

Optimal fiscal policy with recursive preferences

Anastasios G. Karantounias *

March 25, 2017

Abstract

I study the implications of recursive utility for the design of optimal fiscal policy. Standard Ramsey tax-smoothing prescriptions are dramatically altered. The planner taxes less in bad times and more in good times, mitigating the effects of shocks. For standard calibrations, labor tax volatility is orders of magnitude larger than in the time-additive case. At the intertemporal margin, there is a novel incentive for introducing distortions that can lead to an ex-ante capital subsidy. Overall, optimal policy calls for an even stronger use of debt returns as a fiscal absorber, leading to the conclusion that actual fiscal policy is even worse than we thought.

JEL CLASSIFICATION: D80; E62; H21; H63.

KEY WORDS: Ramsey plan, tax smoothing, Epstein-Zin, recursive utility, excess burden, labor tax, capital tax, martingale, fiscal insurance.

*Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree St NE, Atlanta, GA 30309, anastasios.karantounias@atl.frb.org. I am grateful to the Editor (Philipp Kircher) and to three anonymous referees for insightful comments. I am thankful to Roc Armenter, David Backus, Pierpaolo Benigno, R. Anton Braun, Vasco Carvalho, Lawrence Christiano, Lukasz Drozd, Kristopher S. Gerardi, Mikhail Golosov, Jonathan Halket, Lars Peter Hansen, Karen Kopecky, Hanno Lustig, Juan Pablo Nicolini, Demian Pouzo, Victor Rios-Rull, Richard Rogerson, Thomas J. Sargent, Yongseok Shin, Stanley E. Zin, to Christopher Sleet for his discussion, to seminar participants at Carnegie Mellon University, the Einaudi Institute of Economics and Finance, the European University Institute, the Federal Reserve Bank of Atlanta, the Federal Reserve Bank of St Louis, LUISS Guido Carli University, the University of California at Davis, the University of Hong Kong, the University of Oxford, Universitat Pompeu Fabra, the University of Reading, and to conference participants at the 1st NYU Alumni Conference, the CRETE 2011 Conference, the 2012 SED Meetings, the 2012 EEA Annual Congress and the 2013 AEA Meetings. All errors are my own. The views expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System.

1 Introduction

The basic fiscal policy prescription in dynamic, stochastic, frictionless economies is *tax-smoothing*. Labor taxes should be essentially constant and any kind of shock should be absorbed by proper debt management. This result comes from the seminal work of [Lucas and Stokey \(1983\)](#) and [Chari et al. \(1994\)](#) and forms the heart of dynamic Ramsey policy.

In this paper, I show that if we differentiate between risk aversion and intertemporal elasticity of substitution and use the recursive preferences of [Epstein and Zin \(1989\)](#) and [Weil \(1990\)](#), the conventional normative tax-smoothing result breaks down, both qualitatively and – maybe more importantly– quantitatively. Optimal policy generates large surpluses and deficits by prescribing high taxes in good times and low taxes in bad times. For standard calibrations tax volatility can be 5 percentage points which contrasts to either zero or minimal volatility of few basis points in the time-additive expected utility case. Furthermore, in contrast to standard Ramsey results, labor taxes are persistent independent of the stochastic properties of exogenous shocks and capital income should be subsidized.

The coefficients of intertemporal elasticity of substitution and risk aversion are two parameters that are a priori important in shaping dynamic policy. They control the desirability of taxing in the current versus future periods and the aversion towards shocks that hit the government budget. Unfortunately, time-additive expected utility renders the analysis of the implications of these two parameters on optimal policy impossible. Moreover, since the temporal dimension of risk is ignored, questions about the implications of long-run fiscal risks on current tax and debt policies can be answered only in a limited way.

Second, and more crucially from an applied perspective, time-additive expected utility has difficulty in generating realistic asset prices, a failure which has made the empirically more successful recursive preferences the norm in the literature that merges macroeconomics and finance.¹ Any model that matches risk premia raises natural questions about the design of optimal fiscal policy, since returns on the government debt portfolio are a central input in the decision to issue debt. It is natural to speculate that standard tax-smoothing prescriptions will be altered both qualitatively and quantitatively, due to the richer structure of pricing kernels. However, little is known about recursive utility and optimal fiscal policy even in the simplest Ramsey setup. This is the task of the current paper.

Consider first an economy without capital as in [Lucas and Stokey \(1983\)](#). Linear taxes and state-contingent debt are used in order to finance an exogenous stream of stochastic government expenditures. A benevolent planner chooses under commitment the policy that maximizes the utility of the representative household. There are two basic results with time-additive expected utility: First, the labor tax should be *constant* if period utility features constant elasticities. Even

¹The literature is vast. See indicatively [Tallarini \(2000\)](#), [Bansal and Yaron \(2004\)](#), [Piazzesi and Schneider \(2007\)](#), [Hansen et al. \(2008\)](#), [Gourio \(2012\)](#), [Rudebusch and Swanson \(2012\)](#), [Petrosky-Nadeau et al. \(2013\)](#) and [Ai and Bansal \(2016\)](#) among others.

when elasticities are not constant, the volatility of the labor tax is quite small. Second, whenever the labor tax varies, it inherits the stochastic properties of the exogenous shocks. Thus, optimal labor taxes do not constitute a distinct source of persistence in the economy.

As I argued earlier, both of these classic results are overturned in the same economy with recursive preferences. There is a simple, yet powerful intuition for that. Assume that risk aversion is greater than the inverse of the intertemporal elasticity of substitution. In that case, the household sacrifices smoothing over time in order to have a smoother consumption profile over states, becoming effectively *averse* to volatility in continuation utilities. As a response, the planner *attenuates* utility volatility by taxing less in bad times, offsetting therefore the effects of an adverse fiscal shock, and taxing more in good times, mitigating the benefits of a favorable fiscal shock.

What is the mechanism behind this intuition? The entire action is coming from the pricing of state-contingent claims. The planner hedges fiscal risk by issuing state-contingent debt against low spending shocks, to be paid by surpluses, and buys assets against high spending shocks, that are used to finance government deficits. With recursive utility the planner “over-insures” and sells more debt against low spending shocks relative to the expected utility benchmark, and, therefore, taxes more when spending is low. The reason is simple: by issuing more debt against good times the planner depresses continuation utilities and therefore raises the stochastic discount factor. Thus, the price of state-contingent claims that the planner sells *rises*, making state-contingent debt against good times *cheaper*. So more revenue can be raised from debt issuance and the planner can relax the budget constraint, which is welfare-improving. Similarly, by purchasing more assets and taxing less against high spending shocks, the planner raises utility and therefore decreases the stochastic discount factor, relaxing again the government budget constraint.

Hence, the additional curvature of the utility function with respect to the “long-run” *amplifies* fiscal insurance, depressing ultimately risk premia. Optimal policy prescribes high gains for bondholders when government spending is low, that are paid with high taxes, and capital losses when spending is high, which allow large deficits with small taxes. In fact, at high levels of government debt, the over-insurance efforts of the government can lead to a positive conditional covariance of the stochastic discount factor with the returns on the government debt portfolio, implying a *negative* conditional risk premium of government debt. The economics behind this striking result make sense: “good” times with low spending shocks can become “bad” times with very high tax rates. Thus, the household is happy to accept a negative premium for a risky security that pays well when distortionary taxes are high.

With recursive utility a tax rate at a future period affects the entire sequence of one-period stochastic discount factors up to that period, due to the forward-looking nature of continuation utilities. As a result, the planner does not choose future tax rates independently from the past, but designs persistent policies in order to properly affect the entire sequence of prices of state-contingent claims. Furthermore, it is cheaper on average to issue debt and postpone taxation for the future, leading optimally to *back-loading* of tax distortions.

Recursive utility introduces non-trivial complications to the numerical analysis of the Ramsey problem. Value functions appear in the constraints since they affect the pricing of the government debt portfolio, hindering the contraction property, introducing non-convexities and complicating the calculation of the state space. A separate contribution of the paper is to deal with these issues and provide a numerical solution of the optimal taxation problem. In a series of numerical exercises I demonstrate the substantial volatility and persistence of the tax rate and analyze the implications for the debt-to-output ratio.

As a final exercise, I quantify the *optimal* use of debt returns and tax revenues for the absorption of fiscal shocks and contrast it to the empirical findings of [Berndt et al. \(2012\)](#). [Berndt et al. \(2012\)](#) measure how fiscal shocks are absorbed by reductions in debt returns (the debt valuation channel or else *fiscal insurance*) or by increases in tax revenues (the surplus channel) in post-war U.S. data and find evidence of limited but non-negligible fiscal insurance. In contrast, optimal policy in an expected utility economy prescribes that the majority of fiscal risk should be absorbed by reductions in returns. Turning to a recursive utility economy, the debt valuation channel is even more prominent and can surpass 100%, in order to compensate for the fact that the surplus channel becomes essentially *inoperative*, due to the decrease of tax rates when an adverse fiscal shock hits. Thus, if we evaluate actual policy from the normative lens of an economy that generates a higher market price of risk the following conclusion emerges: actual fiscal policy is even worse than we thought.

The basic insights of optimal fiscal policy with recursive utility extend also in an economy with capital as in the setups of [Chari et al. \(1994\)](#) and [Zhu \(1992\)](#). The planner still over-insures and sets high and persistent labor taxes against good shocks. Furthermore, in contrast to the essentially zero ex-ante capital tax result of [Chari et al. \(1994\)](#) and [Zhu \(1992\)](#), there is an incentive to introduce an ex-ante *subsidy*. The reason is simple: the planner again mitigates fiscal shocks and manipulates prices by using a state-contingent subsidy to capital income in bad times and a state-contingent capital tax in good times. Bad times are weighed more though due to high marginal utility and a high marginal product of capital, leading on average to a capital subsidy.

Related literature. The main reference on optimal taxation with time-additive expected utility for an economy without capital is [Lucas and Stokey \(1983\)](#). The respective references for an economy with capital are [Chari et al. \(1994\)](#) and [Zhu \(1992\)](#). The models I examine reduce to the models analyzed in these studies, if I equate the risk aversion parameter to the inverse of the intertemporal elasticity of substitution parameter. Furthermore, the economy with capital reduces to the deterministic economy of [Chamley \(1986\)](#), if I shut off uncertainty.²

Related studies include [Farhi and Werning \(2008\)](#), who analyze the implications of recursive

²It 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 \(1960\)](#). See [Chari and Kehoe \(1999\)](#) for a comprehensive survey of optimal fiscal policy.

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. Of interest is also the work of [Gottardi et al. \(2015\)](#), who study optimal taxation of human and physical capital with uninsurable idiosyncratic shocks and recursive preferences.³

Other studies have analyzed the interaction of fiscal policies and asset prices with recursive preferences from a positive angle. [Gomes et al. \(2013\)](#) build a quantitative model and analyze the implications of fiscal policies on asset prices and the wealth distribution. [Croce et al. \(2012a\)](#) show that corporate taxes can create sizeable risk premia with recursive preferences. [Croce et al. \(2012b\)](#) analyze the effect of exogenous fiscal rules on the endogenous growth rate of the economy. None of these studies though considers *optimal* policy.

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 \(2010\)](#), [Shin \(2006\)](#), [Sleet and Yeltekin \(2006\)](#), [Bhandari et al. \(2016\)](#) or with time-additive expected utility and private information as in [Sleet \(2004\)](#). In studies like [Aiyagari et al. \(2002\)](#), who provided the foundation of the tax-smoothing results of [Barro \(1979\)](#), 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 increases the tax rate when government spending is high. Instead, the opposite happens in the current paper.⁴ More generally, with incomplete markets as in [Aiyagari et al. \(2002\)](#), the planner would like to allocate tax distortions in a constant way across states and dates but he *cannot*, whereas with complete markets and recursive preferences he could in principle follow a constant distortion policy, but does not find it optimal to do so.

The paper is organized as follows. Section 2 lays out an economy without capital and section 3 sets up the Ramsey problem, its recursive formulation and derives the associated optimality conditions. Section 4 is devoted to the analysis of the *excess burden* of distortionary taxation, a multiplier that reflects how tax distortions are allocated across states and dates. The implications for labor taxes are derived in section 5. Detailed numerical exercises are provided in section 6. Section 7 analyzes government debt returns and optimal fiscal insurance. Section 8 extends the analysis to an economy with capital and derives implications for the ex-ante capital tax. Section 9 discusses the case of preference for late resolution of uncertainty. Finally, section 10 concludes and an Appendix follows. A separate Online Appendix provides additional details and robustness exercises.

³There is an extensive literature that studies optimal risk-sharing with recursive utility. See [Anderson \(2005\)](#) and references therein.

⁴Furthermore, with incomplete markets as in [Aiyagari et al. \(2002\)](#), it is typically optimal to front-load distortions in order to create a buffer stock of assets, furnishing a tax rate with a negative drift. In contrast, in the current analysis the tax rate exhibits a positive drift, in order to take advantage of cheaper state-contingent debt. It is interesting to observe that [Sleet \(2004\)](#) also obtains a positive drift in the tax rate in a setup with private information about the government spending needs.

2 Economy without capital

I start the analysis of optimal fiscal policy with recursive utility in an economy without capital as in [Lucas and Stokey \(1983\)](#). In a later section, I extend the analysis to an economy with capital as in [Chari et al. \(1994\)](#) and [Zhu \(1992\)](#) and I derive the implications for capital taxation.

Time is discrete and the horizon is infinite. There is uncertainty in the economy stemming from exogenous government expenditure shocks g . Shocks take values in a finite set. Let $g^t \equiv (g_0, g_1, \dots, g_t)$ denote the partial history of shocks up to time t and let $\pi_t(g^t)$ denote the probability of this history. The initial shock is assumed to be given, so that $\pi_0(g_0) = 1$.

The economy is populated by a representative household that is endowed with one unit of time and consumes $c_t(g^t)$, works $h_t(g^t)$, pays linear labor income taxes with rate $\tau_t(g^t)$ and trades in complete asset markets. Leisure of the household is $l_t(g^t) = 1 - h_t(g^t)$. The notation denotes that the relevant variables are measurable functions of the history g^t . Labor markets are competitive, which leads to an equilibrium wage of unity, $w_t(g^t) = 1$. The resource constraint in the economy reads

$$c_t(g^t) + g_t = h_t(g^t), \forall t, g^t. \quad (1)$$

2.1 Preferences

The representative household ranks consumption and leisure plans following a recursive utility criterion of [Kreps and Porteus \(1978\)](#). I focus on the isoelastic preferences of [Epstein and Zin \(1989\)](#) and [Weil \(1990\)](#) (EZW henceforth), that are described by the utility recursion

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

where $u(c, 1 - h) > 0$. The household derives utility from a composite good that consists of consumption and leisure, $u(c, 1 - h)$, and from the certainty equivalent of continuation utility, $\mu_t \equiv (E_t V_{t+1}^{1-\gamma})^{\frac{1}{1-\gamma}}$. E_t denotes the conditional expectation operator given information at t with respect to measure π . The parameter $1/\rho$ captures the constant intertemporal elasticity of substitution between the composite good and the certainty equivalent, 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 $\rho = \gamma$. This is easily seen by applying the monotonic transformation $v_t \equiv \frac{V_t^{1-\rho} - 1}{(1-\beta)(1-\rho)}$, since the utility recursion (2) becomes

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

where $U(c, 1 - h) \equiv \frac{u^{1-\rho}-1}{1-\rho}$. Recursion (3) implies that the household is *averse* to volatility in future utility when $\rho < \gamma$, whereas it *loves* volatility when $\rho > \gamma$.⁵ Thus, when $\rho < \gamma$, recursive utility adds *curvature* with respect to future risks, a feature that is typically necessary to match asset-pricing data.⁶ For that reason, I assume $\rho < \gamma$ for the main body of the paper, unless otherwise specified. In a later section I consider also the case of $\rho > \gamma$.

When $\rho = 1$, recursion (2) becomes $V_t = u_t^{1-\beta} \mu_t^\beta$. Using the transformation $v_t \equiv \frac{\ln V_t}{1-\beta}$ we get

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

which for $\gamma > 1$ has the interpretation of a *risk-sensitive* recursion with risk-sensitivity parameter $\sigma \equiv (1-\beta)(1-\gamma)$.⁷

It will be useful to define

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

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 $\pi_{t+1}(g_{t+1}|g^t)$, or, in other words, a conditional likelihood ratio. Similarly, define the product of the conditional likelihood ratios as $M_t(g^t) \equiv \prod_{i=1}^t m_i(g^i)$, $M_0 \equiv 1$. 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.

⁵Define the monotonic function $H(x) \equiv \left[(1 + (1-\beta)(1-\rho)x)^{\frac{1-\gamma}{1-\rho}} - 1 \right] / [(1-\beta)(1-\gamma)]$. Recursion (3) can be written as $v_t = U_t + \beta H^{-1}(E_t H(v_{t+1}))$. $H(x)$ is concave for $\rho < \gamma$ and convex for $\rho > \gamma$. The aversion or love of utility volatility correspond respectively to preference for early or late resolution of uncertainty. They contrast to the case of $\rho = \gamma$, which features neutrality to future risks and therefore indifference to the temporal resolution of uncertainty.

⁶See for example Tallarini (2000), Bansal and Yaron (2004), Piazzesi and Schneider (2007) and Epstein et al. (2014).

⁷More 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 of the multiplier preferences and an analysis of the relationship between ambiguity aversion and temporal resolution of uncertainty respectively.

2.2 Competitive equilibrium

Household's problem. Let $\{x\} \equiv \{x_t(g^t)\}_{t \geq 0, g^t}$ stand for the sequence of an arbitrary random variable x_t . The representative household chooses $\{c, h, b\}$ to maximize $V_0(\{c\}, \{h\})$ subject to

$$c_t(g^t) + \sum_{g^{t+1}} p_t(g_{t+1}, g^t) b_{t+1}(g^{t+1}) \leq (1 - \tau_t(g^t)) h_t(g^t) + b_t(g^t),$$

the non-negativity constraint for consumption $c_t(g^t) \geq 0$ and the feasibility constraint for labor $h_t(g^t) \in [0, 1]$, where initial debt b_0 is given. The variable $b_{t+1}(g^{t+1})$ stands for the holdings at history g^t of an Arrow claim that delivers one unit of consumption next period if the state is g_{t+1} and zero units otherwise. This security trades at price $p_t(g_{t+1}, g^t)$ in units of the history-contingent consumption $c_t(g^t)$.

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

$$\lim_{t \rightarrow \infty} \sum_{g^{t+1}} q_{t+1}(g^{t+1}) b_{t+1}(g^{t+1}) \geq 0 \quad (6)$$

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

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

$$b_t(g^t) + g_t = \tau_t(g^t) h_t(g^t) + \sum_{g^{t+1}} p_t(g_{t+1}, g^t) b_{t+1}(g^{t+1}),$$

When $b_t > 0$, the government borrows from the household and when $b_t < 0$, the government lends to the household.

Definition 1. *A competitive equilibrium with taxes is a stochastic process for prices $\{p\}$, an allocation $\{c, h, b\}$ and a government policy $\{g, \tau, b\}$ such that: 1) Given prices $\{p\}$ and taxes $\{\tau\}$, the allocation $\{c, h, b\}$ solves the households's problem. 2) Prices are such so that markets clear, i.e. the resource constraint (1) holds.*

2.3 Household's optimality conditions

The labor supply decision of the household is governed by

$$\frac{U_l(g^t)}{U_c(g^t)} = 1 - \tau_t(g^t), \quad (7)$$

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,

$$\begin{aligned} p_t(g_{t+1}, g^t) &= \beta \pi_{t+1}(g_{t+1}|g^t) \left(\frac{V_{t+1}(g^{t+1})}{\mu_t(V_{t+1})} \right)^{\rho-\gamma} \frac{U_c(g^{t+1})}{U_c(g^t)} \\ &= \beta \pi_{t+1}(g_{t+1}|g^t) m_{t+1}(g^{t+1})^{\frac{\rho-\gamma}{1-\gamma}} \frac{U_c(g^{t+1})}{U_c(g^t)}, \end{aligned} \quad (8)$$

where the second line uses the definition of the conditional likelihood ratio (5). The transversality condition is

$$\lim_{t \rightarrow \infty} \sum_{g^{t+1}} \beta^{t+1} \pi_{t+1}(g^{t+1}) M_{t+1}(g^{t+1})^{\frac{\rho-\gamma}{1-\gamma}} U_c(g^{t+1}) b_{t+1}(g^{t+1}) = 0. \quad (9)$$

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

$$S_{t+1} \equiv \beta \left(\frac{V_{t+1}}{\mu_t} \right)^{\rho-\gamma} \frac{U_{c,t+1}}{U_{ct}} = \beta m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} \frac{U_{c,t+1}}{U_{ct}}. \quad (10)$$

The stochastic discount factor features *continuation values*, scaled by their certainty equivalent μ_t , when $\rho \neq \gamma$. 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 – captured by continuation values– *directly* affects S_{t+1} . Increases in consumption growth at $t + 1$ reduce period marginal utility and therefore the stochastic discount factor in the standard time-additive setup. When $\rho < \gamma$, increases in continuation values act exactly the *same* way; they decrease the stochastic discount factor, because the household dislikes volatility in future utility. This is the essence of the additional “curvature” that emerges with recursive utility.⁸

3 Ramsey problem

I formulate the Ramsey problem under commitment. I follow the primal approach of [Lucas and Stokey \(1983\)](#) and set up a problem of a Ramsey planner, who chooses allocations that satisfy the resource constraint (1) and *implementability* constraints, i.e. constraints that allow the optimal allocation to be implemented as a competitive equilibrium.

⁸[Bansal and Yaron \(2004\)](#) and [Hansen et al. \(2008\)](#) have explored ways of making the continuation value channel quantitatively important in order to increase the market price of risk.

3.1 Implementability constraints

Use (7) and (8) to eliminate labor taxes and equilibrium prices from the household's dynamic budget constraint. This leads to a sequence of implementability constraints:

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

$$U_{ct}b_t = U_{ct}c_t - U_{lt}h_t + \beta E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} U_{c,t+1}b_{t+1}, t \geq 0$$

where $c_t \geq 0$, $h_t \in [0, 1]$ and (b_0, g_0) given. Furthermore, the transversality condition (9) has to be satisfied. The conditional likelihood ratios $m_{t+1} = V_{t+1}^{1-\gamma} / E_t V_{t+1}^{1-\gamma}$, $t \geq 0$, are determined by continuation values that follow 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, as we will see in the next section.

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 (1).*

3.2 Recursive formulation

I follow the methodology of Kydland and Prescott (1980) and break the Ramsey problem in two subproblems: the problem from period one onward and the initial period problem. Let z_t denote debt in (period) marginal utility units, $z_t \equiv U_{ct}b_t$. I represent the policy problem for $t \geq 1$ recursively by keeping track of g – the exogenous shock – and z , the variable that captures the commitment of the planner to his past promises. Note that z is a forward-looking variable that it not inherited from the past. This creates the need to specify $Z(g)$, the space where z lives. The set $Z(g)$ represents the values of debt in marginal utility units that can be generated from an implementable allocation when the initial shock is g and is defined in the Appendix.⁹ Let $V(z_1, g_1)$ denote the value function of the planner's problem from period one onward, where $z_1 \in Z(g_1)$ and assume that shocks follow a Markov process with transition probabilities $\pi(g'|g)$.

Bellman equation. The functional equation that determines the value function V takes the form

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

⁹A separate Online Appendix provides the sequential formulation of the Ramsey problem.

subject to

$$z = U_c c - U_l h + \beta \sum_{g'} \pi(g'|g) \frac{V(z'_{g'}, g')^{\rho-\gamma}}{[\sum_{g'} \pi(g'|g) V(z'_{g'}, g')^{1-\gamma}]^{\frac{\rho-\gamma}{1-\gamma}}} z'_{g'} \quad (11)$$

$$c + g = h \quad (12)$$

$$c \geq 0, h \in [0, 1] \quad (13)$$

$$z'_{g'} \in Z(g'). \quad (14)$$

The planner is maximizing welfare by choosing consumption, labor (and thus effectively the labor tax) and state-contingent debt for next period $z'_{g'}$, subject to the government budget constraint (11) (expressed in terms of allocations), and the resource constraint, (12). The nature of the Ramsey problem is fundamentally changed because, in contrast to time-additive utility, continuation values matter for the determination of the market value of the government debt portfolio, and therefore show up in constraint (11). As such, the dynamic tradeoff of taxing at the current period versus postponing taxation for the future by issuing debt is altered because the planner has to take into account how debt issuance affects equilibrium prices through the “long-run.” This tradeoff is at the heart of next section.

Initial period problem. The value of z_1 that was taken as given in the formulation of the planner’s problem at $t \geq 1$ is chosen optimally at $t = 0$. In this sense, 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. The initial period problem is stated in the Online Appendix.

3.3 Optimality conditions

Let

$$\Omega(c, h) \equiv U_c(c, 1 - h)c - U_l(c, 1 - h)h, \quad (15)$$

with $\Omega_i, i = c, h$ denoting the respective partial derivatives. The variable Ω stands for consumption net of after-tax labor income, in (period) marginal utility units. In equilibrium it is also equal to the *primary surplus* in marginal utility units, $U_c \cdot (\tau h - g)$.

It is easier to derive the optimality conditions by using the transformed value function, $v(z, g) \equiv \frac{V(z, g)^{1-\rho}-1}{(1-\beta)(1-\rho)}$, that corresponds to recursion (3). Let Φ and λ be the respective multipliers on the dynamic implementability constraint and the resource constraint of the transformed problem. The first-order necessary conditions for an interior solution at points of differentiability of the value function are

$$c : \quad U_c + \Phi \Omega_c = \lambda \tag{16}$$

$$h : \quad U_l - \Phi \Omega_h = \lambda \tag{17}$$

$$z'_{g'} : \quad v_z(z'_{g'}, g') + \Phi [1 + (1 - \beta)(\rho - \gamma)v_z(z'_{g'}, g')\eta'_{g'}] = 0, \tag{18}$$

where

$$\eta'_{g'} \equiv V'_{g'}{}^{\rho-1} z'_{g'} - \mu^{\rho-1} \sum_{g'} \pi(g'|g) m'_{g'}{}^{\frac{\rho-\gamma}{1-\gamma}} z'_{g'}. \tag{19}$$

In definition (19) recall that $m'_{g'}$ stands for the conditional likelihood ratio and μ for the certainty equivalent.¹⁰

Equivalently, by using the definition of $m'_{g'}$, we can rewrite the variable $\eta'_{g'}$ as $\eta'_{g'} = V'_{g'}{}^{\rho-1} z'_{g'} - \sum_{g'} \pi(g'|g) m'_{g'} V'_{g'}{}^{\rho-1} z'_{g'}$. So it stands for the conditional innovation of $V'_{g'}{}^{\rho-1} z'_{g'}$ under $\pi_t \cdot M_t$ and takes positive and negative values with an average of zero, $\sum_{g'} \pi(g'|g) m'_{g'} \eta'_{g'} = 0$.

For $\rho = 1$, $\eta'_{g'}$ simplifies to $\eta'_{g'} = z'_{g'} - \sum_{g'} \pi(g'|g) m'_{g'} z'_{g'}$, so it simplifies to the state-contingent debt position in marginal utility units relative to the value of the government debt portfolio. For that reason, I call $\eta'_{g'}$ the government's *relative debt position* in marginal utility units.

The envelope condition takes the form $v_z(z, g) = -\Phi$. Note that $\Phi \geq 0$, so $v_z(z, g) \leq 0$.¹¹ The multiplier Φ is strictly positive if the first-best cannot be achieved, i.e. if the government does not have sufficient assets to support the first-best allocation.

4 Recursive utility and the excess burden of taxation

4.1 Overview of the mechanism

How does the government tax across states and dates and how does it manage its state-contingent debt in a welfare-maximizing way? To fix ideas, I provide here an overview of the mechanism that is supported by the analysis of the optimality conditions and the numerical analysis of later sections.

The government is absorbing spending shocks through its debt portfolio. It achieves that by selling claims to consumption against low spending shocks (good times) and by purchasing claims to consumption against high spending shocks (bad times). In the standard time-additive setup,

¹⁰ $V'_{g'}$ is shorthand for $V(z'_{g'}, g')$. I use the non-transformed value function V in (19) (which is equal to $[1 + (1 - \beta)(1 - \rho)v]^{\frac{1}{1-\rho}}$) as a matter of convenience; it allows a more compact exposition of the first-order conditions.

¹¹I am implicitly assuming that the government has access to lump-sum transfers, so that the dynamic implementability constraint takes the form $z_t \leq \Omega_t + \beta E_t m_{t+1}{}^{\frac{\rho-\gamma}{1-\gamma}} z_{t+1}$.

the size of sales and purchases of state-contingent claims is such so that the tax rate remains essentially constant across states and dates, leading to the typical tax-smoothing result. Note that consumption is high (low), and therefore the stochastic discount factor is low (high) when spending shocks are low (high). So the price of claims sold is low and the price of claims bought is high.

The government has the same motives to use state-contingent debt in order to hedge fiscal shocks in a recursive utility economy. The difference is that there is a novel instrument of affecting the stochastic discount factor, lifetime utilities, which allows the government to make fiscal insurance cheaper. The government achieves that by “over-insuring,” i.e. it sells more claims to consumption against low shocks and purchases more claims to consumption against high shocks. Issuance of more debt against low spending shocks *reduces* continuation utilities and *increases* the stochastic discount factor and, as a result, the price of claims sold. Thus, the current revenue from selling claims to the private sector against a low spending shock next period increases, allowing the relaxation of the government budget and less taxation at the current period. More claims sold against a low shock next period implies that higher taxes have to be levied in the future at that state, in order to repay debt. A similar mechanism holds for high spending states: the government insures against fiscal risk by purchasing more claims to consumption against high spending shocks, which increases the household’s utility, depressing therefore the stochastic discount factor and the price of claims bought. More assets (or less debt) against high shocks implies less taxes contingent on these states of the world.

The mechanism is intuitive and makes economic sense. It simply says that the planner should mitigate the effects of fiscal shocks by taxing more in good times and less in bad times, since this way state-contingent debt against good times becomes cheaper and state-contingent assets against bad times more profitable, due to the additional curvature of recursive utility. Furthermore, this mechanism leads on average to back-loading of tax distortions over time, due to the reduced interest rate cost of debt. Lastly, persistence of optimal tax rates is optimal independent of the persistence of exogenous shocks: the planner changes smoothly the tax rate over time in order to take full advantage of the forward-looking nature of continuation utilities.

4.2 Preliminaries: expected utility and the excess burden

Proceeding now to the specifics of the mechanism, note that the entire action is coming from Φ , the multiplier on the implementability constraint, which captures the cost of an additional unit of debt in marginal utility units. Increases in debt are costly because they have to be accompanied with an increase in distortionary taxation ($\Phi = 0$ when lump-sum taxes are available). I refer to Φ as the *excess burden* of distortionary taxation and interpret it as an indicator of tax *distortions*.

In order to build intuition about the excess burden of taxation, consider first the time-additive expected utility world of [Lucas and Stokey \(1983\)](#) where $\rho = \gamma$. It is easy to see that the optimality condition with respect to z'_g takes the form

$$-v_z(z'_{g'}, g') = \Phi. \quad (20)$$

Optimality condition (20) has a typical marginal cost-marginal benefit interpretation. The left-hand side captures the marginal cost of issuing more debt against g' next period. Selling more claims to consumption at g' is costly because the planner has to increase distortionary taxation in order to repay debt. However, by issuing more debt for next period, the planner can relax the government budget and tax less at the current period. The marginal benefit of relaxing the budget constraint has shadow value Φ , which is the right-hand side of (20).

By using the envelope condition, condition (20) implies that $\Phi'_{g'} = \Phi, \forall g'$, for all values of the state (z, g) . Thus, in a time-additive expected utility economy, the planner sells and buys as many state-contingent claims as necessary, in order to *equalize* the excess burden of taxation across states and dates. This is the formal result that hides behind the *tax-smoothing* intuition in typical frictionless Ramsey models. Furthermore, the constant excess burden is also the source of Lucas and Stokey's celebrated *history-independence* result, since optimal allocations and tax rates can be written solely as functions of the exogenous shocks and the constant Φ .¹²

4.3 Pricing with recursive utility and the excess burden

Turn now to the recursive utility case. The optimality condition with respect to $z'_{g'}$ (18) can be written as:

$$\underbrace{-v_z(z'_{g'}, g')}_{\text{MC of increasing } z'_{g'}} = \Phi \cdot \left[\underbrace{1}_{\text{EU term}} + \underbrace{(1 - \beta)(\rho - \gamma)v_z(z'_{g'}, g')\eta'_{g'}}_{\text{EZW term: price effect of increasing } z'_{g'}} \right]. \quad (21)$$

As before, the left-hand side denotes the marginal cost of issuing more debt tomorrow at g' . The right-hand side of (21) captures the utility benefit from the government's *marginal revenue* from debt issuance. The first term in the right-hand side captures the same increase in revenue as in the time-additive setup, keeping prices *constant*. The second term is novel and is coming from the change in prices due to the increased debt position: an increase in debt reduces utility which *increases* the stochastic discount factor, $(\rho - \gamma)v_z > 0$ for $\rho < \gamma$.

How the planner is going to use this novel price effect of recursive utility depends, according to (21), on the *relative debt position* $\eta'_{g'}$. To see clearly the mechanism, turn into sequence notation, collect the terms that involve v_z , and use the envelope condition in order to rewrite (21) in terms

¹²The excess burden of taxation is also constant in a deterministic economy ($\eta'_{g'} \equiv 0, \forall g$). Thus, apart from the *level* of the constant Φ , there is no essential difference between a deterministic world and a stochastic but time-additive world.

of the *inverse* of Φ (assuming that Φ is not zero) as¹³

$$\frac{1}{\Phi_{t+1}} = \frac{1}{\Phi_t} + (1 - \beta)(\rho - \gamma)\eta_{t+1}, t \geq 0, \quad (22)$$

where $\eta_{t+1} \equiv V_{t+1}^{\rho-1} z_{t+1} - \mu_t^{\rho-1} E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} z_{t+1} = V_{t+1}^{\rho-1} z_{t+1} - E_t m_{t+1} V_{t+1}^{\rho-1} z_{t+1}$. Consider fiscal shocks \hat{g} and \tilde{g} at $t + 1$ such that $\eta_{t+1}(\hat{g}) > 0 > \eta_{t+1}(\tilde{g})$. Then, (22) implies that $\Phi_{t+1}(\hat{g}) > \Phi_t > \Phi_{t+1}(\tilde{g})$ for $\rho < \gamma$. So, in contrast to the time-additive setup, the excess burden of taxation, and therefore the tax rate, varies across states and dates and is higher at states of the world next period against which the relative debt position is positive, and lower at states of the world against which the relative debt position is negative.¹⁴

What is happening here? Exactly the story that we highlighted in the overview of the mechanism. The increase in prices due to the additional curvature of recursive utility is beneficial at states of the world against which the planner issues relatively more debt. In other words, the planner taxes more at states of the world against which it is *cheaper* to issue debt. The opposite happens for states of the world against which he has a relatively small debt position.

Two comments are due. First, note that is not just the debt position (adjusted by marginal utility – and continuation utility when $\rho \neq 1$) but the debt position *relative* to (a multiple of) the market value of the debt portfolio, $E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} z_{t+1}$, that matters for the increase or decrease of the excess burden of taxation across states and dates. The reason for this is coming from the state *non-separabilities* that emerge with recursive utility. In particular, an increase of $z'_{g'}$ may increase the price of the respective claim at g' by reducing utility, but reduces also the certainty equivalent and *decreases* therefore the rest of the prices of state-contingent claims at $\bar{g}, \bar{g} \neq g'$. This is why the *relative* position η_{t+1} captures the net effect of price manipulation through the continuation utility channel.

Second, in the overview of the mechanism we stressed that the government is using state-contingent debt to hedge fiscal shocks by selling claims against low spending shocks and purchasing claims (or selling less claims) against high spending shocks. Thus, we expect to have $b_{t+1}(g_L) > b_{t+1}(g_H)$ for $g_H > g_L$. Assume that $\rho = 1 < \gamma$ and that the same ranking of debt positions holds also for debt in marginal utility units, i.e. $z_{t+1}(g_L) > z_{t+1}(g_H)$. Then, $\eta_{t+1}(g_L) > 0 > \eta_{t+1}(g_H)$, which implies that $\Phi_{t+1}(g_L) > \Phi_t > \Phi_{t+1}(g_H)$. Thus, the excess burden, and therefore the tax rate, increases for low fiscal shocks and decreases for high fiscal shocks, leading to increased surpluses

¹³Otherwise, write the optimality condition as $\Phi_{t+1} = \Phi_t / [1 + (1 - \beta)(\rho - \gamma)\eta_{t+1}\Phi_t]$. Thus, if $\Phi_t = 0$, then $\Phi_{t+i} = 0, i \geq 0$, so the first-best is an absorbing state.

¹⁴The varying excess burden has also implications for the size of z_t over time. It is tempting to deduce that the planner is not only increasing the excess burden for a high-debt state next period ($\eta_{t+1} > 0$), but also issues more state-contingent debt for next period. Formally, the deduction would be $\Phi_{t+1} = -v_z(z_{t+1}, g) > \Phi_t = -v_z(z_t, g) \Rightarrow z_{t+1} > z_t$, which is a statement about the concavity of v at g . This statement cannot be made in general due to the non-convexities of the Ramsey problem, but it turns out to be numerically true.

and deficits. We are going to see explicitly this fiscal hedging when we solve the model numerically.

To conclude this section, the following proposition summarizes the results about the excess burden of taxation.

Proposition 2. 1. *The excess burden is constant across states and dates when $\rho = \gamma$.*

2. *Assume $\rho < \gamma$ and let \hat{g} and \tilde{g} be shocks at $t + 1$ such that $\eta_{t+1}(\hat{g}) > 0 > \eta_{t+1}(\tilde{g})$. Then, the law of motion of the excess burden (22) implies that $\Phi_{t+1}(\hat{g}) > \Phi_t > \Phi_{t+1}(\tilde{g})$.*

3. *(Fiscal hedging and the excess burden) Let $g_H > g_L$ and assume that $\rho = 1 < \gamma$. If $z_{t+1}(g_L) > z_{t+1}(g_H)$, then $\Phi_{t+1}(g_L) > \Phi_t > \Phi_{t+1}(g_H)$.¹⁵*

4.4 Dynamics of the excess burden of taxation

The relative debt position η_t captures the incentives of the planner to increase or decrease the excess burden, given the excess burden of the previous period, i.e. given the past shadow cost of debt and tax promises. To understand the role of the past, consider a change in debt at time t . This change will affect continuation values at t but also at all previous periods, because utilities are forward-looking, i.e. the household at $t - i, i = 1, 2, \dots, t$ is taking into account the entire future stream of consumption and leisure when it prices Arrow claims. As a result, *all* past prices of state-contingent claims $p_i(s_{i+1}, s^i), i = 0, 1, 2, \dots, t - 1$ change with a change in continuation values at time t . This is why the excess burden depends on the sum of the past relative debt positions $\{\eta_i\}_{i=1}^t$, as we can see by solving (22) backwards. Furthermore, we have:

Proposition 3. *(Persistence and back-loading of the excess burden)*

The inverse of Φ_t is a martingale with respect to the continuation-value adjusted measure $\pi_t \cdot M_t$ for $\rho \leq \gamma$. Therefore, Φ_t is a submartingale with respect to $\pi_t \cdot M_t$, $E_t m_{t+1} \Phi_{t+1} \geq \Phi_t$. As a result,

$$E_t \Phi_{t+1} \geq \Phi_t - \text{Cov}_t(m_{t+1}, \Phi_{t+1}), \quad (23)$$

so if $\text{Cov}_t(m_{t+1}, \Phi_{t+1}) \leq 0$, Φ_t is a submartingale with respect to π , $E_t \Phi_{t+1} \geq \Phi_t$.

Proof. Take conditional expectation in (22) 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 $E_t m_{t+1} \eta_{t+1} = 0$. 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

¹⁵Clearly, the corresponding statement for $\rho \neq 1 < \gamma$ is: if $V_{t+1}^{\rho-1}(g_L) z_{t+1}(g_L) > V_{t+1}^{\rho-1}(g_H) z_{t+1}(g_H)$, then $\Phi_{t+1}(g_L) > \Phi_t > \Phi_{t+1}(g_H)$.

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 $E_t m_{t+1} \Phi_{t+1} \geq \Phi_t$. Inequality (23) is derived by the submartingale result and the fact that $E_t m_{t+1} \Phi_{t+1} = \text{Cov}_t(m_{t+1}, \Phi_{t+1}) + E_t \Phi_{t+1}$, since $E_t m_{t+1} = 1$. \square

The martingale result about the inverse of the excess burden of taxation implies *persistence* independent of the stochastic properties of exogenous shocks, contrasting the standard time-additive Ramsey results.¹⁶ Furthermore, the submartingale result shows that the planner wants on “average” to *back-load* tax distortions, in the sense that the excess burden exhibits a positive drift with respect to the continuation-value adjusted measure, independent of $\rho \leq \gamma$. In order to determine the drift with respect to the actual measure that generates uncertainty, π , we need to determine the covariance of the excess burden with the change of measure m_{t+1} . Consider without loss of generality the case of $\rho = 1 < \gamma$. Then, high fiscal shocks, since they reduce utility, are associated with a higher conditional probability mass and therefore a higher m_{t+1} , leading to a positive correlation of m_{t+1} with spending. Furthermore, we expect the excess burden to be negatively correlated with spending. As a result, we expect $\text{Cov}_t(m_{t+1}, \Phi_{t+1}) \leq 0$ and therefore proposition 3 implies a positive drift in Φ_t with respect to π . More intuitively, since the average excess burden of taxation is increasing according to the utility-adjusted beliefs that do *not* assign a lot of probability mass on states of the world with a high excess burden, it will still be increasing on average according to the data-generating process, which puts more weight on exactly these contingencies of a high excess burden. We will explore further the persistence and the back-loading of tax distortions in the numerical exercises section.

5 Optimal labor income taxation

The following proposition exhibits the exact relationship of the excess burden of taxation with the labor tax.

Proposition 4. (*Labor tax*) *The optimal labor tax is*

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

where $\epsilon_{cc} \equiv -U_{cc}c/U_c > 0$ and $\epsilon_{ch} \equiv U_{cl}h/U_c$, i.e. the own and cross elasticity of the period marginal utility of consumption, and $\epsilon_{hh} \equiv -U_{ll}h/U_l > 0$ and $\epsilon_{hc} \equiv U_{lc}c/U_l$, the own and cross elasticity of the period marginal disutility of labor. When $U_{cl} \geq 0$, then $\epsilon_{ch}, \epsilon_{hc} \geq 0$ and $\tau_t \geq 0$.¹⁷

¹⁶In the Online Appendix I discuss why the martingale property is not sufficient to establish convergence results of the inverse of the excess burden with respect to π and I elaborate on candidate convergence points for the utility function of section 6.

¹⁷The labor tax formula holds also for the deterministic and stochastic time-additive case for any period utility U that satisfies the standard monotonicity and concavity assumptions, i.e. without being restricted to $U = (u^{1-\rho} - 1)/(1 - \rho)$, $u > 0$.

Proof. Combine the first-order conditions (16)-(17) to get the optimal wedge in labor supply as $\frac{U_l}{U_c} \cdot \frac{1-\Phi \frac{\Omega_h}{U_l}}{1+\Phi \frac{\Omega_c}{U_c}} = 1$. Calculate the derivatives to get $\Omega_c/U_c = 1 - \epsilon_{cc} - \epsilon_{ch}$ and $\Omega_h/U_l = -1 - \epsilon_{hh} - \epsilon_{hc}$. Use the labor supply condition $U_l/U_c = 1 - \tau$ and rewrite the optimal wedge as $\tau = -\Phi(\Omega_c/U_c + \Omega_h/U_l)/(1 - \Phi\Omega_h/U_l)$. Use now the elasticity formulas to get the result. \square

The formula in proposition 4 expresses the optimal labor tax in terms of the excess burden of taxation Φ_t and the elasticities of the period marginal utility of consumption and disutility of labor. Ceteris paribus, the labor tax varies *monotonically* with the excess burden of taxation, a fact which justifies the interpretation of Φ_t as an indicator of tax distortions.¹⁸ Period elasticities essentially capture the sensitivity to shocks of the marginal rate of substitution of consumption and leisure and the (period) marginal utility component of the pricing kernel. In the time-additive case where $\Phi_t = \Phi$, the *only* variation in the labor tax occurs through variation in period elasticities, so when elasticities are *constant*, optimal policy prescribes *perfect* tax-smoothing.

With recursive utility though, even in the constant period elasticity case, the labor tax varies monotonically with the non-constant excess burden of taxation. Consider for example the composite good u

$$u(c, 1 - h) = \left[c^{1-\rho} - (1 - \rho)a_h \frac{h^{1+\phi_h}}{1 + \phi_h} \right]^{\frac{1}{1-\rho}}, \quad (24)$$

which implies a period utility function with constant elasticities, $U = \frac{c^{1-\rho}-1}{1-\rho} - a_h \frac{h^{1+\phi_h}}{1+\phi_h}$.¹⁹ We get the following proposition:

Proposition 5. (*Labor tax with power utility and constant Frisch elasticity*)

1. *The labor tax follows the law of motion*

$$\frac{1}{\tau_{t+1}} = \frac{1}{\tau_t} + \frac{(1 - \beta)(\rho - \gamma)}{\rho + \phi_h} \eta_{t+1}, t \geq 1. \quad (25)$$

2. *Tax rates across states and dates:*

- *Let $\rho < \gamma$. Let \hat{g} and \tilde{g} be shocks at $t + 1$ such that $\eta_{t+1}(\hat{g}) > 0 > \eta_{t+1}(\tilde{g})$. Then, $\tau_{t+1}(\hat{g}) > \tau_t > \tau_{t+1}(\tilde{g})$.*

¹⁸We have $\frac{\partial \tau}{\partial \Phi}|_{\epsilon_{i,j} \text{ constant}} = \frac{\epsilon_{cc} + \epsilon_{ch} + \epsilon_{hh} + \epsilon_{hc}}{[1 + \Phi(1 + \epsilon_{hh} + \epsilon_{hc})]^2} > 0$, as long as the numerator is positive. $U_{cl} \geq 0$ is sufficient for that.

¹⁹It is assumed that parameters are such so that $c^{1-\rho} - (1 - \rho)a_h \frac{h^{1+\phi_h}}{1 + \phi_h} > 0$, so that $u > 0$ is well defined. For $\rho = 1$, the utility recursion becomes $V_t = \exp \left[(1 - \beta)(\ln c - a_h \frac{h^{1+\phi_h}}{1 + \phi_h}) + \beta \ln \mu_t \right]$. If we want to drop these restrictions on preference parameters, we can just consider risk-sensitive preferences with the particular period utility U .

- Let $\rho = 1 < \gamma$ and assume that shocks take two values, $g_H > g_L$. If $z_{t+1}(g_L) > z_{t+1}(g_H)$, then $\tau_{t+1}(g_L) > \tau_t > \tau_{t+1}(g_H)$.²⁰

3. (Persistence and back-loading of the labor tax) The inverse of the labor tax is a martingale with respect to $\pi_t \cdot M_t$ for $\rho \leq \gamma$. Therefore, τ_t is a submartingale with respect to $\pi_t \cdot M_t$, $E_t m_{t+1} \tau_{t+1} \geq \tau_t$ and

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

If $\text{Cov}_t(m_{t+1}, \tau_{t+1}) \leq 0$, then $E_t \tau_{t+1} \geq \tau_t$.

Proof. The labor tax formula in proposition 4 specializes to

$$\tau_t = \frac{\Phi_t(\rho + \phi_h)}{1 + \Phi_t(1 + \phi_h)}, t \geq 1. \quad (26)$$

The formula shows that the crucial parameter for the period elasticities channel is ρ (and *not* γ), whereas both ρ and γ affect the Ramsey outcome through the law of motion of Φ_t , (22). Taking inverses in (26) delivers $\frac{1}{\tau_t} = \frac{1+\phi_h}{\rho+\phi_h} + \frac{1}{\rho+\phi_h} \frac{1}{\Phi_t}$, so $1/\tau_t$ is an *affine* function of $1/\Phi_t$. Use then (22) to get the law of motion of the labor tax (25). Notice the resemblance of (25) to (22), a fact that leads to the same conclusions about the variation of tax rates across states and dates and (sub)martingale properties as in proposition 3. \square

When we have a period utility function with a power subutility of consumption and constant Frisch elasticity, period elasticities are constant and the labor tax behaves *exactly* as the excess burden of taxation, with the elegant law of motion (25). The entire analysis of section 4 about the variation of the excess burden across states and dates, the positive drift and persistence, can be recast word by word in terms of the labor tax and will not be repeated.

6 Numerical exercises

In this section I provide various numerical exercises in order to highlight three main results of the paper: a) the planner's "over-insurance" that leads to higher tax rates when fiscal shocks are favorable and smaller tax rates when fiscal shocks are adverse, b) the volatility and back-loading of tax distortions, c) the persistence of tax distortions independent of the persistence of exogenous shocks. In a nutshell, the tax rate behaves like a random walk with a positive drift in the short and medium-run, with an increment that is negatively correlated with fiscal shocks. The unconditional volatility of the tax rate is orders of magnitude larger than in the time-additive expected utility case.

²⁰Same comment applies as in footnote 15.

6.1 Solution method

The numerical analysis with recursive preferences is highly non-trivial. There are three complications: At first, the state space where z lives is endogenous, i.e. we have to find values of debt in marginal utility units that can be generated at a competitive equilibrium. Second, the contraction property is impaired due to the presence of the value functions in the implementability constraint, a fact which makes convergence of iterative procedures difficult. Third, there are novel non-convexities in the implementability constraint due to recursive utility. I will illustrate the gist of the numerical method and provide additional details in the Online Appendix.

The way I proceed is as follows. I generate feasible values of z and calculate the respective utility by assuming that the planner follows a constant Φ policy, i.e. I assume that the planner ignores the prescriptions of optimal policy and just equalizes the excess burden of taxation over states and dates. By varying Φ I can generate a *set* of values of z , which I use as a proxy of the state space. The respective value functions are used as a first guess in the numerical algorithm. I implement a double loop: In the inner loop, I fix the value function in the constraint and solve the Bellman equation using grid search. The inner loop is convergent. In the outer loop, I update the value function in the constraint and repeat the inner loop. Although there is no guarantee of convergence of the double loop, this procedure works fairly well. After convergence, I add a final step to improve precision: I employ the output of the double loop as a first guess, fit the value functions with cubic splines and use a continuous optimization routine.

6.2 Calibration

I use the utility function of proposition 5 that delivers perfect tax-smoothing in the time-additive economy and a standard calibration. In particular, let $\rho = 1$ and consider the utility recursion

$$v_t = \ln c_t - a_h \frac{h_t^{1+\phi_h}}{1+\phi_h} + \frac{\beta}{(1-\beta)(1-\gamma)} \ln E_t \exp((1-\beta)(1-\gamma)v_{t+1}), \quad (27)$$

where $\gamma > 1$. The frequency is annual and Frisch elasticity is unitary, $(\beta, \phi_h) = (0.96, 1)$. The atemporal risk aversion is $\gamma = 10$.²¹ I assume that shocks are i.i.d. in order to focus on the persistence generated *endogenously* by optimal policy. Expenditures shocks take two values, $g_L = 0.072$ and $g_H = 0.088$, with probability $\pi = 0.5$. These values correspond to 18% and 22% of average first-best output respectively, or 20.37% and 24.28% of output in the second-best expected

²¹The range of the risk-aversion parameter varies wildly in studies that try to match asset-pricing facts. For example, Tallarini (2000) uses a risk aversion parameter above 50 in order to generate a high market price of risk, whereas Bansal and Yaron (2004) use low values of risk aversion in environments with long-run risks and stochastic volatility. Note that the plausibility of the size of atemporal risk aversion cannot be judged independently from the stochastic processes that drive uncertainty in the economy, since they jointly bear implications for the premium for early resolution of uncertainty. See Epstein et al. (2014) for a thoughtful evaluation of calibration practices in the asset-pricing literature from this angle.

utility economy. So the standard deviation of the share of government spending in output is small and about 2%. I set $a_h = 7.8125$ which implies that the household works on average 40% of its available time if we are at the first-best, or 35.8% of its time in the second-best time-additive economy. Initial debt is zero and the initial realization of the government expenditure shock is low, $g_0 = g_L$.

6.3 Expected utility plan

The time-additive expected utility case of $\gamma = 1$ corresponds to the environment of [Lucas and Stokey \(1983\)](#). The Ramsey plan is history-independent and the tax rate is constant and equal to 22.3% for our calibration. The planner issues zero debt against low shocks, $b_L = 0$, and insures against high spending by buying *assets*, $b_H < 0$. The level of assets corresponds to 3.81% of output. Thus, the debt-to-output ratio has mean -1.91% and standard deviation 1.91%. Whenever there is a low shock, the planner, who has no debt to repay ($b_L = 0$), runs a surplus $\tau h_L - g_L > 0$ and uses the surplus to buy assets against the high shock. The amount of assets is equal to $b_H = (\tau h_H - g_H)/(1 - \beta\pi)$. When the shock is high, the planner uses the interest income on these assets to finance the deficit $\tau h_H - g_H < 0$.²²

6.4 Fiscal hedging, over-insurance and price manipulation

Turning to recursive utility, the left panel in figure 1 plots the difference between the policy functions for z' next period when g' is low and high respectively. The graph shows that the government hedges fiscal shocks by issuing more debt in marginal utility units for the low shock and less for the high shock, $z'_L > z'_H$. Thus, as highlighted in the overview of the mechanism, propositions 2 and 5 imply that tax distortions decrease when fiscal shocks are high, $\Phi'_L > \Phi > \Phi'_H$ and $\tau'_L > \tau > \tau'_H$. The right panel in figure 1 plots the difference in the policy functions in the recursive utility and the expected utility case, $z'_i - z_i^{\text{EU}}, i = L, H$, in order to demonstrate the “over-insurance” property of the optimal plan: the planner is issuing more debt in marginal utility units than he would in the time-additive economy against g_L and less debt that in the time-additive economy against g_H . So the planner is actively taking larger positions in absolute value.²³

To see the price manipulation that takes place with recursive utility, figure 2 contrasts the optimal stochastic discount factor $S(g' = g_i, z, g), i = L, H$, and its components (top and bottom left panels), to the induced stochastic discount factor that pertains to a sub-optimal constant Φ policy (top and bottom right panels). It shows how the planner, by issuing more debt against g_L and increasing the respective tax rate, manages to increase the pricing kernel and therefore the

²²Note that if the initial shock was high, $g_0 = g_H$, we would have $b_H = 0$ and $b_L > 0$. The planner insures against adverse shocks by running a deficit when government expenditures are high, that are financed by debt contingent on a low expenditure shock. When shocks are low, the planner runs a surplus to pay back the issued debt.

²³A virtually identical graph would emerge if we compared the optimal policy functions z'_i with the positions that would be induced in a recursive utility economy with a planner that follows a sub-optimal constant- Φ policy.

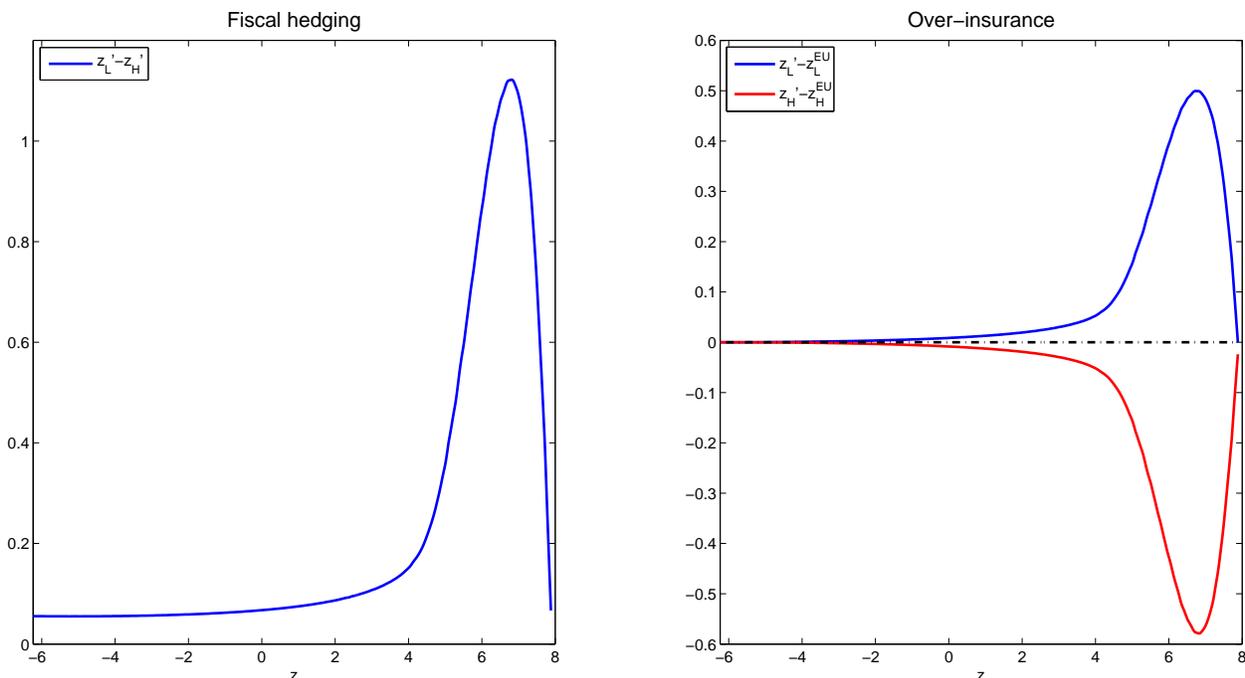


Figure 1: The left panel depicts the difference $z'_L - z'_H$. The difference starts decreasing at high values of z , because the probability of a binding upper bound increases. The right panel compares positions for recursive and time-additive utility. For both graphs the current shock is low, $g = g_L$. A similar picture emerges when $g = g_H$.

price of a claim to consumption, making therefore debt cheaper. Note that the increase in the continuation value part of the stochastic discount factor is naturally reinforced by an increase in the period marginal utility part due to decreased future consumption. Similarly, by issuing less debt or buying more assets against a high fiscal shock, and taxing consequently less, the planner is decreasing the pricing kernel for bad states of the world. The right panels show that following the standard prescription of a constant excess burden does not deliver these beneficial pricing effects.

6.5 Persistence and negative correlation with spending

Consider a simulation of 10,000 sample paths that are 2,000 periods long.²⁴ Table 1 highlights the persistence that propositions 3 and 5 hinted at. The median persistence of the tax rate is very high (0.998), despite the fact that government expenditure shocks are i.i.d., which contrasts to the standard history-independence result of Lucas and Stokey (1983). As expected, the *change* in the tax rates is strongly negatively correlated with government expenditures (-0.99) and therefore with output.²⁵

²⁴In the Online Appendix I provide an example with a sequence of low and high alternating shocks and comparative statics with respect to risk aversion and shock volatility.

²⁵The theory predicts that changes in tax rates are affected by the relative debt position, which is highly negatively correlated with fiscal shocks. In contrast, the level of the tax rate is affected by the cumulative relative debt position

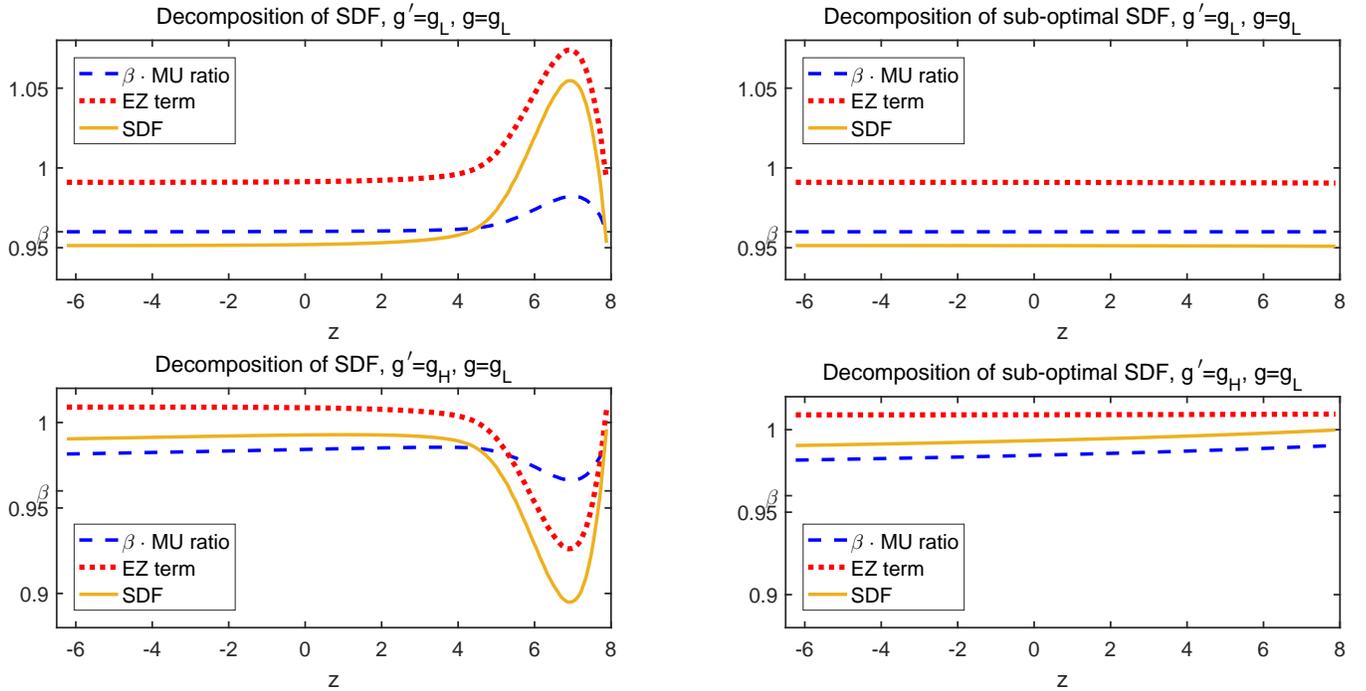


Figure 2: The left panels decompose the optimal stochastic discount factor, when the current shock is low, $g = g_L$. The right panels perform the same exercise assuming that a sub-optimal constant Φ policy is followed. A similar picture emerges when $g = g_H$.

Table 1: Statistics of tax rate sample paths.

	<i>Recursive utility</i>	
	short samples	long samples
Autocorrelation	0.9791	0.9980
Correlation of $\Delta\tau$ with g	-0.9999	-0.9984
Correlation of $\Delta\tau$ with output	-0.9977	-0.9762
Correlation of τ with g	-0.1098	-0.0346
Correlation of τ with output	-0.1793	-0.2418

The table reports median sample statistics across 10000 sample paths of the tax rate. For the time-additive case the respective moments are not well defined since the tax rate is constant. For the recursive utility case the median sample statistics are calculated for short samples (the first 200 periods) and long samples (2000 periods).

6.6 Back-loading and volatility of distortions

Figure 3 plots the mean, standard deviation, the 5th and the 95th percentile of the tax rate and the debt-to-output ratio. It shows that there is a *positive* drift in the tax rate with respect to the $\sum_i \eta_i$, leading overall to a small correlation with government expenditures.

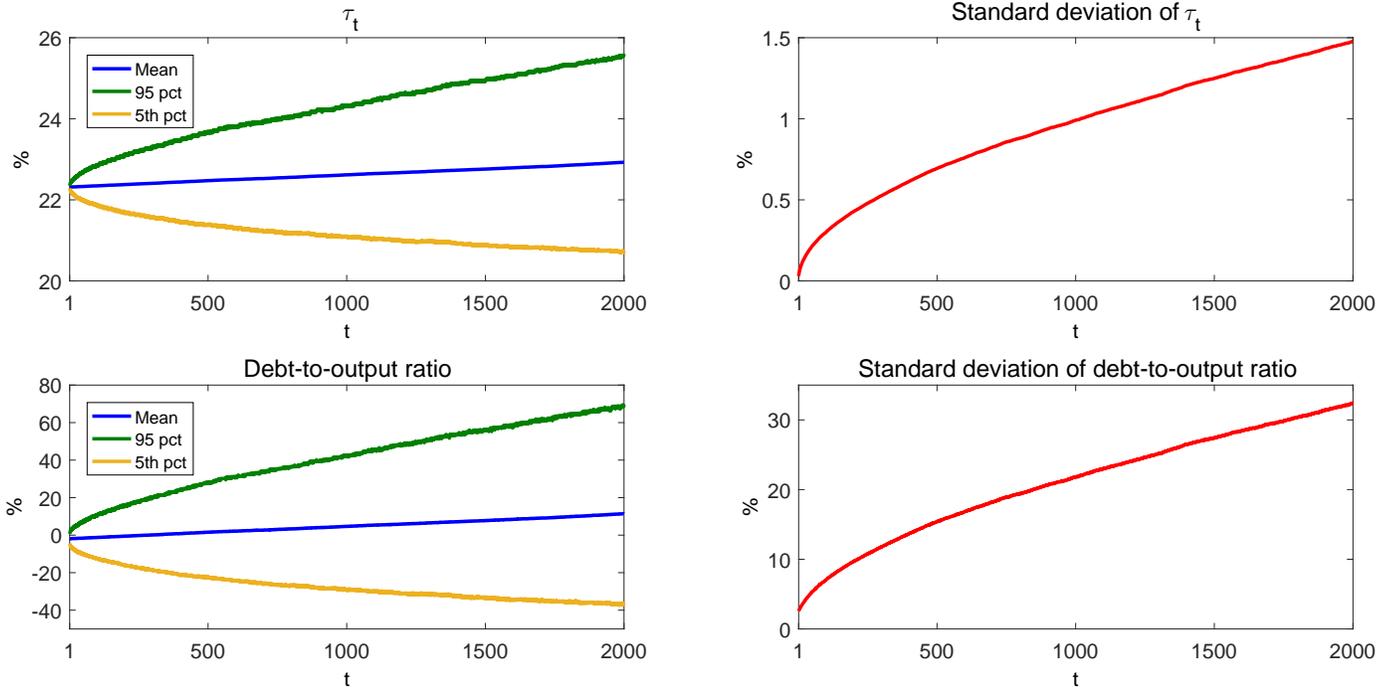


Figure 3: Ensemble moments of the tax rate and the debt-to-output ratio.

data-generating process, which is mirrored also in the debt-to-output ratio. This back-loading of distortions reflects the submartingale results of propositions 3 and 5. The increase in the mean tax rate is slow (about 60 basis points in 2000 periods) but the standard deviation rises to almost 1.5 percentage points. So the distribution of the tax rate is “fanning-out” over time. Similarly, the mean and the standard deviation of the debt-to-output rise to 11 and 32 percentage points respectively at $t = 2000$.

6.7 Stationary distribution

Consider now the long-run. Postponing on average taxes to the future leads progressively to high levels of debt and to the possibility that fiscal hedging becomes limited, due to an upper bound on debt issuance. In that case, the positive drift of the tax rate breaks down and its distribution becomes stationary.^{26,27}

Table 2 reports moments of interest. The tax rate has mean 30.8% and standard deviation

²⁶What breaks down is the martingale result of proposition 3. The optimality condition with respect to z when there is an upper bound on z becomes $\Phi_{t+1}(1 + (1 - \beta)(1 - \gamma)\eta_{t+1}\Phi_t) \leq \Phi_t$. If $1 + (1 - \beta)(1 - \gamma)\eta_{t+1}\Phi_t > 0$, we get $\frac{1}{\Phi_{t+1}} \geq \frac{1}{\Phi_t} + (1 - \beta)(1 - \gamma)\eta_{t+1}$, which implies that $1/\Phi_t$ is a submartingale (and *not* a martingale) with respect to $\pi_t \cdot M_t$. Therefore, the convexity of function f in the proof of proposition 3 is not sufficient anymore to infer $E_t m_{t+1} \Phi_{t+1} \geq \Phi_t$ (we need also f to be monotonically increasing and it is actually decreasing). The same reasoning applies to the tax rate in proposition 5.

²⁷The upper bounds of the state space are rarely visited. The 98th percentile of the debt-to-output ratio is about 400% whereas the upper bound is close to 600%. Several robustness exercises with respect to the size of the state space are provided in the Online Appendix.

Table 2: Moments from the stationary distribution.

Stationary distribution					
	τ in %		b/y in %		
Mean	30.86		181.97		
St. dev.	4.94		104.28		
98th pct	40.6		397.3		
St. dev. of change	0.17		12.72		
Autocorrelation	0.9994		0.9926		
Correlations (τ, b, g)					
Corr($\Delta\tau, g$)	-0.6183	Corr($\Delta b, g$)	-0.7639	Corr($\Delta\tau, b$)	0.0476
Corr($\Delta\tau, \Delta g$)	-0.4383	Corr($\Delta b, \Delta g$)	-0.9070	Corr($\Delta\tau, \Delta b$)	0.7228
Corr(τ, g)	-0.0219	Corr(b, g)	-0.0653	Corr(τ, b)	0.9933

The simulation is 60 million periods long. The first 2 million periods were dropped. Remember that in the expected utility case the tax rate is 22.3% and that the debt-to-output ratio has mean -1.91% and a standard deviation of 1.91%.

close to 5 percentage points. This tax rate is pretty high: it supports debt-to-output ratios that have mean 182% with a standard deviation of 105 percentage points. The conditional volatility of the tax rate and the debt-to-output ratio are small but the unconditional volatility is large due to the extremely high persistence in the long-run.

Such volatility figures are unprecedented in the time-additive expected utility literature that prescribes tax-smoothing. For our period utility function volatility is zero in the time-additive case, but even for a utility function with a varying Frisch elasticity labor tax volatility is minimal: [Chari et al. \(1994\)](#) and [Chari and Kehoe \(1999\)](#) report volatility of 5 to 15 basis points for standard business cycle calibrations.²⁸

7 Optimal debt returns and fiscal insurance

In this section I am taking a deeper look at the theory of debt management with recursive utility. I focus on the use of the *return* of the government debt portfolio as a tool of fiscal insurance.²⁹ To that end, I measure optimal fiscal insurance in simulated data by using the decomposition of

²⁸The same conclusion is drawn if we use the persistent shock specification of [Chari et al. \(1994\)](#): labor tax volatility at the stationary distribution amounts to 5.5 percentage points. What is different with persistent shocks is the increased speed at which the stationary distribution is reached. The mean and the standard deviation of the tax rate increase by 4 and 5 percentage points respectively in 2,000 periods (in contrast to the lower medium-run numbers displayed in figure 3). See the Online Appendix for further details.

²⁹See [Hall and Sargent \(2011\)](#) for the careful measurement of the return of the government debt portfolio and [Hall and Krieger \(2000\)](#) for an analysis of optimal debt returns in the [Lucas and Stokey \(1983\)](#) setup. [Marcet and Scott \(2009\)](#) contrast fiscal insurance in complete and incomplete markets.

Table 3: Returns on government debt portfolio, $R(g', g, z)$.

$\mathbf{R} - \mathbf{1}$ in %	Expected utility at z_{EU}		Recursive utility at z_{EU}		Recursive utility at $E(z)$	
	g_L	g_H	g_L	g_H	g_L	g_H
g_L	5.96	-27.92	6.95	-41.15	5.30	-15.61
g_H	49.47	1.68	69.52	0.81	25.96	3.10

Rows denote current shock. The value of z at the expected utility case is $(z_L, z_H) = (0.7795, 0.5399)$. The average value of z with recursive utility is $E(z) = 2.2626$.

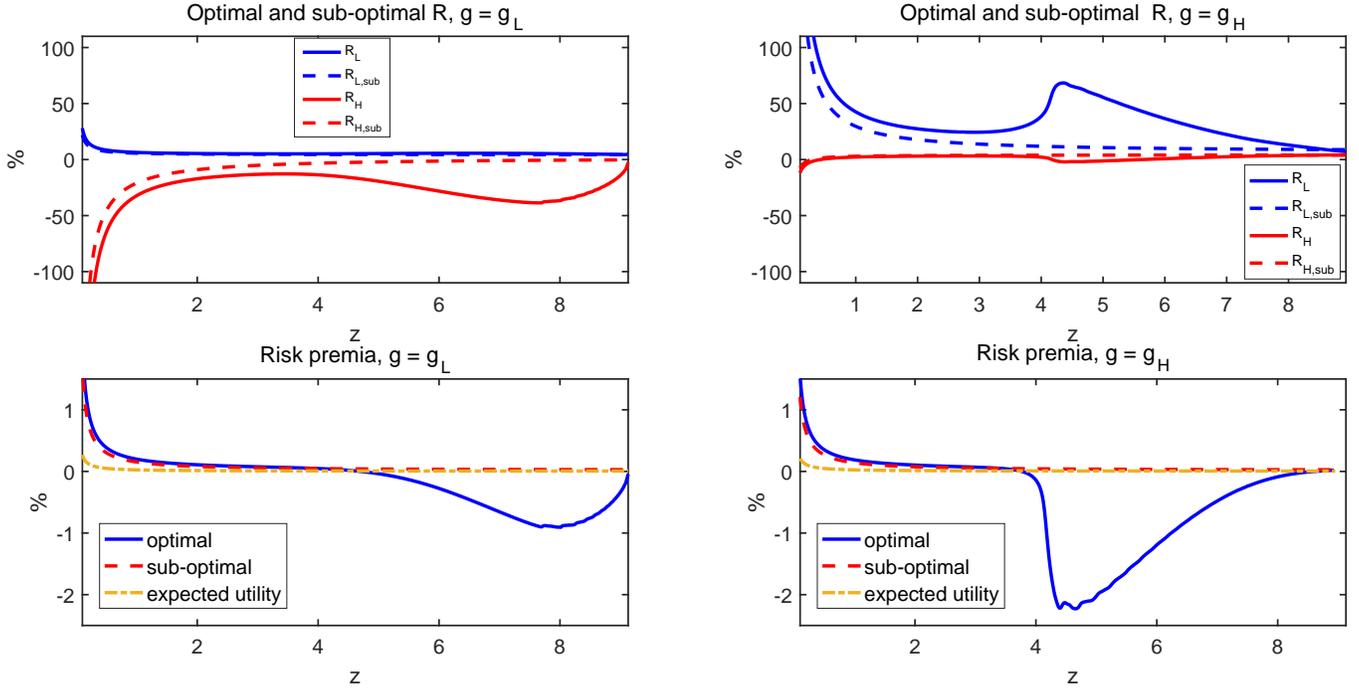


Figure 4: The top panels contrast the optimal $R(g', g, z)$ to the sub-optimal return coming from a constant Φ policy. The bottom panels plot the respective conditional risk premia, where I also include the expected utility risk-premia for comparison.

Berndt et al. (2012) (BLY henceforth) and contrast it to their empirical findings about actual fiscal insurance.

BLY devised a method to quantify fiscal insurance in post-war US data by log-linearizing the intertemporal budget constraint of the government.³⁰ Let the government budget constraint be written as

$$b_{t+1} = R_{t+1} \cdot (b_t + g_t - T_t), \quad (28)$$

³⁰Their exercise follows the spirit of Campbell (1993) – who worked with the household’s budget constraint– and Gourinchas and Rey (2007) –who employed the country’s external constraint.

where $R_{t+1} \equiv b_{t+1}(g^{t+1}) / \sum_{g^{t+1}} p_t(g_{t+1}, g^t) b_{t+1}(g^{t+1})$, the return on the government debt *portfolio*, constructed in the model economy by the state-contingent positions b_{t+1} , and $T_t \equiv \tau_t h_t$, the tax revenues. By construction, we have $\sum_{g^{t+1}} p_t(g_{t+1}, g^t) R_{t+1}(g^{t+1}) = 1$. BLY log-linearize (28) and derive a representation in terms of *news* or *surprises* in the present value of government expenditures, returns and tax revenues,³¹

$$I_{t+1}^g = -\frac{1}{\mu_g} I_{t+1}^R + \frac{1}{\mu_g} I_{t+1}^T, \quad (29)$$

where

$$\begin{aligned} I_{t+1}^g &\equiv (E_{t+1} - E_t) \sum_{i=0}^{\infty} \rho_{\text{BLY}}^i \Delta \ln g_{t+i+1} \\ I_{t+1}^R &\equiv (E_{t+1} - E_t) \sum_{i=0}^{\infty} \rho_{\text{BLY}}^i \ln R_{t+i+1} \\ I_{t+1}^T &\equiv (E_{t+1} - E_t) \sum_{i=0}^{\infty} \rho_{\text{BLY}}^i \mu_T \Delta \ln T_{t+i+1}, \end{aligned} \quad (30)$$

and $(\mu_g, \mu_T, \rho_{\text{BLY}})$ approximation constants. Decomposition (29) captures how a fiscal shock is absorbed: a positive surprise in the growth rate of spending, I_{t+1}^g , is financed by either a negative surprise in (current or future) returns, I_{t+1}^R , or by a positive surprise in (current or future) growth rates of tax revenues, I_{t+1}^T . BLY refer to these types of fiscal adjustment as the *debt valuation* channel and the *surplus* channel respectively. The decomposition can be written in terms of fiscal adjustment *betas*,

$$1 = -\frac{\beta_R}{\mu_g} + \frac{\beta_T}{\mu_g}, \quad \text{where} \quad \beta_R \equiv \frac{\text{Cov}(I_{t+1}^g, I_{t+1}^R)}{\text{Var}(I_{t+1}^g)}, \quad \beta_T \equiv \frac{\text{Cov}(I_{t+1}^g, I_{t+1}^T)}{\text{Var}(I_{t+1}^g)}. \quad (31)$$

The fraction of fiscal shocks absorbed by debt returns and tax revenues are $-\beta_R/\mu_g$ and β_T/μ_g respectively. *Fiscal insurance* refers to the reduction of returns in light of a positive fiscal shock, $\beta_R < 0$.

7.1 Returns and risk premia

For the fiscal insurance exercise I use the Chari et al. (1994) specification of fiscal shocks that captures well the dynamics of government consumption in post-war U.S. data. I set initial debt to

³¹See Berndt et al. (2012) for the derivations and the Online Appendix for the definition of the approximation constants.

50% of first-best output. The utility function and the calibration of the rest of the parameters is the same as in the previous section.

Table 3 provides the conditional returns of the government debt portfolio implied by the Ramsey plan in the expected and recursive utility economy. It shows the essence of debt return management, i.e. the reduction of the return on government debt in bad times in exchange of an increase in return in good times. For example, in the expected utility economy bond-holders suffer capital losses of -28% when there is a switch from a low to a high fiscal shock. They still buy government debt because they are compensated with a high return of 49% when there is a switch back to a low shock. The gains and losses to the bond-holders at the same value of the state variable z with recursive utility are much larger (-41% and 69% respectively), due to the “over-insurance” property. On average though, the government issues larger quantities of debt with recursive utility, which actually makes the size of conditional returns necessary to absorb fiscal shocks *smaller*. This can be seen in the third part of table 3, which displays the conditional returns for EZW utility at the average debt holdings, $E(z)$.

Figure 4 takes a closer look at the returns of the government portfolio. The top panels demonstrate the desire of the government to increase the returns of the debt portfolio for good shocks and decrease it for bad shocks, by contrasting the optimal returns with recursive utility with the sub-optimal returns that are induced by a constant Φ policy. The bottom panels plot the conditional premium of government debt above the risk-free rate for recursive utility (following either optimal or sub-optimal policy) and for expected utility. What is interesting to observe is the fact that for large levels of debt, when over-insurance becomes even more pronounced, the optimal conditional risk premium of government debt becomes *negative*.³²

The reason for government debt becoming a *hedge* is simple: the risk premium over the risk-free rate R_t^F can be expressed as $E_t R_{t+1}/R_t^F - 1 = -Cov_t(S_{t+1}, R_{t+1})$. Debt returns are high when fiscal shocks are low. But optimal policy with recursive utility prescribes large tax rates at exactly these states of the world. As a result, at some point tax rates at good shocks become so high that both consumption and continuation values of agents *fall* (despite shocks being favorable), and therefore S_{t+1} increases. This leads to a positive covariance of the stochastic discount factor with government returns and a negative risk premium. In other words, optimal policy converts “good” times (with low g) to “bad” times with high tax rates (and “bad” times with high g to “good” times with low tax rates). Thus, the household is happy to accept a negative risk premium for a security that pays well when tax rates are so high.³³

³²The unconditional risk premium remains positive. See the Online Appendix for additional information on average returns and the market price of risk.

³³We can see also the change in the ranking of the discount factors in figure 2. For large enough debt we have $S(g_L, z, g) > S(g_H, z, g)$, whereas for low enough debt the opposite holds. With either expected utility, or a sub-optimal constant Φ policy and recursive utility, this does not happen and we always have $S(g_L, z, g) < S(g_H, z, g)$ and a positive conditional risk premium.

Table 4: News to expenditures, returns and revenues.

	Expected utility			Recursive utility		
	I^g	I^R	I^T	I^g	I^R	I^T
I^g	0.91			0.98		
I^R	-1	8.60		-0.79	7.52	
I^T	1	-1	1.37	-0.74	0.53	2.44

Standard deviations (on the diagonal, multiplied by 100) and correlations of the news variables at the stationary distribution. Calibration of shocks as in [Chari et al. \(1994\)](#).

Table 5: Fiscal insurance.

	Expected utility		Recursive utility	
	<i>Valuation channel</i>	<i>Surplus channel</i>	<i>Valuation channel</i>	<i>Surplus channel</i>
Beta	-9.46	1.51	-6.05	-1.85
Current	-9.65	5.28	-6.20	-0.64
Future	0.19	-3.77	0.15	-1.21
Fraction in %	87.85	13.99	180.93	-55.16
Current	89.67	49.08	185.13	-19.13
Future	-1.82	-35.09	-4.20	-36.03

Fiscal adjustment betas and fiscal insurance fractions. The approximation constants are $(\mu_g, \mu_T, \rho_{BLY}) = (10.7654, 11.7654, 0.958)$ and $(\mu_g, \mu_T, \rho_{BLY}) = (3.3462, 4.3462, 0.9525)$ in the time-additive and recursive utility case respectively. The R^2 in the expected utility case is almost 100% for both regressions. For the recursive utility economy the R^2 is 62.44% and 55.07% for the return and revenues regression respectively. The current return beta comes from regressing $\ln R_{t+1} - E_t \ln R_{t+1}$ on news to spending. Similarly, the current tax revenue beta comes from regressing current news to the growth in tax revenues on news to spending.

7.2 Fiscal insurance

Table 4 reports the correlations and the standard deviations of news to government spending, debt returns and tax revenues at the stationary distribution and table 5 reports the respective fiscal adjustment betas and fiscal insurance fractions. For both the expected and the recursive utility case news to optimal returns are pretty volatile and negatively correlated to fiscal shocks. What is important to notice is that news to the growth rate in tax revenues (I^T) are positively correlated with news to fiscal shocks in the expected utility case (absorbing therefore part of the fiscal shock) but *negatively* correlated in the recursive utility case.

Turning to fiscal insurance fractions, about 87% of fiscal risk is absorbed by the debt valuation channel and about 13% by the surplus channel in the expected utility economy. Thus, the debt valuation channel is prominent in the absorption of shocks. Fiscal insurance motives are *amplified*

with recursive utility: the planner is reducing even more returns in the face of adverse shocks, to the point where the tax rate is actually *reduced*, explaining the negative correlation we saw in table 4. As a result, the reliance on the debt valuation channel is even larger and the surplus channel becomes essentially *inoperative*. The fraction of fiscal risk absorbed by reductions in the market value of debt is about 180% (predominantly by a reduction in *current* returns), which allows the government to reduce the growth in tax revenues, leading to a surplus channel of -55% .³⁴

Is actual fiscal insurance even worse than we thought? BLY measure fiscal insurance on post-war U.S. data. They focus on defense spending in order to capture the exogeneity of government expenditures and show that 9% of defense spending shocks has been absorbed by a reduction in returns (mainly through future returns) and 73% by an increase in non-defense surpluses. Thus, there is some amount of fiscal insurance in the data; smaller though than what optimal policy in an expected utility economy would recommend. The current exercise shows that in environments that can generate higher market prices of risk, governments debt returns have to be used to a much *greater* extent as a fiscal shock absorber. Thus, if we were to evaluate actual fiscal policy through the normative prescriptions of the recursive utility economy, the following conclusion emerges: actual fiscal policy is even *worse* than we thought.

8 Economy with capital

Consider now an economy with capital as in Zhu (1992) and Chari et al. (1994) and recursive preferences. To fix ideas, let s capture uncertainty about government expenditure or technology shocks, with the probability of a partial history denoted by $\pi_t(s^t)$. The resource constraint in an economy with capital reads

$$c_t(s^t) + k_{t+1}(s^t) - (1 - \delta)k_t(s^{t-1}) + g_t(s^t) = F(s_t, k_t(s^{t-1}), h_t(s^t)), \quad (32)$$

where δ denotes the depreciation rate, $k_{t+1}(s^t)$ capital measurable with respect to s^t and F a

³⁴The fractions do not add to 100% due to the approximation error coming from log-linearizing (28). The same thing happens with actual fractions from post-war U.S. data (see Berndt et al. (2012) and the respective table in the Online Appendix that reproduces their results). Furthermore, two robustness exercises are provided in the Online Appendix. At first, in order to apply the log-linear methodology of Berndt et al. (2012), I excluded negative debt realizations that amount to 4.4% of the stationary distribution. In the Online Appendix I use a linear approximation of (28) that allows me to include this type of observations. The size of the valuation and surplus channel for both expected and recursive utility remains essentially the same. Second, one may think that the stark contrast between the expected and recursive utility case is coming from the much larger debt and taxes in the latter case, a fact which is reflected in the very different approximation constants across the two economies. In order to control that, I calculate in the Online Appendix the expected utility fiscal insurance fractions by setting initial debt equal to the mean of the recursive utility economy. This leads to similar approximation constants with the recursive utility case, so any difference in the fiscal channels is stemming from the endogenous reaction of returns and tax revenues. The valuation and surplus channel in the expected utility economy become 83% and 17% respectively, so the difference between expected and recursive utility is even starker.

constant returns to scale production function. The representative household accumulates capital, that can be rented at rental rate $r_t(s^t)$, and pays capital income taxes with rate $\tau_t^K(s^t)$. The household's budget constraint reads

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

where $R_t^K(s^t) \equiv (1 - \tau_t^K(s^t)) r_t(s^t) + 1 - \delta$, the after-tax gross return on capital.

I provide the details of the competitive equilibrium and the analysis of the Ramsey problem in the Appendix and summarize here the main results. In short, the completeness of the markets allows the recasting of the household's budget constraint in terms of wealth, $W_t \equiv b_t + R_t^K k_t$, making therefore *wealth in marginal utility units*, $z_t \equiv U_{c_t} W_t$, the relevant state variable for the optimal taxation problem. With this interpretation of z_t , the dynamic implementability constraint remains the same as in an economy without capital. The recursive formulation of the Ramsey problem has (z, k, s) as state variables. The excess burden of taxation Φ captures now the shadow cost of an additional unit of wealth in marginal utility units, $\Phi = -v_z(z, k, s)$, where v denotes the value function. As expected, the excess burden of taxation is not constant anymore across states and dates. In particular, we have:

Proposition 6. *The law of motion of Φ_t in an economy with capital remains the same as in (22), with η_{t+1} defined as in (19), denoting now the relative wealth position in marginal utility units, with an average of zero, $E_t m_{t+1} \eta_{t+1} = 0$. Let \hat{s} and \tilde{s} denote states of the world at $t + 1$ for which $\eta_{t+1}(\hat{s}) > 0 > \eta_{t+1}(\tilde{s})$. Then $\Phi_{t+1}(\hat{s}) > \Phi_t > \Phi_{t+1}(\tilde{s})$, when $\rho < \gamma$. Propositions 3, 4 and 5 go through, so the same conclusions are drawn for the dynamics of the excess burden and the labor tax as in an economy without capital.*

Proposition 6 generalizes our previous results about the excess burden of taxation and the labor tax. Recall that in an economy without capital the planner was taxing more events against which he was issuing relatively more debt in order to take advantage of the positive covariance between debt in marginal utility units and the stochastic discount factor, through the channel of continuation values. Market completeness makes state-contingent wealth in marginal utility units the relevant hedging instrument in an economy with capital. Note also that we allowed technology shocks in the specification of uncertainty in this section, in addition to the typical government expenditure shocks. We expect that the planner hedges adverse shocks, which are high fiscal shocks and low technology shocks with low wealth positions, and favorable shocks, i.e. low fiscal shocks or high technology shocks with high wealth positions. In that case, the planner decreases the labor tax for high spending shocks and low technology shocks, mitigating again the effects of shocks. The opposite happens for favorable shocks.

8.1 Capital taxation

Capital accumulation affects through continuation values the pricing of state-contingent claims, a fact which alters the incentives for taxation at the intertemporal margin. In particular, the optimal accumulation of capital is governed by (details in the Appendix),

$$E_t S_{t+1}^* (1 - \delta + F_{K,t+1}) = 1, \quad \text{where } S_{t+1}^* \equiv \beta m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} \frac{\lambda_{t+1}/\Phi_{t+1}}{\lambda_t/\Phi_t}, \quad (33)$$

where λ_t stands for the multiplier on the resource constraint (32) in the recursive formulation of the second-best problem.

I will call S_{t+1}^* the *planner's* stochastic discount factor. The discount factor S_{t+1}^* captures how the planner discounts the *pre-tax* capital return $1 - \delta + F_{K,t+1}$ at the *second-best* allocation. S_{t+1}^* contrasts to the market stochastic discount factor, $S_{t+1} \equiv \beta m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} U_{c,t+1}/U_{c,t}$, which prices *after-tax* returns, $E_t S_{t+1} R_{t+1}^K = 1$. In a first-best world with lump-sum taxes available, we identically have $S_{t+1}^* \equiv S_{t+1}$. At the second-best, the difference in the two discount factors $S_{t+1} - S_{t+1}^*$ is useful in summarizing the optimal wedge at the intertemporal margin, in the form of the *ex-ante* tax rate on capital income.

In particular, as is well known from [Zhu \(1992\)](#) and [Chari et al. \(1994\)](#), only the non-state contingent ex-ante capital tax $\bar{\tau}_{t+1}^K(s^t)$ can be uniquely determined by the second-best allocation. This tax is defined as $\bar{\tau}_{t+1}^K \equiv (E_t S_{t+1} (1 - \delta + F_{K,t+1}) - 1) / E_t S_{t+1} F_{K,t+1}$, which by (33) becomes

$$\bar{\tau}_{t+1}^K = \frac{E_t [S_{t+1} - S_{t+1}^*] (1 - \delta + F_{K,t+1})}{E_t S_{t+1} F_{K,t+1}}. \quad (34)$$

Thus, the sign of the ex-ante capital tax is determined by the numerator in (34), i.e. the non-centered covariance of the two discount factors with the pre-tax capital return. The difference $S_{t+1} - S_{t+1}^*$ can be expressed in terms of differences in the inverse of the excess burden of taxation and differences in the own and cross elasticity of the marginal utility of consumption, which leads to the following proposition about capital taxation.³⁵

Proposition 7. (*Capital taxation criterion*) *The ex-ante tax rate on capital income $\bar{\tau}_{t+1}^K, t \geq 1$ is positive (negative) iff*

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

³⁵As it was the case with the labor tax in footnote 17, the capital tax criterion applies also for the deterministic and stochastic time-additive case for any standard U .

with weights $\zeta_{t+1} \equiv S_{t+1}(1 - \delta + F_{K,t+1})/E_t S_{t+1}(1 - \delta + F_{K,t+1})$. If $\epsilon_{cc} + \epsilon_{ch}$ is constant, then any capital taxation comes from variation in the excess burden Φ_t .

Proof. See Appendix. □

The ex-ante capital tax furnishes by construction the same present discounted value of tax revenues as any vector of feasible state-contingent capital taxes. As such, it *averages* intertemporal distortions across states next period, with weights ζ_{t+1} that depend on the stochastic discount factor and the pre-tax capital return. The distortions at each state next period depend on *both* the change in the elasticity of the marginal utility of consumption (the expected utility part) *and* the change in the excess burden of taxation (the novel recursive utility part).

Time-additive economy. Assume that we are either in a deterministic economy or in a stochastic but time-additive economy with $\rho = \gamma$. In both cases Φ_t is constant and the capital taxation criterion of proposition 7 depends only on the change in period elasticities. For the deterministic case, capital income is taxed (subsidized) if the sum of the own and cross elasticity is increasing (decreasing). A necessary and sufficient condition for a zero capital tax at *every* period from period two onward is a *constant* sum of elasticities, $\epsilon_{cc} + \epsilon_{ch}$, 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 only at the deterministic *steady state*, where the constancy of the consumption-labor allocation delivers constant elasticities. This delivers the steady-state results of Chamley (1986) and Judd (1985). In the stochastic case of Chari et al. (1994) and Zhu (1992), the sign of the ex-ante capital tax depends on the *weighted* average of the change in elasticities.³⁶

Recursive utility. The full version of the capital tax criterion in proposition 7 applies when $\rho \neq \gamma$. To focus on the novel effects of recursive utility, consider the case of constant period elasticities and assume that $\rho < \gamma$. For an example in this class, let the composite good be

$$u(c, 1 - h) = [c^{1-\rho} - (1 - \rho)v(h)]^{\frac{1}{1-\rho}}, \quad v', v'' > 0, \quad (35)$$

that delivers a period utility $U = (u^{1-\rho} - 1)/(1 - \rho)$, which is separable between consumption and leisure and isoelastic in consumption.³⁷ Chari et al. (1994) and Zhu (1992) have demonstrated that these preferences deliver a *zero* ex-ante capital tax from period two onward. This is easily interpreted in terms of proposition 7, since $\epsilon_{cc} = \rho$ and $\epsilon_{ch} = 0$.

With recursive preferences though, even in the constant period elasticity case, there is a novel source of taxation coming from the willingness of the planner to take advantage of the pricing

³⁶Variation in $\epsilon_{cc} + \epsilon_{ch}$ is a necessary condition for a non-zero ex-ante capital tax, but is not sufficient anymore since the weighted average can still in principle deliver a zero tax.

³⁷The same comments as in footnote 19 apply. The constant Frisch elasticity case is obviously a member of this class.

effects of continuation values. By using the law of motion of the excess burden of taxation (22) to substitute η_{t+1} for the change in $1/\Phi_t$, the criterion becomes

$$\bar{\tau}_{t+1}^K > (<) 0 \quad \text{iff} \quad E_t \zeta_{t+1} \eta_{t+1} > (<) 0, \quad \text{when } \rho < \gamma. \quad (36)$$

Thus, the capital taxation criterion depends on the weighted average of the relative wealth positions η_{t+1} . To understand the logic behind the criterion, note that the *change* in the excess burden of taxation determines the sign of distortions at the intertemporal margin. States where there are positive relative wealth positions ($\eta_{t+1} > 0$), make the planner increase the excess burden of taxation, $\Phi_{t+1} > \Phi_t$. This raises the labor tax and leads to a planner's discount factor that is smaller than the market discount factor, $S_{t+1}^* < S_{t+1}$, which we can think of as introducing a *state-contingent* capital tax.³⁸ To understand the intuition, a positive state-contingent capital tax reduces capital accumulation and therefore utility. In a recursive utility world this increases the price of the respective Arrow claim and the value of state-contingent wealth. And this appreciation of the value of wealth is beneficial when wealth positions are relatively large ($\eta_{t+1} > 0$). In the opposite case of $\eta_{t+1} < 0$ the planner is decreasing the labor tax and has the incentive to put a state-contingent capital subsidy ($S_{t+1}^* > S_{t+1}$). The ex-ante capital tax depends on the weighing of the positive versus the negative intertemporal distortions.

8.2 Ex-ante subsidy

To gain more insight about the sign of the ex-ante capital tax, we need to understand the behavior of the weights ζ_{t+1} . Consider the separable preferences in (35) and let $\rho = 1 < \gamma$. Then, by using the property that $E_t m_{t+1} \eta_{t+1} = 0$ and the definition of ζ_{t+1} , the capital tax criterion simplifies to

$$\bar{\tau}_{t+1}^K > (<) 0 \quad \text{iff} \quad \text{Cov}_t^M(c_{t+1}^{-1} \cdot (1 - \delta + F_{K,t+1}), z_{t+1}) > (<) 0.$$

Thus, we can express the criterion in terms of the conditional *covariance* (with respect to the continuation-value adjusted measure M) of the marginal utility weighted pre-tax capital return with the wealth positions in marginal utility units, z_{t+1} .³⁹ Assume for example that the only shocks in the economy are fiscal shocks and that they take two values, $g_H > g_L$. We expect that the negative income effect of a fiscal shock reduces consumption and makes the household work more, leading to a smaller capital-labor ratio. As a result, we expect marginal utility and the marginal

³⁸The difference in the two discount factors for the separable preferences (35) can be written as $\frac{S_{t+1} - S_{t+1}^*}{S_{t+1}} = \frac{(1-\beta)(\gamma-\rho)\eta_{t+1}}{1/\Phi_{t+1}-\rho}$. See the Appendix for details.

³⁹A more complicated covariance criterion emerges when $\gamma > \rho \neq 1$: $\bar{\tau}_{t+1}^K > (<) 0$ iff $\text{Cov}_t^M(V_{t+1}^{\rho-1} \cdot U_{c,t+1} \cdot (1 - \delta + F_{K,t+1}), V_{t+1}^{\rho-1} z_{t+1}) > (<) 0$.

product of capital to increase at adverse fiscal shocks. Thus, when the government hedges fiscal shocks by taking small positions against high shocks, $z_H < z_L$, the covariance will be *negative*, leading to an ex-ante capital *subsidy*.⁴⁰ Intuitively, the planner mitigates the effects of fiscal shocks by using a state-contingent capital subsidy at g_H and a state-contingent capital tax at g_L . But since adverse fiscal shocks are weighed more, we have an ex-ante subsidy to capital income. The Online Appendix provides a detailed example in an economy with a simplified stochastic structure (deterministic except for one period) that confirms this analysis.

9 Discussion: the case of $\rho > \gamma$

Consider now the case of $\rho > \gamma$. The direction of inequalities in propositions 2, 5 and 6 is obviously reversed.

Proposition 8. (*Desire to smooth over time stronger than desire to smooth over states*) Assume that $\rho > \gamma$, so that the household loves volatility in future utility. Then, $\Phi_{t+1}(\hat{g}) < \Phi_t < \Phi_{t+1}(\tilde{g})$ when \hat{g}, \tilde{g} are such so that $\eta_{t+1}(\hat{g}) > 0 > \eta_{t+1}(\tilde{g})$. Similarly, in proposition 5 we have $\tau_{t+1}(\hat{g}) < \tau_t < \tau_{t+1}(\tilde{g})$ when $\eta_{t+1}(\hat{g}) > 0 > \eta_{t+1}(\tilde{g})$. The same reversion of the direction of inequalities for Φ_t holds also in an economy with capital, as in proposition 6. Proposition 7 goes through, but the direction of inequalities is reversed in (36): $\bar{\tau}_{t+1}^K > (<) 0$ iff $E_t \zeta_{t+1} \eta_{t+1} < (>) 0$.

Proposition 8 shows that the planner varies the excess burden over states and dates in an opposite way when $\rho > \gamma$. The reason is simple. Increases in continuation utility *increase* the stochastic discount factor when the household loves volatility in future utility (instead of decreasing it). Issuance of additional state-contingent debt reduces the stochastic discount factor, making debt more expensive. Thus, the planner finds it optimal to “under-insure” in comparison to expected utility, selling less claims against good times and buying less claims against bad times. This is accompanied with smaller taxes in good times and higher taxes in bad times, amplifying the effects of fiscal shocks. Following the discussion in the previous section, there is an ex-ante capital *tax* instead of a subsidy, since bad times (which are weighed more) carry now a higher excess burden.

The martingale and submartingale results of propositions 3, 5 and 6 hold also for $\rho > \gamma$, so the persistence and back-loading results with respect to $\pi_t \cdot M_t$ go through. The back-loading with respect to the physical measure goes through as well: the excess burden of taxation is now positively correlated with government spending. But the agent loves volatility in utility, and therefore places more probability mass on high-utility, low-spending shocks. Thus, we have again $\text{Cov}_t(m_{t+1}, \Phi_{t+1}) \leq 0$ and a positive drift with respect to the data-generating process.⁴¹

⁴⁰The covariance is $\text{Cov}^M = E_t m_{t+1} c_{t+1}^{-1} (1 - \delta + F_{K,t+1}) \eta_{t+1}$. Let subscripts denote if we are at the high or low shock and suppress time subscripts. By assumption we have $c_H < c_L$, $F_{K,H} > F_{K,L}$, $\eta_H < 0$ and $\eta_L > 0$. Therefore, $c_H^{-1}(1 - \delta + F_{K,H}) > c_L^{-1}(1 - \delta + F_{K,L})$. The covariance is $\text{Cov}^M = c_L^{-1}(1 - \delta + F_{K,L}) \pi_L m_L \eta_L + c_H^{-1}(1 - \delta + F_{K,H}) \pi_H m_H \eta_H$. But $c_L^{-1}(1 - \delta + F_{K,L}) \eta_L < c_H^{-1}(1 - \delta + F_{K,H}) \eta_L$, since $\eta_L > 0$. Therefore, $\text{Cov}^M < c_H^{-1}(1 - \delta + F_{K,H}) [\pi_L m_L \eta_L + \pi_H m_H \eta_H] = 0$, since $E m \eta = 0$.

⁴¹In the Online Appendix I provide a full-blown quantitative exercise by setting $\rho = 1$ and $\gamma = 0$ for the constant

10 Concluding remarks

Dynamic fiscal policy revolves around the proper use of government debt returns in order to minimize the welfare loss of distortionary taxation. The asset-pricing literature has taught us that it is crucial to differentiate between smoothing over time and smoothing over states if our models aspire to generate realistic asset prices. Naturally, models that match better asset prices raise obvious questions about the structure of the government debt portfolio and the resulting tax policy.

The analysis in this paper shows that if we respect these two distinct attitudes and analyze optimal policy with recursive preferences, the tax-smoothing prescriptions of the dynamic Ramsey literature are shattered. Optimal labor taxes become substantially volatile and persistent. The planner mitigates the effects of fiscal shocks by taxing more in good times and less in bad times. Debt returns should be used to an even greater degree as a fiscal shock absorber, indicating that actual fiscal policy is even worse than we thought. Lastly, there is a novel incentive for an introduction of an intertemporal wedge, that can lead to ex-ante capital subsidies.

I have focused on time and risk in otherwise standard economies of the dynamic Ramsey tradition. An analysis beyond the representative agent framework as in [Werning \(2007\)](#), [Bassetto \(2014\)](#) or [Bhandari et al. \(2015\)](#), or an exploration of different timing protocols like lack of commitment, are worthy directions for future research.

Frisch utility function (27) with the same i.i.d. specification of shocks as in the baseline exercise. These are the RINCE preferences of [Farmer \(1990\)](#). The correlation of tax rates with government spending is highly positive and the autocorrelation of the tax rate close to unity, whereas the positive drift is small and discernible only in the long-run for this parametrization.

A Economy without capital

A.1 State space

At first, define

$$\begin{aligned}
A(g_1) &\equiv \left\{ (z_1, V_1) \mid \exists \{c_t, h_t\}_{t \geq 1}, \{z_{t+1}, V_{t+1}\}_{t \geq 1}, \text{ with } c_t \geq 0, h_t \in [0, 1] \right. \\
&\quad \text{such that:} \\
&\quad z_t = \Omega(c_t, h_t) + \beta E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} z_{t+1}, t \geq 1 \\
&\quad V_t = \left[(1-\beta)u(c_t, 1-h_t)^{1-\rho} + \beta \mu_t (V_{t+1})^{1-\rho} \right]^{\frac{1}{1-\rho}}, t \geq 1 \\
&\quad c_t + g_t = h_t, t \geq 1 \\
&\quad \text{where } m_{t+1} \text{ defined as in (5)} \\
&\quad \left. \text{and the transversality condition holds, } \lim_{t \rightarrow \infty} E_1 \beta^t \left(\frac{M_{t+1}}{M_1} \right)^{\frac{\rho-\gamma}{1-\gamma}} z_{t+1} = 0. \right\}
\end{aligned}$$

The set $A(g_1)$ stands for the set of values of z and V at $t = 1$ that can be generated by an implementable allocation when the shock is g_1 . From $A(g)$ we get the state space as $Z(g) \equiv \{z \mid \exists (z, V) \in A(g)\}$.

A.2 Transformed Bellman equation

Let $v(z, g) \equiv \frac{V(z, g)^{1-\rho} - 1}{(1-\beta)(1-\rho)}$. The Bellman equation takes the form

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

subject to the transformed implementability constraint

$$z = U_c c - U_h h + \beta \sum_{g'} \pi(g'|g) \frac{[1 + (1-\beta)(1-\rho)v(z'_g, g')]^{\frac{\rho-\gamma}{1-\rho}}}{[\sum_{g'} \pi(g'|g) [1 + (1-\beta)(1-\rho)v(z'_g, g')]^{\frac{1-\gamma}{1-\rho}}]^{\frac{\rho-\gamma}{1-\gamma}}} z'_g$$

and to (12)-(14). Recall that $m'_{g'} \equiv \frac{V(z'_g, g')^{1-\gamma}}{\sum_{g'} \pi(g'|g) V(z'_g, g')^{1-\gamma}} = \frac{[1 + (1-\beta)(1-\rho)v(z'_g, g')]^{\frac{1-\gamma}{1-\rho}}}{\sum_{g'} \pi(g'|g) [1 + (1-\beta)(1-\rho)v(z'_g, g')]^{\frac{1-\gamma}{1-\rho}}}$.

B Economy with capital

B.1 Competitive equilibrium

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 $r_t = F_K(s^t)$.

The first-order condition with respect to an Arrow security is the same as in (8). The labor supply condition is $U_l/U_c = (1 - \tau)w$. The Euler equation for capital is

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

Conditions (8) and (B.1) deliver the no-arbitrage condition $\sum_{s_{t+1}} p_t(s_{t+1}, s^t) R_{t+1}^K(s^{t+1}) = 1$. The transversality conditions are

$$\lim_{t \rightarrow \infty} E_0 \beta^t M_t^{\frac{\rho-\gamma}{1-\gamma}} U_{ct} k_{t+1} = 0 \quad \text{and} \quad \lim_{t \rightarrow \infty} E_0 \beta^{t+1} M_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} U_{c,t+1} b_{t+1} = 0$$

B.2 Ramsey problem

Define wealth as $W_t(s^t) \equiv b_t(s^t) + R_t^K(s^t)k_t(s^{t-1})$. 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. The household's budget constraint in terms of W_t becomes

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

Eliminate $\{\tau_t, p_t\}$ and multiply with U_{ct} to get $U_{ct} W_t = U_{ct} c_t - U_{lt} h_t + \beta E_t m_{t+1}^{\frac{\rho-\gamma}{1-\gamma}} U_{c,t+1} W_{t+1}$, which leads to the same implementability constraint for $z_t \equiv U_{ct} W_t$. At $t = 0$ we have $U_{c0} W_0 = U_{c0} c_0 - U_{l0} h_0 + \beta E_0 m_1^{\frac{\rho-\gamma}{1-\gamma}} z_1$, where $W_0 \equiv [(1 - \tau_0^K) F_K(s_0, k_0, h_0) + 1 - \delta] k_0 + b_0$, and $(k_0, b_0, \tau_0^K, s_0)$ given.

B.3 Transformed Bellman equation with capital

Let $v(z, k, s) \equiv \frac{V(z, k, s)^{1-\rho} - 1}{(1-\beta)(1-\rho)}$. The Bellman equation takes the form

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

subject to

$$z = \Omega(c, h) + \beta \sum_{s'} \pi(s'|s) \frac{[1 + (1 - \beta)(1 - \rho)v(z'_{s'}, k', s')]^{\frac{\rho - \gamma}{1 - \rho}}}{[\sum_{s'} \pi(s'|s)[1 + (1 - \beta)(1 - \rho)v(z'_{s'}, k', s')]^{\frac{1 - \gamma}{1 - \rho}}]^{\frac{\rho - \gamma}{1 - \gamma}}} z'_{s'} \quad (\text{B.2})$$

$$c + k' - (1 - \delta)k + g_s = F(s, k, h) \quad (\text{B.3})$$

$$c, k' \geq 0, h \in [0, 1] \quad (\text{B.4})$$

The values $(z'_{s'}, k')$ have to belong to the proper state space, i.e. it has to be possible that they can be generated by a competitive equilibrium with taxes that starts at (k, s) .

B.4 First-order necessary conditions

$$c : \quad U_c + \Phi \Omega_c = \lambda \quad (\text{B.5})$$

$$h : \quad -U_l + \Phi \Omega_h = -\lambda F_H \quad (\text{B.6})$$

$$k' : \quad \lambda = \beta \sum_{s'} \pi(s'|s) m'_{s'}{}^{\frac{\rho - \gamma}{1 - \gamma}} v_k(z'_{s'}, k', s') [1 + (1 - \beta)(\rho - \gamma)\eta'_{s'} \Phi] \quad (\text{B.7})$$

$$z'_{s'} : \quad v_z(z'_{s'}, k', s') + \Phi [1 + (1 - \beta)(\rho - \gamma)v_z(z'_{s'}, k', s')\eta'_{s'}] = 0. \quad (\text{B.8})$$

The relative *wealth* position $\eta'_{s'}$ is defined as in (19) (with a value function V that also depends on capital now), so we again have $\sum_{s'} \pi(s'|s) m'_{s'} \eta'_{s'} = 0$. The envelope conditions are

$$v_z(z, k, s) = -\Phi \quad (\text{B.9})$$

$$v_k(z, k, s) = \lambda(1 - \delta + F_K). \quad (\text{B.10})$$

The envelope condition (B.9) together with (B.8) delivers the same law of motion of Φ_t as in (22), leading to the same results as in proposition 3. Combine (B.5) and (B.6) and use the fact that $(1 - \tau)F_H = U_l/U_c$ to get the same labor tax results as in propositions 4 and 5. Turn into sequence notation, use the law of motion of Φ_t (22) to replace $1 + (1 - \beta)(\rho - \gamma)\eta_{t+1}\Phi_t$ in (B.7) with the ratio Φ_t/Φ_{t+1} and the envelope condition (B.10) to eliminate v_k to finally get (33).

B.5 Proof of proposition 7

The first-order condition with respect to consumption for $t \geq 1$ is $U_{ct} + \Phi_t \Omega_{ct} = \lambda_t$. Thus, $1/\Phi_t + \Omega_{ct}/U_{ct} = \lambda_t/(\Phi_t U_{ct}) > 0$. Write the planner's discount factor as $S_{t+1}^* = S_{t+1} \frac{\lambda_{t+1}/(\Phi_{t+1} U_{c,t+1})}{\lambda_t/(\Phi_t U_{ct})} = S_{t+1} \frac{1/\Phi_{t+1} + \Omega_{c,t+1}/U_{c,t+1}}{1/\Phi_t + \Omega_{ct}/U_{ct}}, t \geq 1$. Remember that $\Omega_c/U_c = 1 - \epsilon_{cc} - \epsilon_{ch}$. Thus,

$$S_{t+1} - S_{t+1}^* = \frac{\frac{1}{\Phi_t} - \frac{1}{\Phi_{t+1}} + \epsilon_{cc,t+1} + \epsilon_{ch,t+1} - \epsilon_{cc,t} - \epsilon_{ch,t}}{\frac{1}{\Phi_t} + 1 - \epsilon_{cc,t} - \epsilon_{ch,t}} \cdot S_{t+1}, t \geq 1. \quad (\text{B.11})$$

The denominator is positive. Use (B.11) in the numerator of (34), simplify and normalize ζ_{t+1} so that $E_t \zeta_{t+1} = 1$ to get the criterion for capital taxation.

References

- Ai, Hengjie and Ravi Bansal. 2016. Risk Preferences and The Macro Announcement Premium. Mimeo, Duke University.
- 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.
- Barro, Robert J. 1979. On the Determination of the Public Debt. *Journal of Political Economy* 87 (5):940–71.
- Bassetto, Marco. 2014. Optimal Fiscal Policy with Heterogenous Agents. *Quantitative Economics* 5 (3):675–704.
- Berndt, Antje, Hanno Lustig, and Sevin Yeltekin. 2012. How Does the U.S. Government Finance Fiscal Shocks? *American Economic Journal: Macroeconomics* 4:69–104.
- Bhandari, Anmol, David Evans, Mikhail Golosov, and Thomas J. Sargent. 2015. Taxes, debts, and redistributions with aggregate shocks. Mimeo, New York University.
- . 2016. Fiscal Policy and Debt Management with Incomplete Markets. *Quarterly Journal of Economics* Forthcoming.
- Campbell, John Y. 1993. Intertermoral asset pricing without consumption data. *American Economic Review* 83 (3):487–512.
- Chamley, Christophe. 1986. Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica* 54 (3):607–622.
- Chari, V.V. and Patrick J. Kehoe. 1999. Optimal fiscal and monetary policy. *Handbook of Macroeconomics* 1:1671–1745.
- 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.
- Croce, Mariano M., Howard Kung, Thien T. Nguyen, and Lukas Schmid. 2012a. Fiscal Policies and Asset Prices. *The Review of Financial Studies* 25 (9):2635–2672.
- Croce, Mariano M., Thien T. Nguyen, and Lukas Schmid. 2012b. The market price of fiscal uncertainty. *Journal of Monetary Economics* 59 (5):401–416.
- 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.
- Epstein, Larry G., Emmanuel Farhi, and Tomasz Strzalecki. 2014. How Much Would You Pay to Resolve Long-Run Risk? *American Economic Review* 104 (9):2680–97.

- Farhi, Emmanuel. 2010. Capital Taxation and Ownership when Markets are Incomplete. *Journal of Political Economy* 118 (5):908–948.
- Farhi, Emmanuel and Iván Werning. 2008. Optimal savings distortions with recursive preferences. *Journal of Monetary Economics* 55 (1):21–42.
- Farmer, Roger E.A. 1990. Rince Preferences. *The Quarterly Journal of Economics* 105 (1):43–60.
- Gomes, Francisco, Alexander Michaelides, and Valery Polkovnichenko. 2013. Fiscal Policy and Asset Prices with Incomplete Markets. *The Review of Financial Studies* 26 (2):531–566.
- Gottardi, Piero, Atsushi Kajii, and Tomoyuki Nakajima. 2015. Optimal Taxation and Debt with Uninsurable Risks to Human Capital Accumulation. *American Economic Review* 105 (11):3443–70.
- Gourinchas, Pierre-Olivier and Hélène Rey. 2007. International Financial Adjustment. *Journal of Political Economy* 115 (4):665–703.
- Gourio, François. 2012. Disaster Risk and Business Cycles. *American Economic Review* 102 (6):2734–66.
- Hall, George J. and Stephan Krieger. 2000. The Tax Smoothing Implications of the Federal Debt Paydown. *Brookings Papers on Economic Activity* 31:253–302.
- Hall, George J. and Thomas J. Sargent. 2011. Interest Rate Risk and Other Determinants of Post-WWII US Government Debt/GDP Dynamics. *American Economic Journal: Macroeconomics* 3 (3):192–214.
- 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.

- Marcet, Albert and Andrew Scott. 2009. Debt and deficit fluctuations and the structure of bond markets. *Journal of Economic Theory* 144 (2):473–501.
- Petrosky-Nadeau, Nicolas, Lu Zhang, and Lars-Alexander Kuehn. 2013. Endogenous disasters and asset prices. Mimeo, Carnegie Mellon University.
- Piazzesi, Monika and Martin Schneider. 2007. Equilibrium Yield Curves. *NBER Macroeconomics Annual* 389–442. Edited by Daron Acemoglu, Kenneth Rogoff, and Michael Woodford.
- Rudebusch, Glenn D. and Eric T. Swanson. 2012. The Bond Premium in a DSGE Model with Long-Run Real and Nominal Risks. *American Economic Journal: Macroeconomics* 4 (1):105–143.
- Shin, Yongseok. 2006. Ramsey meets Bewley: Optimal Government Financing with Incomplete Markets. Mimeo, Washington University in St. Louis.
- Sleet, Christopher. 2004. Optimal taxation with private government information. *Review of Economic Studies* 71 (4):1217–1239.
- 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.
- Werning, Iván. 2007. Optimal fiscal policy with redistribution. *Quarterly Journal of Economics* 122 (3):925–967.
- Zhu, Xiaodong. 1992. Optimal fiscal policy in a stochastic growth model. *Journal of Economic Theory* 58 (2):250–289.