

Karantounias, Anastasios G.

Working Paper

Optimal fiscal policy with recursive preferences

Working Paper, No. 2013-7

Provided in Cooperation with:

Federal Reserve Bank of Atlanta

Suggested Citation: Karantounias, Anastasios G. (2013) : Optimal fiscal policy with recursive preferences, Working Paper, No. 2013-7, Federal Reserve Bank of Atlanta, Atlanta, GA

This Version is available at:

<https://hdl.handle.net/10419/101004>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Optimal Fiscal Policy with Recursive Preferences

Anastasios G. Karantounias

Working Paper 2013-7

September 2013

Abstract: I study optimal capital and labor income taxation in a business cycle model with the recursive preferences of Epstein and Zin (1989) and Weil (1990). In contrast to the case of time-additive expected utility, I find that it is no longer optimal to make the welfare cost of distortionary taxes constant over states and dates. This dramatically alters standard taxation prescriptions: optimal policy calls for taxation at the intertemporal margin, variation of taxation at the intratemporal margin, and persistence of labor taxes independent of the stochastic properties of exogenous shocks. Ignoring the distinction between smoothing over time and smoothing over states is not an innocuous assumption for optimal policy.

JEL classification: D80; E62; H21; H63

Key words: Ramsey plan, Epstein-Zin, recursive utility, risk-sensitive preferences, labor tax, capital tax, martingale

The first draft of this paper was published in April 2010. The author thanks David Backus, Pierpaolo Benigno, Vasco Carvalho, Lawrence Christiano, Kristopher S. Gerardi, Lars Peter Hansen, Juan Pablo Nicolini, Demian Pouzo, Victor Rios-Rull, Richard Rogerson, Thomas J. Sargent, and Stanley E. Zin. He also thanks seminar participants at the University of California at Davis, Carnegie Mellon University, Universitat Pompeu Fabra, the European University Institute, LUISS Guido Carli University, and the University of Hong Kong, and conference participants at the first New York University Alumni Conference, the CRETE 2011 Conference, the 2012 SED Meetings, the 2012 EEA Annual Congress, and the 2013 AEA meetings. He is grateful to Christopher Sleet for his discussion at the AEA Meetings in San Diego. The views expressed here are the author's and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Any remaining errors are the author's responsibility.

Please address questions regarding content to Anastasios Karantounias, Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree Street, N.E., Atlanta, GA 30309-4470, 404-498-8825, anastasios.karantounias@atl.frb.org.

Federal Reserve Bank of Atlanta working papers, including revised versions, are available on the Atlanta Fed's website at frbatlanta.org/pubs/WP/. Use the WebScriber Service at frbatlanta.org to receive e-mail notifications about new papers.

1 Introduction

This paper provides a theoretical analysis of normative fiscal policy in a business cycle model with the recursive preferences of [Epstein and Zin \(1989\)](#) and [Weil \(1990\)](#). 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, where a benevolent planner chooses taxes and debt under commitment in order to maximize the utility of the representative household.

Time and risk play a central role in optimal fiscal policy. The policymaker has to decide whether to tax in the current period or postpone taxation by issuing debt, and how to respond to shocks that affect the government budget constraint. These two dimensions of the policy problem are encoded in asset prices, which are central to the policymaker since they inform him about the desirability of debt and therefore the extent to which he should resort to distortionary taxation.

As is well known from [Epstein and Zin \(1989\)](#) and [Weil \(1990\)](#), standard time-additive expected utility, by forcing the coefficient of risk aversion to be equal to the inverse of the coefficient of intertemporal elasticity of substitution, fails to make a distinction between smoothing over time and smoothing over states and imposes indifference to the temporal resolution of uncertainty. This feature, besides being theoretically unappealing, may also result in an artificially low market price of risk, a fact that has popularized the use of recursive preferences in the macro-finance literature.¹ However, the implications of the distinction between time and risk for the analysis of optimal policy are not known. This is the task of the current paper.

The basic lessons about optimal taxation over the business cycle with time-additive expected utility come from the work of [Chari et al. \(1994\)](#) and [Zhu \(1992\)](#). These studies prescribe for essentially *zero* ex-ante taxation of capital income and essentially *constant* labor taxes. An additional result is that labor taxes inherit the stochastic properties of exogenous shocks, and thus optimal taxes do not constitute a distinct source of persistence in the economy.

These results are overturned in an economy with recursive preferences. There is a novel incentive for taxation at the intertemporal margin and for variation of taxation at the intratemporal margin. Moreover, labor taxes persist independently of the stochastic properties of the exogenous shocks.

The crucial element in the analysis is the *excess burden* of distortionary taxation, a multiplier that captures how the planner wants to allocate distortions over states and dates. The optimal policy prescription with time-additive expected utility is to spread the welfare distortions over states and dates equally by making this multiplier *constant*. Instead, I find that the excess burden of taxation with recursive preferences is *time-varying* and *persistent*, which is the reason

¹See for example [Tallarini \(2000\)](#), [Bansal and Yaron \(2004\)](#) and [Hansen et al. \(2008\)](#).

why the results of [Chari et al. \(1994\)](#) and [Zhu \(1992\)](#) do not hold.

I provide explicit formulas for optimal capital and labor taxes that clarify the role of preference parameters and decompose the incentives for taxation in terms of elasticities of *period* marginal utilities and the time-varying excess burden. In particular, optimal ex-ante capital taxation depends on the *change* of elasticities and the *change* of the excess burden of taxation, whereas optimal labor taxation depends on the *level* of elasticities and the *level* of the excess burden of taxation. Therefore, with time-additive expected utility, a case which is nested in the paper's framework, the ex-ante capital tax is zero and the labor tax constant when period elasticities are constant. Instead, with recursive preferences, these are exactly the cases where the entire action at the intertemporal and intratemporal margin is coming from the time-varying excess burden of taxation.

The reason why the planner does not want to make the excess burden of distortionary taxation constant becomes clear if we consider the dynamic tradeoffs involved. Facing a fiscal shock, the planner has to decide whether to tax or to postpone taxation to the future by issuing state-contingent debt. An increased amount of debt 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 it 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 excess burden of distortionary taxation. This is true for *both* a deterministic time-additive economy *and* a stochastic time-additive expected utility economy.

However, with recursive preferences, the planner must also consider the impact of increased debt positions on an additional component of the stochastic discount factor, namely continuation values, and acknowledge the resulting price effects. Increased debt leads to a reduction in utility, which increases the prices of state-contingent securities when there is preference for early resolution of uncertainty. This is beneficial in states where the government has a relatively large state-contingent debt position, since selling claims to consumption next period at a high price allows a lower tax rate in the current period. But it is harmful in states where the government has a relatively small state-contingent debt position (or assets), since the government buys claims to consumption at a high price, forcing it to tax more in the current period.

The fact that, with a preference for early resolution of uncertainty, an increase in debt becomes *less* costly, makes the planner *shift* the excess burden and therefore tax distortions towards contingencies for which he issues debt and away from contingencies where he buys assets.² Moreover, government insurance with complete markets typically involves hedging adverse shocks like high government expenditure shocks or low technology shocks by purchasing state-contingent assets, and issuing state-contingent debt for favorable shocks like low

²The opposite holds with a preference for late resolution of uncertainty. Preference for early resolution of uncertainty is typically needed to capture asset-pricing facts, making it empirically the most relevant case.

government expenditures shocks or high technology shocks. Therefore, the incentive to shift tax distortions towards contingencies for which debt is issued results in a *negative* correlation of changes in labor taxes with government expenditures and a *positive* correlation with technology shocks.

I show that the inverse of the excess burden of distortionary taxation is a martingale with respect to a continuation-value adjusted measure, inducing *persistence* to the policy variables and to the optimal allocation. This is in contrast to the results of [Lucas and Stokey \(1983\)](#) who show that in an economy without capital the labor tax, if it varies, inherits the stochastic properties of the exogenous shocks. [Chari et al. \(1994\)](#) have shown quantitatively that this result survives in an economy with capital. With recursive preferences though, optimal labor taxes become persistent independent of the properties of the exogenous shocks.

Policy-induced persistence has also implications for the debt positions of the government. Since optimal policy calls for larger taxes when debt is issued and smaller taxes when assets are purchased, and since taxes are persistent, there is a corresponding relatively large increase in absolute value in the present value of future surpluses and thus an increase in absolute value in debt.

Related literature. The main references on optimal taxation with time-additive expected utility are [Chari et al. \(1994\)](#) and [Zhu \(1992\)](#). The model I build reduces to the standard business cycle model analyzed in these studies, if I equate the risk aversion parameter to the inverse of the intertemporal elasticity of substitution parameter, and to the deterministic economy of [Chamley \(1986\)](#) and [Judd \(1985\)](#), if I shut off uncertainty.³ Whenever appropriate, I show how the analysis specializes to an economy without capital as in [Lucas and Stokey \(1983\)](#).

Other related studies include [Farhi and Werning \(2008\)](#), who analyze the implications of recursive preferences for private information setups and [Karantounias \(2013\)](#), who analyzes optimal taxation in an economy without capital, in a setup where the household entertains fears of misspecification but the fiscal authority does not.

I follow a recursive formulation along the lines of [Kydland and Prescott \(1980\)](#) by keeping track of wealth in marginal utility units (or debt in marginal utility units in an economy without capital), a state variable that is *redundant* in the time-additive expected utility case, assigning therefore a novel role to wealth. This formulation succinctly summarizes the effects of recursive preferences in terms of the excess burden of taxation. There are similarities in spirit with the analysis of risk-sharing under recursive preferences, as in [Anderson \(2005\)](#), which leads to *time-varying* Pareto weights.

Another relevant line of research is the analysis of optimal taxation with time-additive expected utility and restricted asset markets as in [Aiyagari et al. \(2002\)](#), [Farhi \(2010\)](#) and [Sleet and](#)

³It is worth noting that [Chamley](#) demonstrated the generality of the zero capital tax result at the deterministic steady state by using the preferences of [Koopmans \(1960\)](#).

[Yeltekin \(2006\)](#). In these studies the lack of insurance markets also causes the planner to allocate distortions in a time-varying and persistent way. However, the lack of markets implies that the planner tries to decrease taxes and debt when good shocks are realized and increase taxes and debt when bad shocks are realized. Instead, the opposite happens in the current paper with preference for early resolution of uncertainty.⁴ More generally, with incomplete markets, the planner would like to make the excess burden of taxation constant but he *cannot*, whereas with complete markets and recursive preferences he could in principle make it constant, but does not find it optimal to do so.

The paper is organized as follows. Section 2 lays out the economy and section 3 sets up the Ramsey problem, its recursive formulation and derives the associated optimality conditions. The heart of the paper resides in section 4, which analyzes the excess burden of distortionary taxation. The implications for capital and labor taxes are derived in section 5 and a simple illustration is provided in section 6. Finally, section 7 concludes and an Appendix follows.

2 Economy

Time is discrete and the horizon is infinite. As [Chari et al. \(1994\)](#) and [Zhu \(1992\)](#) do, I use the stochastic growth model with distortionary labor and capital income taxation, with the exception of 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, that is captured by s . The variable s takes values in a finite or countable set. Let $s^t = (s_0, s_1, \dots, s_t)$ denote the partial history of shocks till time t and let $\pi_t(s^t)$ denote the probability of this history. The initial shock is assumed to 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 denotes that consumption and leisure at time t are measurable functions of the history s^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 . V_t follows the recursion

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

⁴Furthermore, with incomplete markets, the excess burden of taxation is a martingale with respect to the risk-adjusted measure in contrast to the martingale characterization in this paper.

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 π , and $\phi(\cdot)$ is an increasing and concave 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 $W(\cdot)$.

I focus my analysis on the isoelastic preferences of [Epstein and Zin \(1989\)](#) and [Weil \(1990\)](#) (EZW henceforth), and use a constant elasticity of substitution time aggregator and a power utility certainty equivalent. In particular, EZW preferences take the form

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

where u is assumed to be positive. The parameter $1/\rho$ captures the intertemporal elasticity of substitution between the composite good and the certainty equivalent of continuation utility, whereas the parameter γ represents risk aversion with respect to atemporal gambles in continuation values. 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 ($\rho = \gamma$) there is indifference to the temporal resolution of uncertainty.

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 ρ -transformation.⁵ 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 refer to U and $U_i, i = c, l$ as *period utility* and *period marginal utility of consumption and leisure* 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 ρ -transformation for $\rho = 1$, $v_t = \frac{\ln V_t}{1-\beta}$, we get the recursion

$$v_t = \ln u_t + \frac{\beta}{(1 - \beta)(1 - \gamma)} \ln E_t \exp [(1 - \beta)(1 - \gamma)v_{t+1}], \quad (4)$$

which for $\gamma > 1$ has the interpretation of a *risk-sensitive* recursion with risk-sensitivity pa-

⁵ Applying the respective γ -transformation $f(V) \equiv \frac{V^{1-\gamma}-1}{(1-\beta)(1-\gamma)}$ on (2) delivers the representation used in [Weil \(1990\)](#).

parameter $\sigma \equiv (1 - \beta)(1 - \gamma)$.⁶

It will be useful to define

$$m_{t+1} \equiv \frac{V_{t+1}^{1-\gamma}}{E_t V_{t+1}^{1-\gamma}}, t \geq 0, \quad (5)$$

with $m_0 \equiv 1$. For $\rho = 1$, the corresponding definition is $m_{t+1} = \frac{\exp[(1-\beta)(1-\gamma)v_{t+1}]}{E_t \exp[(1-\beta)(1-\gamma)v_{t+1}]}$. 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

$$M_t(s^t) \equiv \prod_{i=1}^t m_i(s^i), \quad (6)$$

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 π . I refer to $\pi_t \cdot M_t$ as the *continuation-value* adjusted measure.

2.2 Competitive equilibrium

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 at the beginning of period $t + 1$ as function of information at t . The resource constraint in the economy takes the form

$$c_t(s^t) + k_{t+1}(s^t) - (1 - \delta)k_t(s^{t-1}) + g_t(s^t) = F(s_t, k_t(s^{t-1}), h_t(s^t)). \quad (7)$$

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 rate δ and can be rented at rental rate $r_t(s^t)$, pays proportional labor income taxes with rate $\tau_t(s^t)$, capital income taxes with rate $\tau_t^K(s^t)$, and trades in complete asset markets. The household's problem is to 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

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

⁶More generally, in the case of risk-sensitive preferences, the period utility function is not restricted to be logarithmic and the recursion takes the form $v_t = U_t + \frac{\beta}{\sigma} \ln E_t \exp(\sigma v_{t+1})$, $\sigma < 0$. There is an intimate link between the risk-sensitive recursion and the multiplier preferences of Hansen and Sargent (2001) that capture the decision maker's fear of misspecification of the probability model π . See Strzalecki (2011) and Strzalecki (2013) for a decision-theoretic treatment and an analysis of the relationship of ambiguity aversion with the temporal resolution of uncertainty.

subject to

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

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 are given. The variable $b_{t+1}(s^{t+1})$ stands for the holdings at history s^t of an Arrow security that delivers one unit of consumption next period if the state is s_{t+1} and zero units otherwise. This security is traded at the price $p_t(s_{t+1}, s^t)$ in units of the history-contingent consumption $c_t(s^t)$.

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

$$\lim_{t \rightarrow \infty} \sum_{s^t} q_t(s^t) [k_{t+1}(s^t) + \sum_{s_{t+1}} p_t(s_{t+1}, s^t) b_{t+1}(s^{t+1})] \geq 0, \quad (8)$$

where $q_t(s^t) \equiv \prod_{i=0}^{t-1} p_i(s_{i+1}, s^i)$, with the normalization $q_0 \equiv 1$. 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 as

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

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

$$\sum_{s_{t+1}} p_t(s_{t+1}, s^t) R_{t+1}^K(s^{t+1}) = 1. \quad (9)$$

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

The no-Ponzi game condition (8) together with the no-arbitrage condition (9) allows us to derive the intertemporal budget constraint of the household at $t = 0$,

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

Firms. A price-taking firm operates the constant returns to scale technology. 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 budget constraint of the government takes the form

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

where

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

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 asymptotic condition that leads to the intertemporal budget constraint of the government is

$$\lim_{t \rightarrow \infty} \sum_{s^{t+1}} q_{t+1}(s^{t+1}) b_{t+1}(s^{t+1}) \leq 0. \quad (10)$$

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 $\{g, \tau, \tau^K, b\}$ such that:*

- *Given prices $\{p, w, r\}$ and taxes $\{\tau, \tau^K\}$, 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 and the asymptotic condition (10) hold.*

2.3 Characterization of competitive equilibrium

I am now in the position to characterize the competitive equilibrium. The labor supply decision is governed by

$$\frac{U_l(s^t)}{U_c(s^t)} = (1 - \tau_t(s^t)) w_t(s^t), \quad (11)$$

which equates the marginal rate of substitution between consumption and leisure with the after-tax wage. The first-order condition with respect to an Arrow security equates its price to the household's intertemporal marginal rate of substitution,

$$\begin{aligned}
p_t(s_{t+1}, s^t) &= \beta \pi_{t+1}(s_{t+1}|s^t) \left(\frac{V_{t+1}(s^{t+1})}{\mu_t(V_{t+1})} \right)^{\rho-\gamma} \frac{U_c(s^{t+1})}{U_c(s^t)} \\
&= \beta \pi_{t+1}(s_{t+1}|s^t) m_{t+1}(s^{t+1})^{\frac{\rho-\gamma}{1-\gamma}} \frac{U_c(s^{t+1})}{U_c(s^t)},
\end{aligned} \tag{12}$$

where the second line uses the definition of the conditional likelihood ratio (5). The change of measure M_t allows also a concise expression for the price of an Arrow-Debreu contract at $t = 0$, $q_t(s^t) = \beta^t \pi_t(s^t) M_t(s^t)^{\frac{\rho-\gamma}{1-\gamma}} \frac{U_c(s^t)}{U_c(s_0)}$.

Turning to the first-order condition with respect to capital, we have

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

which, together with (12), delivers the no-arbitrage condition.⁷ Furthermore, at the optimum the two asymptotic conditions (8) and (10) have to hold with equality, which lead to two transversality conditions with respect to capital and Arrow securities

$$\lim_{t \rightarrow \infty} \sum_{s^t} \beta^t \pi_t(s^t) M_t(s^t)^{\frac{\rho-\gamma}{1-\gamma}} U_c(s^t) k_{t+1}(s^t) = 0 \tag{13}$$

$$\lim_{t \rightarrow \infty} \sum_{s^{t+1}} \beta^{t+1} \pi_{t+1}(s^{t+1}) M_{t+1}(s^{t+1})^{\frac{\rho-\gamma}{1-\gamma}} U_c(s^{t+1}) b_{t+1}(s^{t+1}) = 0 \tag{14}$$

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

$$S_{t+1} \equiv \beta \left(\frac{V_{t+1}}{\mu_t} \right)^{\rho-\gamma} \frac{U_{c,t+1}}{U_{ct}} = \beta m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} \frac{U_{c,t+1}}{U_{ct}}. \tag{15}$$

The disentanglement of risk aversion and intertemporal elasticity of substitution ($\rho \neq \gamma$) introduces *continuation values* scaled by their certainty equivalent μ_t into the stochastic discount factor. As a result, besides caring for the short-run ($U_{c,t+1}/U_{ct}$), the household cares also for the “long-run”, in the sense that the entire sequence of future consumption and leisure will directly affect its intertemporal marginal rate of substitution today.⁸

⁷In order to derive the household’s optimality conditions we need to calculate the derivatives of the utility index with respect to consumption and labor. The derivative with respect to c_{t+i} can be calculated recursively from the relationship $\frac{\partial V_t}{\partial c_{t+i}} = \frac{\partial V_t}{\partial \mu_t} \frac{\partial \mu_t}{\partial V_{t+1}} \frac{\partial V_{t+1}}{\partial c_{t+i}}$, $i \geq 1$. Similarly for labor. Therefore, we have $\frac{\partial V_0}{\partial c_t} = (1-\beta)V_0^\rho \beta^t \pi_t M_t^{\frac{\rho-\gamma}{1-\gamma}} U_{ct}$ and $\frac{\partial V_0}{\partial h_t} = -(1-\beta)V_0^\rho \beta^t \pi_t M_t^{\frac{\rho-\gamma}{1-\gamma}} U_{lt}$. For the ρ -transformation we have $\frac{\partial v_0}{\partial c_t} = \beta^t \pi_t M_t^{\frac{\rho-\gamma}{1-\gamma}} U_{ct}$ and $\frac{\partial v_0}{\partial h_t} = -\beta^t \pi_t M_t^{\frac{\rho-\gamma}{1-\gamma}} U_{lt}$.

⁸Bansal and Yaron (2004) and Hansen et al. (2008) have explored ways of making the intertemporal profile of consumption quantitatively important in order to increase the market price of risk.

It is crucial to understand how continuation values affect the stochastic discount factor in the case of recursive preferences. Consider the derivative of the stochastic discount factor with respect to V_{t+1} ,

$$\frac{\partial S_{t+1}}{\partial V_{t+1}} = (\rho - \gamma)\beta \frac{U_{c,t+1}}{U_{ct}} m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} V_{t+1}^{-1} [1 - \pi_{t+1}(s_{t+1}|s^t)m_{t+1}].$$

The sign of the derivative of S_{t+1} depends on the preference for *early* or *late* resolution of uncertainty, ($\pi_{t+1}(s_{t+1}|s^t)m_{t+1} < 1$ according to the change of measure). An *increase* in continuation value leads to a *decrease* in the stochastic discount factor in the case of preference for early resolution of uncertainty ($\rho < \gamma$). Therefore, the price of an Arrow security *decreases*. To state it differently, an agent who would have high utility at s_{t+1} would require a higher return in order to hold a claim to one unit of consumption at this contingency. In the case of preference for late resolution of uncertainty ($\rho > \gamma$), an increase in continuation value would lead to an *increase* in the price of the state-contingent claim and therefore to a decrease in the required return.

States next period are interconnected through the certainty equivalent μ_t , which depends positively on continuation values ($\partial\mu_t/\partial V_{t+1} = \pi_{t+1}(s_{t+1}|s^t)(V_{t+1}/\mu_t)^{-\gamma} > 0$). Therefore, an increase in continuation value at state $\bar{s}_{t+1} \neq s_{t+1}$ will affect the stochastic discount factor through the certainty equivalent and therefore the price of an Arrow security at state s_{t+1} . To see that, let $\bar{V}_{t+1} \equiv V_{t+1}(\bar{s}_{t+1}, s^t)$ and compute the derivative of the stochastic discount factor with respect to \bar{V}_{t+1} to get

$$\frac{\partial S_{t+1}}{\partial \bar{V}_{t+1}} = (\gamma - \rho)\beta \frac{U_{c,t+1}}{U_{ct}} m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} \bar{V}_{t+1}^{-1} \pi_{t+1}(\bar{s}_{t+1}|s^t) \bar{m}_{t+1},$$

where \bar{m}_{t+1} corresponds to the likelihood ratio at \bar{s}_{t+1} . Thus, an increase in continuation value at $\bar{s}_{t+1} \neq s_{t+1}$, by increasing the certainty equivalent μ_t , is *increasing* (decreasing) the price of an Arrow security at s_{t+1} when there is preference for early (late) resolution of uncertainty.

Both of these mechanisms of affecting asset prices through continuation values are important for understanding the Ramsey plan. We will see later that they provide the means of optimally affecting the market value of the household's wealth (Arrow securities plus capital) and, as a result, the dynamic tradeoffs that the planner is facing when he tries to minimize the welfare cost of distortionary taxation.

3 Ramsey problem

I formulate 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 (7) 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 consists of the holdings of Arrow securities b_t and capital wealth $R_t^K k_t$,

$$W_t(s^t) \equiv b_t(s^t) + R_t^K(s^t)k_t(s^{t-1}). \quad (16)$$

Given the above definition, we can recast the household's budget constraint in terms of wealth. In particular, note that

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

by using the no-arbitrage condition (9). Therefore, the household's dynamic budget constraint becomes

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

Using now (11) and (12) to eliminate labor taxes and equilibrium prices and multiplying by marginal utility of consumption leads to

$$U_{ct}W_t = \Omega_t + \beta E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} U_{c,t+1}W_{t+1}, \quad (17)$$

where

$$\Omega_t \equiv U_{ct}[c_t - (1 - \tau_t)w_t h_t] = U_{ct}c_t - U_{ht}h_t. \quad (18)$$

The variable Ω_t stands for consumption net of after-tax labor income, in *period* marginal utility of consumption units. Note that Ω_t is a function of consumption and labor only, $\Omega_t = \Omega(c_t, h_t)$. We can summarize this discussion in terms of a proposition:

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

$$\begin{aligned} U_{ct}W_t &= \Omega_t + \beta E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} U_{c,t+1}W_{t+1}, t \geq 1 \\ U_{c0}W_0 &= \Omega_0 + \beta E_0 m_1^{\frac{\rho-\gamma}{1-\gamma}} U_{c,1}W_1 \end{aligned}$$

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 (13) and (14) have to be satisfied. The conditional likelihood ratios m_{t+1} , $t \geq 0$, defined in (5), are endogenously determined by continuation values that follow the recursion (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 convenient for a recursive formulation. Note that the initial tax on capital income τ_0^K acts like a lump-sum tax. I am abstracting from this channel for tax revenues and take τ_0^K as exogenously given.

A consumption-labor-capital allocation (c, h, k) that satisfies the resource constraint (7) and the constraints of proposition 1 can be implemented as a competitive equilibrium by recovering prices $\{p, w, r\}$, tax rates $\{\tau, \tau^K\}$ 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 τ , whereas there is *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 (16). 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 and will be the subject of the subsequent analysis,⁹

⁹The ex-ante tax rate is associated with the ex-post tax rate through the relationship

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

$$\bar{\tau}_{t+1}^K(s^t) \equiv \frac{\sum_{s_{t+1}} p_t(s_{t+1}, s^t) (r_{t+1}(s^{t+1}) + 1 - \delta) - 1}{\sum_{s_{t+1}} p_t(s_{t+1}, s^t) r_{t+1}(s^{t+1})}. \quad (19)$$

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 constraint (7).*

3.2 Recursive formulation

I follow the methodology of [Kydland and Prescott \(1980\)](#) and break the Ramsey problem in two subproblems, which are stated in the Appendix: the problem from period one onward and the initial period problem. For that purpose, let z_t denote the household's wealth in period marginal utility units, $z_t \equiv U_{ct} W_t$, and rewrite the dynamic implementability constraint (17) as

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

It will be useful for later purposes to define $\omega_t \equiv E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} z_{t+1}$. The variable ω appears in the right-hand side of the dynamic implementability constraint and is instrumental in the interpretation of the Ramsey plan. It can be roughly thought of as the *market value of the household's wealth*, since $\omega_t = \frac{U_{ct}}{\beta} E_t S_{t+1} W_{t+1}$.

I represent the commitment problem from period one onward recursively by keeping track of the natural state variables (k_t, s_t) and wealth in marginal utility units z_t , that captures the commitment of the planner to his past promises. Note that wealth in marginal utility units is a forward-looking variable that it not inherited from the past. This creates the need to specify $Z(s, k)$, the space where z lives. The set $Z(s, k)$ represents the values of wealth in marginal utility units that can be *generated* from an implementable allocation when the initial shock is s and capital k and is defined in the Appendix. Let $V(z_1, k_1, s_1)$ denote the value function of the planner's problem from period one onward, where $z_1 \in Z(s_1, k_1)$. Assume that the exogenous shocks follow a Markov process with transition probabilities $\pi(s'|s)$.

Bellman equation. The functional equation that determines the value function $V(\cdot)$ takes the form

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

subject to

$$z = \Omega(c, h) + \beta \sum_{s'} \pi(s'|s) \frac{V(z'_{s'}, k', s')^{\rho-\gamma}}{[\sum_{s'} \pi(s'|s) V(z'_{s'}, k', s')^{1-\gamma}]^{\frac{\rho-\gamma}{1-\gamma}}} z'_{s'} \quad (20)$$

$$c + k' - (1 - \delta)k + g_s = F(s, k, h) \quad (21)$$

$$c, k' \geq 0, h \in [0, 1] \quad (22)$$

$$z'_{s'} \in Z(s', k') \quad (23)$$

The variables k' and $z'_{s'}$ stand for capital and wealth in marginal utility units. The notation $z'_{s'}$ captures the fact that the planner is choosing wealth next period at state s' , which is a consequence of the complete markets assumption.

The nature of the Ramsey problem is fundamentally changed because, in contrast to the case of time-additive utility, the value function shows up in the dynamic implementability constraint and in particular in the determination of the market value of wealth ω . This is due to the fact that continuation values determine the stochastic discount factor as we saw earlier. Thus, there is an essential *non-linearity* in the state variable z in the dynamic implementability constraint.

Initial period problem. The initial value of the forward-looking variable z_1 that was taken as *given* in the formulation of the planner's problem from period one onward is chosen *optimally* in order to maximize the utility of the household at $t = 0$. In this sense, the variable z is a pseudo-state variable, i.e. a jump variable that is treated as a state variable in order to capture the commitment of the planner to the optimal plan devised at the initial period. Furthermore, 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 tax rate on initial capital income τ_0^K . As a result, the overall value of the Ramsey problem and the initial period policy functions (c_0, h_0, k_1, z_1) depend on $(b_0, k_0, s_0, \tau_0^K)$.

3.3 Optimality conditions

It turns out that is easier to derive the optimality conditions of the problem by using the ρ -transformation of the value function, $v(z, k, s) \equiv \frac{V(z, k, s)^{1-\rho}-1}{(1-\beta)(1-\rho)}$. The transformed Bellman equation is stated in the Appendix.

Let Φ and λ be the multipliers on the dynamic implementability constraint and the resource constraint respectively of the transformed problem and let $m'_{s'}$ denote the conditional likelihood ratio, which obviously depends on the value function. Note that at the optimal solution the multipliers will be functions of the state, $\Phi = \Phi(z, k, s)$ and $\lambda = \lambda(z, k, s)$. The first-order necessary conditions for an interior solution at points of differentiability of the value function

are

$$c : U_c + \Phi \Omega_c = \lambda \quad (24)$$

$$h : -U_l + \Phi \Omega_h = -\lambda F_H \quad (25)$$

$$k' : \lambda = \beta \sum_{s'} \pi(s'|s) m_{s'}^{\frac{\rho-\gamma}{1-\gamma}} v_k(z'_{s'}, k', s') + \beta \Phi \frac{\partial \omega}{\partial k'} \quad (26)$$

$$z'_{s'} : \pi(s'|s) m_{s'}^{\frac{\rho-\gamma}{1-\gamma}} v_z(z'_{s'}, k', s') + \Phi \frac{\partial \omega}{\partial z'_{s'}} = 0 \quad (27)$$

The variables $\Omega_i, i = c, h$ stand for the partial derivatives of Ω with respect to consumption and labor. The derivatives of the market value of wealth ω with respect to capital and wealth in marginal utility units in (26) and (27) take the form

$$\frac{\partial \omega}{\partial k'} = (1 - \beta)(\rho - \gamma) \sum_{s'} \pi(s'|s) m_{s'}^{\frac{\rho-\gamma}{1-\gamma}} v_k(z'_{s'}, k', s') \eta'_{s'} \quad (28)$$

$$\frac{\partial \omega}{\partial z'_{s'}} = \pi(s'|s) m_{s'}^{\frac{\rho-\gamma}{1-\gamma}} [1 + (1 - \beta)(\rho - \gamma) v_z(z'_{s'}, k', s') \eta'_{s'}], \quad (29)$$

where

$$\eta'_{s'} \equiv V_{s'}^{\rho-1} z'_{s'} - \mu^{\rho-1} \sum_{s'} \pi(s'|s) m_{s'}^{\frac{\rho-\gamma}{1-\gamma}} z'_{s'}. \quad (30)$$

The variable $\eta'_{s'}$ is crucial for the interpretation of the Ramsey plan and will be analyzed in detail later. The variable μ stands for the certainty equivalent and $V'_{s'}$ is shorthand for $V(z'_{s'}, k', s')$.¹⁰ Using (28) and (29), the first-order conditions with respect to k' and $z'_{s'}$ are finally becoming

$$k' : \lambda = \beta \sum_{s'} \pi(s'|s) m_{s'}^{\frac{\rho-\gamma}{1-\gamma}} v_k(z'_{s'}, k', s') [1 + (1 - \beta)(\rho - \gamma) \eta'_{s'} \Phi] \quad (31)$$

$$z'_{s'} : v_z(z'_{s'}, k', s') + \Phi [1 + (1 - \beta)(\rho - \gamma) v_z(z'_{s'}, k', s') \eta'_{s'}] = 0. \quad (32)$$

The initial period optimality conditions which determine the initial allocation (c_0, h_0, k_1) and the optimal initial value of the state variable z_1 as functions of $(b_0, k_0, s_0, \tau_0^K)$ are stated in the Appendix.

¹⁰I use the non-transformed value function V (which is equal to $[1 + (1 - \beta)(1 - \rho)v]^{\frac{1}{1-\rho}}$) in the definition of $\eta'_{s'}$ as matter of convenience, since it allows a more compact exposition of the first-order conditions.

Envelope conditions. The envelope conditions for the endogenous state variables are

$$v_z(z, k, s) = -\Phi \quad (33)$$

$$v_k(z, k, s) = \lambda(1 - \delta + F_K). \quad (34)$$

Note that $\Phi \geq 0$ and that $\lambda > 0$ so that $v_z(z, k, s) \leq 0$ and $v_k(z, k, s) > 0$.¹¹

Economy without capital. It is easy to see how the analysis specializes to an economy without capital as in [Lucas and Stokey \(1983\)](#). In that case, the resource constraint reads $c_t(s^t) + g_t(s^t) = a_t(s^t)h_t(s^t)$, where $a_t(s^t)$ captures the technology shock. Following the same steps as in the economy with capital, under Markov shocks we can express the commitment problem recursively by keeping as a state variable (z, s) and solving for a value function $v(z, s)$. Since there is no capital, z_t is defined as *debt* in marginal utility of consumption units, $z_t \equiv U_{ct}b_t$ and Ω_t , which is consumption net of after-tax labor income in marginal utility units, is also equal to the primary surplus of the government in marginal utility units. Furthermore, ω captures the market value of the *government* portfolio of Arrow securities in an economy without capital. So the recursive formulation goes through with a dynamic implementability constraint as in (20), a resource constraint $c + g_s = a_s h$ and sets Z that depend only s . At time zero the value of the pseudo-state variable z_1 is optimally chosen, given the initial realization of shocks s_0 and the initial government debt b_0 . As a result, the respective first-order conditions for period one onward are (24), (25), (32), with $F_H = a_s$.

4 Excess burden of distortionary taxation

The main object of the analysis is the multiplier Φ , which captures the shadow cost of the constraints that the competitive equilibrium imposes in the second-best world. In particular, as the envelope condition (33) shows, Φ captures the cost of an additional unit of wealth in marginal utility units. It is a cost, because increases in wealth (debt and capital) have to be accompanied by an increase in distortionary taxation. In a first-best world with lump-sum taxes available, Φ would be *zero*. For that reason, I refer to it as the *excess burden* of distortionary taxation.

¹¹I am implicitly assuming that the government has access to lump-sum transfers, so that the dynamic implementability constraint takes the form $z_t \leq \Omega_t + \beta\omega_t$. Furthermore, note that the definition of the sets $Z(s, k)$ can be sharpened by considering as a lower bound negative positions z that involve no transfers and support the first-best allocation.

4.1 Price effect of continuation value

Changes in $z'_{s'}$ affect through continuation values the market value of the household's wealth portfolio ω . To facilitate the analysis, I will drop the "marginal utility" qualifier and will refer to $z'_{s'}$ simply as wealth.

Consider the first-order condition (32), which captures the dynamic tradeoffs that the planner is facing when $z'_{s'}$ is increased, and decompose it into three terms:

$$\underbrace{v_z(z'_{s'}, k', s')}_{\text{MC of increasing } z'_{s'}} + \underbrace{\Phi}_{\text{MB of relaxing IC}} + \underbrace{(1 - \beta)(\rho - \gamma)v_z(z'_{s'}, k', s')\eta'_{s'}}_{\text{EZW term: price effect of } z'_{s'} (+/-)} \Phi = 0. \quad (35)$$

The first-order condition equates the marginal costs and benefits of increasing $z'_{s'}$. An increase of $z'_{s'}$ at shock s' has a marginal cost since it is associated with more taxation in the future ($v_z(z'_{s'}, k', s') < 0$) 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 (35) and the *only* two terms that are relevant if we were in the time-additive expected utility world of Chari et al. (1994) and Zhu (1992) where $\rho = \gamma$. In that case, the optimality condition would reduce to

$$-v_z(z'_{s'}, k', s') = \Phi,$$

which – by using the envelope condition with respect to z (33)– implies that $\Phi'_{s'} = \Phi$ for all values of the state (z, k, s) . Thus, in the case of time-additive expected utility, the planner would optimally make the excess burden of distortionary taxation *constant*. This is the formal result that hides behind the common intuition that the policymaker should spread welfare distortions among states and dates.

However, when we go to the recursive utility case ($\rho \neq \gamma$), an increase of $z'_{s'}$ has a novel effect, which I will call the *price effect* of continuation value, since it affects the market value of the household's wealth ω through the value function. This effect is captured by the third term in (35) and can be either positive or negative, depending on the preference for early (late) resolution of uncertainty ($\rho < (>)\gamma$) and the sign of $\eta'_{s'}$. In order to see the economics clearly, consider again the derivative of ω with respect to z , (29),

$$\frac{\partial \omega}{\partial z'_{s'}} = \pi(s'|s)m_{s'}^{\frac{\rho-\gamma}{1-\gamma}} [1 + (1 - \beta)(\rho - \gamma)v_z(z'_{s'}, k', s')\eta'_{s'}].$$

The first term represents the marginal increase in the market value of wealth by increasing the wealth position, keeping m constant. This is the marginal benefit of increasing $z'_{s'}$, that is present in the time-additive case. The second term denotes the novel marginal change in ω due to the presence of the value function in the stochastic discount factor. Use the definition of η'_s in (30), and decompose the second term as

$$(1 - \beta)(\rho - \gamma)v_z(z'_{s'}, k', s')V_{s'}'^{\rho-1}z'_{s'} - (1 - \beta)(\rho - \gamma)v_z(z'_{s'}, k', s')\mu^{\rho-1}\omega. \quad (36)$$

Consider first the case of preference for early resolution of uncertainty ($\rho < \gamma$). An increase in $z'_{s'}$ reduces continuation value at s' ($v_z(z'_{s'}, k', s') < 0$) and, since $\rho < \gamma$, it *increases* the price of a claim to consumption next period at s' ($(\rho - \gamma)v_z(z'_{s'}, k', s') > 0$). This increase in price is beneficial or costly to the planner depending on the term $V_{s'}'^{\rho-1}z'_{s'}$. Positive wealth positions $z'_{s'} > 0$ imply a benefit since a higher value of these positions *relaxes* the implementability constraint. However, as we noted previously in the analysis of the stochastic discount factor with recursive preferences, a decrease in continuation value at s' will decrease the certainty equivalent and therefore will decrease the price of claims at states $\bar{s}' \neq s'$ when $\rho < \gamma$. This explains the minus in the second term in (36). The benefit or cost of this reduction in asset prices depends on the overall value of the market value of wealth ω . Therefore, the *net* benefit or cost to the planner of affecting the market value of wealth through continuation values depends on the variable $\eta'_{s'} = V_{s'}'^{\rho-1}z'_{s'} - \mu^{\rho-1}\omega$, which explains why I use a separate notation for it.

In particular, note that $\eta'_{s'}$ has the following property:

Lemma 1. (*Innovation property*)

$$\sum_{s'} \pi(s'|s)m'_{s'}\eta'_{s'} = 0$$

Proof. Use the definition of η to get

$$\begin{aligned} \sum_{s'} \pi(s'|s)m'_{s'}\eta'_{s'} &= \sum_{s'} \pi(s'|s)m'_{s'}V_{s'}'^{\rho-1}z'_{s'} - \mu^{\rho-1} \underbrace{\sum_{s'} \pi(s'|s)m'_{s'}}_{=1} \sum_{s'} \pi(s'|s)m'_{s'} \frac{V_{s'}'^{\rho-\gamma}}{\mu^{1-\gamma}} z'_{s'} \\ &= \mu^{\rho-1} \left[\sum_{s'} \pi(s'|s)m'_{s'} \underbrace{\left(\frac{V_{s'}'}{\mu} \right)^{\rho-1}}_{=m'_{s'} \frac{1}{1-\gamma}} z'_{s'} - \omega \right] \\ &= \mu^{\rho-1} \left[\sum_{s'} \pi(s'|s)m'_{s'} \frac{V_{s'}'^{\rho-\gamma}}{\mu^{1-\gamma}} z'_{s'} - \omega \right] = \mu^{\rho-1}[\omega - \omega] = 0. \end{aligned}$$

□

Therefore, $\eta'_{s'}$ can take both positive and negative values. Note that we could see the result above by rewriting $\eta'_{s'}$ as $\eta'_{s'} = V_{s'}^{\rho-1} z'_{s'} - \sum_{s'} \pi(s'|s) m'_{s'} V_{s'}^{\rho-1} z'_{s'}$, so $\eta'_{s'}$ could be interpreted as the conditional *innovation* of $V_{s'}^{\rho-1} z'_{s'}$ under the continuation-value adjusted measure $\pi_t \cdot M_t$. In the case of $\rho = 1$, there is a sharper interpretation of $\eta'_{s'}$ as the conditional innovation of wealth $z'_{s'}$ under the mentioned measure, $\eta'_{s'} = z'_{s'} - \omega$, so its sign would capture if the position $z'_{s'}$ is above or below the market value of the household's wealth portfolio ω . For that reason, I will be referring to $\eta'_{s'}$ as the innovation in wealth or *net wealth position*, even for $\rho \neq 1$.

To summarize, when $\eta'_{s'} > 0$, which can be written as $z'_{s'} > m_{s'}^{\frac{1-\rho}{1-\gamma}} \omega$, then there is a net marginal *benefit* and therefore a *positive* price effect of continuation value. In contrast, when $\eta'_{s'} < 0$, i.e. when $z'_{s'} < m_{s'}^{\frac{1-\rho}{1-\gamma}} \omega$, then there is a marginal *cost* and a *negative* price effect. What is the economic intuition behind this result? For ease of exposition, I will refer to the case of a positive net wealth position as a situation of net *debt*, and to the case of a negative net wealth position as a situation of net *assets*. Consider a situation of net debt. In this case, the increase in asset prices induced by the decrease in continuation value is beneficial because these are situations where the planner is *selling* securities on net. Therefore, the planner benefits marginally by increasing the price at which he sells, leading to a reduction of the return on his liabilities. In the opposite case where we have net assets, the planner is *buying* on net securities at s' . As a result, an increase in their price entails a *cost*, since the return on these assets is becoming low.

In the case of preference for late resolution of uncertainty ($\rho > \gamma$), the direction of the results is *reversed*, since a decrease in continuation value leads to a *decrease* in the price of state-contingent claims. As a result, there is now a marginal cost when $\eta'_{s'} > 0$, and a marginal benefit when $\eta'_{s'} < 0$. The reason behind this outcome is the same though. The planner still has an incentive to reduce the returns when there is net debt and increase returns when there are net assets. However, since an increase in the wealth position $z'_{s'}$ decreases prices, it bears a benefit at states where the planner is purchasing instead of selling securities.

To see the implications of the price effect on the excess burden of taxation, use the envelope condition with respect to $z'_{s'}$ and rewrite the optimality condition in terms of the inverse of Φ (assuming that Φ is not zero, i.e. the state (z, k, s) is not such so that the first-best allocation can be supported) as¹²

$$\frac{1}{\Phi'_{s'}} = \frac{1}{\Phi} + (1 - \beta)(\rho - \gamma)\eta'_{s'},$$

which, using also the optimality condition for the optimal z_1 and turning into sequence notation, becomes

¹²Otherwise, write the optimality condition in sequence notation as $\Phi_{t+1} = \Phi_t / [1 + (1 - \beta)(\rho - \gamma)\eta_{t+1}\Phi_t]$. Thus, if $\Phi_t = 0$, then $\Phi_{t+i} = 0$, $i \geq 0$.

$$\frac{1}{\Phi_{t+1}} = \frac{1}{\Phi_t} + (1 - \beta)(\rho - \gamma)\eta_{t+1}, t \geq 0. \quad (37)$$

We can make statements now about the allocation of the excess burden of taxation among both states and dates. Consider again a preference for early resolution of uncertainty ($\rho < \gamma$), a state s' where $\eta'_{s'} > 0$ and a state \tilde{s}' where $\eta'_{\tilde{s}'} < 0$. Then $\Phi'_{s'} > \Phi > \Phi'_{\tilde{s}'}$. Thus, the fact that the planner can make debt positions less costly and asset positions more profitable makes the planner *shift* taxation (in the sense of the excess burden of taxation) to states where he issues debt on net and away from states where he holds assets on net ($\Phi'_{s'} > \Phi'_{\tilde{s}'}$). This is the source of the variation over states in the excess burden.

Furthermore, as far as the *dynamics* are concerned, we see that for $\eta'_{s'} > 0$, the excess burden of taxation increases with respect to the current one, $\Phi'_{s'} > \Phi$, whereas for $\eta'_{\tilde{s}'} < 0$ it decreases ($\Phi'_{\tilde{s}'} < \Phi$). Thus, a sequence of positive (negative) net wealth positions leads to an increasing (decreasing) sequence of excess burden of taxation over time. When $\rho > \gamma$, the opposite result holds ($\Phi'_{s'} < \Phi < \Phi'_{\tilde{s}'}$), leading the planner to shift taxation *away* from states where there is net debt towards states where there are net assets.

Relationship to shocks. We typically expect that the planner *hedges* government expenditure shocks by taking a negative net wealth position $\eta'_{s'} < 0$ (net assets) for states s' that involve high expenditure shocks, allowing therefore running a deficit and a positive net wealth position $\eta'_{s'} > 0$ (net debt) for s' that involve low expenditure shocks, that is paid back by surpluses. If that is the case, then for $\rho < \gamma$, the excess burden of taxation decreases with high expenditure shocks and increases with low expenditure shocks. This translates, as we will see later, to lower labor taxes for high expenditure shocks and higher labor taxes for low expenditure shocks. In the same vein, we may expect that for high technology shocks the planner is taking a positive net position whereas for low technology shocks he assumes a negative net position. Therefore, and again for $\rho < \gamma$, the excess burden of taxation is increasing for high technology shocks and decreasing for low ones. To conclude, we expect the change in the excess burden of taxation to be *negatively* correlated with government expenditures and *positively* correlated with technology shocks.

Deterministic versus stochastic world. In a deterministic world the net wealth position η_{t+1} is identically equal to zero ($\eta_{t+1} \equiv 0$), which leads to a *constant* excess burden of taxation $\bar{\Phi}$. Furthermore, as analyzed earlier, even under uncertainty but with $\rho = \gamma$ we have $\Phi_t = \Phi_0 \forall t \geq 0$, so the excess burden of taxation would again stay constant at its initial value (albeit different than the relevant one in the deterministic case). Thus, as far as Φ is concerned, there is no *essential* difference between a deterministic world and a stochastic world where $\rho = \gamma$.

Persistence. The net wealth position η_t captures the incentives of the planner to reduce or increase the returns on claims, given the excess burden of taxation of the previous period and therefore becomes the determinant of the *conditional* time-variation of Φ_t . The law of motion (37) indicates that the inverse of the excess burden of taxation at t depends on the *cumulative* net wealth positions $\eta_i, i = 1, \dots, t$, a property which is explained by the fact that all past prices of state-contingent claims change with a change in continuation values at time t . Thus, the excess burden of taxation and therefore the allocation depend on the *past*. Note furthermore that, if there is an *absorbing* state, then η_{t+1} becomes identically zero after the absorbing state is reached, and therefore Φ_t stays *permanently* at the level that it reaches when the absorbing state is hit. More generally, we have

Proposition 2. (*Martingale Characterization*)

The inverse of Φ_t is a martingale with respect to the continuation-value adjusted measure $\pi_t \cdot M_t$ and therefore Φ_t is a submartingale with respect to $\pi_t \cdot M_t$.

Proof. Take conditional expectation in (37) to get

$$E_t m_{t+1} \frac{1}{\Phi_{t+1}} = \frac{1}{\Phi_t} E_t m_{t+1} + (1 - \beta)(\rho - \gamma) E_t m_{t+1} \eta_{t+1} = \frac{1}{\Phi_t},$$

since $E_t m_{t+1} = 1$ and $E_t m_{t+1} \eta_{t+1} = 0$ by lemma 1. Thus $1/\Phi_t$ is a martingale with respect to $\pi_t \cdot M_t$. Furthermore, since the function $f(x) = 1/x$ is convex for $x > 0$, an application of the conditional version of 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 $x_t = 1/\Phi_t$ and use the martingale result to finally get $E_t m_{t+1} \Phi_{t+1} \geq \Phi_t$. \square

The martingale result about the inverse of the excess burden of taxation can be interpreted loosely as an indication of *persistence*. A natural question that arises is about the behavior of Φ_t under the physical measure π . Using the fact that $\text{Cov}_t(m_{t+1}, \Phi_{t+1}) = E_t m_{t+1} \Phi_{t+1} - E_t \Phi_{t+1}$ (since $E_t m_{t+1} = 1$) and the submartingale property of Φ_t , we get that

$$E_t \Phi_{t+1} \geq \Phi_t - \text{Cov}_t(m_{t+1}, \Phi_{t+1}).$$

In principle, we cannot sign definitely the conditional covariance of the excess burden of taxation with the increment to the continuation-value adjusted measure. We can get an idea about its expected sign if we consider the $\rho = 1$ and $\gamma > 1$ case (or more generally risk-sensitive preferences). Conditional on Φ_t , we expect a negative correlation of Φ_{t+1} with government expenditure shocks and a positive correlation with technology shocks. But high government expenditure shocks, since they provide low utility, are associated with a higher conditional probability mass and therefore a higher $m_{t+1} = \exp[(1 - \beta)(1 - \gamma)v_{t+1}] / E_t \exp[(1 - \beta)(1 - \gamma)v_{t+1}]$,

whereas high technology shocks, by increasing utility, lead to a lower m_{t+1} . Therefore, for both types of shocks, we expect the conditional covariance of m_{t+1} and Φ_{t+1} to be *negative*. In that case, the excess burden of taxation exhibits a positive drift also with respect to the physical measure π .

4.2 Recursive versus sequential formulation

Readers accustomed to optimal taxation problems with complete markets may wonder how the excess burden of taxation can be time-varying when there is a *unique* intertemporal budget constraint. In this section I illustrate the sequential formulation of the problem in order to make clear where this result is coming from. The details are relegated to the Appendix.

The intertemporal budget constraint of the household, after expressing tax rates and asset prices in terms of marginal rates of substitution, becomes

$$\sum_{t=0}^{\infty} \sum_{s^t} \beta^t \pi_t(s^t) M_t(s^t)^{\frac{\rho-\gamma}{1-\gamma}} \Omega(c_t(s^t), h_t(s^t)) = U_{c_0} W_0.$$

Since there is a unique intertemporal budget constraint, we can assign a multiplier $\bar{\Phi}$ and proceed with the analysis of the problem. However, the nature of the problem with recursive preferences is altered. Consider for example an increase in consumption. This will not *only* affect $\Omega(c_t, h_t)$ (as in the time-additive case) but it will also affect the unconditional likelihood ratio M_t through continuation values.

In the Appendix I show how to treat this problem by appending additional “implementability” constraints to the Ramsey problem that describe utility recursions and the law of motion of M_t .¹³ The connection between the excess burden of taxation Φ_t and multipliers on continuation values is analyzed, making clear that the time-varying multiplier Φ_t captures the shadow value of the additional implementability constraints that arise even in a complete markets setup. The benefit of the recursive formulation of the commitment problem, besides illuminating obviously that z is the relevant state variable, is to clearly summarize the effects of continuation values in terms of a *time-varying* Φ_t . This allows a clean comparison with the time-additive expected utility case. There are obvious similarities in spirit with the optimal risk-sharing literature with recursive preferences, which finds *time-varying* Pareto weights.¹⁴

¹³In the case of the multiplier preferences of Hansen and Sargent (2001), it is natural to think of the utility recursions as implementability constraints since they correspond to optimality conditions of the malevolent alter-ego of the household, that minimizes the household’s utility subject to a penalty. See Karantounias (2013). This minimization procedure would also emerge naturally if we expressed recursive utility as the variational utility of Geoffard (1996).

¹⁴See for example Anderson (2005) and references therein.

4.3 The novel role of wealth

The time-varying nature of Φ_t is also the deeper reason for keeping z as a state variable:

Proposition 3. (“Wealth matters”) EZW preferences contribute an additional state variable z_t to the optimal taxation problem. With time-additive expected utility $\rho = \gamma$, it would be sufficient to keep track only of (k, s) , because the excess burden of taxation is constant.

Proof. See Appendix. □

The proof relies on the fact that wealth is necessary for the determination of the solution only through the excess burden of taxation. Since, for the case of expected utility, this cost is constant, wealth becomes *redundant* as a state variable, as long as the planner takes into account how the excess burden of taxation affects the household’s utility.¹⁵ Another way to state the above result is that, given a constant excess burden of taxation, wealth (or debt in an economy without capital) is a *by-product* of the planner’s problem. In contrast, with recursive utility, wealth resumes its importance as a state variable, since it becomes an instrument of directly affecting asset prices through the enriched stochastic discount factor.

5 Capital and labor income taxation

In this section I analyze the taxation of the intertemporal and the intratemporal margin.

5.1 Intertemporal wedge: capital tax

Turning to the optimal choice of capital k' , it is clear that capital has a novel effect on ω through the stochastic discount factor. In particular, consider the optimality condition with respect to capital (26), which is rewritten here for convenience,

$$\underbrace{\lambda}_{\text{MC of increasing } k'} = \underbrace{\beta \sum_{s'} \pi(s'|s) m_{s'}^{\frac{\rho-\gamma}{1-\gamma}} v_k(z'_{s'}, k', s')}_{\text{MB of increasing } k'} + \underbrace{\beta \Phi \frac{\partial \omega}{\partial k'}}_{\text{EZW term: price effect of } k' (+/-)}$$

The optimality condition equates the marginal cost of increasing k' by one unit in 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 ω . As can be seen from the derivative of ω with respect to capital (28), for $\rho = \gamma$, this marginal benefit or

¹⁵Similarly, in an economy without capital and time-additive expected utility, we would only need to keep track of the exogenous shocks, which would be a manifestation of the *history-independence* result of [Lucas and Stokey \(1983\)](#).

cost is *absent*, $\partial\omega/\partial k' \equiv 0$. For the recursive utility case though, an increase in capital increases continuation value $v_k > 0$ and therefore decreases prices for $\rho < \gamma$. What matters again is the net wealth position $\eta'_{s'}$. For $\eta'_{s'} > 0$ there would be now a marginal cost, whereas for $\eta'_{s'} < 0$ there would be a marginal benefit. Since capital is not state-contingent, the actual benefit or cost depends on the discounted expected value of the product of v_k and $\eta'_{s'}$, as seen from (28).

Use now sequence notation and the law of motion of Φ_t (37) to replace $1 + (1 - \beta)(\rho - \gamma)\eta_{t+1}\Phi_t$ in (31) with the ratio Φ_t/Φ_{t+1} . This allows us to see explicitly the dependence of the optimal capital decision on the time-varying Φ_t ,

$$\lambda_t = \beta E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} v_k(z_{t+1}, k_{t+1}, s_{t+1}) \frac{\Phi_t}{\Phi_{t+1}}. \quad (38)$$

Furthermore, using the envelope condition with respect to capital (34) to eliminate v_k delivers

$$E_t S_{t+1}^* (1 - \delta + F_{K,t+1}) = 1, \quad (39)$$

where

$$S_{t+1}^* \equiv \beta m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} \frac{\lambda_{t+1}/\Phi_{t+1}}{\lambda_t/\Phi_t}. \quad (40)$$

I will call S_{t+1}^* the *planner's* stochastic discount factor. The variable S_{t+1}^* captures how the planner discounts the *pre-tax* gross return on capital at the *second-best* allocation. S_{t+1}^* contrasts to the market stochastic discount factor $S_{t+1} \equiv \beta m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} U_{c,t+1}/U_{c,t}$. In a first-best world with lump-sum taxes available, we have identically $S_{t+1}^* \equiv S_{t+1}$. The planner's 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 (19) and use (39) to get

$$\bar{\tau}_{t+1}^K = \frac{E_t [S_{t+1} - S_{t+1}^*] (1 - \delta + F_{K,t+1})}{E_t S_{t+1} F_{K,t+1}}. \quad (41)$$

Thus, there is a positive (negative) tax rate on capital income if the numerator of (41) 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 planner's 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 the pre-tax return on capital that is larger (smaller) than the respective non-

centered covariance of the planner's stochastic discount factor with the pre-tax return, leads to a positive (negative) ex-ante tax rate $\bar{\tau}_{t+1}^K > 0$ ($\bar{\tau}_{t+1}^K < 0$).

The planner's stochastic discount factor S_{t+1}^* is associated with the market stochastic discount factor S_{t+1} as follows:

$$\begin{aligned} S_{t+1}^* &= S_{t+1} \frac{\lambda_{t+1}/(\Phi_{t+1}U_{c,t+1})}{\lambda_t/(\Phi_t U_{c,t})} \\ &= S_{t+1} \frac{\frac{1}{\Phi_{t+1}} + \frac{\Omega_{c,t+1}}{U_{c,t+1}}}{\frac{1}{\Phi_t} + \frac{\Omega_{c,t}}{U_{c,t}}}, t \geq 1 \end{aligned}$$

where I used the optimality condition with respect to consumption (24) in the second line. Thus, we have

$$\frac{S_{t+1} - S_{t+1}^*}{S_{t+1}} = \frac{\frac{1}{\Phi_t} - \frac{1}{\Phi_{t+1}} + \frac{\Omega_{c,t}}{U_{c,t}} - \frac{\Omega_{c,t+1}}{U_{c,t+1}}}{\frac{1}{\Phi_t} + \frac{\Omega_{c,t}}{U_{c,t}}}. \quad (42)$$

The difference in the two discount factors and, consequently, a wedge at the intertemporal margin, can be attributed to time variation in Ω_c/U_c and to time variation in Φ_t . Note that Ω_c/U_c can be expressed in terms of elasticities,

$$\frac{\Omega_c}{U_c} = 1 - \epsilon_{cc} - \epsilon_{ch}, \quad (43)$$

where $\epsilon_{cc} \equiv -U_{cc}c/U_c > 0$ and $\epsilon_{ch} \equiv U_{cl}h/U_{cl}$ i.e. the *own* and *cross* elasticity of the period marginal utility of consumption with respect to consumption and labor.¹⁶ Using these elasticities and the expression for the difference in the two discount factors (42) allows us a *decomposition* of the incentives for an intertemporal wedge in terms of *changes* in $1/\Phi_t$ and *changes* in ϵ_{cc} and ϵ_{ch} :

Proposition 4. (*Sources of capital taxation*) *The ex-ante tax rate on capital income $\bar{\tau}_{t+1}^K$, $t \geq 1$ is positive (negative) iff*

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

If $\epsilon_{cc} + \epsilon_{ch}$ is constant, then the only reason for taxing the intertemporal margin comes from variation in the excess burden of taxation Φ_t .

¹⁶Each elasticity is multiplied with minus unity.

Proposition 4 includes as subcases the deterministic economy and the economy of Chari et al. (1994) and Zhu (1992). Furthermore, the criterion for capital taxation holds also for the deterministic and stochastic time-additive case for any period utility U satisfying the standard monotonicity and concavity assumptions, i.e. without being restricted to $U = (u^{1-\rho} - 1)/(1 - \rho)$, $u > 0$. The criterion holds also for the risk-sensitive preferences with parameter $\sigma < 0$, without the logarithmic restriction on the period utility function, i.e. without being confined to treat the risk-sensitive preferences as a subcase of EZW utility for $\rho = 1$.¹⁷

Deterministic economy. Assume that we are in a deterministic economy. Then Φ_t is *constant* and the formula for capital taxation simplifies to¹⁸

$$\tau_{t+1}^K > (<) 0 \quad \text{iff} \quad \epsilon_{cc,t+1} + \epsilon_{ch,t+1} > (<) \quad \epsilon_{cc,t} + \epsilon_{ch,t}.$$

Thus capital income is taxed (subsidized) if the sum of the own and cross period elasticities is increasing (decreasing). A necessary and sufficient condition for a zero capital tax at *every* period from period two onward in the deterministic economy is that the sum of the elasticities of the period marginal utility of consumption is *constant* (which implies that $S_{t+1}^* = S_{t+1}$). If the period utility function is such so that the elasticities are *not* constant for each period, then there is zero tax on capital income at the deterministic *steady state*, where the constancy of the consumption-labor allocation delivers constant elasticities. This delivers the zero-tax result of Chamley (1986) and Judd (1985).

Stochastic economy, time-additive expected utility. Turn now to the stochastic case of Chari et al. (1994) and Zhu (1992) with time-additive expected utility ($\rho = \gamma$), where Φ_t is again constant. The formula in proposition 4 becomes

$$\bar{\tau}_{t+1}^K > (<) 0 \quad \text{iff} \quad E_t S_{t+1} (1 - \delta + F_{K,t+1}) \left[\epsilon_{cc,t+1} + \epsilon_{ch,t+1} - \epsilon_{cc,t} - \epsilon_{ch,t} \right] > (<) 0$$

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 factor, the pre-tax gross return on capital and the change in the sum of period elasticities, due to the fact that only the ex-ante tax rate is determined by the allocation. Note that variation in the sum

¹⁷In that case, conditional likelihood ratios read $m_{t+1} = \exp(\sigma v_{t+1})/E_t \exp(\sigma v_{t+1})$, the law of motion of Φ_t remains the same by replacing $(1 - \beta)(\rho - \gamma)$ with σ in (37) and the planner's stochastic discount factor becomes $S_{t+1}^* = \beta m_{t+1} \frac{\lambda_{t+1}/\Phi_{t+1}}{\lambda_t/\Phi_t}$. The optimality condition with respect to capital reads $\lambda_t = \beta E_t m_{t+1} v_k(z_{t+1}, k_{t+1}, s_{t+1}) [1 + \sigma \eta_{t+1} \Phi_t]$, where $\eta_{t+1} = z_{t+1} - E_t m_{t+1} z_{t+1}$.

¹⁸Without uncertainty there is obviously no difference between the ex-ante and ex-post tax rate on capital income, so I am dropping the bar notation for that case.

of period elasticities is a *necessary* condition for an intertemporal wedge, since a constant sum of period elasticities implies $S_{t+1}^* = S_{t+1}$. It is not a sufficient condition, since the integration necessary for the determination of the ex-ante tax rate could still in principle deliver a zero tax.

Stochastic economy, $\rho \neq \gamma$. For the case of EZW preferences, the full version of the formula in proposition 4 applies. The change in *both* the excess burden of taxation *and* the sum of elasticities determines the intertemporal wedge. Consider the case of constant period elasticities, which would deliver a zero tax in the case when the distinction between time and risk is absent, as analyzed in the previous paragraph. With recursive preferences though, there is a novel source of taxation coming from changes in Φ_t . States where there are debt positions on net ($\eta_{t+1} > 0$), lead to a higher excess burden of taxation in the case of preference for early resolution of uncertainty, $\Phi_{t+1} > \Phi_t$, and therefore to a planner's discount factor that is smaller than the market discount factor, $S_{t+1}^* < S_{t+1}$, as can be seen from (42). So an increase in Φ_t captures an incentive to introduce a *positive* intertemporal wedge. In contrast, at the states where there are assets on net ($\eta_{t+1} < 0$), we have $\Phi_{t+1} < \Phi_t$ and therefore a planner's discount factor that is larger than the market discount factor, $S_{t+1}^* > S_{t+1}$, leading to an incentive to introduce a *negative* wedge. The final answer about the sign of the ex-ante tax rate depends on the integration of the product of the market discount factor, the pre-tax return and the change in $1/\Phi_t$, so $\bar{\tau}_{t+1}^K > (< 0)$ if and only if $E_t S_{t+1} (1 - \delta + F_{K,t+1}) (1 - \beta) (\gamma - \rho) \eta_{t+1} > (<) 0$. The smaller (larger) in absolute value the negative net wealth positions η_{t+1} are relative to the respective positive ones, the more probable it is to have an ex-ante tax (subsidy) in the case of preference for early resolution of uncertainty.

5.2 Intratemporal wedge: labor tax

Turning to the intratemporal wedge, we can eliminate λ and combine the first-order conditions with respect to consumption and labor (24)-(25) so as to derive the optimal wedge in labor supply

$$\frac{U_l}{U_c} \cdot \frac{1 - \Phi \frac{\Omega_h}{U_l}}{1 + \Phi \frac{\Omega_c}{U_c}} = F_H. \quad (44)$$

Expressing also the term Ω_h/U_l in terms of elasticities delivers

$$\frac{\Omega_h}{U_l} = -1 - \epsilon_{hh} - \epsilon_{hc}, \quad (45)$$

where $\epsilon_{hh} \equiv -U_{ll}h/U_l > 0$ 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. The period elasticities are helpful for

deriving an explicit formula for the optimal labor tax. In particular,¹⁹

Proposition 5. (*Labor tax*) *The optimal labor tax is*

$$\tau_t = \Phi_t \frac{\epsilon_{cc,t} + \epsilon_{ch,t} + \epsilon_{hh,t} + \epsilon_{hc,t}}{1 + \Phi_t(1 + \epsilon_{hh,t} + \epsilon_{hc,t})}, \quad t \geq 1.$$

Proof. Use the labor supply condition $U_l/U_c = (1 - \tau)F_H$ in order to express (44) in terms of the labor tax as $\tau = -\Phi(\Omega_c/U_c + \Omega_h/U_l)/(1 - \Phi\Omega_h/U_l)$. Use now the elasticity formulas (43) and (45) to get the result. \square

The same comment about the applicability of the capital taxation criterion in proposition 4 can be made for the labor tax formula in proposition 5, i.e. it holds also for the deterministic, stochastic time-additive case and the risk-sensitive case, for more general period utility functions U . The formula holds also for the respective economies without capital.

The formula in proposition 5 expresses the optimal labor tax in terms of curvature properties of the period marginal utility of consumption and leisure and in terms of the excess burden of taxation Φ_t . As is clear from the formula, the variation of the labor tax is due either to variation in the period elasticities *or* due to variation in Φ_t , a stochastic process that follows the law of motion (37). Note also that when $U_{cl} \geq 0$, the cross elasticities become non-negative $\epsilon_{ch}, \epsilon_{hc} \geq 0$ and, as a result, the labor tax is positive. Furthermore, in the case of constant elasticities, the formula shows that the labor tax varies monotonically with the excess burden of taxation.²⁰

To conclude, there is an intertemporal wedge if there is a *change* in period elasticities and a *change* in the excess burden of taxation whereas the intratemporal wedge depends on the *level* of the period elasticities and the *level* of the excess burden of taxation. Thus, focusing on the time-varying excess burden of taxation –the element which recursive utility contributes– we infer that the capital tax depends on the net wealth position η_t , whereas the labor tax depends on the cumulative net wealth position $\sum_i \eta_i$.

The analysis of capital and labor taxation up to this point has not taken a stance on the composite good u . In the following section, I provide examples of u that deliver explicit formulas for the period elasticities $\epsilon_{ij}, i, j = c, h$.

5.3 Examples of aggregators u

Consider two specifications of the composite good u .

¹⁹The optimal labor tax at $t = 0$ is different due the presence of initial debt and capital,

$$\tau_0 = \Phi_0 \frac{\epsilon_{cc} + \epsilon_{ch} + \epsilon_{hh} + \epsilon_{hc} + (1 - \tau_0^K) \frac{F_{KH}}{F_H} k_0 - (\epsilon_{cc} + \epsilon_{hc}) c_0^{-1} W_0}{1 + \Phi_0 (1 + \epsilon_{hh} + \epsilon_{hc} - \epsilon_{hc} c_0^{-1} W_0)}. \quad (46)$$

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

²⁰We have $\frac{\partial \tau}{\partial \Phi} = \frac{\epsilon_{cc} + \epsilon_{ch} + \epsilon_{hh} + \epsilon_{hc}}{[1 + \Phi(1 + \epsilon_{hh} + \epsilon_{hc})]^2} > 0$, as long as the numerator is positive. $U_{cl} \geq 0$ is sufficient for that.

5.3.1 Power utility in consumption and separability (zero capital tax for expected utility)

Consider preferences that are separable between consumption and leisure and isoelastic in consumption with parameter ρ ,²¹

$$u(c, 1 - h) = [c^{1-\rho} - (1 - \rho)v(h)]^{\frac{1}{1-\rho}}, \quad (47)$$

where $v', v'' > 0$. For specification (47) we get the own elasticities $\epsilon_{cc} = \rho$ and $\epsilon_{hh} = v''(h)h/v'(h) > 0$ 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 $\varepsilon_F(h) = 1/\epsilon_{hh} = 1/\frac{v''(h)h}{v'(h)} > 0$.

Chari et al. (1994) and Zhu (1992) have shown that these preferences imply a *zero* ex-ante tax rate on capital income from period two onward. This is easily interpreted in terms of the formula in proposition 4, since the own and the cross elasticity of the marginal utility of consumption are *constant*, leading therefore to $S_{t+1}^* = S_{t+1}$ and $\bar{\tau}_{t+1}^K = 0$. For the recursive utility case though, this is a case where taxation of the intertemporal margin is *only* due to variation in the excess burden of taxation.

The labor tax formula in proposition 5 specializes to

$$\tau_t = \Phi_t \frac{\rho + \frac{1}{\varepsilon_F(h_t)}}{1 + \Phi_t(1 + \frac{1}{\varepsilon_F(h_t)}), t \geq 1, \quad (48)$$

which provides a convenient interpretation of the elasticity variation in terms of the Frisch elasticity. Some remarks are due here. Consider for a moment the constant Φ case of the deterministic or stochastic but time-additive economy and let the Frisch elasticities be such that $\varepsilon_F(h) > \varepsilon_F(h'), h \neq h'$. Then the respective tax rates will be $\tau < \tau'$, i.e. the optimal tax is negatively related to the elasticity of labor supply.²² High labor is typically associated with high government expenditure shocks, so if the Frisch elasticity is inversely related to labor, we expect the tax rate to be high for high government expenditures shocks.

Turning now to the recursive utility case and in particular to $\rho < \gamma$, then, as we analyzed earlier, the excess burden of taxation increases (decreases) for a net debt (net asset) position. As a result we expect that the tax rate will increase (decrease). Therefore, if the planner hedges high government expenditures with a net asset position, then, on the one hand, it has an incentive

²¹It is assumed that parameters are such that $c^{1-\rho} - (1 - \rho)v(h) > 0$. Otherwise, we can just consider the risk-sensitive case and drop the non-negativity restriction. We can see in a clearer way the separability between consumption and leisure by applying the ρ -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 $V_t = \exp[(1 - \beta)(\ln c - v(h)) + \beta \ln \mu_t]$.

²²The derivative of the tax rate with respect to the Frisch elasticity is $\frac{\partial \tau}{\partial \varepsilon_F} = -\frac{\Phi}{\varepsilon_F^2} \cdot \frac{1 + \Phi(1 - \rho)}{(1 + \Phi(1 + \frac{1}{\varepsilon_F}))^2}$. The sign of $1 + \Phi(1 - \rho)$ can be found from the first-order condition with respect to c , (24), which becomes for these preferences $c^{-\rho}(1 + \Phi(1 - \rho)) = \lambda > 0$. So $1 + \Phi(1 - \rho) > 0$ even for $\rho > 1$. Thus, $\partial \tau / \partial \varepsilon_F < 0$.

to decrease the tax rate in order to increase returns, and on the other hand, it has the incentive to increase the tax rate due to the low Frisch elasticity at these contingencies, leading to two opposite forces coming from the two sources of taxation at the intratemporal margin.

Constant Frisch elasticity. The differences in optimal taxation with recursive utility become stark when $v(h) = a_h \frac{h^{1+\phi_h}}{1+\phi_h}$, which, implies a *constant* Frisch elasticity $\epsilon_F = 1/\phi_h$. For the time-additive case, not only we have a zero capital tax but also a *constant* labor tax, leading to *perfect* tax smoothing, $\tau = \frac{\Phi(\rho+\phi_h)}{1+\Phi(1+\phi_h)}$. For $\rho \neq \gamma$ though, we have:

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

1. *The labor tax takes the form*

$$\frac{1}{\tau_{t+1}} = \frac{1}{\tau_t} + \frac{(1-\beta)(\rho-\gamma)}{\rho+\phi_h} \eta_{t+1}, t \geq 1. \quad (49)$$

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 $\tau_{t+1} > \tau_t$ (because Φ_{t+1} increases)*
- *if $\eta_{t+1} < 0$, then $\tau_{t+1} < \tau_t$ (because Φ_{t+1} decreases)*

3. (*Martingale Characterization*) *The inverse of the labor tax $1/\tau_t$ is a martingale with respect to the measure $\pi_t \cdot M_t$ and therefore, τ_t is a submartingale with respect to $\pi_t \cdot M_t$. Furthermore, $E_t \tau_{t+1} \geq \tau_t - \text{Cov}_t(m_{t+1}, \Phi_{t+1})$.*

Proof. The labor tax becomes $\tau_t = \Phi_t(\rho + \phi_h)/(1 + \Phi_t(1 + \phi_h))$ with inverse

$$\frac{1}{\tau_t} = \frac{1 + \phi_h}{\rho + \phi_h} + \frac{1}{\rho + \phi_h} \frac{1}{\Phi_t}.$$

Note that $1/\tau_t$ is an *affine* function of $1/\Phi_t$. Use the law of motion of Φ_t in (37) to write the law of motion of the labor tax as in (49). Notice the close resemblance of the law of motion of the labor tax (49) to the law of motion of the excess burden of distortionary taxation (37), a fact that leads to similar conclusions about monotonicity and martingale-like properties as in proposition 2 and the discussion thereafter. \square

The above period utility function is an example of the constant elasticity case. As a result, the labor tax, instead of being constant (or more generally instead of inheriting the stochastic properties of the exogenous shocks), inherits the stochastic properties of Φ_t .

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

Preferences (47) do not allow for balanced growth in a deterministic version of our economy, unless $\rho = 1$. Consider now a Cobb-Douglas aggregator

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

where $\theta \in (0, 1)$. These preferences are a workhorse for the real business cycle literature and, unless $\rho = 1$, feature non-separabilities between consumption and leisure. In contrast to the separable preferences, there is variation in the sum of the elasticities of the marginal utility of consumption, leading to taxation of capital even in the expected utility case ($\rho = \gamma$).²³ The variation due to the period elasticity channel is absent if $\rho = 1$ and we would have only the novel variation in Φ_t as a channel for the taxation of capital. The labor tax for $t \geq 1$ becomes

$$\tau_t = \Phi_t \frac{1 + \frac{h_t}{1-h_t}}{1 + \Phi_t \left[1 + \theta(1 - \rho) + (\theta + (1 - \theta)\rho) \frac{h_t}{1-h_t} \right]},$$

which for $\rho = 1$ simplifies further to formula (48) with $\varepsilon_F(h) = (1 - h)/h$.

The formulas for both examples of the composite good show that the crucial parameter for the period elasticities channel is ρ (and not γ), whereas both ρ and γ affect the Ramsey outcome through Φ_t .

6 A simple illustration

The evolution of the excess burden and therefore the analytic formulas for taxes presented in earlier sections hinge on the sign of the net wealth positions η_{t+1} , i.e. on the government's insurance against adverse shocks. In this section, I assume a simplified stochastic structure - deterministic except for one period- with only fiscal shocks. This setup is rich enough to capture both the variation in the excess burden and its dynamics. Details are relegated to the Appendix.

²³The own and cross elasticities of period marginal utility of consumption are $\epsilon_{cc} = 1 - \theta(1 - \rho)$ and $\epsilon_{ch} = (1 - \theta)(1 - \rho)h/(1 - h)$. The elasticities of the period marginal disutility of labor are $\epsilon_{hh} = (\theta + (1 - \theta)\rho)h/(1 - h)$ and $\epsilon_{hc} = \theta(1 - \rho)$. The dependence of ϵ_{ch} on labor h vanishes when $\rho = 1$ (and ϵ_{ch} becomes zero), leading to zero taxation of capital income in the time-additive expected utility case. The Frisch elasticity is $\varepsilon_F(h) = \frac{1-h}{h} \frac{1-\theta(1-\rho)}{\rho}$.

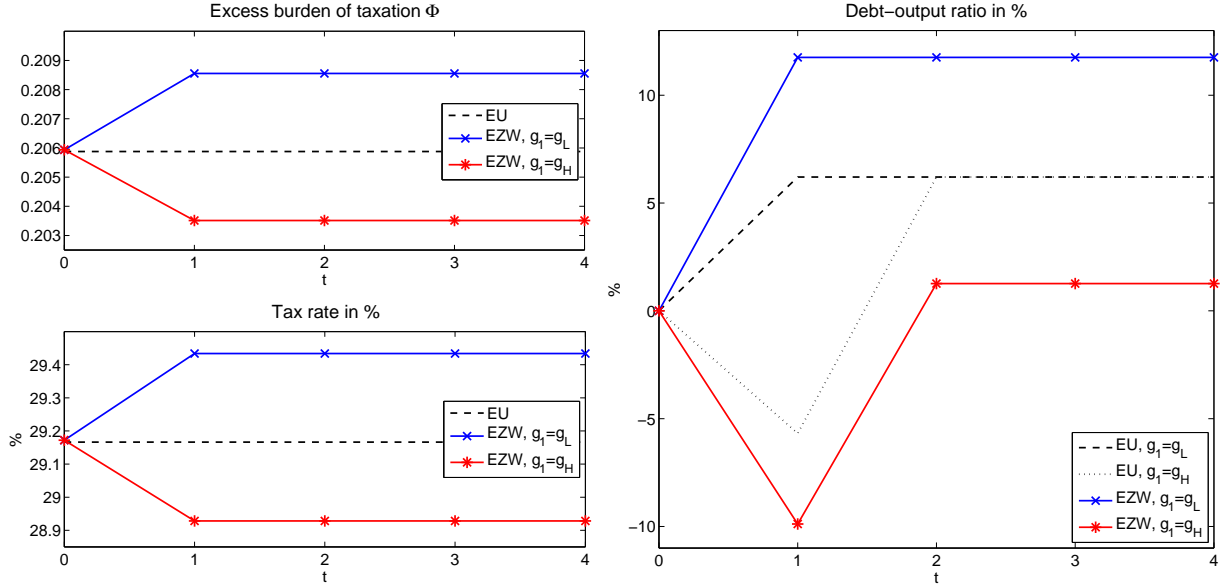


Figure 1: The left panel shows the evolution of the excess burden of taxation and the labor tax along the two histories for the expected utility (EU, $\gamma = 1$) and the recursive utility (EZW, $\gamma = 10$) case. The right panel shows the respective debt positions normalized by output.

6.1 Economy without capital

Assume there is an economy without capital and only fiscal shocks, with resource constraint $c_t + g_t = h_t$, zero initial debt, and consider the case of $\rho = 1$ and $\gamma > 1$.²⁴ Assume that government expenditures take the value g_L with certainty except for period one. At period one we have $g_1 = g_H$ with probability π and $g_1 = g_L$ with probability $1 - \pi$, where $g_H > g_L$. I use superscripts for the endogenous variables in order to denote if we are in the high-shock history ($g_1 = g_H$) or in the low-shock history ($g_1 = g_L$). For example, $c_t^i, i = H, L$, denotes consumption at period $t \geq 1$ when the shock at $t = 1$ is high or low respectively. Similar notation holds for the rest of the variables.

With time-additive expected utility, we are in the environment of [Lucas and Stokey \(1983\)](#), where consumption is function of the shock g and the constant $\bar{\Phi}$, $c_t = c(g_t, \bar{\Phi})$. Therefore, consumption is the same for all nodes of the tree except for the high-shock realization at $t = 1$, i.e. $c_t = c(g_L, \bar{\Phi}) \forall t \neq 1$, $c_1^L = c(g_L, \bar{\Phi})$ and $c_1^H = c(g_H, \bar{\Phi})$. As a result, consumption, and correspondingly labor and the tax rate, revert after period one to their initial value after a high shock realization.

Instead, with recursive utility, we can think of consumption as function of the shock and the time-varying Φ_t , $c_t = c(g_t, \Phi_t)$. In order to determine Φ_t , we need to figure out the debt positions of the government at the beginning of $t = 1$. Note at first that, due to the absence

²⁴Or more generally, consider risk-sensitive preferences with parameter $\sigma < 0$. The analysis of this example does not depend on the functional form of U , as long as U is such so that it is optimal to impose a tax instead of a subsidy.

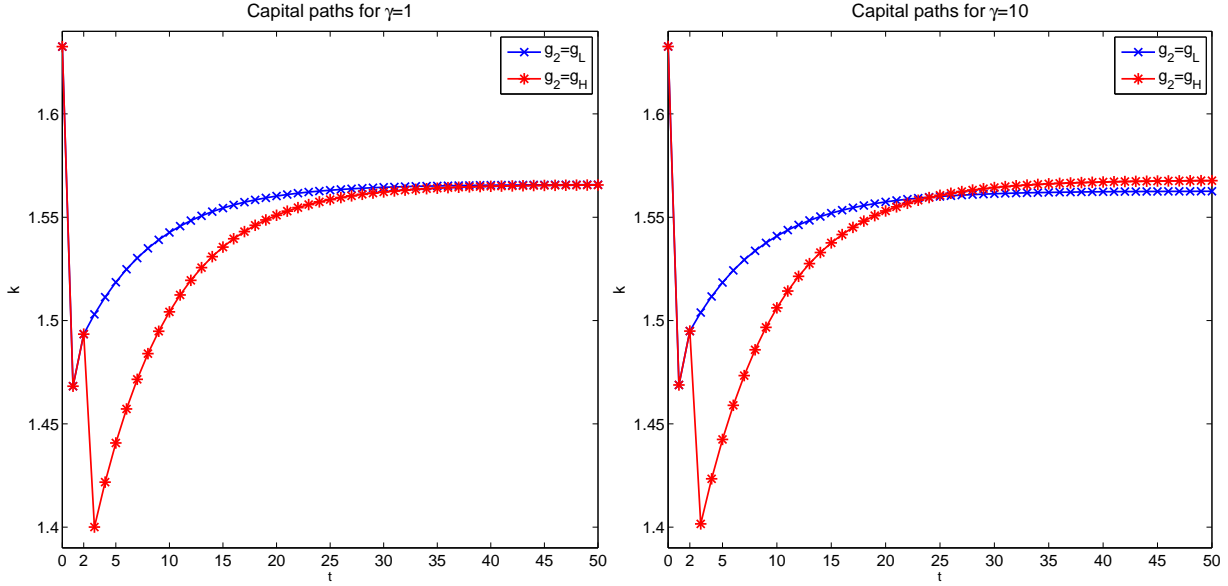


Figure 2: The left panel depicts the capital paths for the expected utility case ($\gamma = 1$) for the high- and low-shock history, and the right panel the respective ones for the recursive utility case ($\gamma = 10$). In the recursive utility case, capital is converging to two different steady states depending on the realization of the shock at $t = 2$.

of uncertainty after $t = 1$, the excess burden of taxation will stay *permanently* at the value it assumed at $t = 1$. Thus $\Phi_t^H = \Phi_1^H, \forall t \geq 1$, $\Phi_t^L = \Phi_1^L, \forall t \geq 1$. Furthermore, if $g_L = 0$ it is easy to show that the government runs a surplus at $t = 0$, borrows for contingency $g_1 = g_L$ ($b_1^L > 0$) and uses the initial surplus and the proceeds from borrowing in order to hedge $g_1 = g_H$ with assets $b_1^H < 0$. If the high shock hits, the government runs a primary deficit that is financed with assets b_1^H and with additional issuance of debt $b_2^H > 0$. For every period after period $t = 2$, the government runs a constant surplus, which is used to pay interest on a constant amount of debt, $b_t^H = b_2^H, t \geq 2$. If the low shock hits at period $t = 1$, then the government issues a constant amount of debt each period $b_t^L = b_1^L, t \geq 1$, and runs a surplus to finance the interest payments. Due to this fiscal hedging, the debt positions in marginal utility units become $z_1^H \equiv U_c(c_1^H, 1 - h_1^H)b_1^H < 0$ and $z_1^L \equiv U_c(c_1^L, 1 - h_1^L)b_1^L > 0$ and therefore the net positions become negative for the high shock and positive for the low shock ($\eta_1^H < 0$ and $\eta_1^L > 0$). As a result, the excess burden of taxation is falling for the high shock and increasing for the low shock, $\Phi_1^H < \Phi_0 < \Phi_1^L$.

If we also assume a period utility function with constant Frisch elasticity as in proposition 6, then it is clear that also the tax rate falls for the high shock, increases for the low shock, $\tau_1^H < \tau_0 < \tau_1^L$, and stays permanently at these values for all periods after period one, whereas with expected utility the tax rate would remain the same for all $t \geq 0$. Figure 1 illustrates this example for a member of this class of utility functions in the case of a positive low shock, $g_L > 0$.²⁵ The net debt positions are still negative for the high shock and positive for the low

²⁵ $U = \ln c - a_h h^{1+\phi_h} / (1 + \phi_h)$ with $(\beta, \gamma, \phi_h) = (0.96, 10, 1)$ and corresponding risk-sensitivity parameter

shock, so the same picture emerges. The figure shows also that the positions of the government (normalized by output) become larger in absolute value in the case of recursive utility, which is to be expected since the permanent decrease or increase of the tax rate leads to a higher present value of future surpluses (deficits) in absolute value.

6.2 Economy with capital

Consider now an economy with capital and a similar stochastic structure as before, but let the uncertainty take place at $t = 2$, so $g_2 = g_H$ with probability π and $g_2 = g_L$ with probability $1 - \pi$, whereas $g_t = g_L, \forall t \neq 2$.^{26,27} Since there is no uncertainty before and after $t = 2$, we have $\Phi_1 = \Phi_0$ and $\Phi_t^i = \Phi_2^i, i = H, L, t \geq 2$, so the excess burden of taxation will remain permanently at the values it assumed at $t = 2$. I consider the same period utility function as in the example without capital, which implies a zero ex-ante capital tax at the second period, a zero capital tax for $t \geq 3$ and a constant labor tax for $t \geq 1$ for the time-additive case.²⁸

In an economy with capital the relevant positions are wealth in marginal utility units (and not just debt). Again, for the high shock at $t = 2$, the planner takes a negative position and for the low shock a positive position, which leads to $\Phi_2^H < \Phi_2^L$ and a permanently lower (higher) labor tax at the high-shock (low-shock) history.²⁹ Since the economy becomes *deterministic* after $t = 2$, the capital tax will be zero for $t \geq 3$, but the ex-ante tax rate for $t = 2$ is not necessarily zero anymore and for the particular illustration it turns out to be a small subsidy.³⁰ Figure 2 focuses on the respective capital paths for the expected utility case and the recursive utility case. What is worth noting is that, since the change in the excess burden of taxation (labor tax) is permanent, there are two steady states depending on what value government expenditures took at $t = 2$. For the high-shock history, which is associated with a lower labor tax, there is a

$\sigma \equiv (1 - \beta)(1 - \gamma)$. I calibrate a_h so that the household works 0.4 of its total time (normalized to unity) when there are no distortionary taxes and $g = g_L$. The shock g_L is set to 0.25 of the first-best output(labor) which is equal to 0.4. The high shock is $g_H = 1.5 \cdot g_L$ with $\pi = 0.5$.

²⁶I let the shock materialize at $t = 2$ because I want to disentangle the effect of initial debt and capital from the effect of uncertainty on the optimal taxation problem.

²⁷The presence of initial wealth (which would be absent if we had zero initial debt, full depreciation and an initial tax rate on capital income of 100%) alters the taxation incentives for labor income at $t = 0$ and capital income at $t = 1$. In particular, the planner has an incentive to increase initial consumption in order to reduce initial wealth in marginal utility units. By subsidizing initial labor income and taxing capital income at $t = 1$, he is able to achieve that.

²⁸The production function is $F = k^\alpha h^{1-\alpha}$. The parameters for the illustration are $(\beta, \gamma, \phi_h, \alpha, \delta, \tau_0^K, b_0) = (0.96, 10, 1, 1/3, 0.08, 0.3, 0)$ with a total endowment of time normalized to unity. The parameter a_h is set so that the household works 0.4 of its time at the first-best steady state. The initial capital is set to 0.9 of the first-best steady state capital. The size of g_L is set so that the share of government expenditures in the first-best steady state output is 0.2. The high shock is $g_H = 2 \cdot g_L$ and $\pi = 0.5$. Due to the minimal uncertainty, I use a relatively large high shock for both illustrations. This would not be necessary in a fully stochastic setup.

²⁹For the time-additive case ($\gamma = 1$) the labor taxes are $\tau_0 = -15.84\%$ and $\tau_t = 29.55\%, t \geq 1$. For the recursive utility case ($\gamma = 10$) we have $\tau_0 = -15.89\%$ and $\tau_t^H = \tau_2^H = 29.32\%, < \tau_1 = 29.59\% < \tau_t^L = \tau_2^L = 29.90\%, t \geq 2$.

³⁰When $\gamma = 1$ capital taxes are $\tau_1^K = 331.77\%, \bar{\tau}_2^K = 0$ and $\tau_t^K = 0, t \geq 3$. When $\gamma = 10$ we have $\tau_1^K = 332.16\%, \bar{\tau}_2^K = -0.2874\%$ and $\tau_t^K = 0, t \geq 3$. I follow Chari et al. (1994) and do not impose upper bounds on capital taxation, allowing capital taxes to exceed 100%.

steady state with higher capital holdings and labor, whereas for the low-shock history, which is associated with a higher labor tax, the steady state involves lower capital holdings and labor.

7 Concluding remarks

Dynamic optimal taxation entails the notions of time and risk. The analysis in this paper shows that when the attitudes towards these two notions are distinct, the conventional normative tax-smoothing results are substantially altered. There is an incentive to tax capital income, the labor tax is time-varying and optimal policy generates endogenously persistence independent of the stochastic properties of the exogenous shocks.

These results indicate that ignoring this distinction is not an innocuous assumption. I have focused on time and risk in an otherwise standard business cycle model. An analysis beyond the representative agent framework, or an exploration of different timing protocols like lack of commitment, are worthy directions for future research.

More broadly, the asset pricing literature like [Bansal and Yaron \(2004\)](#) and [Hansen et al. \(2008\)](#), has used recursive preferences in conjunction with persistent exogenous risk, in order to make the intertemporal profile of consumption quantitatively important for asset pricing purposes. It is interesting to observe that from an optimal policy perspective, *merely* the fact that the agent cares for the “long-run”, causes the planner to optimally introduce persistence to policy variables in an otherwise frictionless environment.

A Recursive formulation

A.1 Two subproblems

Define

$$\begin{aligned}
 Z(s_1, k_1) \equiv & \left\{ z_1 \mid \exists \{c_t, h_t\}_{t \geq 1}, \{k_{t+1}, z_{t+1}, V_{t+1}\}_{t \geq 1}, \text{ with } c_t, k_{t+1} \geq 0 \text{ and } h_t \in [0, 1] \right. \\
 & \text{such that:} \\
 & z_t = \Omega(c_t, h_t) + \beta E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} z_{t+1}, t \geq 1 \\
 & c_t + k_{t+1} - (1 - \delta)k_t + g_t = F(s_t, k_t, h_t), t \geq 1 \\
 & \text{where } m_{t+1} \text{ defined as in (5) with } V_{t+1}, t \geq 1 \text{ following recursion (2)} \\
 & \left. \text{and the transversality condition } \lim_{t \rightarrow \infty} E_1 \beta^t \left(\frac{M_{t+1}}{M_1} \right)^{\frac{\rho-\gamma}{1-\gamma}} z_{t+1} = 0 \right\} \quad (\text{A.1})
 \end{aligned}$$

The set $Z(s_1, k_1)$ stands for the set of values of z_t at $t = 1$, that can be generated by an implementable allocation, when the initial shock is s_1 and the capital inherited from period zero is k_1 . So the state space is endogenous. It could be potentially calculated by following the recursive procedure of [Kydland and Prescott \(1980\)](#).

Problem 1. Let $z_1 \in Z(s_1, k_1)$. The problem from period one onward is

$$V(z_1, k_1, s_1) \equiv \max V_1(\{c_t, h_t\}_{t \geq 1})$$

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 and capital, the feasibility constraint on labor, the transversality condition $\lim_{t \rightarrow \infty} E_0 \beta^t M_t^{\frac{\rho-\gamma}{1-\gamma}} z_t = 0$,³¹ where m_{t+1} denotes the conditional likelihood ratio that is generated by the utility recursion (2).

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

$$\bar{V}_0(b_0, k_0, s_0, \tau_0^K) \equiv \max_{c_0, h_0, k_1, z_1, s_1} \left[(1 - \beta)u(c_0, 1 - h_0)^{1-\rho} + \beta \left[\sum_{s_1} \pi_1(s_1 | s_0) V(z_1, s_1, k_1, s_1)^{1-\gamma} \right]^{\frac{1-\rho}{1-\gamma}} \right]^{\frac{1}{1-\rho}}$$

³¹Strictly speaking, this condition (which is implied by (13) and (14)) is sufficient for the exhaustion of the intertemporal budget constraint of the household.

subject to

$$U_{c_0} \left[\left((1 - \tau_0^K) F_K(s_0, k_0, h_0) + 1 - \delta \right) k_0 + b_0 \right] = \Omega(c_0, h_0) + \beta \sum_{s_1} \pi_1(s_1|s_0) \frac{V(z_{1,s_1}, k_1, s_1)^{\rho-\gamma}}{\left[\sum_{s_1} \pi_1(s_1|s_0) V(z_{1,s_1}, k_1, s_1)^{1-\gamma} \right]^{\frac{\rho-\gamma}{1-\gamma}}} z_{1,s_1} \quad (\text{A.2})$$

$$c_0 + k_1 - (1 - \delta)k_0 + g_0 = F(s_0, k_0, h_0) \quad (\text{A.3})$$

$$c_0, k_1 \geq 0, h_0 \in [0, 1], \quad (\text{A.4})$$

$$z_{1,s_1} \in Z(s_1, k_1) \quad (\text{A.5})$$

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

The notation z_{1,s_1} denotes the value of the state variable z_1 at s_1 . The overall value of the Ramsey problem $\bar{V}(\cdot)$ depends on the initial conditions $(b_0, k_0, s_0, \tau_0^K)$, which is why I use a different notation for the initial value function.

A.2 Transformed Bellman equation

Given the ρ -transformation of the value function, $v(z, k, s) \equiv \frac{V(z, k, s)^{1-\rho}-1}{(1-\beta)(1-\rho)}$, the Bellman equation takes the form

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

subject to the transformed dynamic implementability constraint

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

and to (21)-(23). The market value of the household's wealth takes the form $\omega = \sum_{s'} \pi(s'|s) m'_{s'} \frac{\rho-\gamma}{1-\gamma} z'_{s'}$, where $m'_{s'}$ stands for the conditional likelihood ratio,

$$m'_{s'} \equiv \frac{V(z'_{s'}, k', s')^{1-\gamma}}{\sum_{s'} \pi(s'|s) V(z'_{s'}, k', s')^{1-\gamma}} = \frac{\left[1 + (1 - \beta)(1 - \rho)v(z'_{s'}, k', s') \right]^{\frac{1-\gamma}{1-\rho}}}{\sum_{s'} \pi(s'|s) \left[1 + (1 - \beta)(1 - \rho)v(z'_{s'}, k', s') \right]^{\frac{1-\gamma}{1-\rho}}}$$

A.3 Initial period optimality conditions

Use the ρ -transformation of the time zero problem and let Φ_0 and λ_0 denote the multipliers on the initial period implementability constraint and the resource constraint respectively. The initial period optimality conditions are:

$$c_0 : \quad U_{c0} + \Phi_0 \left[\Omega_{c0} - U_{cc,0} W_0 \right] = \lambda_0 \quad (\text{A.6})$$

$$h_0 : \quad -U_{l0} + \Phi_0 \left[\Omega_{h0} + U_{cl,0} W_0 - U_{c0} (1 - \tau_0^K) F_{KH,0} k_0 \right] = -\lambda_0 F_{H0} \quad (\text{A.7})$$

$$k_1 : \quad \lambda_0 = \beta \sum_{s_1} \pi(s_1|s_0) m_{1,s_1}^{\frac{\rho-\gamma}{1-\gamma}} v_k(z_{1,s_1}, k_1, s_1) \left[1 + (1-\beta)(\rho-\gamma)\eta_{1,s_1} \Phi_0 \right] \quad (\text{A.8})$$

$$z_{1,s_1} : \quad v_z(z_{1,s_1}, k_1, s_1) + \Phi_0 \left[1 + (1-\beta)(\rho-\gamma)v_z(z_{1,s_1}, k_1, s_1)\eta_{1,s_1} \right] = 0, \quad (\text{A.9})$$

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 and η_{1,s_1} defined as in (30). The initial period first-order conditions for an economy without capital for the variables (c_0, h_0, z_{1,s_1}) , are (A.6),(A.7) and (A.9) with $W_0 = b_0$, $F_{H0} = a_{s_0}$, $F_{KH} \equiv 0$.

B Sequential formulation

Let $X_t \equiv M_t^{\frac{\rho-\gamma}{1-\gamma}}$, $X_0 \equiv 1$. Let v refer to the ρ -transformation of the utility criterion. The Ramsey problem is

$$\max v_0(\{c\}, \{h\})$$

subject to

$$\sum_{t=0}^{\infty} \beta^t \sum_{s^t} \pi_t(s^t) X_t(s^t) \Omega(c_t(s^t), h_t(s^t)) = U_{c0} W_0 \quad (\text{B.1})$$

$$c_t(s^t) + k_{t+1}(s^t) - (1 - \delta)k_t(s^{t-1}) + g_t(s^t) = F(s_t, k_t(s^{t-1}), h_t(s^t)) \quad (\text{B.2})$$

$$X_{t+1}(s^{t+1}) = m_{t+1}(s^{t+1})^{\frac{\rho-\gamma}{1-\gamma}} X_t(s^t), \quad X_0 \equiv 1 \quad (\text{B.3})$$

$$v_t(s^t) = U(c_t(s^t), 1 - h_t(s^t)) + \beta \frac{\left[\sum_{s_{t+1}} \pi_{t+1}(s_{t+1}|s^t) \left[1 + (1-\beta)(1-\rho)v_{t+1}(s^{t+1}) \right]^{\frac{1-\gamma}{1-\rho}} \right]^{\frac{1-\rho}{1-\gamma}} - 1}{(1-\beta)(1-\rho)}, \quad t \geq 1 \quad (\text{B.4})$$

where $W_0 \equiv R_0^K k_0 + b_0$, $(b_0, k_0, s_0, \tau_0^K)$ given, and $m_{t+1} = \frac{[1+(1-\beta)(1-\rho)v_{t+1}]^{\frac{1-\gamma}{1-\rho}}}{E_t[1+(1-\beta)(1-\rho)v_{t+1}]^{\frac{1-\gamma}{1-\rho}}}$.

Assign multipliers $\bar{\Phi}$, $\beta^t \pi_t \lambda_t$, $\beta^t \pi_t \nu_t$ and $\beta^t \pi_t \xi_t$ on (B.1), (B.2), (B.3) and (B.4) respectively. The derivatives of the utility function are $\frac{\partial v_0}{\partial c_t} = \beta^t \pi_t X_t U_{c_t}$ and $\frac{\partial v_0}{\partial h_t} = -\beta^t \pi_t X_t U_{l_t}$. The first-order necessary conditions are

$$c_t, t \geq 1 : \quad X_t(s^t)U_c(s^t) + \bar{\Phi}X_t(s^t)\Omega_c(s^t) + \xi_t(s^t)U_c(s^t) = \lambda_t(s^t) \quad (\text{B.5})$$

$$h_t, t \geq 1 : \quad -X_t(s^t)U_l(s^t) + \bar{\Phi}X_t(s^t)\Omega_h(s^t) - \xi_t(s^t)U_l(s^t) = -\lambda_t(s^t)F_H(s^t) \quad (\text{B.6})$$

$$k_{t+1}(s^t), t \geq 0 : \quad \lambda_t(s^t) = \beta \sum_{s_{t+1}} \pi_{t+1}(s_{t+1}|s^t) \lambda_{t+1}(s^{t+1}) [1 - \delta + F_K(s^{t+1})] \quad (\text{B.7})$$

$$X_t(s^t), t \geq 1 : \quad \nu_t(s^t) = \bar{\Phi}\Omega_t(s^t) + \beta \sum_{s_{t+1}} \pi_{t+1}(s_{t+1}|s^t) m_{t+1}(s^{t+1})^{\frac{\rho-\gamma}{1-\gamma}} \nu_{t+1}(s^{t+1}) \quad (\text{B.8})$$

$$v_t(s^t), t \geq 1 : \quad \xi_t(s^t) = (1 - \beta)(\rho - \gamma)X_t(s^t)\phi_t(s^t) + m_t(s^t)^{\frac{\rho-\gamma}{1-\gamma}} \xi_{t-1}(s^{t-1}), \quad (\text{B.9})$$

where

$$\phi_t(s^t) \equiv V_t(s^t)^{\rho-1} \nu_t(s^t) - \mu_t(s^t)^{\rho-1} \sum_{s_t} \pi_t(s_t|s^{t-1}) m_t(s^t)^{\frac{\rho-\gamma}{1-\gamma}} \nu_t(s^t),$$

and $\xi_0 \equiv 0$. The optimality conditions with respect to the initial consumption-labor allocation are (A.6) and (A.7).

I will show now the mapping between the sequential formulation and the recursive formulation in the text and in particular the relationship between the time-varying Φ_t and ξ_t . Solve at first (B.8) forward to get

$$\nu_t = \bar{\Phi} E_t \sum_{i=0}^{\infty} \beta^i \frac{X_{t+i}}{X_t} \Omega_{t+i}$$

and therefore $\nu_t = \bar{\Phi} U_{c_t} W_t = \Phi_{z_t}$, i.e. ν_t – the shadow value to the planner of an increase in X_t – is equal to wealth (in marginal utility terms) times the cost of taxation $\bar{\Phi}$. Thus, ϕ_t –the “innovation” in the multiplier ν_t – is equal to a multiple of η_t , $\phi_t = \bar{\Phi} \eta_t$. Furthermore, define the scaled multiplier $\tilde{\xi}_t \equiv \xi_t / X_t$, $\tilde{\xi}_0 \equiv 0$ and note that it follows the law of motion

$$\begin{aligned}
\tilde{\xi}_t &= (1 - \beta)(\rho - \gamma)\phi_t + \tilde{\xi}_{t-1} \\
&= (1 - \beta)(\rho - \gamma) \sum_{i=1}^t \phi_i = (1 - \beta)(\rho - \gamma) \sum_{i=1}^t \eta_i \bar{\Phi}
\end{aligned}$$

Turn now to the multiplier in the text which, when solved backwards, delivers $\Phi_t = \Phi_0 / (1 + (1 - \beta)(\rho - \gamma) \sum_{i=1}^t \eta_i \Phi_0)$, where Φ_0 is the multiplier on the initial period implementability constraint. Thus, by setting $\Phi_0 = \bar{\Phi}$ we have

$$\Phi_t = \frac{\bar{\Phi}}{1 + \tilde{\xi}_t}, \quad (\text{B.10})$$

or, in terms of the non-scaled ξ_t , $\Phi_t = \bar{\Phi} X_t / (X_t + \xi_t)$. Therefore, the time-varying excess burden of taxation captures the additional channel of continuation utilities in intertemporal marginal rates of substitution. Consider now the multipliers λ_t in the sequential formulation and their relationship to their counterparts in the text. Given (B.10), their relationship is $\lambda_t = (X_t + \xi_t) \lambda_t^R$, where λ_t^R stands for the multipliers for the recursive formulation. Thus,

$$\begin{aligned}
\frac{\lambda_{t+1}}{\lambda_t} &= \frac{X_{t+1} + \xi_{t+1}}{X_t + \xi_t} \frac{\lambda_{t+1}^R}{\lambda_t^R} = \frac{X_{t+1}}{X_t} \frac{\frac{\bar{\Phi} X_t}{X_t + \xi_t}}{\frac{\bar{\Phi} X_{t+1}}{X_{t+1} + \xi_{t+1}}} \frac{\lambda_{t+1}^R}{\lambda_t^R} \\
&= m_{t+1}^{\frac{\rho - \gamma}{1 - \gamma}} \frac{\lambda_{t+1}^R}{\lambda_t^R} \frac{\Phi_t}{\Phi_{t+1}},
\end{aligned}$$

which delivers the optimality condition with respect to capital (39).

C Proof of proposition 3

Proof. For $\rho = \gamma$, the Bellman equation becomes

$$v(z, k, s) = \max_{c \geq 0, h \in [0, 1], k' \geq 0, z'_s \in Z(s', k')} [U(c, 1 - h) + \beta \sum_{s'} \pi(s'|s) v(z'_s, k', s')]$$

subject to $z = \Omega(c, h) + \beta \sum_{s'} \pi(s'|s) z'_s$ and the resource constraint (21). Since the value function is not appearing anymore in the dynamic implementability constraint, the first-order condition with respect to z'_s delivers $-v_z(z'_s, k', s') = \Phi$. Thus, after solving for the optimal initial value of the state z_1 , we get a constant $\bar{\Phi}$. The rest of the first-order conditions become

$$\begin{aligned}
c : \quad & U_c + \bar{\Phi}\Omega_c = \lambda \\
h : \quad & -U_l + \bar{\Phi}\Omega_h = -\lambda F_H(s, k, h) \\
k' : \quad & \lambda = \beta \sum_{s'} \pi(s'|s) \lambda'_{s'} (1 - \delta + F_{K,s'}),
\end{aligned}$$

where I already used the envelope condition with respect to capital, (34). Note that if we knew $\bar{\Phi}$, then the optimality conditions above together with the resource constraint (21) 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 is necessary for the determination of the solution only through the excess burden of taxation, i.e. through the shadow cost that additional wealth is imposing. Since for the case of expected utility this cost is *constant* to some value $\bar{\Phi}$ independent of (z, k, s) , as long as the planner takes into account how $\bar{\Phi}$ affects the household's utility, he does not need to keep track of z , but only of the state variable (k, s) (or s in an economy without capital). In order to see that clearly, proceed as in Chari et al. (1994) and solve the problem

$$v^{CCK}(k, s; \bar{\Phi}) = \max_{c \geq 0, h \in [0, 1], k' \geq 0} [\bar{U}(c, h; \bar{\Phi}) + \beta \sum_{s'} \pi(s'|s) v^{CCK}(k', s'; \bar{\Phi})]$$

subject to the resource constraint (21), where $\bar{U}(c, h; \bar{\Phi}) \equiv U(c, 1 - h) + \bar{\Phi}\Omega(c, h)$. Note that the period return function is augmented so that it takes into account the relevant part of the implementability constraint Ω 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.³²

□

D Details about the illustrations

D.1 Economy without capital

I am assuming that the period utility function is such so that it is always optimal to tax ($U_{cl} \geq 0$ is sufficient for that as we saw in the text) and consider the risk-sensitive case in order to show that the analysis does not depend on the functional form of U , except for the above restriction.

³²A full characterization requires solving the initial period problem

$$\max_{\{c_0, h_0, k_1\}} [\bar{U}(c_0, h_0; \bar{\Phi}) - \bar{\Phi}U_{c0}W_0 + \beta \sum_{s_1} \pi(s_1|s_0) v^{CCK}(k_1, s_1; \bar{\Phi})]$$

subject to the initial resource constraint. Note that the initial policy functions are functions of $\bar{\Phi}$, the initial conditions (b_0, k_0, s_0) and the fixed τ_0^K . The final step is to adjust appropriately $\bar{\Phi}$ so that the intertemporal implementability constraint is satisfied.

The two values that the excess burden of taxation is taking at period one can be found from $\Phi_1^i = \frac{\Phi_0}{1 + \sigma \eta_1^i \Phi_0}$, $i = H, L$, where $\eta_1^i = z_1^i - \omega_0$, $i = H, L$, and $\omega_0 = \pi m_1^H z_1^H + (1 - \pi) m_1^L z_1^L$, the market value of the government's portfolio which is the mean of z_1 under the continuation-value adjusted measure. The likelihood ratios depend on continuation values as $m_1^i = \exp(\sigma V_1^i) / (\pi \exp(\sigma V_1^H) + (1 - \pi) \exp(\sigma V_1^L))$, $i = H, L$. In order to determine the relationship between Φ_1^H and Φ_1^L we need to find the debt positions of the government.

Consider consumption which, by using (44) and the resource constraint, can be expressed as function of the shock and the time-varying Φ_t , $c_t = c(g_t, \Phi_t)$ (so labor is $h_t = h(g_t, \Phi_t) = c(g_t, \Phi_t) + g_t$ and the tax rate $\tau_t = 1 - U_l/U_c = \tau(g_t, \Phi_t)$). At the initial period we have $c_0 = c(g_L, \Phi_0)$. Since Φ_t does *not* revert to Φ_0 , for the high-shock history we have $c_t^H = c_2^H = c(g_L, \Phi_1^H)$, $\forall t \geq 2$ and for the low-shock history $c_t^L = c_2^L = c(g_L, \Phi_1^L)$, $\forall t \geq 1$. For $g_1 = g_H$ we have $c_1^H = c(g_H, \Phi_1^H)$. Thus, continuation values are $V_1^H = U(c_1^H, 1 - h_1^H) + \beta V_2^H$, $V_2^H = U(c_2^H, 1 - h_2^H) / (1 - \beta)$, $V_1^L = U(c_1^L, 1 - h_1^L) / (1 - \beta)$. The tax rates are $\tau_0 = \tau(g_L, \Phi_0)$, $\tau_1^H = \tau(g_H, \Phi_1^H)$, $\tau_t^H = \tau_2^H = \tau(g_L, \Phi_1^H)$, $t \geq 2$, $\tau_t^L = \tau_1^L = \tau(g_L, \Phi_1^L)$, $t \geq 1$.

For the high-shock history, the government budget constraints read $b_1^H = \tau_1^H h_1^H - g_H + p_1^H b_2^H$ and $b_t^H = \tau_2^H h_2^H - g_L + p_t^H b_{t+1}^H$, $t \geq 2$. Prices are $p_1^H = \beta U_c(c_2^H, 1 - h_2^H) / U_c(c_1^H, 1 - h_1^H)$, $p_t^H = \beta$, $t \geq 2$. Thus $b_t^H = b_2^H = \frac{\tau_2^H h_2^H - g_L}{1 - \beta}$, $t \geq 2$. For the low-shock history, we have $p_t^L = \beta$, $t \geq 1$ and bond positions $b_t^L = b_1^L = \frac{\tau_1^L h_1^L - g_L}{1 - \beta}$, $t \geq 1$.

Consider now the initial period. The budget constraint reads $0 = \tau_0 h_0 - g_L + p_0^H b_1^H + p_1^L b_1^L$, where $p_0^H = \beta \pi m_1^H U_c(c_1^H, 1 - h_1^H) / U_c(c_0, 1 - h_0)$ and $p_1^L = \beta (1 - \pi) m_1^L U_c(c_1^L, 1 - h_1^L) / U_c(c_0, 1 - h_0)$. If $g_L = 0$, we have $b_1^L > 0$ and since at the initial period the government is running a surplus $\tau_0 h_0$, we have $b_1^H < 0$. Furthermore, in this case we also have $b_2^H > 0$ and therefore from the budget constraint at $g_1 = g_H$ we conclude that the government is running a deficit, $\tau_1^H h_1^H - g_H < 0$.

For the utility function of the example we have $\Omega(c, h) = 1 - a_h h^{1 + \phi_h}$ and $\tau_t = \tau(\Phi_t) = \Phi_t(1 + \phi_h) / (1 + \Phi_t(1 + \phi_h))$, which holds for $t \geq 0$ since the initial debt is equal to zero. We can solve for the optimal allocation and the excess burden of taxation by using a double loop:

- *Inner loop*: Fix Φ_0 and make a guess for Φ_1^i , $i = H, L$. Given the guess calculate the induced debt positions, the continuation utilities and the innovations η_1 . Use η_1 and update the guess for Φ_1^i from the law of motion for the excess burden of taxation and iterate till convergence.
- *Outer loop*: After convergence of the inner loop is achieved, calculate the initial allocation and the residual from the initial budget constraint, adjust Φ_0 and go back to the inner loop in order to find the corresponding Φ_1^i . Stop when Φ_0 is such so that the initial budget constraint is satisfied.

D.2 Economy with capital

The period utility function is the same as in the example without capital. The labor tax formula holds only for $t \geq 1$ due to the presence of the initial wealth W_0 . The procedure to solve the problem involves again a double loop for the determination of $\Phi_2^i, i = H, L$ and Φ_0 . Here are the details:

- *Inner loop:* Fix Φ_0 and make a guess for (Φ_2^H, Φ_2^L) . Given these two values of the excess burden of taxation, the problem from period $t = 3$ onward for both histories behaves as a deterministic Ramsey taxation problem, but with different Φ 's depending on the high- or low-shock history. For the high-shock history, for $t \geq 3$ solve the "CCK" Bellman equation referred to in the proof of proposition 3,

$$v(k) = \max_{c, h, k'} \bar{U}(c, 1 - h; \Phi_2^H) + \beta v(k')$$

subject to $c + k' - (1 - \delta)k + g_L = k^\alpha h^{1 - \alpha}$, with the modified return function $\bar{U}(c, 1 - h; \Phi_2^H) = \ln c - a_h \frac{h^{1 + \phi_h}}{1 + \phi_h} + \Phi_2^H(1 - a_h h^{1 + \phi_h})$. For the low-shock history, for $t \geq 2$, solve the same Bellman equation but with the return function $\bar{U}(c, 1 - h; \Phi_2^L)$.

To determine the wealth positions z_2^i and the respective innovations that allow the update of the guesses for Φ_2^i , proceed as follows: Fix k_3^H and consider the respective Euler equation:

$$\frac{1}{c_2^H} = \beta \frac{1}{c_3^H} \left[1 - \delta + \alpha \left(\frac{k_3^H}{h_3^H} \right)^{\alpha - 1} \right]$$

Given k_3^H and the policy functions we found from solving the Bellman equation, the right-hand side is known, determining therefore c_2^H . Furthermore, use the intratemporal wedge condition for $g_2 = g_H$ to get $h_2^H = \left[\frac{(1 - \tau_2^H)(1 - \alpha)}{a_h c_2^H} \right]^{\frac{1}{\alpha + \phi_h}} k_2^{\frac{\alpha}{\alpha + \phi_h}}$, where $\tau_2^H = \tau(\Phi_2^H)$. Plug the expression for labor in the resource constraint at $g_2 = g_H$, $c_2^H + k_3^H - (1 - \delta)k_2 + g_H = k_2^\alpha (h_2^H)^{1 - \alpha}$ to get one equation in the unknown k_2 and use a non-linear solver to determine it. Furthermore, use the policy functions for $t \geq 3$ to determine V_3^H and z_3^H . Utilities are calculated with the original period utility function (and *not* with the modified \bar{U}). Finally, use (c_2^H, h_2^H) to get $V_2^H = U(c_2^H, 1 - h_2^H) + \beta V_3^H$ and $z_2^H = \Omega(c_2^H, h_2^H) + \beta z_3^H$. Use now the policy functions for the low-shock history to determine V_2^L and z_2^L at k_2 . Having the utility values and the wealth positions at $t = 2$ allows us to calculate $m_2^i, i = H, L$, the market value of the wealth portfolio $\omega_1 = \pi m_2^H z_2^H + (1 - \pi) m_2^L z_2^L$ and therefore the innovations

$\eta_2^i = z_2^i - \omega_1, i = H, L$, given the guess for Φ_2^i . Use the innovations η_2^i to update the guess for $\Phi_2^i, \Phi_2^i = \frac{\Phi_0}{1 + \sigma \eta_2^i \Phi_0}, i = H, L$ and iterate till convergence.

- *Outer loop:* After we reach convergence for Φ_2^i , calculate the rest of the allocation for $t = 0, 1$ given the initial Φ_0 . In particular, the Euler equation for k_2 is

$$\frac{1}{c_1 \Phi_0} = \beta \pi m_2^H \frac{1}{c_2^H \Phi_2^H} \left[1 - \delta + \alpha \left(\frac{k_2}{h_2^H} \right)^{\alpha-1} \right] + \beta (1 - \pi) m_2^L \frac{1}{c_2^L \Phi_2^L} \left[1 - \delta + \alpha \left(\frac{k_2}{h_2^L} \right)^{\alpha-1} \right].$$

The right-hand side is known, which delivers c_1 . Express now labor at $t = 1$ as $h_1 = \left[\frac{(1-\tau_1)(1-\alpha)}{a_h c_1} \right]^{\frac{1}{\alpha+\phi_h}} k_1^{\frac{\alpha}{\alpha+\phi_h}}, \tau_1 = \tau(\Phi_0)$ and use this expression to solve for k_1 from the resource constraint. Calculate furthermore $z_1 = \Omega(c_1, h_1) + \beta \omega_1$. The initial period requires a different treatment due to the presence of initial wealth $W_0 = b_0 + [(1 - \tau_0^K) \alpha (k_0/h_0)^{\alpha-1} + 1 - \delta] k_0$. Use the Euler equation for capital to get the initial value of the multiplier $\lambda_0, \lambda_0 = \frac{\beta}{c_1} [1 - \delta + \alpha (k_1/h_1)^{\alpha-1}]$. Then use the first-order conditions for (c_0, h_0) , (A.6)-(A.7) and the resource constraint at $t = 0$ to get a system in three unknowns (c_0, h_0, k_0) to be solved with a non-linear solver. Update Φ_0 by calculating the residual in the initial budget constraint, $I \equiv \Omega(c_0, h_0) + \beta z_1 - \frac{1}{c_0} W_0$. If $I > (<) 0$ decrease (increase) Φ_0 and go back to the inner loop to redetermine $\Phi_2^i, i = H, L$ given the new Φ_0 . Stop when the initial budget constraint holds, $I = 0$.

The solution method for the outer loop is based on a fixed value k_3^H , which delivers in the end an initial value of capital k_0 . I experimented with k_3^H so that the endogenous initial capital takes the value mentioned in the text.

There is plethora of methods for solving the Bellman equation. I use the envelope condition method of [Maliar and Maliar \(2013\)](#). I approximate the value function with a 5th degree polynomial in capital and I use 100 grid points. Furthermore, since the steady-state capital depends on Φ_2^i , I re-adjust the bounds of the state space for each calculation of the value function in order to focus on the relevant part of the state space. For the high-shock history, I set the lower bound as $\underline{K} = 0.95 \cdot \min(k_3^H, k_{ss}^H)$ and the upper bound $\bar{K} = 1.05 \cdot \max(k_3^H, k_{ss}^H)$. In the same vain, for the low-shock history, I set $\underline{K} = 0.95 \cdot \min(k_2, k_{ss}^L)$ and $\bar{K} = 1.05 \cdot \max(k_2, k_{ss}^L)$. The variables $k_{ss}^i, i = H, L$ denote the respective steady states.

References

- Aiyagari, S. Rao, Albert Marcet, Thomas J. Sargent, and Juha Seppala. 2002. Optimal Taxation without State-Contingent Debt. *Journal of Political Economy* 110 (6):1220–1254.
- Anderson, Evan W. 2005. The dynamics of risk-sensitive allocations. *Journal of Economic Theory* 125 (2):93–150.
- 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.
- 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.
- Farhi, Emmanuel. 2010. Capital Taxation and Ownership when Markets are Incomplete. *Journal of Political Economy* 118 (5):908–948.
- Farhi, Emmanuel and Ivan Werning. 2008. Optimal savings distortions with recursive preferences. *Journal of Monetary Economics* 55 (1):21–42.
- Geoffard, Pierre-Yves. 1996. Discounting and optimizing: Capital accumulation problems as variational minmax problems. *Journal of Economic Theory* 69:53–70.
- Hansen, Lars Peter and Thomas J. Sargent. 2001. Robust Control and Model Uncertainty. *American Economic Review* 91 (2):60–66.
- Hansen, Lars Peter, John C. Heaton, and Nan Li. 2008. Consumption Strikes Back? Measuring Long-Run Risk. *Journal of Political Economy* 116 (2):260–302.
- Judd, Kenneth L. 1985. Redistributive taxation in a simple perfect foresight model. *Journal of Public Economics* 28 (1):59–83.
- Karantounias, Anastasios G. 2013. Managing pessimistic expectations and fiscal policy. *Theoretical Economics* 8 (1):193–231.
- Koopmans, Tjalling C. 1960. Stationary ordinary utility and impatience. *Econometrica* 28 (2):287–309.

- 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.
- Maliar, Lilia and Serguei Maliar. 2013. Envelope Condition Method versus Endogenous Grid Method for Solving Dynamic Programming Problems. *Economic Letters* 120:262–266.
- Sleet, Christopher and Sevin Yeltekin. 2006. Optimal taxation with endogenously incomplete debt markets. *Journal of Economic Theory* 127 (1):36–73.
- Strzalecki, Tomasz. 2011. Axiomatic Foundations of Multiplier Preferences. *Econometrica* 79 (1):47–73.
- . 2013. Temporal Resolution of Uncertainty and Recursive Models of Ambiguity Aversion. *Econometrica* 81 (3):1039–1074.
- Tallarini, Thomas D. 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.