beta\mu^{1-\rho}]^{\frac{\rho}{1-\rho}} (1 - \beta)u^{-\rho} u_h + \Phi \Omega_h(c, h) = -\lambda F_H(s, k, h) \\
k' : & \quad [(1 - \beta)u^{1-\rho} + \beta\mu^{1-\rho}]^{\frac{\rho}{1-\rho}} \beta \mu^{\gamma-\rho} \sum_{s'} \pi(s'|s) H(z'_{s'}, k', s')^{-\gamma} H_K(z'_{s'}, k', s') + \beta \Phi \frac{\partial \omega}{\partial k'} = \lambda \\
z'_{s'} : & \quad [(1 - \beta)u^{1-\rho} + \beta\mu^{1-\rho}]^{\frac{\rho}{1-\rho}} \beta \mu^{\gamma-\rho} \pi(s'|s) H(z'_{s'}, k', s')^{-\gamma} H_z(z'_{s'}, k', s') + \beta \Phi \frac{\partial \omega}{\partial z'_{s'}} = 0
\end{aligned}$$

Furthermore, the derivatives of  $\omega$  take the form

$$\begin{aligned}
\frac{\partial \omega}{\partial k'} &= (\rho - \gamma) \sum_{s'} \pi(s'|s) m_{s'}^{\prime \frac{\rho-\gamma}{1-\gamma}} H_K(z'_{s'}, k', s') H(z'_{s'}, k', s')^{-1} \eta'_{s'} \\
\frac{\partial \omega}{\partial z'_{s'}} &= \pi(s'|s) (\rho - \gamma) m_{s'}^{\prime \frac{\rho-\gamma}{1-\gamma}} H_z(z'_{s'}, k', s') H(z'_{s'}, k', s')^{-1} \eta'_{s'} + \pi(s'|s) m_{s'}^{\prime \frac{\rho-\gamma}{1-\gamma}}
\end{aligned}$$

where  $\eta'_{s'}$  and  $m'_{s'}$  are defined in the text. Use now the fact that at the optimal solution  $H(z, k, s)^{1-\rho} = (1 - \beta)u^{1-\rho} + \beta\mu^{1-\rho}$  and the tilde notation (35) to get equations (37-42) in the text.

## B Sequential formulation

[Add the sequential formulation here.]

## C Numerical computation

[Details of the numerical solution]

### References

- Bansal, Ravi and Amir Yaron. 2004. Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles. *Journal of Finance* LIX (4):1481–1509.
- Barillas, Francisco and Jesus Fernandez-Villaverde. 2007. A generalization of the endogenous grid method. *Journal of Economic Dynamics and Control* 31 (8):2698–2712.
- Benigno, Pierpaolo and Michael Woodford. 2006. Optimal taxation in an RBC model: A linear-quadratic approach. *Journal of Economic Dynamics and Control* 30 (9-10):1445–1489.
- Berndt, Antje, Hanno N. Lustig, and Sevin Yeltekin. 2010. How does the U.S. government finance fiscal shocks? Mimeo, Carnegie Mellon University and UCLA.
- Chamley, Christophe. 1986. Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica* 54 (3):607–622.
- Chari, V.V., Lawrence J. Christiano, and Patrick J. Kehoe. 1994. Optimal Fiscal Policy in a Business Cycle Model. *Journal of Political Economy* 102 (4):617–652.
- Epstein, Larry G. and Stanley E. Zin. 1989. Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica* 57 (4):937–969.
- Hansen, Lars Peter and Thomas J. Sargent. 2001. Robust Control and Model Uncertainty. *American Economic Review* 91 (2):60–66.
- Hansen, Lars Peter, John Heaton, and Nan Li. 2006. Consumption Strikes Back: Measuring Long Run Risk. University of Chicago.
- Judd, Kenneth L. 1985. Redistributive taxation in a simple perfect foresight model. *Journal of Public Economics* 28 (1):59–83.
- Kreps, D. M. and E. L. Porteus. 1978. Temporal Resolution of Uncertainty and Dynamic Choice. *Econometrica* 46:185–200.

- Kydland, Finn E. and Edward C. Prescott. 1980. Dynamic optimal taxation, rational expectations and optimal control. *Journal of Economic Dynamics and Control* 2:79–91.
- Lucas, Robert Jr. and Nancy L. Stokey. 1983. Optimal fiscal and monetary policy in an economy without capital. *Journal of Monetary Economics* 12 (1):55–93.
- Strzalecki, Tomasz. 2010a. Axiomatic Foundations of Multiplier Preferences. *Econometrica* Forthcoming.
- . 2010b. Temporal Resolution of Uncertainty and Recursive Models of Ambiguity Aversion. Mimeo, Harvard University.
- Thomas D. Tallarini, Jr. 2000. Risk-sensitive real business cycles. *Journal of Monetary Economics* 45 (3):507–532.
- Weil, Philippe. 1990. Non-Expected Utility in Macroeconomics. *Quarterly Journal of Economics* CV (1):29–42.
- Zhu, Xiaodong. 1992. Optimal fiscal policy in a stochastic growth model. *Journal of Economic Theory* 58 (2):250–289.