Optimal Taxation of Entrepreneurial Capital Under Private Information

Stefania Albanesi*
Columbia University

October 2005
PRELIMINARY AND INCOMPLETE

Abstract

This paper studies optimal taxation of entrepreneurial capital and financial assets in economies with private information. Returns to entrepreneurial capital are risky and depend on entrepreneurs' effort, which is not observed. The presence of idiosyncratic risk in capital returns implies that the intertemporal wedge on capital that characterizes constrained-efficient allocations can be positive or negative. The properties of optimal marginal taxes on entrepreneurial capital depend on the sign of the intertemporal wedge. If the wedge is positive, the marginal capital tax should be decreasing in capital returns, while the opposite is true when the wedge is negative. Optimal taxes on other assets should be set according to their correlation with risky productive capital. The intertemporal wedge associated with an asset is greater than the one associated with entrepreneurial capital as long as their correlation is less than one. The optimal tax system tends to reduce the variance of capital returns after tax relative to before tax, while the opposite is true for other assets. If entrepreneurs are allowed to sell shares of their capital to outside investors, returns to externally owned capital are subject to double taxation- at the level of the entrepreneur and at the level of the outside investors.

*I wish to thank V.V. Chari, Pierre-Andre' Chiappori, Chris Edmonds and Victor Rios-Rull for helpful conversations. I am grateful to seminar participants at NYU and Columbia University and to conference participants at the SED Annual Meeting, the SAET meeting, and at the NBER Summer Institute for useful comments. Mailing Address: 1022 IAB, West 118th Street, New York, NY 10027, E-mail: sa2310@columbia.edu.

1. Introduction

This paper studies optimal taxation of entrepreneurial capital and financial assets in economies with private information. The optimal setting of capital income taxes is a central question in macroeconomics. Previous studies of optimal capital taxes have abstracted from idiosyncratic risk in capital returns and incentive problems in the operation of capital that are potentially prevalent for entrepreneurial capital. Yet, entrepreneurial capital accounts for a substantial fraction of economy-wide capital and household wealth in the US. Moscowitz and Vissing-Jorgensen (2002) identify entrepreneurial capital with private equity. Based on the Survey of Consumer Finances they document that the total value of private equity is similar in magnitude to public equity. Households who hold entrepreneurial capital invest on average more than 70% of their wealth in a single private company which they actively manage. Gentry and Hubbard (2000) estimate that about 9% of US households can be classified as entrepreneurial. Based on the Survey of Consumer Finances, entrepreneurial households own 38% of assets of the household sector and 39% of net worth. Macroeconomic studies of optimal capital income taxation have also devoted little attention to the taxation of financial assets different from capital and to whether capital should be taxed at the firm or at the investor level.

This paper considers economies where capital returns are risky and depend positively on agents' effort. This implies that capital is agent specific and generates idiosyncratic risk in capital returns. The risky productive capital is identified with entrepreneurial capital, given the active role played by owners in the determination of capital returns. Higher effort increases the probability of high capital returns thus increasing the expected returns from capital. The returns from entrepreneurial capital are observable but entrepreneurial effort is private information. Entrepreneurs are allowed to trade financial assets, with returns possibly correlated with their idiosyncratic uncertainty, and can also sell shares of their capital to outside investors.

We characterize optimal taxes, following the dynamic public finance approach, reviewed in Kocherlakota (2005a). Under this approach, which extends Mirrlees's (1971) analysis to a dynamic framework, optimal taxes implement the constrained-efficient allocation as competitive equilibrium for a particular decentralized trading arrangement. The only a priori restrictions imposed on the tax system stem from the informational constraints. Hence, it is essential to characterize the constrained-efficient allocation for this environment to derive the optimal tax system.

The constrained-efficient allocation solves the problem of a planner who allocates investment and consumption across time and states to maximize the agents' ex ante lifetime utility, subject to a resource constraint and an incentive compatibility constraint. This allocation displays a wedge between the marginal benefit on an additional unit of entrepreneurial capital and the private marginal cost, given by the marginal utility of current consumption. The intertemporal wedge, which stems from the effects of capital on the agents' incentive to exert effort, can be *positive or negative.* This result stands in contrast with findings for private information economies with idiosyncratic risk in labor earnings. As shown in Golosov, Kocherlakota and Tsyvinski (2003), building on Diamond and Mirrlees (1978) and Rogerson (1985), in such economies the intertemporal wedge is always positive. An additional unit of capital has an adverse effect on incentives, since it reduces the dependence of consumption on effort or labor supply, and tightens the incentive compatibility constraint on effort. Hence, an agent's optimal level of effort is always decreasing in their holdings of capital. Instead, with idiosyncratic risk in capital returns, the optimal level of effort may be increasing in capital. In this case, more capital relaxes the incentive compatibility constraint on effort, and intertemporal wedge is negative.

The dependence of the distribution of capital returns on effort drives the incentive effects of capital. Specifically, at high effort, the mean and the variance of capital returns are greater than under low effort. We can then rely on intuition from portfolio theory to understand the role of capital. For given variance of capital returns, capital's higher expected return at high effort generates a substitution effect, which tends to increase capital at high effort relative to low effort, and an opposing wealth effect. The substitution effect dominates when the coefficient of relative risk aversion is smaller than one. For given expected capital returns, the increase in variance induces agents to reduce their exposure, thus reducing capital holdings at high effort relative to low effort. On the other hand, a precautionary motive tends to increase optimal capital holdings at high effort relative to low effort. For preferences in the CRRA class, the first effect dominates if the coefficient of relative risk aversion is smaller than one, while the second effect dominates otherwise.

Intertemporal substitution governs the effect of expected capital returns, whereas the response to the variance of returns is driven by risk aversion, and they have opposing effects on how optimal capital holdings vary as a function of effort. To disentangle intertemporal substitution and risk aversion, we consider a recursive utility version of the model with Kreps-Portheus preferences. We find that if risk aversion is greater than the elasticity of intertemporal substitution, the intertemporal wedge is negative. This is consistent with intuition. When risk aversion is high, the precautionary motive is strong, thus making it more likely that agents will increase their holdings of capital at high effort in response to higher return variance. When intertemporal substitution is low, it is more likely that agents will increase capital holdings when the expected rate of return on capital is high, thus making effort increasing in capital.

To study optimal taxes, we first consider a decentralized arrangement in which entrepreneurs choose investment and effort and individual investment is observable. Entrepreneurs can also trade risk-free bonds and risky securities, possibly correlated with their idiosyncratic capital returns. The optimal tax system conditions tax payments on the observable history of capital returns. The properties of the optimal marginal tax on entrepreneurial capital depends on the sign of the intertemporal wedge. When the intertemporal wedge is positive, the marginal tax rate on capital is decreasing in capital returns, while the opposite is true when the intertemporal wedge is negative. The intertemporal wedge on the risk-free bond is always positive and higher than the intertemporal wedge on capital, and the marginal tax on bonds is decreasing in capital returns. Irrespective of the sign of the intertemporal wedge on productive capital, the marginal capital tax is always lower than the marginal tax on bonds in the bad state, while the opposite is true in the good state. The optimal marginal tax on risky securities depends on the *correlation* of their returns with idiosyncratic uncertainty. If the correlation between the returns from a risky security and capital is less than one, the intertemporal wedge on the security is greater than the intertemporal wedge on capital and the marginal tax on the security is higher than the marginal tax on capital in the bad state. The intertemporal wedge on any security decreases with the correlation with the security's return and idiosyncratic risk. Hence, capital is subsidized relative to risk-free bonds and risky securities in the bad state. The optimal marginal tax on capital tends to reduce the variance of after tax returns on capital relative to returns before tax, while the opposite is true for risk-free bonds and risky securities.

We also allow entrepreneurs to sell shares of their capital and buy equity in other entrepreneurs' capital. Each entrepreneur can be viewed as a firm, so that this arrangement introduces a market for private equity, and the amount of capital invested can be interpreted as firm size. Entrepreneurs can also purchase risk-free bonds. The optimal tax system for this market structure embeds a prescription for optimal *double taxation* of entrepreneurial capital- at the firm level, through the marginal tax on earnings, and at the investor level, through a marginal tax on stocks returns. Specifically, it is necessary that the tax on earnings be "passed on" to stock investors via a corresponding reduction in dividend distributions to avoid equilibria in which entrepreneurs sell all their capital to outside investors. In such equilibria, an entrepreneur exerts no effort and thus it is impossible to implement the constrained-efficient allocation. Moreover, marginal taxation of dividends received by outside investors is necessary to preserve incentives for the usual reasons. Hence, earnings from entrepreneurial capital are subject to double taxation.

This paper is related to Albanesi and Sleet (2005) and Kocherlakota (2005b) who also study the optimal taxation of capital income and labor income. They consider environments with idiosyncratic labor earnings risk, where skill or ability is private information. They study fiscal implementations with one asset and assume that asset returns are uncorrelated with idiosyncratic risk. Grochulski and Piskorski (2005) study optimal wealth taxes in economies with risky human capital, where human capital and idiosyncratic skills are private information. Farhi and Werning (2005) study optimal estate taxation in an overlapping generation economy with private information. They also find that the intertemporal wedge is positive if agents discount the future at a higher rate than the planner. Cagetti and De Nardi (2004) explore the effects of tax reforms in a quantitative model of entrepreneurship with endogenous borrowing constraints.

The plan of the paper is as follows. Section 1 present the economy and studies constrained-efficient allocations and the incentive effects of capital. Section 2 investigates optimal taxes. Section 3 concludes.

2. A Model

The baseline model builds on Sandmo (1970) and Levhari and Srinivasan (1969). The economy is populated by a continuum of unit measure of ex ante identical agents. Each agent lives for two periods. Agents consume in both periods and exert effort in the first period[1]. Their lifetime utility given by:

$$U = u(c_0) - v(e) + \beta u(c_1),$$

with $\beta \in (0,1)$, $u' > 0$, $u'' < 0$, $v' > 0$, $v'' > 0$, and $\lim_{c \to 0} u'(c) = \infty$.

Agents are endowed with K_0 units of the consumption good at time 0. The economy is endowed with an investment technology. The return on investment $R(K_1)$ depends on an agent's effort:

$$R(K_1) = K_1(1+x),$$

where x is the random net return on capital. We assume that $e \in \{0,1\}$. The stochastic process for x is:

$$x = \begin{cases} \overline{x} \text{ with probability } \pi(e), \\ \underline{x} \text{ with probability } 1-\pi(e), \end{cases} \tag{1}$$

with $\bar{x} > \underline{x}$ and $\pi(1) > \pi(0)$. The first assumption implies that $E_1(x) > E_0(x)$, where E_e denotes the expectation operator for probability distribution $\pi(e)$. The second assumption implies $V_1(x) \geq V_0(x)$, where V_e denotes that variance for probability distribution $\pi(e)$.

We assume effort is private information, while the realized value of x, as well as its distribution, and K_0 and K_1 are public information.

The planner maximizes the agent's lifetime utility by choice of K_1 and of consumption at time 1, conditional on the realized value of x, $c_1(x)$. Given that effort is private information, the planner faces an incentive compatibility constraint.

The planning problem is:

$$U^*(K_0) = \max_{e \in \{0,1\}, K_1 \in [0,K_0],\ c_0, c_1(x) \geq 0} u(c_0) - v(e) + \beta E_e u(c_1(x)) \qquad \text{(Problem 1)}$$

[1] The solution is the same if agent exerts effort at time 1.

subject to

$$c_0 + K_1 \leq K_0,\ \ E_e c_1(x) \leq K_1 E_e (1+x), \tag{2}$$

$$\beta E_1 u(c_1(x)) - \beta E_0 u(c_1(x)) \geq v(1) - v(0), \tag{3}$$

where E_e denotes the expectation operator with respect to the probability distribution $\pi(e)$. We refer to the allocation $\{e, K_1, c_0, c_1(\underline{x}), c_1(\bar{x})\}$ that solves this problem as constrained-efficient.

Proposition 1. *An allocation* $\{e^*, K_1^*, c_0^*, c_1^*(\underline{x}), c_1^*(\bar{x})\}$ *that solves Problem* 1 *with* $e^* = 1$ *satisfies:*

$$\frac{u'(c_1^*(\underline{x}))}{u'(c_1^*(\bar{x}))} = \frac{\left[1+\mu\frac{(\pi(1)-\pi(0))}{\pi(1)}\right]}{\left[1-\mu\frac{(\pi(1)-\pi(0))}{(1-\pi(1))}\right]} > 1, \tag{4}$$

$$u'(c_0^*)\, E_1\left[\frac{1}{u'(c_1^*(x))}\right] = \beta E_1(1+x), \tag{5}$$

where $\mu > 0$ *is the multiplier on the incentive compatibility constraint (3).*

Condition (4) implies that $c_1^*(\bar{x}) > c_1^*(\underline{x})-$ there is *partial insurance*. Using (5), we can derive the intertemporal profile of constrained-efficient allocations. Let the intertemporal wedge on risky productive capital be defined as:

$$\mathrm{IW}_K = \beta E_1 u'(c_1^*(x))(1+x) - u'(c_0^*).$$

Since the utility function is strictly concave, (5) implies $IW \neq 0$. To investigate the sign of the wedge, we rewrite:

$$\mathrm{IW}_K = \beta Cov_1\left(u'(c_1^*(x)), 1+x\right) + \mu(\pi(1)-\pi(0))\beta\left[u'(c_1^*(\underline{x})) - u'(c_1^*(\bar{x}))\right]E_1(1+x), \tag{6}$$

where Cov_1 denotes the covariance conditional on $e = 1$. Partial insurance, $\mu > 0$, and $\pi(1) - \pi(0) > 0$ imply that the second term on the right hand side of (6) is positive, while $Cov_1(u'(c_1^*(x)), 1+x) < 0$. Hence, the intertemporal wedge can be positive or negative.

In dynamic moral hazard models, higher wealth in a subsequent period influences the agent's attitude towards the risky distribution of outcomes in subsequent periods, as discussed in Rogerson (1985), which in turn affects incentives to exert effort. The intertemporal wedge measures the social cost of transferring wealth to a future period that stems from these incentive effects. In standard repeated moral hazard models, such as Rogerson (1985), higher wealth always has an adverse effect on incentives, because it reduces the dependence of current consumption on the current realization of uncertainty, and therefore on effort. Hence, the intertemporal wedge is positive and the agent is savings constrained along the optimal allocation. In the model described here, higher capital increases the agent's expected utility and provides insurance by increasing utility in the bad state, but it also increases the expected returns from effort. Hence, a higher level of capital does not necessarily have an adverse effect on incentives. The sign of the wedge depends on the balance between these two effects, which we now examine in more detail.

2.1. The Incentive Effects of Capital and the Intertemporal Wedge

To relate the sign of the intertemporal wedge to the incentive effects of capital, we consider the agents' lifetime utility maximization problem:

$$\left\{\hat{e}, \hat{K}_1\right\} = \arg \max_{K_1 \in [0, K_0],\ e \in \{0,1\}} U(e, K_1) - v(e),$$

where

$$U(e, K_1) \equiv u(K_0 - K_1) + \pi(e) u(K_1(1+\bar{x})) + (1 - \pi(e)) u(K_1(1+\underline{x})).$$

The agent's Euler equation is:

$$U_{K_1} = -u'\left(K_0 - \hat{K}_1\right) + E_{\hat{e}} u'\left(\hat{K}_1(1+x)\right)(1+x) = 0. \tag{7}$$

This equation clarifies that the optimal level of capital depends on effort. This complementarity between capital and effort determines the incentive effects of capital. To see this, differentiate U_{K_1} with respect to e, to obtain:

$$\frac{\Delta U_{K_1}}{\Delta e} \equiv (\pi(1) - \pi(0)) \left[u'(\hat{c}_1(\bar{x}))(1+\bar{x}) - u'(\hat{c}_1(\underline{x}))(1+\underline{x})\right]. \tag{8}$$

This term is the discrete analogue of the off-diagonal term of the Hessian matrix in the agent's lifetime decision problem. Totally differentiating (7) yields:

$$\frac{\Delta e}{\Delta K_1} = \frac{-U_{K_1 K_1}}{\frac{\Delta U_{K_1}}{\Delta e}}.$$

The numerator of this expression is positive by concavity of u, hence, the expressions $\frac{\Delta e}{\Delta K_1}$ and $\frac{\Delta U_{K_1}}{\Delta e}$ have the same sign.

Manipulating (6), we obtain:

$$\mathrm{IW}_K = \mu(\pi(1) - \pi(0)) \left[u'(c_1^*(\underline{x}))(1+\underline{x}) - u'(c_1^*(\overline{x}))(1+\overline{x})\right]. \tag{9}$$

Evaluating (8) at the constrained-efficient allocation, it follows that:

$$sign\left\{\frac{\Delta e}{\Delta K_1}(e^*, K_1^*)\right\} = sign\left\{\frac{\Delta U_{K_1}}{\Delta e}(e^*, K_1^*)\right\} = sign\left\{-\mathrm{IW}_K\right\}.$$

Hence, if there is a positive complementarity between capital and effort at the constrained-efficient allocation, that is $\frac{\Delta U_{K_1}}{\Delta e}(e^*, K_1^*) > 0$, the intertemporal wedge is negative. Conversely, if the complementarity between capital and effort is negative, $\frac{\Delta U_{K_1}}{\Delta e}(e^*, K_1^*) < 0$, the intertemporal wedge is positive. A positive/negative complementarity between capital and effort imply that if an agent will find it optimal to reduce/increase investment if she lowers her effort. This means that the intertemporal wedge is positive/negative when more capital tightens/relaxes the incentive compatibility constraint.

The analogue of (8) for a riskless asset, B_1, with gross return $(1 + r)$ is:

$$\frac{\Delta U_{B_1}}{\Delta e} \equiv (\pi(1) - \pi(0)) \left[u'(\hat{c}_1(\bar{x})) - u'(\hat{c}_1(\underline{x}))\right](1 + r).$$

This expression is negative as long as $\hat{c}_1(\bar{x}) > \hat{c}_1(\underline{x})$. Hence, an agent choosing $e = 0$ will always choose a higher value of B_1 relative to an agent choosing $e = 1$. It follows that higher levels of a risk free bond always tighten the incentive compatibility constraint, so that:

$$\mathrm{IW}_B = (\pi(1) - \pi(0)) \left[u'(c_1^*(\underline{x})) - u'(c_1^*(\bar{x}))\right](1 + r) > 0.$$

Assuming $r = E_1(x)$, we can write:

$$\begin{aligned} E_1 u'(c_1^*(x)) E_1(1 + x) - u'(c_0^*) &= \mu(\pi(1) - \pi(0))\{u'(c_1^*(\underline{x})) - u'(c_1^*(\overline{x}))\} E_1(1 + x) \qquad (10) \\ &> E_1 u'(c_1^*(x))(1 + x) - u'(c_0^*). \end{aligned}$$

Hence, the intertemporal wedge on risky productive capital is always smaller than the wedge on a riskless asset with the same expected rate of return.

2.1.1. The Complementarity between Capital and Effort

The fact that the distribution of capital returns depends on effort implies that capital under high effort can be thought of as a different asset than capital under low effort. Under high effort, capital has higher expected returns than capital under low effort, since $E_1(x) > E_0(x)$, and if $\pi(1) < 1 - \pi(0)$, capital returns under high effort have higher variance than under low effort, so that $V_1(x) > V_0(x)$. Then, we can use intuition from portfolio theory to interpret the incentive effects of capital. For given variance of capital returns, the effect of high effort on capital's expected return generates a negative substitution effect on time 0 consumption, which tends to increase capital at high effort, and a positive wealth effect on time 0 consumption that tends to reduce capital at high effort. The substitution effect dominates when the coefficient of relative risk aversion is smaller than one for standard preferences, as shown in Gollier (2001).

For given expected return, an increase in the variance of returns has two countervailing effects on the optimal level of capital at high effort, as discussed in Sandmo (1970). On one hand, the increase in the variance of capital returns makes an agent less inclined to expose herself to the possibility of loss and thus reduces capital holdings. On the other hand, greater riskiness makes it necessary to prevent low consumption in the bad state, which increases optimal capital holdings at high effort. So the substitution effect of increased variance in capital returns on time 0 consumption is positive, while the wealth effect is negative. Levhari and Srinivasan (1969) show that for preferences in the CRRA class, the substitution effect dominates if the coefficient of relative risk aversion is smaller than one, and capital holdings decrease in response to increased variance of returns, while the wealth effect dominates and capital holding increase if it is greater or equal to 1. Hence, the rate of return effect and the variance effect seem to have opposing effects on how the optimal capital level varies as a function of effort, at least for utility functions in the CRRA class.

If we restrict attention to utility functions in the CRRA class, with $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$, so that $\frac{u'(c)c}{u(c)} = 1 - \sigma$, from (8) we can derive:

$$\begin{aligned} \frac{\Delta U_{K_1}}{\Delta e}\hat{K}_1 &= (\pi(1) - \pi(0))\,\beta\left[u'(\hat{c}_1(\bar{x}))(1+\bar{x}) - u'(\hat{c}_1(\underline{x}))(1+\underline{x})\right]K_1 \qquad (11) \\ &> (\pi(1) - \pi(0))\,\beta\left[u'(\hat{c}_1(\bar{x}))\,\hat{c}_1(\bar{x}) - u'(\hat{c}_1(\underline{x}))\,\hat{c}_1(\underline{x})\right] \\ &= (\pi(1) - \pi(0))(1-\sigma)\,\beta\left[u(\hat{c}_1(\bar{x})) - u(\hat{c}_1(\underline{x}))\right], \end{aligned}$$

if $(1+\bar{x})\,\hat{K}_1 > \hat{c}_1(\bar{x}) > \hat{c}_1(\underline{x}) > (1+\underline{x})\,\hat{K}_1$. This implies that if $1 > \sigma$, the intertemporal wedge will be negative if $(1+\overline{x})K_1 \geq c_1^*(\overline{x}) \geq c_1^*(\underline{x}) \geq (1+\underline{x})K_1$. However, this condition does not restrict the sign of the intertemporal wedge for $\sigma \geq 1$, which is the empirical relevant case.

This result suggests that it is intertemporal substitution and the difference in the expected returns from capital at high and low effort that drive the sign of the intertemporal wedge for this class of preferences. However, expected utility preferences do not enable to disentangle the role of risk aversion or intertemporal substitution. Backus, Routledge, Zin (2004) display examples illustrating that risk aversion or intertemporal substitution may matter in different classes of models with recursive utility. In a version of Weil's (1989) precautionary savings model with Kreps-Portheus preferences, only risk aversion matters for the presence of a precautionary saving motive. On the other hand, in a version of the Merton-Samuelson portfolio choice model, intertemporal substitution is the only parameter that determines the presence of a precautionary motive. In the Weil model, income is the only source of risk, while asset returns risk is present in a portfolio choice model. Angeletos (2005) also displays a model with asset returns risk in which the intertemporal substitution is the only parameter that matters for precautionary savings.

The model described above has elements of portfolio choice, since by choice of effort the agent determines the distribution of capital returns. In addition, higher capital increases consumption in the bad state, irrespective of the distribution of capital returns, thus providing insurance. To sort out the role of intertemporal substitution and risk aversion for the optimal wedge, we study a version of the model with recursive utility.

Assume agents' preferences over lifetime consumption are given by:

$$U_0\left(\{c_0, c_1\}; e\right) = \left[c_0^{1-\sigma} + \beta U_1\left(e\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}},$$

where

$$U_1\left(e\right) = \left[E_e c_1^{1-\alpha}\right]^{\frac{1}{1-\alpha}},$$

is the certainty equivalent of time 1 consumption. Here, α is the coefficient of relative risk aversion and σ is the elasticity of intertemporal substitution. We write the planner's problem assuming high effort is implemented. We will denote with $U_0\left(e\right)$ lifetime utility at effort e. The planner solves the problem:

$$\max_{K_1 \in [0, K_0],\ c_1(x),\ e \in \{0,1\}} U_0\left(e\right) - v\left(e\right)$$

subject to

$$E_e c_1\left(x\right) = K_1 E_e\left(1 + x\right),$$

$$U_0\left(1\right) - U_0\left(0\right) \geq v\left(1\right) - v\left(0\right).$$

The first order necessary conditions for the planning problem imply[2]:

$$\left[\frac{c_1^*\left(\bar{x}\right)}{c_1^*\left(\underline{x}\right)}\right]^{-a} = \frac{1 + \mu\bar{A}}{1 + \mu\underline{A}}, \tag{12}$$

$$\begin{aligned} \mathrm{IW}_K &= \beta E_1\left(1 + x\right) c_1\left(x\right)^{-\alpha} U_1\left(1\right)^{\alpha - \sigma} - c_0^{-\sigma} \qquad (13) \\ &= \mu\bar{A}\beta c_1^*\left(\underline{x}\right)^{-\alpha} U_1\left(1\right)^{\alpha-\sigma}\left\{\left(1 + E_1\left(x\right)\right)\frac{\bar{A}^{-1} + \mu}{A^{-1} + \mu}\right. \\ &\quad \left. - \left[\pi\left(1\right)\left(1 + \bar{x}\right)\left(\frac{c_1^*\left(\bar{x}\right)}{c_1^*\left(\underline{x}\right)}\right)^{-\alpha}\frac{\underline{A}}{\bar{A}} - \left(1 - \pi\left(1\right)\right)\left(1 + \underline{x}\right)\right]\right\}, \end{aligned}$$

where $\bar{A} = \left(1 - \frac{1-\pi(0)}{1-\pi(1)}\left[\frac{U_1(0)}{U_1(1)}\right]^{\alpha-\sigma}\left[\frac{U_0(0)}{U_0(1)}\right]^{\sigma}\right)$, $\underline{A}= \left(1 - \frac{\pi(0)}{\pi(1)}\left[\frac{U_1(0)}{U_1(1)}\right]^{\alpha-\sigma}\left[\frac{U_0(0)}{U_0(1)}\right]^{\sigma}\right)$ and $A = \left[1 - \frac{U_0(0)^{\sigma}}{U_0(1)^{\sigma}}\right]$. Since $\bar{A} < \underline{A}$, equation (12) implies partial insurance, that is $c_1^*\left(\bar{x}\right) > c_1^*\left(\underline{x}\right)$ for $\alpha > 0$. Equation (13) is the expression for the intertemporal wedge. The sign of this expression is ambiguous, hence, the intertemporal wedge can be positive or negative. We show the following.

[2]The derivation is in the appendix.

Proposition 2. *If $\alpha > \sigma$, risk aversion is greater than intertemporal substitution, the intertemporal wedge is negative.*

The proof can be found in the appendix. The proposition shows that if risk aversion is greater than intertemporal substitution, the intertemporal wedge on risky productive capital is negative. To interpret this finding, recall that the intertemporal wedge is negative when an agent who exerts low effort reduces her capital holdings. At lower effort, expected capital returns and the variance of capital returns are lower. Lower expected capital returns generate a negative wealth effect on time 0 consumption that determines an increase in capital holdings via the intertemporal substitution channel and a positive substitution effect that reduces capital holdings via the intertemporal substitution channel. With no uncertainty and with $\sigma = \alpha$, the substitution effects that decreases capital holdings via the intertemporal substitution channel dominates for $\sigma < 1$. On the other hand, lower variance of capital returns determines a substitution effect that increases capital holdings and a wealth effect that reduces capital holdings (e.g. reduces the need to increase consumption in the bad state) via the risk aversion channel. Hence, when risk aversion is high, capital holdings will tend to be higher at low effort. These results from standard preferences suggest that with recursive utility, capital holdings should be lower at low effort for low values of the elasticity of intertemporal substitution and for high risk aversion. This is confirmed by proposition 2. For $\sigma < \alpha$, the intertemporal substitution affect tends to reduce capital holdings at low effort and the risk aversion effect tends to reduce capital holdings at low effort, and the intertemporal wedge is negative.

The expression for the intertemporal wedge (13) clearly implies that if the intertemporal wedge on a risk free asset is always positive, since $\bar{A} < \underline{A}$. This adverse effect on incentives of a riskless asset is independent of the assumptions on preferences[3]. In addition, the proof of proposition 2 can be applied to any asset with return $r(x)$, where $r(\bar{x}) > r(\underline{x})$. It follows that the intertemporal wedge is negative for $\alpha > \sigma$, for any asset that is positively correlated with entrepreneurial capital.

The result in proposition 2 may be important in light of empirical estimates of risk aversion and intertemporal substitution using microeconomic data. Attanasio and Vissing-Jorgensen (2003) use CEX data to show that, for stock holders, the elasticity of intertemporal substitution is likely to be above, but close to, one, whereas relative risk aversion estimates vary between $5 - 10$ for reasonable assumptions on the correlation between consumption growth and stock returns. Previous studies by Attanasio and Weber (1989) and Campbell (1996) had also found, albeit less conclusively, that risk aversion is significantly larger than intertemporal substitution. In light of this evidence,

[3]It follows that in standard private information economies with labor income risk, such as the ones considered in Albanesi and Sleet (2005) and Kocherlakota (2005), the extention to Kreps-Portheus preferences would not alter the prescriptions for optimal asset taxation.

proposition 2 suggests that in empirically relevant ranges for intertemporal substitution and risk aversion, the intertemporal wedge on entrepreneurial capital is indeed negative.

To investigate the properties of optimal allocations in more detail, we now turn to numerical examples.

2.2. Numerical Examples

We restrict attention to utility functions in the CRRA class and assume, with $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$, $v(e) = \gamma e^{1/\gamma}$, $\gamma > 0$. In addition, we assume $\pi(e) = a + be$, with $a \geq 0$, $b > 0$ and $2a + b \leq 1$, so that $V_1(x) \geq V_0(x)$. Hence, the parameter b represents the impact of effort on capital returns. For our benchmark parameterization, we fix $\gamma = 0.1111$, corresponding to a compensated effort supply elasticity of 0.1, a value in line with micro evidence for males. We interpret x as percentage earnings on entrepreneurial capital, which we identify as private equity. We parameterize the distribution of x with the distribution of earnings conditional on survival for private equity in Moscowitz and Vissing-Jorgensen (2002). This corresponds to $\underline{x} = 25\%$ and $\bar{x} = 75\%$ and $\pi(50\%|e = 1) = 0.75$[4]. We set $a = 0$, so that at low effort capital is risk free, and $b = 0.5$. This corresponds to $E_1 x = 0.5$, $E_0 x = 0.25$, $SD_1(x) = 0.25$, where SD_e denotes the standard deviation, conditional on effort e. We consider several other parameterizations of the capital returns distribution to check robustness. We set $K_0 = 1$ for all the parameterizations considered.

Equation (9) shows that the sign of the intertemporal wedge depends on the spread in consumption across states at the constrained-efficient allocation, as well as on the spread in capital returns across states $(\overline{x} - \underline{x})$, which determines the differential $E_1(x) - E_0(x)$ and $V_1(x) - V_0(x)$. If the spread in consumption across states is small or the variance of capital returns is large, the intertemporal wedge on capital is more likely to be negative, other things equal. If the spread in consumption across states at the constrained-efficient allocation is small, the complementarity between capital and effort will mostly be driven by the higher expected rate of return on capital implied by higher effort, and thus is likely to be positive. On the other hand, if the spread in consumption across states is large, then risk exposure is likely to be the dominant force in the complementarity between capital and effort. Since at high values of σ the constrained-efficient allocation is likely to display less spread in consumption across states, the intertemporal wedge will likely be negative for high values of σ.

Given the importance of the parameter σ for the sign of the intertemporal wedge, we compute the optimal allocation and the intertemporal wedge as a function of σ. Our

[4] The average returns to private equity, including capital gains and earnings, are estimated from SCF data to be 12.3, 17.0 and 22.2 percent per year in the time periods 1990-1992, 1993-1995, 1996-1998, as reported in Moskowitz and Vissing-Jorgensen (2002).

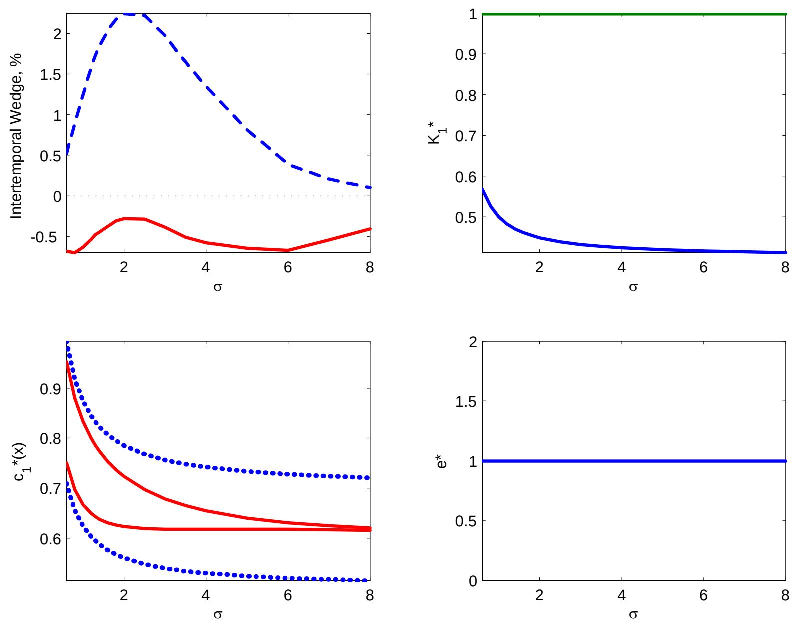


Figure 1: Benchmark parameterization

findings for the benchmark parameterization are displayed in figure 1. The intertemporal wedge on K_1 is the solid line, while the dashed line corresponds to the intertemporal wedge for a risk-free asset with the same expected return as capital. The intertemporal wedge for capital is negative for all values of σ, while the intertemporal wedge on a riskless asset with the same expected return as capital is always positive. Investment is decreasing in σ. In the third panel, we plot constrained-efficient consumption (solid line) and total capital earnings, $K_1^* (1 + x)$ (dotted line), in each state. The amount of insurance embedded in the constrained-efficient allocation increases with σ. Hence, the spread in optimal consumption across states decreases with σ. This contributes to a negative value of the intertemporal wedge as σ increases.

To explore robustness, we also compute the constrained-efficient allocation for higher spread in capital returns. We set $\{\underline{x}, \bar{x}\} = \{0, 1\}$, so that $E_1 x = 0.5$ and $SD_1 x = 0.5$. As shown in figure 2, the intertemporal wedge on productive capital is negative and large in absolute value relative to the previous case. However, since the expected value of capital returns is the same as for figure 1, K_1^*, $c_1^* (x)$ and the intertemporal wedge on the riskless asset are also the same.

Figure 3 shows the properties of the optimal allocation for a smaller spread in capital returns. In particular, $\{\underline{x}, \bar{x}\} = \{0.3, 0.7\}$, which implies $E_1 x = 0.5$, $SD_1 (x) = 0.2$,

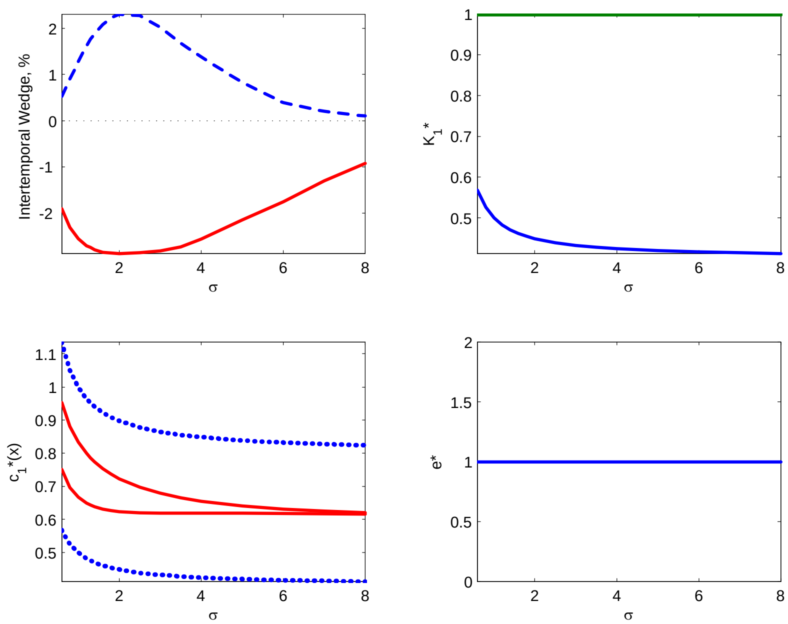


Figure 2: Larger spread in capital returns.

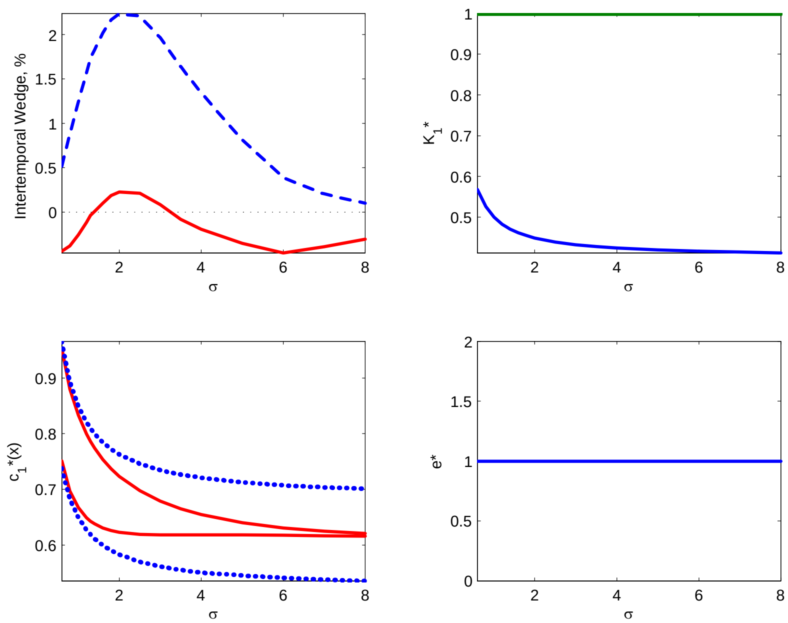


Figure 3: Smaller spread in capital returns.

$E_0(x) = 0.3$. From (9), we know that a smaller spread in capital returns increases the value of the intertemporal wedge on capital. The intertemporal wedge on capital in this case is positive for intermediate values of σ.

Lastly, we consider a parameterization with a higher cost of effort. Specifically, we set $\gamma = 0.15$, which corresponds to a compensated effort supply elasticity of 18%. All other parameters are as in the benchmark parameterization. The incentive compatibility constraint implies that the spread in consumption across states in the constrained-efficient allocation must be larger, with a higher cost of effort. Hence, the intertemporal wedge is now positive for intermediate values of σ. As in the previous examples, the intertemporal wedge turns negative for high values of σ, when the spread in consumption across states is smaller.

3. Optimal Taxes

We now consider how to implement constrained-efficient allocations in a setting where agents can trade in competitive markets. We explore different market structures, or decentralized trading arrangements. A market structure specifies the distribution of own-

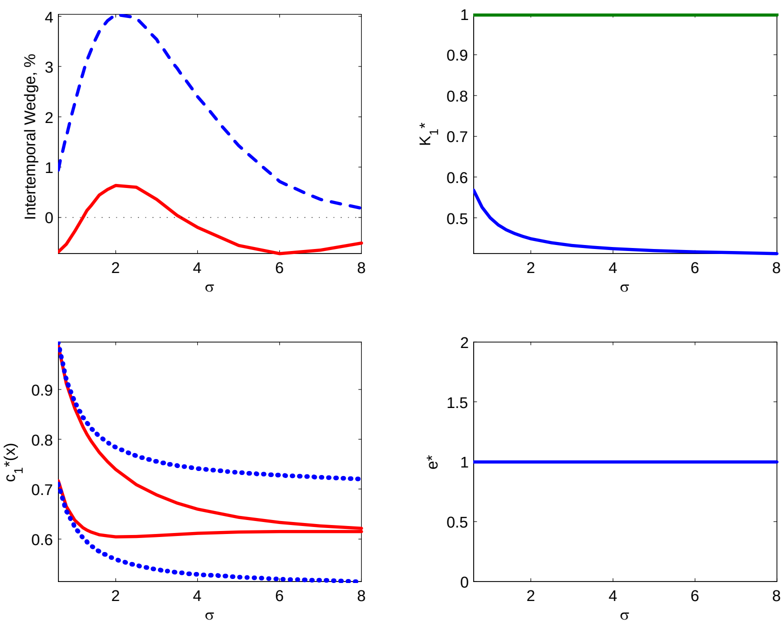


Figure 4: Higher cost of effort.

ership rights, the feasible trades between agents and the information structure. Agents are subject to taxes that influence their budget constraints. The only ex ante constraint imposed on the tax system is that it must specify transfers that are conditioned only on agents' observable characteristics. A tax system implements the constrained-efficient allocation if such an allocation arises as the competitive equilibrium outcome under this tax system for a particular market structure. We examine three different market structures. In the simplest one, agents choose capital and effort. We then allow agents to accumulate capital and trade bonds. Here, we first consider risk-free bonds and then allow for trade in securities with returns that are correlated with idiosyncratic uncertainty. Lastly, we allow entrepreneurs to sell shares in their own capital. In addition, they accumulate capital and trade risk-free bonds. The first market structure is very restrictive but can used to derive the properties of optimal marginal capital taxes. These properties carry over to the more complex market structures. The main result for the market structure in which agents can accumulate capital and trade risk-free securities is the optimality of differential asset taxation. The main result for the environment where entrepreneurs can issue shares in their own capital is the optimality of double taxation of capital earnings.

3.1. Optimal Capital Taxes

The first market structure we consider is one in which agents independently choose capital holding as well as effort. To focus on the properties of optimal marginal capital taxes, we rule out any other means of transferring wealth intertemporally. Decisions occur as follows. Agents are endowed with initial capital K_0 and choose K_1 at the beginning of period 0. Their choice is observable by the government. Agents then choose effort, which is not observed by the government or other agents. At the beginning of period 1, x is realized and observed. Then, the government collects taxes and agents consume. The tax system is represented by the tax function $T(K_1, x)$. We restrict attention to functions T that are differentiable almost everywhere in their first argument and satisfy $E_1 T(K_1, x) = 0$.

An agent's problem is:

$$\left\{\hat{e}, \hat{K}_1\right\}(K_0, T) = \arg \max_{K_1 \in [0, K_0],\ e \in \{0,1\}} U(e, K_1; T) - v(e), \qquad \text{(Problem 2)}$$

where

$$U(e, K_1; T) = u(K_0 - K_1) + \beta E_e u(K_1(1+x) - T(K_1, x))$$

and E_e denotes expectations taken with respect to the probability distributions $\pi(e)$.

The agent's intertemporal Euler equation is:

$$u'(K_0 - K_1) = \beta E_{\hat{e}}\left[u'(K_1(1+x) - T(K_1, x))(1 + x - T_k(K_1, x))\right],$$

where $\hat{e}$ is an agent's optimal choice of effort.

Setting $T(K_1^*, x)$ to satisfy:

$$c_1^*(x) = K_1^*(1+x) - T(K_1^*, x), \tag{14}$$

ensures that, if K_1^* is chosen, high effort will be optimal for the agent. Evaluating the Euler equation at $\{1, K_1^*\}$, we can write:

$$u'(K_0 - K_1^*) = \beta E_1 \left[u'(c_1^*(x))(1 + x - T_k(K_1^*, x)) \right]. \tag{15}$$

We consider separable tax systems of the form: $T(K_1, x) = \rho(x) + \tau_K(x) K_1$. The restrictions on $T(K_1^*, x)$ implied by (14) and.(15) do not fully pin down the properties of the marginal asset tax and do not ensure that the constrained-efficient allocation is optimal for the agent. To see this, let $\tau_K(\bar{x}) = \tau_K(\underline{x}) = \tau$, so that the marginal capital tax does not depend on x. Let $\bar{\rho}(x)$ and $\bar{\tau}$ the values of ρ and τ that satisfy (14) and (15), and let $\bar{T}(K_1, x) = \bar{\rho}(x) + \bar{\tau} K_1$. To check whether $e = 1$ and K_1^* are optimal for the agent under this tax system, we can examine the approximate Hessian of Problem 2. A sufficient condition for the determinant of the Hessian to be positive is that the off-diagonal term is small in absolute value. This term is given by:

$$\begin{aligned} \frac{\Delta U_{K_1}}{\Delta e}\left(\left(e^*, K_1^*; \bar{T}\right)\right) \equiv \; & (\pi(1) - \pi(0))\beta\left[u'(\bar{c}_1^*)(1+\bar{x}) - u'(\underline{c}_1^*)(1+\underline{x})\right] \\ & - \bar{\tau}(\pi(1) - \pi(0))\beta\left[u'(\bar{c}_1^*) - u'(\underline{c}_1^*)\right]. \end{aligned}$$

As discussed in section 2.1, the first term in $\frac{\Delta U_{K_1}}{\Delta e}$ has the opposite sign of the intertemporal wedge. Since the sign of $\bar{\tau}$ is the same as the sign of the intertemporal wedge:

$$sign\left\{\frac{\Delta e}{\Delta K_1}\left(e^*, K_1^*; \bar{T}\right)\right\} = sign\left\{-\mathrm{IW}_K\right\}.$$

If the absolute value of $\frac{\Delta U_{K_1}}{\Delta e}$ is sufficiently large, agents may conduct profitable "joint deviations" in which they set $e = 0$ and $K_1 \neq K_1^*$. If the intertemporal wedge is negative (positive), a deviating agent will reduce (increase) her capital holdings relative to K_1^*.

How can marginal asset taxes be set to prevent such a joint deviation? Following Albanesi and Sleet (2005), we show the following.

Proposition 3. *A the tax system $T^*(K_1, x)$, in the class $T(K_1, x) = \tau_K(x)K_1 + \rho(x)$, that satisfies:*

$$c_1^*(x) = K_1^*(1+x) - \rho^*(x), \tag{16}$$

$$\frac{u'(c_0^*)}{\beta u'(c_1^*(x))} = 1 + x - \tau_K^*(x), \tag{17}$$

implements the constrained-efficient allocation.

Proof. Suppose not. By definition, $\left\{\hat{e}, \hat{K}_1\right\}(K_0, T^*)$ satisfies:

$$u'\left(K_0 - \hat{K}_1\right) = \beta E_{\hat{e}} u'\left(\hat{K}_1(1+x) - T^*\left(\hat{K}_1, x\right)\right)(1 + x - \tau_K^*(x)).$$

If $\left\{\hat{e}, \hat{K}_1\right\}(K_0, T^*) \neq (1, K_1^*)$:

$$1 = E_{\hat{e}} \frac{u'\left(\hat{K}_1(1+x) - T^*\left(\hat{K}_1, x\right)\right)}{u'\left(K_0 - \hat{K}_1\right)} \frac{u'(c_0^*)}{u'(c_1^*(x))}.$$

Then, irrespective of $\hat{e}$, $\hat{K}_1 > K_1^*$ implies $\frac{u'\left(\hat{K}_1(1+x)-T^*\left(\hat{K}_1,x\right)\right)}{u'\left(K_0-\hat{K}_1\right)} < \frac{u'\left(c_0^*\right)}{u'\left(c_1^*(x)\right)}$, while $\hat{K}_1 < K_1^*$ implies $\frac{u'\left(\hat{K}_1(1+x)-T^*\left(\hat{K}_1,x\right)\right)}{u'\left(K_0-\hat{K}_1\right)} > \frac{u'\left(c_0^*\right)}{u'\left(c_1^*(x)\right)}$. This is a contradiction. Hence, at $\hat{K}_1 = K_1^*$. Then, by incentive compatibility (16) implies $\hat{e} = 1$. ■

The following corollary characterizes the properties of the optimal tax system. The average marginal capital tax is zero. The marginal capital tax is decreasing in capital returns, and thus appears regressive, if the intertemporal wedge is positive, while it is increasing in capital returns and thus appears regressive is the intertemporal wedge is negative.

Corollary 4. *The tax system $T^*(K_1, x)$ defined by (16) and (17) implies:*
i) $E_1 \tau_K^*(x) = 0$;
ii) $sign\,(\tau_K^*(\bar{x}) - \tau_K^*(\underline{x})) = sign\,(-IW_K)$.

Proof. By (17):

$$E_1\left[1 + x - \frac{u'(c_0^*)}{u'(c_1^*(x))}\right] = E_1 \tau_K^*(x),$$

which from (5) implies $E_1 \tau_K^*(x) = 0 = 0$. (17) also implies:

$$u'(c_1^*(\overline{x}))\tau_K^*(\overline{x}) - u'(c_1^*(\underline{x}))\tau_K^*(\underline{x}) = u'(c_1^*(\overline{x}))(1+\overline{x}) - u'(c_1^*(\underline{x}))(1+\underline{x}).$$

Since:

$$sign\left[u'(c_1^*(\overline{x}))(1+\overline{x}) - u'(c_1^*(\underline{x}))(1+\underline{x})\right] = sign\,(\text{-IW}_K)$$

and $u'(c_1^*(\overline{x})) < u'(c_1^*(\underline{x}))$, it follows that $sign\,(\tau_K^*(\bar{x}) - \tau_K^*(\underline{x})) = sign\,(-\text{IW}_K)$. ■

Remark 5. *We have assumed here that the level of capital held by agents is observed by the government. In this setting, however, the government need not know the level of K_1 to implement the the constrained-efficient allocation, since the marginal tax on* *capital does not depend on the level of capital. This will also hold in a model with more than two periods, if marginal taxes are allowed to depend on the history of realizations of idiosyncratic capital returns, as in Kocherlakota (2005). If on the other hand, marginal taxes are constrained to depend on the current level of capital returns only, as in Albanesi and Sleet (2005), then marginal capital taxes will depend on the level of* K_1*, which summarizes an agent's history. The assumption that* K_1 *is observable is motivated by concerns regarding administrative feasibility. Since capital is agent specific and directly held, it is only feasible to collect capital taxes if the level of capital is observed. For assets that are traded on financial markets, it is possible for the government to collect taxes at the source, without directly observing the quantities of assets held by each individual agent. We turn to this case in the next implementation.*

3.2. Optimal Differential Asset Taxation

We now allow agents to trade risk-free bonds. Agents' initial capital is K_0 and their initial endowment of bonds is $B_0 \geq 0$.The timing of events is as follows. At time 0, agents choose K_1 and they purchase, at price q, bonds, $B_1 \geq 0$, that pay one unit of consumption at time 1. Agents then exert effort. At time 1, capital returns x are realized, the government levies taxes and the agents consume. We assume that K_1 and B_1 are observed by the government. The tax system is represented by the function: $T(B_1, K_1, x)$.

The agents' problem is:

$$\left\{\hat{e}, \hat{K}_1, \hat{B}_1\right\}(B_0, K_0, T) = \arg \max_{K_1 \in [0, K_0],\ B_1 \geq \bar{B},\ e \in \{0,1\}} U(e, K_1, B_1; T) - v(e), \tag{Problem 3}$$

where

$$U(e, K_1, B_1; T) = u(K_0 + B_0 - K_1 - B_1) + E_e u(K_1(1+x) + B_1(1+r) - T(K_1, B_1, x)),$$

subject to $K_0 + B_0 - K_1 - qB_1 \geq 0$ and $K_1(1+x) + B_1 - T(B_1, K_1, x) \geq 0$ for $x \in X$. Here, the debt limit $\bar{B}$ is imposed to ensure that an agent's problem is well defined. The natural debt limit for tax systems in the class $T(B_1, K_1, x) = \rho(x) + \tau_B(x) B_1 + \tau_K(x) K_1$ is $\bar{B} = -\frac{[K_1(1+\underline{x}-\tau_K(\underline{x}))-\rho(\underline{x})]}{1+r-\tau_B(\underline{x})}$. This limit ensures that agents will be able to pay back all outstanding debt in the low state.

We restrict attention to a tax system of the form: $T(B_1, K_1, x) = \rho(x) + \tau_B(x) B_1 + \tau_K(x) K_1$ and show the following.

Proposition 6. *A tax system* $T^*(B_1, K_1, x)$*, in the class* $T(B_1, K_1, x) = \rho(x) +$

$\tau_B(x) B_1 + \tau_K(x) K_1$, and an initial bond allocation B_0^* that satisfy:

$$1 - \tau_B^*(x) = \frac{qu'(c_0^*)}{\beta u'(c_1^*(x))}, \quad (18)$$

$$1 + x - \tau_K^*(x) = \frac{u'(c_0^*)}{\beta u'(c_1^*(x))}, \quad (19)$$

$$c_1^*(x) = K_1^*(1+x) - \tau_K(x) K_1^* - \rho^*(x), \quad (20)$$

and

$$c_0^* = B_0^* + K_0 - K_1^* - B_1^*, \quad (21)$$

for some $B_1^* \geq \bar{B}$ and some $q < 1$ implements the constrained-efficient allocation.

Proof. We want to show that

$$\left\{\hat{e}, \hat{K}_1, \hat{B}_1\right\}(B_0, K_0, T^*) = (1, K_1^*, B_1^*),$$

for some $B_1^* \geq 0$ and for given q. The agents' intertemporal Euler equations evaluated at the constrained efficient allocation are:

$$U_B(\hat{e}) = -u'(c_0^*) q + E_{\hat{e}} u'(c_1^*(x))(1 - \tau_B^*(x)) = 0,$$

$$U_K(\hat{e}) = -u'(c_0^*) + E_{\hat{e}} u'(c_1^*(x))(1 + x - \tau_K^*(x)) = 0,$$

where U_x denotes a partial derivative with respect to the variable x. Suppose that $\left\{\hat{e}, \hat{K}_1, \hat{B}_1\right\}(B_0, K_0, T^*) \neq (1, K_1^*, B_1^*)$. We consider two classes of deviations, local deviations at an interior solution of the agents problem and large deviations where either K_1 or B_1 is not interior. For local deviations, note that at T^* :

$$1 = E_{\hat{e}} \frac{u'(\hat{c}_1(x))}{u'(\hat{c}_0)} \frac{u'(c_0^*)}{u'(c_1^*(x))},$$

irrespective of the value of $\hat{e}$. It follows that if either $\hat{K}_1 \gtrless K_1^*$ or $\hat{B}_1 \gtrless B_1^*$, then (18) and (19) imply $\frac{u'(\hat{c}_1(x))}{u'(\hat{c}_0)} \lesseqgtr \frac{u'(c_0^*)}{u'(c_1^*(x))}$ irrespective of the value of $\hat{e}$, a contradiction at (21). The local sufficient conditions for optimality are also satisfied irrespective of the value of $\hat{e}$. To see this, consider the sub-optimization problem associated with the choice of B_1 and K_1 for given e. The elements of the Hessian, H_U, for this problem are:

$$U_{BB}(\hat{e}) = q^2 u''(c_0^*) + E_{\hat{e}} u''(c_1^*(x))(1 - \tau_B^*(x))^2 \leq 0,$$

$$U_{KK}(\hat{e}) = u''(c_0^*) + E_{\hat{e}} u''(c_1^*(x))(1 + x - \tau_K^*(x))^2 \leq 0,$$

$$U_{BK}(\hat{e}) = u''(c_0^*)\, q + E_{\hat{e}} u''(c_1^*(x))\,(1-\tau_B^*(x))\,(1+x-\tau_K^*(x)).$$

Under (18)-(19), $|H_U| = \left(1-2q+q^2\right) u''(c_0^*)\, E_{\hat{e}} u''(c_1^*(x)) \left(\frac{qu'(c_0^*)}{\beta u'(c_1^*(x))}\right)^2 > 0$ for $q \neq 1$. We now consider large deviations. Since $K_1 = K_0$ and $B_1 = 0$ can never be optimal given the Inada conditions, we only need to consider the following deviation: $\hat{K}_1 = 0$ and $\hat{B}_1 > 0$. This deviation is budget feasible for the agent. If such a deviation is optimal it satisfies:

$$-u'\left(K_0 - q\hat{B}_1\right) q + E_{\hat{e}} u'\left(\hat{B}_1 + \rho^*(x)\right)(1-\tau_B^*(x)) = 0, \tag{22}$$

$$-u'\left(K_0 - q\hat{B}_1\right) q + E_{\hat{e}} u'\left(\hat{B}_1 + \rho^*(x)\right)(1+x-\tau_K^*(x)) < 0. \tag{23}$$

However, by (18) and (19), $(1-\tau_B^*(x)) = (1+x-\tau_K^*(x))$, and (22) and (23) clearly cannot hold at the same time irrespective of the value of $\hat{e}$. Then, at K_1^*, B_1^* are globally optimal irrespective of the value of $\hat{e}$. At at K_1^*, B_1^*, $\rho^*(x)$ implies $\hat{e} = 1$ since the constrained-efficient allocation is incentive compatible. The fact there are no profitable large deviations implies that global sufficient conditions are satisfied. ■

The tax system T^* removes the complementarity between effort and investment and effort and bond holdings and it guarantees that the necessary and sufficient conditions for the *joint* global optimality of K_1^* and $B_1^* = 0$ are satisfied at all effort levels.

Corollary 7. *The tax system $T^*(K_1, B_1, x)$, the bond price $q = 1/E_1(1+x)$ and allocation $\{1, B_1^*, K_1^*\}$ with $B_1^* = 0$ constitute a competitive equilibrium for the market economy with initial capital K_0 and initial bond holdings $B_0 = 0$.*

Proof. By Proposition 6, for any $q < 1$ and $B_1^* \geq 0$, the allocation $\{1, B_1^*, K_1^*\}$ solves the agents' optimization problem in the market economy. In addition, (20) and (21) imply that the resource constraint are satisfied at $t = 0$ and $t = 1$ and that the government budget constraint is also satisfied. ■

We now derive implications for asset taxes.

Proposition 8. *The tax system $T^*(B_1, K_1, x)$ defined by (18), (19) and (20) implies:*

i) $E_1\tau_K^*(x) = 0$;

ii) sign $(\tau_K^*(\bar{x}) - \tau_K^*(\underline{x})) =$ sign $(-IW_K)$;

iii)$E_1\tau_B^*(x) = 0$;

iv) $\tau_B^*(\bar{x}) < \tau_B^*(\underline{x})$;

v) $\tau_B^*(\underline{x}) > \tau_K^*(\underline{x})$ and $\tau_B^*(\bar{x}) < \tau_K^*(\bar{x})$;

vi) the intertemporal wedge associated with the risk-less bond is greater than the intertemporal wedge associated with risky productive capital.

Proof. Results i) and ii) can be shown as in Corollary 4. Result iii) follows from the planner's Euler equation, since:

$$E_1\tau_B^*(x) = 1 - qE_1\left(\frac{u'(c_0^*)}{\beta u'(c_1^*(x))}\right),$$

and $q = 1/E_1(1+x)$. Result iv) follows directly from (18) and $u'(c_1^*(\bar{x})) < u'(c_1^*(\underline{x}))$. To show result v) note that (18) and (19) imply $\tau_B^*(x) - \tau_K^*(x) = E_1x - x$. The intertemporal wedge associated with each asset is:

$$\begin{aligned} E_1u'(c_1^*(x))(1+x) - u'(c_0^*) &= E_1u'(c_1^*(x))\tau_K^*(x), \\ E_1u'(c_1^*(x))E_1(1+x) - u'(c_0^*) &= E_1u'(c_1^*(x))\tau_B^*(x). \end{aligned}$$

Result vi) follows directly. ■

In Corollary 7, we construct a competitive equilibrium in which $q = 1/E_1(1+x)$, which implies, as shown in proposition 8 that $E_1\tau_B^*(x) = 0$. We can construct other competitive equilibria, in which $q \neq 1/E_1(1+x)$. Let $0 < \hat{q} < 1$. Then, (19) and the planner's Euler equation imply:

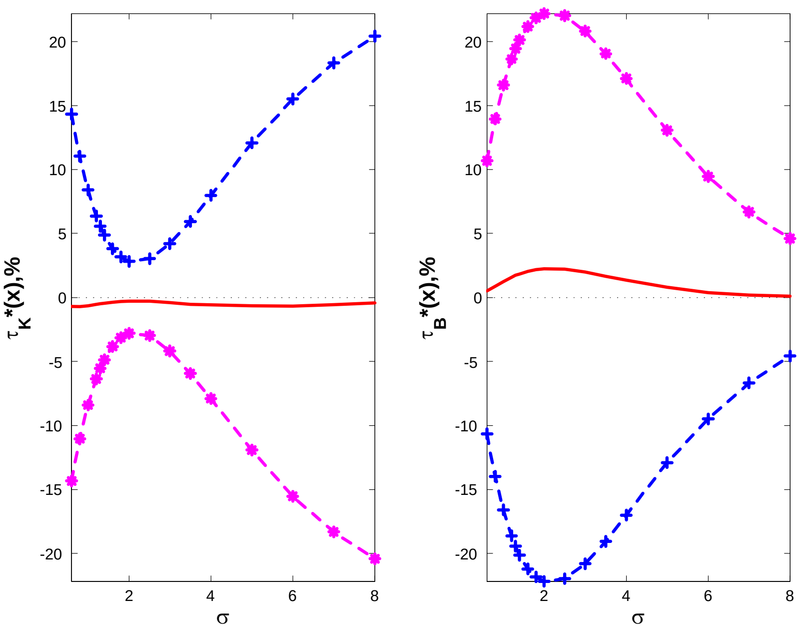
$$E_1\tau_B^*(x) = 1 - \hat{q}E_1\left(\frac{u'(c_0^*)}{\beta u'(c_1^*(x))}\right) = 1 - \hat{q}E_1(1+x).$$

So that: $E_1\tau_B^*(x) > 0$ for $\hat{q} < 1/E_1(1+x)$ and $E_1\tau_B^*(x) < 0$ for $\hat{q} > 1/E_1(1+x)$. Hence, there is an indeterminacy in the equilibrium bond price and the level of the marginal bond tax. This indeterminacy does not affect the dependence of marginal bond taxes on x, which is pinned down by (19).

The asset positions K_1 and B_1 need not be observable to implement the constrained–efficient allocation, as long as x is observable, because taxes on assets are linear. As discussed in Remark 5, if marginal asset taxes can be made dependent on the realization of individual risk, it is possible to implement the constrained-efficient allocation with no hidden saving in a market structure where asset positions are not observed by the government[5]. This implementation requires an administrative arrangement in which taxes are collected at the source. It is natural to think of such a system for financial assets, such as bonds, traded on a centralized market, if a market institution or intermediaries are clearing the transactions. Here, we maintain the assumption that B_1 is observable to the government for symmetry with section 3.1.

[5] Bizer and DiMarzo (1999) derive a related result in a standard moral hazard model in which agents may borrow. They show that as long as debt repayments can be made state contingent, it is possible to implement the constrained-efficient allocation with observable savings, even if borrowing is unobserved by the principal, who designs the incentive-compatible transfer (salary) policy. In their setting, it is important that agents borrow, rather than save. Only in this case can the return be made state contingent. This requirement does not hold in our implementation, since the government can set marginal bond taxes to be state contingent.

The optimal marginal capital and bond taxes for the benchmark parameterization discussed in section 2.2 are plotted in figure **??**, for $r = E_1(x)$. The solid line in each panel corresponds to the intertemporal wedge. The dashed-asterix line corresponds to marginal taxes in state $\underline{x}$, whereas the dashed-cross line corresponds to optimal marginal taxes in state $\bar{x}$. The marginal tax on capital, displayed in the left panel, is negative in the low state and positive in the good state, while the opposite is true for the marginal tax on bonds. Hence, the marginal capital tax is increasing in earnings, while the marginal bond tax is decreasing in earnings. Despite the fact that wedges are quite small in percentage terms, the magnitude of marginal taxes is significant. The capital tax ranges from 3 to 21% in absolute value as a function of σ, while the bond tax ranges from 5 to 24% in absolute value.



Optimal capital and bond taxes, benchmark parameterization.

In figure 5, we report optimal marginal capital and bond taxes for the parameterization with a larger cost of effort. In this case, the intertemporal wedge on productive capital is positive for intermediate values of σ. When IW_K is positive, the marginal capital tax is decreasing in capital returns, but the marginal tax on capital in the bad state is always lower than the marginal tax on bonds in the bad state.

The main finding in the fiscal implementation for the market structure with riskless bonds is the *optimality of differential asset taxation.* There are two aspects of this result. First, the intertemporal wedge on a riskless asset is always greater than the intertemporal wedge on entrepreneurial capital. Second, entrepreneurial capital should

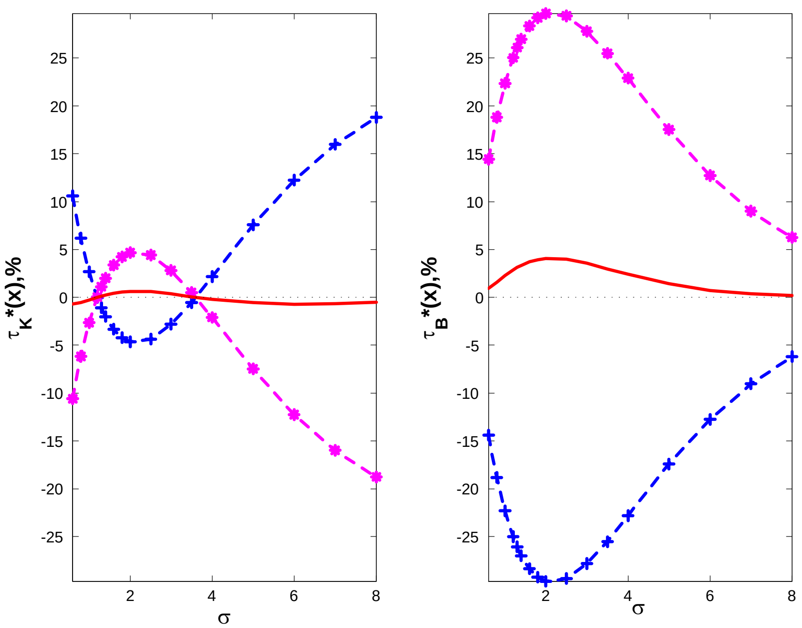


Figure 5: Optimal capital and bond taxes, larger cost of effort.

be subsidized relatively to a riskless asset in the bad state, irrespective of the sign of the intertemporal wedge. The optimal tax system reduces after tax spread in capital returns, whereas it increases the after tax spread in the returns to the riskless bond.

Results are similar for risky securities with returns that are correlated with x. Consider a security with return $r(x) > 0$ for $x = \underline{x}, \bar{x}$, in zero net supply. Letting the candidate tax system be given by $T(S_1, K_1, x) = \tau_K(x) K_1 + \tau_S(x) S_1 + \rho(x)$.Set $\tau_K^*(x)$ and $\rho^*(x)$ as in (19) and (20) for $S_1^* = 0$. Set marginal taxes on the security according to:

$$1 + r(x) - \tau_S^*(x) = \frac{qu'(c_0^*)}{\beta u'(c_1^*(x))}. \tag{24}$$

The equilibrium price of the security is $q = \frac{E_1\left(1+r(x)-\tau_S^*(x)\right)}{E_1\left(1+x-\tau_K^*(x)\right)}$[6]. Then, (24) implies $E_1\tilde{r}(x) = E_1 x$, where $1 + \tilde{r}(x) = \frac{1+r(x)}{q}$.

The intertemporal wedge on the risky security is:

$$\mathrm{IW}_S = E_1 u'(c_1^*(x))(1 + \tilde{r}(x)) - u'(c_0^*),$$

Following the usual reasoning:

$$sign\{\mathrm{IW}_S\} = sign\left\{u'(c_1^*(\underline{x}))(1 + \tilde{r}(\underline{x})) - u'(c_1^*(x))(1 + \tilde{r}(\bar{x}))\right\}.$$

Then, if $Cov_1(\tilde{r}(x), x) \leq 0$, the intertemporal wedge on the risky security is positive. The intertemporal wedge can be positive or negative if $Cov_1(\tilde{r}(x), x) > 0$. The following result holds.

Proposition 9. *If $Cov_1(\tilde{r}(x), x) > 0$ and $V_1(x) > V_1(\tilde{r}(x))$, then:*

$$E_1 u'(c_1^*(x))(1 + \tilde{r}(x)) - u'(c_0^*) > E_1 u'(c_1^*(x))(1 + x) - u'(c_0^*),$$

$$\tau_S^*(\bar{x}) - \tau_K^*(\bar{x}) < 0 \text{ and } \tau_S^*(\underline{x}) - \tau_K^*(\underline{x}) > 0.$$

Proof. This follows from:

$$\begin{aligned} E_1 u'(c_1^*(x))(1 + \tilde{r}(x)) - E_1 u'(c_1^*(x))(1 + x) &= Cov_1\left(u'(c_1^*(x)), \tilde{r}(x)\right) - Cov_1\left(u'(c_1^*(x)), x\right) \\ &= Cov_1\left(u'(c_1^*(x)), \tilde{r}(x) - x\right). \end{aligned}$$

$Cov_1(u'(c_1^*(x)), \tilde{r}(x) - x) > 0$ if $\tilde{r}(x) - x$ is decreasing in x, or $Cov_1(\tilde{r}(x) - x, x) < 0$. By the definition of covariance and by the fact that $E_1 x = E_1\tilde{r}(x)$:

$$Cov_1(\tilde{r}(x) - x, x) = E_1\tilde{r}(x)x - E_1 x^2 = Cov_1(\tilde{r}(x), x) - V_1(x). \tag{25}$$

[6] As in the case with risk-free bonds, the equilibrium expected return on this security is not separately pinned down from $E_1\tau_s^*(x)$.

By $V_1(x) > V_1(\tilde{r}(x))$ and $Cov_1(\tilde{r}(x), x) > 0$, $0 < Corr_1(\tilde{r}(x), x) < 1$. Then:

$$Cov_1(\tilde{r}(x), x) - V_1(x) = Sd_1(x)\left[Corr_1(\tilde{r}(x), x)\, Sd_1(\tilde{r}(x)) - Sd_1(x)\right] < 0.$$

In addition, $\tau_S^*(x) - \tau_K^*(x) = \tilde{r}(x) - x$. Since $\tilde{r}(x) - x$ is decreasing in x and $E_1\tilde{r}(x) = E_1 x$, $\tau_S^*(\bar{x}) - \tau_K^*(\bar{x}) < 0$ and $\tau_S^*(\underline{x}) - \tau_K^*(\underline{x}) > 0$. ■

If $Cov_1(\tilde{r}(x), x) > 0$ and $V_1(x) > V_1(\tilde{r}(x))$, $Corr_1(\tilde{r}(x), x) \in (0, 1)$, where $Corr_e$ denotes the correlation conditional on $\pi(e)$. The proposition states that a security positively correlated with capital with lower variance of returns has a higher intertemporal wedge than capital. An agent would be willing to hold such a security instead of capital, since it is associated with lower earnings risk. However, this has an adverse effect on incentives. This motivates the higher intertemporal wedge and the fact that $\tau_S^*(x) - \tau_K^*(x)$ is decreasing in x, which implies that capital is subsidized with respect to the risky security in the bad state. Then, the correlation of an asset's returns with idiosyncratic capital risk determines the asset's incentive effects and the properties of optimal marginal taxes on the asset.

3.3. Optimal Double Taxation of Capital Income

We now allow entrepreneurs to sell shares of their capital and buy shares of other entrepreneurs' capital. In this setting, each entrepreneur can be interpreted as a private company, so that this arrangement introduces a market for private equity. The amount of capital invested by an entrepreneur can be interpreted as the size of their firm.

An entrepreneur's budget constraint in each period is :

$$c_0 = K_0 - K_1 - B_1 - \int_{i\in[0,1]} S_1(i)\, di + sK_1,$$

$$c_1(x) = K_1(1+x) - sK_1(1+d(x)) + \int_{i\in[0,1]} (1+D(i))\, S_1(i)\, di + B_1(1+r) - T\left(K_1, s, B_1, \{S_1\}_i, x\right),$$

where $s \in [0,1]$ is the fraction of ownership sold, $S_1(i)$ is the value of shares in company i purchased, $d(x)$ dividends distributed to shareholders and $D(i, \tilde{x})$ denotes dividends earned from stocks of company i if the realized returns are $\tilde{x}$ for $\tilde{x} \in X$. The size of each entrepreneur's firm is simply given by their capital stock at the end of period 0. The value of a firm in terms of consumption goods, net of dividend payments is 1. Gross stock returns for an entrepreneur with private equity portfolio $\{S_1(i)\}_i$ are given by $\int_{i\in[0,1]}(1+D(i))\, S_1(i)\, di$, where $D(i)$ denotes expected dividends from firm i[7]. Entrepreneurs choose B_1, K_1, $\{S_1(i)\}_i$ as well as effort at time 0, taking as given the

[7] This holds since x is i.i.d. across entrepreneurs and the law of large numbers hold, given that i is distributed on a continuum.

distribution policy, dividend processes for other companies, r and taxes. At time 1, x is realized, dividends are distributed, the government collects taxes and the entrepreneurs consume. The tax system, $T(K_1, s, B_1, \{S_1\}_i, x)$, is conditional on variables that are assumed to be public information.

We consider candidate tax systems of the form $T(K_1, B_1, \{S_1\}_i, x) = \tau_P(x)(1+x)K_1 + \tau_B(x)B_1 + \tau_s(x)\int_i S_1(i)\,di + \rho(x)$. Here, $\tau_P(x)$ is a marginal tax on gross profits. This candidate tax system does not condition taxes on the fraction of equity issued by entrepreneurs. The marginal tax on stock returns, $\tau_S(x)$, depends on the agent holding the stock and is the same for all stocks. We assume that the distribution policy $d(x)$ is exogenous[8].

An entrepreneur's Euler equations are:

$$-u'(c_0)(1-s) + \beta E_{\hat{e}}\left[(1+x)(1-\tau_P(x)) - (1+d(x))s\right]u'(c_1(x)) = 0, \tag{26}$$

$$-u'(c_0) + \beta E_{\hat{e}}\left[1 + r - \tau_B(x)\right]u'(c_1(x)) = 0, \tag{27}$$

$$-u'(c_0) + \beta E_{\hat{e}}\left[1 + D(i) - \tau_S(x)\right]u'(c_1(x)) = 0, \tag{28}$$

$$\left[u'(c_0) - \beta E_{\hat{e}}(1+d(x))u'(c_1(x))\right]K_1 \begin{cases} = 0 \text{ for } s \in (0,1) \\ \leq 0 \text{ for } s = 0 \\ > 0 \text{ for } s = 1. \end{cases}. \tag{29}$$

We define a competitive equilibrium for this trading structure and then consider how to implement the constrained-efficient allocation. Since all entrepreneurs are ex ante identical, we restrict attention to symmetric equilibria in which $sK_1 = S_1(i)$, effort is constant across entrepreneurs. Consequently, $D(i) = E_{\hat{e}}D(i, \tilde{x})$, and $D(i,x) = d(x)$ for all i.

Definition 10. *A (symmetric) competitive equilibrium is an allocation for the entrepreneurs* $\left\{\hat{K}_1, \hat{s}, \hat{B}_1, \left\{\hat{S}_1(i)\right\}_i, \hat{e}, \hat{c}_1(x)\right\}$ *with* $\hat{s} \in [0,1)$, *an interest rate* r, *a distribution policy* $\hat{d}(x)$ *and a dividend process* $\hat{D}(i,x)$ *for* $i \in [0,1]$, $x \in X$, *and a tax system* $T(K_1, s, B_1, \{S_1\}_i, x)$, *such that:*

i) the allocation $\left\{\hat{K}_1, \hat{s}, \hat{B}_1, \left\{\hat{S}_1(i)\right\}_i, \hat{e}, \hat{c}_1(x)\right\}$ *solves the entrepreneurs' problem, for given* $\hat{d}(x)$, $\hat{D}(i, \tilde{x})$, r *and* T;

ii) the dividend process is consistent with the distribution policy, $\hat{d}(x) = \hat{D}(i,x)$ *for all* i *and* $x \in X$;

ii) the stock market clears, $\hat{s}\hat{K}_1 = \hat{S}_1(i)$;

iii) the resource constraint is satisfied in each period.

---

[8] In the following section, we allow shareholders to optimally select the distribution policy, subject to an incentive compatibility constraint.

The restriction $\hat{s} \in [0,1)$ stems from the fact that an entrepreneur's optimal choice of capital is undetermined at $s = 1$, since her utility does not depend on K_1 in this case. Moreover, at $s = 1$, an entrepreneur does not have any gains from effort, so $e = 0$. The optimal choice of s is a function of the distribution policy. The following result holds.

Proposition 11. *In any competitive equilibrium with* $s \in (0,1)$, $E_{\hat{e}} (1 + d(x)) u'(c_1(x)) = E_{\hat{e}} (1 + x)(1 - \tau_P(x)) u'(c_1(x))$.

Proof. Suppose that the distribution policy is $\hat{d}(x)$ and that $E_{\hat{e}} (1 + d(x)) u'(c_1(x)) \neq E_{\hat{e}} (1 + x)(1 - \tau_P(x)) u'(c_1(x))$ for some tax system where (26) holds with equality at $\hat{\tau}_P(x)$. Denote the corresponding competitive equilibrium allocation with $\left\{\hat{K}_1, \hat{s}, \hat{B}_1, \left\{\hat{S}_1(i)\right\}_i, \hat{e}, \hat{c}_1(x)\right\}$, with $\hat{K}_1 > 0$. If $E_{\hat{e}} (1 + d(x)) u'(c_1(x)) > E_{\hat{e}} (1 + x)(1 - \tau_P(x)) u'(c_1(x))$, for some $0 < \hat{s} < 1$, we can write:

$$
\begin{aligned}
0 &= -u'(\hat{c}_0)(1 - \hat{s}) + \beta E_{\hat{e}} \left[(1 + x)(1 - \hat{\tau}_P(x)) - \left(1 + \hat{d}(x)\right)\hat{s}\right] u'(\hat{c}_1(x)) \\
&< -(1 - \hat{s}) \left[u'(\hat{c}_0) - \beta E_{\hat{e}} \left(1 + \hat{d}(x)\right) u'(\hat{c}_1(x))\right],
\end{aligned}
$$

which implies $0 > u'(\hat{c}_0) - \beta E_1 \left(1 + \hat{d}(x)\right) u'(\hat{c}_1(x))$. But by (29), $\hat{s} = 0$. Contradiction. Similarly, if $E_{\hat{e}} (1 + d(x)) u'(c_1(x)) < E_{\hat{e}} (1 + x)(1 - \tau_P(x)) u'(c_1(x))$ for some $0 < \hat{s} < 1$:

$$
\begin{aligned}
0 &= -u'(\hat{c}_0)(1 - \hat{s}) + \beta E_{\hat{e}} \left[(1 + x)(1 - \hat{\tau}_P(x)) - \left(1 + \hat{d}(x)\right)\hat{s}\right] u'(\hat{c}_1(x)) \\
&> -(1 - \hat{s}) \left[u'(\hat{c}_0) - \beta E_{\hat{e}} \left(1 + \hat{d}(x)\right) u'(\hat{c}_1(x))\right].
\end{aligned}
$$

Then, $u'(\hat{c}_0) - \beta E_1 \left(1 + \hat{d}(x)\right) u'(\hat{c}_1(x)) > 0$, which by (29) implies $\hat{s} = 1$. Contradiction. ■

Based on this result, we consider the distribution policy: $1 + d^*(x) = (1 + x)(1 - \tau_P(x))$, which satisfies the restriction in proposition 11. Dividends per share are simply given by after tax profits. This prescription should be interpreted as part of the share issuing agreement, and is taken as given by the entrepreneurs and the shareholders. Then: $1 + D^*(i) = E_{\hat{e}} (1 + x)(1 - \tau_P(x)) = 1 + E_{\hat{e}} x - E_{\hat{e}} \tau_P(x) - E_{\hat{e}} \tau_P(x) x$. If we restrict attention to competitive equilibria in which $r = E_1 x$, in any competitive equilibrium with $B_1 \geq 0$ and $S_1(i) > 0$ for all i, it must be that $\tau_S(x) - \tau_B(x) = -E_{\hat{e}} \tau_P(x) - E_{\hat{e}} \tau_P(x) x$ for $x \in X$[9].

[9] As discussed in section 3.2, the equilibrium rate of return on bonds is not pinned down separately from the expected marginal bond tax.

Consider a candidate tax system where the marginal profit tax is $\tau_P^*(x)$, defined by:

$$(1+x)(1-\tau_P^*(x)) = \frac{u'(c_0^*)}{\beta u'(c_1^*(x))}. \tag{30}$$

Equation (30) is the entrepreneurs' Euler equation for K_1, evaluated at the constrained-efficient allocation at distribution policy $d^*(x)$. It implies that $\tau_P^*(x)$ is independent from s. Impose that $\tau_B^*(x)$ satisfies (19). In addition, set $\tau_S^*(x)$ so that:

$$1 + D^*(i) - \tau_S^*(x) = \frac{u'(c_0^*)}{\beta u'(c_1^*(x))}, \tag{31}$$

where $D^*(i) = E_{\hat{e}}x - E_{\hat{e}}\tau_P^*(x) - E_{\hat{e}}\tau_P^*(x)x$. Lastly, let $\rho^*(x)$ satisfy:

$$\begin{aligned} c_1^*(x) &= K_1^*(1+x)(1-\tau_P^*(x)) - s^*K_1^*(1+d^*(x)) + (1+E_1(x)-\tau_B^*(x))B_1^* \\ &+ \left[\int_{i\in[0,1]}(1+D^*(i)-\tau_S^*(x))S_1^*(i)\,di\right] - \rho^*(x), \end{aligned} \tag{32}$$

for some $s^* \in [0,1]$ and B_1^*, with $S_1^* = s^*K_1^*$. This tax system implies that (29) is satisfied as an equality at any $s^* \in [0,1)$ for distribution policy $d^*(x)$.

Proposition 12. *The tax system* $T^*(K_1, B_1, \{S_1\}_i, x) = \tau_P^*(x)(1+x)K_1 + \tau_B^*(x)B_1 + \tau_s^*(x)\int_i S_1(i)\,di + \rho^*(x)$, *where* $\tau_P^*(x)$, $\tau_B^*(x)$, $\tau_S^*(x)$ *and* $\rho^*(x)$ *satisfy (30), (19), (31) and (32), respectively, implements the constrained-efficient allocation at interest rate* $r = E_1x$ *with distribution policy* $1 + d^*(x) = (1+x)(1-\tau_P^*(x))$ *and dividend process* $D^*(i)$ *for all* i. *The allocation* $\left\{K_1^*, s^*, 0, \{S_1^*(i)\}_i, 1, c_1^*(x)\right\}$ *with* $s^*K_1^* = S_1^*(i)$ *for all* i *and* $s^* \in [0,1)$, *the tax system* $T^*(K_1, B_1, \{S_1\}_i, x)$, *the interest rate* $r = E_1x$, *the distribution policy* $d^*(x)$ *and the dividend process* $D^*(i,x)$ *constitute a competitive equilibrium.*

The proof is in the appendix and proceeds as the one for proposition 6. The setting of marginal taxes ensures that the entrepreneurs' first order necessary conditions and local second order necessary conditions for an interior solution to their problem are satisfied. In addition, it ensures that the allocation is globally optimal because it rules out any corner solutions to the entrepreneurs' investment and portfolio problems, irrespectively of the level of effort. Lastly, the setting of $\rho^*(x)$ ensures high effort is optimal at the appropriate level of capital and portfolio choices.

The properties of the optimal tax system can be derived from (30)-(32). First:

$$E_1\tau_P^*(x) = 1 - E_1\left[\frac{u'(c_0^*)}{\beta(1+x)u'(c_1^*(x))}\right],$$

so that $E_1\tau_P^*(x) > 0$ if $IW_K > 0$ and $E_1\tau_P^*(x) < 0$ if $IW_K < 0$. In addition, $\tau_P^*(\bar{x}) - \tau_P^*(\underline{x}) < 0$ when $IW_K > 0$ and $\tau_P^*(\bar{x}) - \tau_P^*(\underline{x}) > 0$ when $\mathrm{IW}_K < 0$ from:

$$\frac{u'(c_0^*)}{\beta(1+\underline{x})\,u'(c_1^*(\underline{x}))} - \frac{u'(c_0^*)}{\beta(1+\bar{x})\,u'(c_1^*(\bar{x}))} = \tau_P^*(\bar{x}) - \tau_P^*(\underline{x}),$$

since $\mathrm{IW}_K\tilde{\ }(1+\underline{x})\,u'(c_1^*(\underline{x})) - (1+\bar{x})\,u'(c_1^*(\bar{x}))$. Lastly, by (31):

$$1 + E_1x - E_1\tau_P^*(x) - E_1x\tau_P^*(x) - E_1\tau_S^*(x) = E_1\frac{u'(c_0^*)}{\beta u'(c_1^*(x))}.$$

This implies $E_1\tau_S^*(x) = -E_1\tau_P^*(x) - E_1x\tau_P^*(x) = -E_1\tau_P^*(x)\,E_1(1+x) - Cov_1(x, \tau_P^*(x))$. If $IW_K \gtrless 0$, $Cov_1(x, \tau_P^*(x)) \lesseqgtr 0$ and $E_1\tau_P^*(x) \gtrless 0$, as discussed above. Hence, the sign of $E_1\tau_S^*(x)$ is typically ambiguous.

The optimal tax system described in proposition 12 embeds a prescription for *optimal double taxation of income from entrepreneurial capital*: at the firm level thought τ_P^*, and at the investor level, through τ_S^*. It is *necessary* that the tax on earnings be "passed on" to stock investors via the distribution policy and that dividend distributions be taxed at the level of the investors- hence, that business income be subject to double taxation- in any equilibrium with $s^* \in (0,1)$. To see this, first note that taxation of stock income is required to ensure that agents choose S_1^*, for the usual reasons. Further, proposition 11 implies that it is necessary for expected discounted dividends to be equal to expected discounted after tax earnings on externally owned capital for an interior value of s to be optimal for an entrepreneur. A corollary of proposition 11 is that in any competitive equilibrium where the marginal tax on earnings satisfies (30), $s \in (0,1)$ requires that the distribution policy satisfies $1 + d(x) = (1+x)(1-\tau_P(x))$. In particular, for the tax system that implements the constrained-efficient allocation with $s^* \in (0,1)$, $1 + d^*(x) = (1+x)(1-\tau_P^*(x))$. This implies that the taxes paid on retained earnings at the firm level are passed on to investors, by a corresponding reduction in distributed earnings. Hence, distributed earnings are subject to double taxation.

If $1 + d^*(x) < (1+x)(1-\tau_P^*(x))$, an entrepreneur will find it optimal to set $s = 1$. But at $s = 1$, the constrained-efficient allocation cannot be implemented if entails high effort on part of entrepreneurs, since an entrepreneur's earnings do not depend on effort. If $1 + d^*(x) > (1+x)(1-\tau_P^*(x))$, then $s^* = 0$. In this case, the constrained-efficient allocation is implemented by the tax system and the competitive equilibrium allocation coincides with that of the market structure with risk-free bonds.

The distribution policy $1+\tilde{d}(x) = 1+x$, which corresponds to no taxation at the firm level for distributed earnings, is of particular interest. Under this distribution policy, (26) can be written as:

$$-\left[u'(c_0) - \beta E_1(1+x)\,u'(c_1(x))\right](1-s) - \beta E_1\left[(1+x)\,\tau_P(x)\right]u'(c_1(x)) = 0.$$

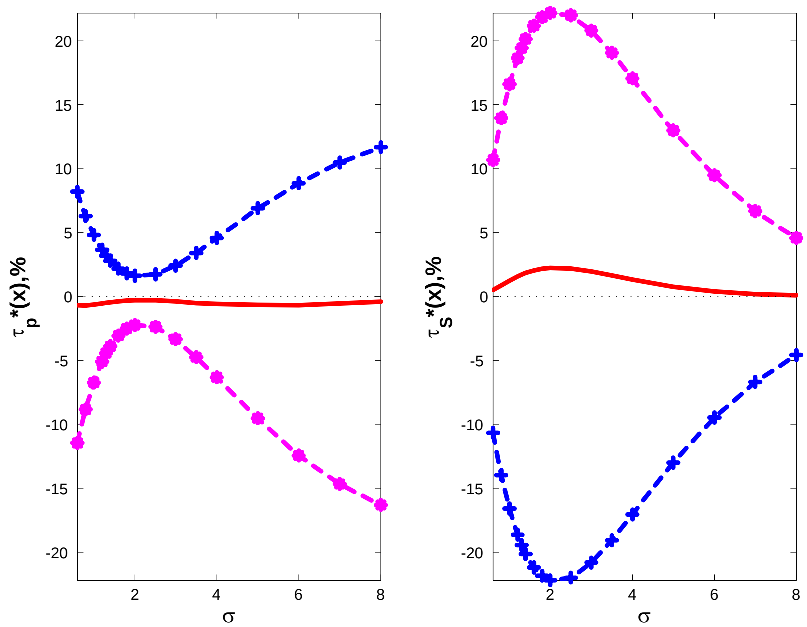


Figure 6: Optimal marginal earnings taxes, benchmark parameterization.

Evaluating this expression at the constrained-efficient allocation for $\tau_P^*(x)$ that solves (30), it follows that $u'(c_0^*) - \beta E_1 (1+x) u'(c_1^*(x)) < 0$, so that $s = 0$ from (29), if the intertemporal wedge is positive. If the intertemporal wedge is negative, $u'(c_0^*) - \beta E_1 (1+x) u'(c_1^*(x)) > 0$, so that $s = 1$.

The optimal tax system does not pin down the equilibrium value of s^* and the equilibrium portfolio composition for this economy. The restriction $s^* \in [0, 1)$ implies that for any value of s^* in this range, the tax system ensures that entrepreneurs find it optimal to choose K_1^*.

For the benchmark parameterization, the optimal marginal earnings taxes are:

Remark 13. *The constrained-efficient allocation can equivalently be implemented with a marginal tax on tax on capital $\tau_K(x)$ that satisfies (18) and with distribution policy: $1 + d(x) = 1 + x - \tau_K(x)$ and dividend process $1 + D(i, \tilde{x}) = 1 + \tilde{x} - \tau_K^*(\tilde{x})$, so that $1 + D(i) = 1 + E_1(x)$, since $E_1 \tau_K^*(x) = 0$. The optimality of double taxation of entrepreneurial earnings can be derived with a similar reasoning.*

3.3.1. Private Equity with Incentive Compatible Dividends

TBA

4. Concluding Remarks

The analysis of optimal taxation with entrepreneurial capital conducted in this paper has three main predictions. The first is that the intertemporal wedge on entrepreneurial capital can be positive or negative. It is negative when risk aversion is greater than intertemporal substitution with Kreps-Portheus preferences and when the coefficient of relative risk aversion is either smaller than one or sufficiently high with expected utility. A negative intertemporal wedge signals that more capital has a positive effect on incentives. This can occur since the returns from effort are increasing in capital. When the intertemporal wedge is negative the optimal marginal tax on capital rises with earnings. The second prediction is that the marginal tax on financial assets depends on the *correlation* of their returns with idiosyncratic uncertainty. If the correlation between an asset's returns and capital is less than one, the intertemporal wedge on the asset is greater than the intertemporal wedge on capital and the marginal tax on the asset is higher than the marginal tax on capital in bad states. The third prediction is that income from entrepreneurial capital should be taxed at the firm level and again when it accrues to outside investors in the form of stock returns. Hence, double taxation of capital income is optimal.

The contribution of this analysis is twofold. First, we characterize the properties of constrained-efficient allocations in private information economies with idiosyncratic capital returns. This class of environments has not been studied in the recursive contracting literature[10]. Second, we derive the properties of optimal taxes on entrepreneurial capital as well as on other financial assets. We also consider whether entrepreneurial capital earnings distributed to outside investors should be taxed at the firm level. This generates a theory of optimal differential asset taxation and provides a foundation for the double taxation of capital earnings.

The empirical public finance literature has documented substantial differences in the tax treatment of different forms of capital income. Specifically, interest income is taxed at a higher rate than stock returns, as discussed in Gordon (2003), while dividends are taxed at a higher rate than accrued capital gains. Personal and corporate tax rates on capital income are also different. As documented by Gordon and Slemrod (1988), the higher marginal tax rate on interest income is a stable property of empirical tax systems in many industrialized economies. Poterba (2002) has documented a strong response of household portfolio composition to this differential tax treatment. Auerbach (2002) finds that firms's investment decisions appear to be sensitive to the taxation of dividend income at the personal level and their choice of organization form is responsive to the differential between corporate and personal tax rates.

[10] An exception is Kahn and Ravikumar (1999). They focus on an implementation with financial intermediaries and rely on numerical simulations. They do no provide an analytical characterization of the wedges associated with the constrained-efficient allocation.

In the economy studied in this paper, the optimal tax system implements the constrained-efficient allocation by influencing portfolio choice and sales of private equity by entrepreneurs. Differential tax treatment of different asset classes is essential to achieve this goal. The same logic applies to company stock held by top executives. A quantitative version of the model can be used to provide an assessment of empirical tax systems. We leave these extensions for future work.

5. Appendix

Proof of Proposition 1. Letting μ be the multiplier on the incentive compatibility constraint and λ the one on the resource constraint, the first order necessary conditions for the planning problem at $e = 1$ are:

$$-u'(K_0 - K_1) + \lambda E_1 (1 + x) = 0,$$

$$(1 - \pi(1)) \beta u'(c_1(\underline{x})) - \mu(\pi(1) - \pi(0)) \beta u'(c_1(\underline{x})) - \lambda(1 - \pi(1)) = 0,$$

$$\pi(1) \beta u'(c_1(\bar{x})) - \mu(\pi(0) - \pi(1)) \beta u'(c_1(\bar{x})) - \lambda \pi(1) = 0.$$

At $e = 0$, the same first order necessary conditions hold with $\mu = 0$. If $e^* = 1$ is optimal, the first order conditions can be simplified to yield (4) and (5). ■

5.1. The Intertemporal Wedge with Recursive Preferences

Assuming high effort is implemented, the first order conditions are:

$$-c_0^{-\sigma} U_0(1)^{\sigma} + \lambda E_1(1 + x) + \mu c_0^{-\sigma} [U_0(0)^{\sigma} - U_0(1)^{\sigma}] = 0, \tag{33}$$

$$\begin{aligned} 0 &= \pi(1) \beta c_1(\bar{x})^{-\alpha} U_1(1)^{\alpha-\sigma} U_0(1)^{\sigma} - \lambda \pi(1) \\ &\quad -\mu \beta c_1(\bar{x})^{-\alpha} \left[\pi(0) U_1(0)^{\alpha-\sigma} U_0(0)^{\sigma} - \pi(1) U_1(1)^{\alpha-\sigma} U_0(1)^{\sigma}\right], \end{aligned} \tag{34}$$

$$\begin{aligned} 0 &= (1 - \pi(1)) \beta c_1(\underline{x})^{-\alpha} U_1(1)^{\alpha-\sigma} U_0(1)^{\sigma} - \lambda(1 - \pi(1)) \\ &\quad -\mu \beta c_1^{-\alpha}(\underline{x}) \left[(1 - \pi(0)) U_1(0)^{\alpha-\sigma} U_0(0)^{\sigma} - (1 - \pi(1)) U_1(1)^{\alpha-\sigma} U_0(1)^{\sigma}\right]. \end{aligned} \tag{35}$$

Combining (34) and (35) yields (12). Multiplying (34) and (35) by $(1 + \bar{x})$ and $(1+\underline{x})$ respectively, and combining with (33):

$$\begin{aligned} IW_K &= \beta E_1(1 + x) c_1(x)^{-\alpha} U_1(1)^{\alpha-\sigma} - c_0^{-\sigma} \\ &= -\mu \bar{A} \left\{\beta(1 + \underline{x}) c_1^{-\alpha}(\underline{x}) U_1(1)^{\alpha-\sigma} - c_0^{-\sigma} \frac{A}{\bar{A}}\right\} \\ &\quad -\bar{A} \mu \beta \pi(1) U_1(1)^{\alpha-\sigma} \left\{(1 + \bar{x}) c_1(\bar{x})^{-\alpha} \frac{A}{\bar{\bar{A}}} - (1 + \underline{x}) c_1^{-\alpha}(\underline{x})\right\}. \end{aligned} \tag{36}$$

Combining (33) and (35) obtains:

$$E_1 (1+x) \beta c_1 (\underline{x})^{-\alpha} = \frac{[1+\mu A]}{\left\{1+\mu \bar{A}\right\}} c_0^{-\sigma}.$$

Substituting this expression in (36) obtains (13).

Proof of Proposition 2. Using (12) to substitute for $\frac{c_1^{-\alpha}(\bar{x})}{c_1^{-\alpha}(\underline{x})}$ in (13):

$$\begin{aligned} IW_K &= \mu \bar{A} \beta c_1 (\underline{x})^{-\alpha} U_1 (1)^{\alpha-\sigma} \left\{ (1+E_1(x)) \frac{\bar{A}^{-1}+\mu}{A^{-1}+\mu} - (1+\underline{x}) \right\} \\ & \quad - \bar{A} \mu \beta \pi (1) U_1 (1)^{\alpha-\sigma} c_1^{-\alpha} (\underline{x}) \left\{ -(1+\underline{x}) + (1+\bar{x}) \left(\frac{\bar{A}^{-1}+\mu}{\underline{A}^{-1}+\mu} \right) \right\} \end{aligned}$$

Manipulating this expression obtains:

$$IW_K = \mu \frac{\bar{A}}{\underline{A}^{-1}+\mu} \beta c_1 (\underline{x})^{-\alpha} U_1 (1)^{\alpha-\sigma} \left\{ \begin{array}{c} (1+\underline{x}) (1-\pi(1)) \frac{\underline{A}^{-1}}{A^{-1}+\mu} \left(\bar{A}^{-1} - \left[1 - \frac{U_0(0)^\sigma}{U_0(1)^\sigma} \right]^{-1} \right) \\ \pi (1) (1+\bar{x}) \frac{\bar{A}^{-1}+\mu}{A^{-1}+\mu} \left(\underline{A}^{-1} - \left[1 - \frac{U_0(0)^\sigma}{U_0(1)^\sigma} \right]^{-1} \right) \end{array} \right\}.$$

Note that $\bar{A}^{-1} - A^{-1} < 0$ if $\frac{(1-\pi(0))U_1(0)^{\alpha-\sigma}}{(1-\pi(1))U_1(1)^{\alpha-\sigma}} < 1$, which is satisfied at $\alpha > \sigma$ since $1 < \frac{(1-\pi(0))}{(1-\pi(1))} < \frac{U_1(1)^{\alpha-\sigma}}{U_1(0)^{\alpha-\sigma}}$. In addition, $\underline{A}^{-1} - A^{-1} < 0$ if $\frac{\pi(0)U_1(0)^{\alpha-\sigma}}{\pi(1)U_1(1)^{\alpha-\sigma}} < 1$ or $\frac{\pi(0)}{\pi(1)} < \frac{U_1(1)^{\alpha-\sigma}}{U_1(0)^{\alpha-\sigma}}$. This is satisfied for $\alpha > \sigma$, since $\frac{U_1(1)}{U_1(0)} > 1$ and $\frac{\pi(0)}{\pi(1)} < 1$. ■

5.2. Optimal Taxes

Proof of Proposition 12. TBA. ■

References

[1] Albanesi, Stefania and Christopher Sleet. 2005. "Dynamic Optimal Taxation with Private Information", forthcoming, *The Review of Economic Studies.*

[2] Angeletos, George-Marios. 2005. "Uninsured Idiosyncratic Investment Risk and Aggregate Saving", manuscript, MIT.

[3] Attanasio, Orazio and Annette Vissing-Jorgensen, 2003. "Stock-Market Participation, Intertemporal Substitution, and Risk-Aversion", *American Economic Review,* Vol. 93, No. 2.

[4] Attanasio, Orazio and Guglielmo Weber, 1989. "Intertemporal Substitution, Risk-Aversion and the Intertemporal Euler Equation for Consumption", Economic Journal, 99(395).

[5] Auerbach, Alan, 2002. "Taxation and Corporate Financial Policy", *Handbook of Public Economics*, Vol. 3, North-Holland.

[6] Backus, David K., Bryan R. Routledge, and Stanley E. Zin. 2005. "Exotic Preferences for Macroeconomists", NBER Macroeconomics Annual.

[7] Bitler, Marianne Tobias Moscowitz, and Annette Vissing-Jorgensen. 2005. "Testing Agency Theory with Entrepreneur Effort and Wealth", forthcoming, *Journal of Finance*.

[8] Bizer, David and Peter De Marzo. 1999. "Optimal Incentive Contracts when Agents Can Save, Borrow and Default", *Journal of Financial Intermediation* 8, 241-260.

[9] Cagetti, Marco and Mariacristina De Nardi. 2004. "Taxation, Entrepreneurship, and Wealth", Federal Reserve Bank of Minneapolis, Staff Report 340.

[10] Campbell, John. 1996. "Understanding Risk and Return", *Journal of Political Economy*, 104(2), pp. 298-345.

[11] Diamond, Peter and James Mirrlees. 1978. A model of social insurance with variable retirement. *Journal of Public Economics* 10:295-336.

[12] Farhi, Emanuel and Ivan Werning. 2005. "Inequality, Social Discounting and Progressive Estate Taxation", NBER WP.

[13] Golosov, Mikhail Narayana Kocherlakota, Aleh Tsyvinski. 2003. "Optimal Indirect and Capital Taxation", *The Review of Economic Studies*, Volume 70, Issue 3, Page 569.

[14] Grochulski, Borys and Tomasz Piskorski. 2005. "Optimal Wealth Taxes with Risky Human Capital", manuscript, NYU.

[15] Hassett, Kevin and Glenn Hubbard, 2002. "Tax Policy and Business Investment", *Handbook of Public Economics*, Vol. 3, North-Holland.

[16] Gordon, Roger, 2003. "Taxation of Interest Income", manuscript, UCSD.

[17] Gordon, Roger and Joel Slemrod. 1988. “Do We Collect Any Revenue from Taxing Capital Income?” In Tax Policy and the Economy, Vol. 2, ed. Lawrence Summers. Cambridge: MIT Press, pp. 89-130.

[18] Judd, Kenneth. 1985. "Redistributive Taxation in a Perfect Foresight Model," *Journal of Public Economics* 28, 59-84.

[19] Khan, Aubhik and B. Ravikumar. 1999. "Growth and Risk-Sharing with Private Information", *Journal of Monetary Economics*.

[20] Kocherlakota, Narayana. 2005a. "Advances in Dynamic Optimal Taxation", paper presented at the World Congress of the Econometric Society, August 2005.

[21] Kocherlakota, Narayana. 2005b. "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation, " forthcoming, *Econometrica*.

[22] Levhari, D. and T. N. Srinivasan. 1969. "Optimal Savings Under Uncertainty", *The Review of Economic Studies* Vol. 36, No. 2, 153-163.

[23] Mirrlees, James. 1971. An exploration in the theory of optimum income taxation. *The Review of Economic Studies* 38:175-208.

[24] Moskovitz, Tobias and Annette Vissing-Jorgensen. 2002. "The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle?", *The American Economic Review*, Vol. 92, No. 4.

[25] Poterba, James. 2002. "Taxation, Risk-Taking, and Household Portfolio Behavior", *Handbook of Public Economics*, Vol. 3, North-Holland.

[26] Rogerson, William. 1985. Repeated moral hazard. *Econometrica* 53:69-76.

[27] Sandmo, Agnar. 1970. "The Effect of Uncertainty on Savings Decisions", *The Review of Economic Studies*, Vol. 37, No. 3, 353-360.

[28] Weil, Philippe, 1993, “Precautionary savings and the permanent income hypothesis,” *The Review of Economic Studies* 60, 367-383.