Optimal capital taxation for time-nonseparable preferences

Koehne, Sebastian and Kuhn, Moritz

18 March 2013

Optimal capital taxation for time-nonseparable preferences

Sebastian Koehne* Moritz Kuhn†

March 18, 2013

Abstract

Fundamental to any theory of capital taxation is a description of individual savings behavior. Successful descriptions of savings behavior have often resorted to the habit formation hypothesis. This paper studies the effect of habit formation on optimal capital taxes in a dynamic Mirrleesian model. We make three distinct contributions. First, we decompose intertemporal wedges (implicit capital taxes) for general time-nonseparable preferences into a wealth effect, a complementarity effect, and a future incentive effect. Second, we provide conditions under which intertemporal wedges are positive. Third, we derive a recursive formulation of constrained efficient allocations and evaluate the quantitative impact of habit formation. In a model parameterized to the U.S. economy, habit formation reduces average intertemporal wedges by about 40 percent compared to the time-separable case. Moreover, intertemporal wedges are close to zero for the largest part of the working life.

Keywords: optimal taxation; intertemporal wedge; habit formation; recursive contracts; new dynamic public finance.

JEL Classification: D82, E21, H21

*Corresponding author at: Institute for International Economic Studies (IIES), Stockholm University, SE-10691 Stockholm, Sweden, Phone: +46 8 16 35 64, sebastian.koehne@iies.su.se

†University of Bonn, Department of Economics, D-53113 Bonn, Germany, Phone: +49 228 73 62096, mokuhn@uni-bonn.de

1 Introduction

Fundamental to any theory of capital taxation is a description of individual savings behavior. Successful descriptions of savings behavior have often resorted to the habit formation hypothesis. This hypothesis states that consumption is complementary across time—consuming a lot today makes individuals hungrier for consumption tomorrow. Habit formation has reconciled theory and evidence for several important aspects of savings behavior, such as the equity premium puzzle, the relationship between savings and growth, wealth inequality, or reactions to monetary policy shocks.[1] Despite the success of habit formation in explaining savings decisions under uncertainty, the implications for the optimal taxation of capital are unknown. The present paper aims to fill this gap. We motivate capital taxation in the tradition of Diamond and Mirrlees (1978) and the recent literature on 'New Dynamic Public Finance' (surveyed by Kocherlakota, 2010). Private information, in combination with uncertainty, implies that the accumulation of assets today affects the incentive to supply labor in the future. Capital taxes serve to correct this externality.

Our paper makes three contributions. First, we provide a decomposition of intertemporal wedges (implicit capital taxes) for general time-nonseparable preferences into a standard *wealth effect*, a *complementarity effect*, and a *future incentive effect* (Proposition 1). Second, we derive theoretical conditions under which intertemporal wedges are positive (Proposition 2). Finally, we extend the recursive contracting approach to habit formation economies and evaluate the quantitative importance of habit formation for intertemporal wedges. For a model parameterized to the U.S. economy, average capital tax rates decrease by about 40 percent compared to the case of time-separable preferences. Moreover, capital tax rates are close to zero except at the very end of the working life.

Our model is a standard dynamic Mirrlees model of optimal taxation generalized to the case of time-nonseparable preferences. Agents face dynamic shocks to their abilities to generate labor income. Labor income is publicly observed, but abilities and labor supply are private information. The only restrictions imposed on life-time consumption utility are monotonicity and concavity. In addition to habit formation, this setup allows for alternative forms of time-

[1] Abel (1990), Constantinides (1990), and Campbell and Cochrane (1999) study asset pricing under habit formation, Ryder and Heal (1973) and Carroll, Overland, and Weil (2000) explore habit formation in a growth model, Diaz, Pijoan-Mas, and Rios-Rull (2003) analyze precautionary savings and wealth inequality, Fuhrer (2000) studies reactions to monetary policy. Messinis (1999) provides a review of the literature.

nonseparability, including models where the consumption good is durable. In this environment, we characterize the solution of the social planning problem in terms of intertemporal wedges. As common in this literature, positive intertemporal wedges represent *implicit* taxes on capital and indicate that decentralizations of the social planning allocation must correct individual capital returns downward in one way or another.[2]

By decomposing intertemporal wedges, we show that optimal capital taxes for habit formation preferences are determined by three forces. First of all, saving should be taxed because the agent has a better incentive to supply labor in the next period if he starts the next period with lower wealth *(wealth effect)*. This force is well-known from the standard time-separable model. Second, saving should be taxed, because enhancing present consumption due to habit formation makes high consumption in the next period more attractive, which reinforces next period's labor supply incentives *(complementarity effect)*. Finally, saving should be subsidized, because boosting next period's consumption due to habit formation improves labor supply incentives in the remaining periods *(future incentive effect)*. Habit formation thus changes the capital taxation motive in countervailing ways, and its impact will depend on the relative magnitude of the last two components. If consumption is substitutable rather than complementary across time (as in models with durable consumption), the signs of the complementarity effect and the future incentive effect are reversed. In both cases, the impact of time-nonseparability on optimal capital taxes is theoretically ambiguous.

We then study how far the perturbation approach of Rogerson (1985) can be extended to identify the sign of intertemporal wedges for time-nonseparable preferences. We show that intertemporal wedges are positive at the very end of the working life for very general models of time-nonseparable preferences, including standard formulations of habit formation and durable consumption. Technically, our finding replaces the well-known Jensen's inequality argument by a more general result on the positivity of the covariance of two monotonic functions. For earlier periods, the perturbation approach becomes infeasible, however.[3]

The quantitative part of the paper evaluates the impact of habit formation on intertemporal wedges in a model parameterized to the U.S. economy. To the best of our knowledge, this is

[2]The decentralization of optimal allocations is not unique; see Golosov, Kocherlakota, and Tsyvinski (2003), Kocherlakota (2005), Albanesi and Sleet (2006), Golosov and Tsyvinski (2006), Werning (2011), Gottardi and Pavoni (2011), Abraham, Koehne, and Pavoni (2012), among others.

[3]For the same reason, the Inverse Euler equation does not generalize to the time-nonseparable case; see Grochulski and Kocherlakota (2010).

the first quantitative analysis of optimal dynamic taxes in a habit formation environment. As a methodological contribution, we extend the recursive contracting approach and show that standard formulations of habit formation can be dealt with by adding one additional state variable to the planner's recursive problem. We find that habit formation has a significantly negative effect on intertemporal wedges. Compared to the time-separable case, average intertemporal wedges drop by about 40 percent. The effect is even stronger if we exclude the very end of the working life. For more than the first three quarters of the working life, intertemporal wedges are virtually zero.

Habit formation reduces intertemporal wedges, because the negative future incentive effect more than outweighs the positive complementarity effect. Intuitively, habit formation creates a link between present consumption-saving decisions and future incentive problems. Encouraging consumption in any period increases future habit levels and thereby raises the agent's marginal utility in the future. This has a positive effect on labor supply incentives. The negative future incentive effect arises because future habit levels respond more strongly to next period's consumption than to present consumption. This creates a motive for subsidizing saving in order to increase next period's consumption and thereby relax future incentive problems. Since incentive problems aggravate over time (incentives rely more on costly immediate consumption rewards and less on future promises), the future incentive effect creates a strong force for saving subsidies with habit formation preferences. This force is nearly as big as the motive to tax saving in order to relax the incentive problem in the immediately following period. As a consequence, intertemporal wedges are close to zero over much of the working life. They become positive only close to retirement when the future incentive effect vanishes.

The paper proceeds as follows. The next section discusses the related literature. Section 3 sets up our general theoretical model. Section 4 decomposes intertemporal wedges into three parts. By means of two stylized examples, we show that nonpositive intertemporal wedges can emerge both from intertemporal complementarity as well as from intertemporal substitutability of consumption. We then provide sufficient conditions for intertemporal wedges to be positive. Section 5 studies the quantitative importance of habit formation for intertemporal wedges, while Section 6 concludes.

2 Related literature

With very few exceptions, most existing studies of dynamic taxation problems work with time-separable preferences for simplicity and tractability. The contribution closest to ours is by Grochulski and Kocherlakota (2010), who explore a taxation framework with general time-nonseparable preferences similar to the present paper. They show that social security systems (with history-dependent taxes and transfers upon retirement) can be used to decentralize optimal allocations when preferences are time-nonseparable. Regarding intertemporal wedges, they construct an insightful example of a 3-period habit formation model with private information only in the final period and show that the intertemporal wedge in the initial period is negative. Our decomposition shows that the future incentive effect is responsible for this result. However, we also reveal that incentive problems in the immediate future create countervailing forces due to the wealth and the complementarity effect. Our quantitative analysis therefore finds that, even though it is possible to construct theoretical cases where wedges are negative, those cases are not representative of typical taxation environments.

Farhi and Werning (2008) analyze optimal savings distortions for a class of time-nonseparable and state-nonseparable recursive preferences. There are at least three main differences between their contribution and the present one. First, different from standard models of habit formation, the preferences explored by Farhi and Werning (2008) do not have an expected utility representation. Second, proportional variations of consumption do not affect the incentives to supply labor in their model. This feature allows disentangling the roles of risk aversion and intertemporal substitution, but it abstracts from a number of effects that determine optimal allocations for alternative cases of time-nonseparability. In particular, forces calling for negative intertemporal wedges are absent in their environment. Finally, our quantitative exercise studies constrained efficient allocations, while Farhi and Werning (2008) explore partial and general equilibrium effects of a reform that distorts the consumption profile, but not labor supply.

The present paper is also related to the dynamic contracting literature on effort persistence, which studies private information problems with a *production technology* that is time-nonseparable; see Mukoyama and Sahin (2005), Kwon (2006), Jarque (2010), and Hopenhayn and Jarque (2010). In contrast to the present model, the Inverse Euler equation remains valid in that framework. Hence, time-nonseparable technologies and time-nonseparable preferences change optimal allocations in fundamentally different ways.

Finally, the paper builds on the extensive literature on time-nonseparable preferences. This literature has evolved around two main concepts. First of all, there is the hypothesis of *habit formation*. This concept goes back to the theory of adaptation formalized in the psychological literature by Helson (1964). Habit formation postulates that individuals compare their current consumption to a historical reference level, and derive utility both from consumption per se and from consumption growth. Frederick and Loewenstein (1999) review the substantial body of empirical research supporting this hypothesis. For instance, workers' self-reported well-being is often closely related to recent changes in pay, but not so much to absolute levels of pay (Clark, 1999). Ravina (2007) finds strong support for habit formation based on micro level consumption data. For a review of habit formation in the macroeconomic literature see Messinis (1999).

Complementary to the habit formation literature, a second line of research focuses on short-run substitution effects typically referred to as *local substitution*. Using high frequency aggregate data on consumption and asset prices, Dunn and Singleton (1986), Eichenbaum and Hansen (1990), and Heaton (1993, 1995) find evidence that consumption is substitutable over short periods of time (weeks, months, quarters). These findings can be micro founded by assuming that consumption goods are partly durable and subject to adjustment costs. The theoretical framework in the present paper is flexible enough to allow for both habit formation and local substitution effects.

3 A dynamic taxation model with time-nonseparable preferences

This section sets up a dynamic Mirrleesian model of optimal taxation with time-nonseparable preferences. The notation largely follows Grochulski and Kocherlakota (2010).

The economy consists of a risk-neutral principal/planner and a unit measure of risk-averse agents facing dynamic skill shocks. The planner observes the output of each agent in every period, but does not observe hours (labor input) and skill levels. Time is discrete and indexed by $t = 1, 2, \ldots, T$, with $T \in \mathbb{N} \cup \{\infty\}$.

3.1 Preferences

To focus on time-nonseparability, we consider preferences that are additively separable between consumption and labor.[4] The agents have identical von-Neumann-Morgenstern preferences and maximize the expected value of

$$U(c_1, \ldots, c_T) - V(l_1, \ldots, l_T) \tag{1}$$

where $c_t \in \mathbb{R}_+$ denotes consumption and $l_t \in \mathbb{R}_+$ represents hours worked at date t, $t = 1, \ldots, T$. Labor disutility $V(l_1, \ldots, l_T)$ is increasing in each argument and weakly convex. Consumption utility $U(c_1, \ldots, c_T)$ is twice continuously differentiable, increasing in each argument and concave. As usual, we use subscripts to denote partial derivatives.

Notice that this setup allows for consumption complementarities across time, $U_{c_t,c_s} > 0$ for $t \neq s$ (as in models of habit formation), as well as consumption substitutabilities $U_{c_t,c_s} < 0$ for $t \neq s$ (as in models of durable consumption), and combinations thereof. The setup also includes the time-separable case, $U_{c_t,c_s} = 0$ for $t \neq s$, of course.[5]

3.2 Skills

Agents differ with respect to their skills. An agent with hours l_t and skill θ_t generates $y_t = \theta_t l_t$ units of output in period t. Output is publicly observable, but hours and skill are private information.

At time zero, a skill path $\theta^T = (\theta_1, \ldots, \theta_T)$ is drawn for each agent. Paths are drawn independently across agents according to a probability measure μ on the set $\Theta^T := \Theta \times \cdots \times \Theta$, where Θ is a finite subset of the positive reals. Denote the expectation operator with respect to μ by $\mathbb{E}[\,\cdot\,]$. At the beginning of period t, every agent learns his current skill θ_t. The information of a given agent in period t is thus his individual history $\theta^t = (\theta_1, \ldots, \theta_t)$. As usual, the notation $\mathbb{E}_t[\,\cdot\,] := \mathbb{E}[\,\cdot\,|\theta^t]$ represents expectations conditional on that history. Similarly, conditional covariances are denoted by $\mathrm{cov}_t(\,\cdot\,,\,\cdot\,)$.

[4] Nonseparable preferences between consumption and labor are explored by Golosov, Tsyvinski, and Werning (2006).

[5] Besides, the model allows for time-nonseparabilities in the preferences over hours, but this will be irrelevant for the questions addressed in this paper. As usual in this literature, our results emerge from the analysis of the consumption allocation that optimally implements a given labor plan. For this question, time-nonseparabilities in V play no role. The details of V become important, however, if one wants to understand the dynamics of optimal labor plans.

3.3 Technology

The planner has access to a linear technology for intertemporal transfers which allows to transform $x \in \mathbb{R}$ units of date-t output into $R_t x$ units of output at date $t+1$. The gross rate of return is deterministic and nonnegative: $R_t > 0$ for all t. It would not be difficult to endogenize the rate of return by introducing an explicit production function that depends on capital and labor. Yet, this would only complicate the notation and generate no new insights. We therefore follow Grochulski and Kocherlakota (2010) and let the rate of return be exogenous.

3.4 Allocations

An *allocation* is a sequence $(\mathbf{c}, \mathbf{y}) = (c_t, y_t)_{t=1,\ldots,T}$ of consumption plans $c_t : \Theta^T \to \mathbb{R}_+$ and output plans $y_t : \Theta^T \to \mathbb{R}_+$ such that, for any period t, c_t and y_t are functions of period-t information. That is, c_t and y_t are θ^t-measurable.

At the beginning of every period, the planner assigns consumption and output to each agent according to the agent's skill report. A reporting strategy is a mapping $\sigma : \Theta^T \to \Theta^T$ such that the period-t component σ_t is θ^t-measurable for all t. Denote the truth-telling strategy by σ^*, with $\sigma^*(\theta^T) = \theta^T$ for all θ^T, and denote the set of all reporting strategies by Σ . Since skills are privately observed, the planner must ensure that all agents reveal their information truthfully. Hence, allocations must satisfy the incentive compatibility constraint

$$\begin{aligned} &\mathbb{E}\left[U(c_1(\sigma^*), \ldots, c_T(\sigma^*)) - V\left(\frac{y_1(\sigma^*)}{\theta_1}, \ldots, \frac{y_T(\sigma^*)}{\theta_T}\right)\right] \\ \geq\ &\mathbb{E}\left[U(c_1(\sigma), \ldots, c_T(\sigma)) - V\left(\frac{y_1(\sigma)}{\theta_1}, \ldots, \frac{y_T(\sigma)}{\theta_T}\right)\right]. \end{aligned} \tag{2}$$

An allocation that satisfies (2) for all reporting strategies $\sigma \in \Sigma$ is called *incentive compatible.*

3.5 Optimal allocations

The social planner seeks to provide a given level U_1 of ex-ante welfare at minimal costs. Hence, an allocation $(\mathbf{c}, \mathbf{y})$ is called *optimal* if it minimizes costs

$$\min_{\mathbf{c},\mathbf{y}} \mathbb{E}\left[\sum_{t=1}^{T} \frac{c_t - y_t}{R_1 \cdots R_{t-1}}\right] \tag{3}$$

subject to the constraints that $(\mathbf{c}, \mathbf{y})$ is incentive compatible and generates welfare U_1,

$$\mathbb{E}\left[U(c_1, \ldots, c_T) - V\left(\frac{y_1}{\theta_1}, \ldots, \frac{y_T}{\theta_T}\right)\right] = U_1. \tag{4}$$

4 Theoretical analysis of intertemporal wedges

As is well known in the dynamic public finance literature, the decentralization of optimal allocations is not unique. Hence, the robust insights from the present analysis are not about explicit tax instruments, but about optimal tax distortions or *wedges*. We show that intertemporal wedges (implicit capital taxes) for time-nonseparable preferences consist of three, not necessarily positive, components. We then provide conditions under which intertemporal wedges are positive.

It is helpful to formalize some concepts. As usual, the shadow rate of return (between periods t and $t+1$) of a given consumption allocation $\big(c_1(\theta^1), \ldots, c_T(\theta^T)\big)$ is defined as the interest rate at which the agent is indifferent between saving and not saving. Formally, the shadow rate of return is

$$\frac{\mathbb{E}_t[U_{c_t}(c_1, \ldots, c_T)]}{\mathbb{E}_t[U_{c_{t+1}}(c_1, \ldots, c_T)]}. \tag{5}$$

We are interested in the difference between the technological rate of intertemporal transformation R_t and the agent's shadow rate of return. It is convenient to write this difference in relative terms and define the *intertemporal wedge* between periods t and $t+1$ as

$$\tau_t^K := 1 - \frac{\mathbb{E}_t[U_{c_t}(c_1, \ldots, c_T)]}{R_t \mathbb{E}_t[U_{c_{t+1}}(c_1, \ldots, c_T)]}. \tag{6}$$

Note that τ_t^K is a random variable that depends on the date-t history θ^t as indicated by the conditional expectations operator $\mathbb{E}_t[\,\cdot\,]$. If the intertemporal wedge τ_t^K is positive, then the marginal rate of transformation R_t exceeds the individual shadow rate of return, and the allocation features an implicit tax on capital. If τ_t^K is negative, we have an implicit subsidy.

4.1 Decomposition of intertemporal wedges

The following result shows that intertemporal wedges for time-nonseparable preferences have three components. For the proof of Proposition 1 and all further proofs see Appendix A.

Proposition 1 (Decomposition)**.** *Let* $(\mathbf{c}, \mathbf{y})$ *be an optimal allocation. For any* $t \in \{1, \ldots, T-1\}$, *the intertemporal wedge equals* $\tau_t^K = \gamma_t (A_t + B_t + C_t)$, *where* γ_t *is positive and where*

$$A_t = -\mathrm{cov}_t \left(R_t U_{c_{t+1}} \, , \, \frac{1}{U_{c_{t+1}}} \right) \geq 0, \tag{7}$$

$$B_t = \mathrm{cov}_t \left(U_{c_t} \, , \, \frac{1}{U_{c_{t+1}}} \right) \gtreqless 0, \tag{8}$$

$$C_t = R_t - \mathbb{E}_t \left[\frac{U_{c_t}}{U_{c_{t+1}}} \right] \gtreqless 0. \tag{9}$$

Intuitively, an intertemporal wedge emerges whenever saving has social effects that are not internalized by the agent. The wedge captures the distortion to the agent's savings margin necessary to align it with the social savings margin. In the present model, the wedge arises because the agent does not internalize the impact of saving on the incentive problem.

Consider the following hypothetical situation: the agent, after working in period t and receiving the transfer $c_t(\theta^t)$, saves one unit of consumption at gross return R_t. Then three effects change the agent's preferences over future states, and thereby the incentive to supply labor (or, put differently, the incentive to report truthfully) in the future. Each effect is associated with one component identified in Proposition 1.

First of all, there is the standard *wealth effect*. Saving one consumption unit at time t yields R_t extra consumption units at time $t+1$, which raises the agent's utility by $R_t U_{c_{t+1}}$. This expression varies negatively with the realization of c_{t+1}, which means that states with low c_{t+1} become relatively more attractive, and thus the agent's incentive to supply labor in period $t+1$ is reduced. The term

$$A_t = -\mathrm{cov}_t \left(R_t U_{c_{t+1}} \, , \, \frac{1}{U_{c_{t+1}}} \right) \tag{10}$$

captures this effect. The first variable, $R_t U_{c_{t+1}}$, expresses the utility gain of R_t extra consumption units at $t+1$. The second variable, $1/U_{c_{t+1}}$, is a monotonic function of c_{t+1}, so the covariance is nonpositive and hence A_t is nonnegative. Moreover, it is easy to see that A_t is positive unless c_{t+1} is constant almost everywhere. Finally, note that the effect picked up by A_t is unrelated to potential time-nonseparabilites of preferences. The term A_t thus captures the component of the intertemporal wedge that is well-known from models with time-separable preferences.

The second component of the intertemporal wedge is the *complementarity effect*. Saving in period t reduces the agent's consumption in period t and thereby diminishes the agent's utility by U_{c_t}. This changes the relative preference over states at time $t+1$ depending on whether consumption is complementary or substitutable over time. When consumption is complementary over time (as in habit formation models), the cross-derivative of U with respect to c_t and c_{t+1} is positive, which implies that the term

$$B_t = \mathrm{cov}_t \left(U_{c_t} \, , \, \frac{1}{U_{c_{t+1}}} \right) \tag{11}$$

contributes positively to the intertemporal wedge. Intuitively, reducing consumption in period t increases the relative attractiveness of low consumption in period $t+1$. This goes in the same direction as the wealth effect and generates an additional motive for taxing savings. When consumption is substitutable across time (as in models with durable consumption), the previous argument is reversed and the term B_t becomes negative, because the substitutability of consumption leads to a beneficial effect of saving on the incentive problem; see Example 1 below.

Finally, the intertemporal wedge has the component

$$C_t = R_t \left(1 - \mathbb{E}_t \left[\frac{U_{c_t}}{R_t U_{c_{t+1}}} \right] \right). \tag{12}$$

This term is a residual that picks up all reasons for distorting the savings margin that are not covered by the two previous components. The formula itself is hard to interpret, but we note that the Inverse Euler equation (Rogerson, 1985) implies that C_t equals zero for time-separable preferences.[6] More generally, we have the following result.

Lemma 1. *Let* $(\mathbf{c}, \mathbf{y})$ *be an optimal allocation. Suppose that, for some* $t \in \{1, \ldots, T-1\}$, *marginal utilities* $U_{c_t}(c_1, \ldots, c_T)$ *and* $U_{c_{t+1}}(c_1, \ldots, c_T)$ *are* θ^{t+1}*-measurable. Then* $C_t = 0$ *almost everywhere.*

Lemma 1 implies that C_t equals zero if there is full insurance in periods $t+2, \ldots, T$ or if preferences in periods $t+2, \ldots, T$ do not depend on c_t and c_{t+1}. In particular, C_t equals zero in the penultimate period. When C_t differs from zero, it captures a distortion to the agent's

[6] If $U(c_1, \ldots, c_T) = \sum_t \beta^{t-1} u(c_t)$, we can write $C_t = u'(c_t) R_t \left(1/u'(c_t) - \mathbb{E}_t \left[1/(\beta R_t u'(c_{t+1}))\right]\right)$. The term in parentheses is equal to zero by the Inverse Euler equation from Rogerson (1985).

savings margin motivated by *future incentive effects*. Intuitively, by distorting the decision between c_t and c_{t+1}, the planner manipulates preferences in periods $t+2$ and later in order to relax the incentive problem at those dates.

It is difficult to determine the sign of the future incentive effect analytically, given that the existence of this effect invalidates the perturbation approach proposed by Rogerson (1985).[7] Yet, the idea underlying the effect is simple. When consumption is complementary between dates $t+1$ and $t+2$, then a high level of consumption at $t+1$ makes high consumption at $t+2$ relatively more attractive. This helps with the incentive problem at $t+2$ and generates a motive for subsidizing saving at t as in Example 2 below. Similar motives arise if consumption at date $t+1$ is complementary with consumption at other future dates.[8] The future incentive effect therefore provides a rationale for subsidizing saving in cases with consumption complementarities (as in habit formation models) and counteracts the positive complementarity effect. With consumption substitutability (as in models with durable consumption), the signs of both effects are reversed. In both cases, the overall effect arising from time-nonseparability remains ambiguous.

We now present two simple examples that highlight the roles of the complementarity effect and the future incentive effect, respectively. Moreover, we show that both effects can lead to nonpositive intertemporal wedges.

Example 1 (Intertemporal substitutability)**.** Let $T = 2$. In the first period, the agent's skill θ_1 is deterministic. In the second period, the skill is distributed in the set $\{\theta_L, \theta_H\}$, where θ_L and $\theta_H > \theta_L$ both have nonzero probability. Allocations take the form $(c_1, c_{2i}, y_1, y_{2i})_{i=L,H}$, where the index $i \in \{L, H\}$ refers to the agent's skill type in the second period. To make the problem nontrivial, we assume $y_{2H} > y_{2L}$.

Suppose that the principal can save and borrow at an interest rate of zero, i.e., $R = 1$. Moreover, suppose for a moment that the agent's preferences over consumption are *time-separable*, i.e., $U(c_1, c_2) = u(c_1) + u(c_2)$, where u is increasing and concave. Then, using the well-known result from Rogerson (1985), at any optimal allocation the agent remains with a residual motive

[7]For the same reason, the Inverse Euler equation does not generalize to the time-nonseparable case; see Grochulski and Kocherlakota (2010).

[8]Our simple reasoning abstracts from potential complementarities between date t and dates $t+2, \ldots, T$, but since the degree of complementarity typically diminishes over time, consumption at date $t+1$ tends to be more crucial for future incentives.

to save:

$$\frac{\mathbb{E}[U_{c_1}(c_1, c_{2i})]}{\mathbb{E}[U_{c_2}(c_1, c_{2i})]} = \frac{u'(c_1)}{\mathbb{E}[u'(c_{2i})]} < 1. \tag{13}$$

Now consider the following case of time-nonseparable preferences: $U(c_1, c_2) = u(c_1 + c_2)$, where u is increasing and concave. Since the cross derivative of U is negative, consumption is substitutable over time. It is straightforward that at any (not necessarily optimal) allocation, we have the following identity:

$$\frac{\mathbb{E}[U_{c_1}(c_1, c_{2i})]}{\mathbb{E}[U_{c_2}(c_1, c_{2i})]} = \frac{\mathbb{E}[u'(c_1 + c_{2i})]}{\mathbb{E}[u'(c_1 + c_{2i})]} = 1. \tag{14}$$

Hence, in contrast to the time-separable case, the agent has no residual motive to save/borrow at an optimal allocation, and the intertemporal wedge is zero: $\tau_1^K = A_1 + B_1 + C_1 = 0$. Note that the identity $U_{c_1} = U_{c_2}$ implies $C_1 = 0$. Hence, the future incentive effect is absent in this example. The intertemporal wedge consists of a positive wealth effect A_1 and a negative complementarity effect B_1, which exactly offset each other.

To see why the wealth effect and the complementarity effect are equally large in this example, note that a marginal increase in second period wealth increases the agent's utility by $U_{c_2}(c_1, c_2)$. A marginal reduction of first period consumption reduces the agent's utility by $U_{c_1}(c_1, c_2)$. The total effect of saving is thus $U_{c_2}(c_1, c_2) - U_{c_1}(c_1, c_2) = 0$, which is independent of c_2. As a consequence, distorting the agent's savings margin does not affect the incentive problem and the intertemporal wedge is zero.

Example 2 (Intertemporal complementarity / habit formation)**.** The following example is taken from Grochulski and Kocherlakota (2010).

Modify the setup from Example 1 by adding an initial period indexed with $t = 0$ and consider consumption preferences of the form

$$U(c_0, c_1, c_2) = u(c_0) + u(c_1) + u(c_2 - c_1), \tag{15}$$

where u is an increasing, concave function. In the final period, the agent does not derive utility from consumption per se, but from consumption relative to the reference level c_1. Intuitively, the agent develops a consumption habit in this case. Since the cross derivative U_{c_1,c_2} is positive, consumption is complementary between dates 1 and 2.

As in the previous example, skills are stochastic only in the final period. Allocations thus take the form $(c_0, c_1, c_{2i}, y_0, y_1, y_{2i})_{i=L,H}$. Grochulski and Kocherlakota (2010) show that, under the condition $y_{2H} > y_{2L}$, any optimal allocation satisfies

$$\frac{\mathbb{E}[U_{c_0}(c_0, c_1, c_{2i})]}{\mathbb{E}[U_{c_1}(c_0, c_1, c_{2i})]} = \frac{u'(c_0)}{u'(c_1) - \mathbb{E}[u'(c_{2i} - c_1)]} > 1, \tag{16}$$

which means that the agent is left with a residual motive to borrow at date $t = 0$. In other words, the intertemporal wedge between periods 0 and 1 is negative: $\tau_0^K < 0$. Our decomposition reveals that a negative future incentive effect is responsible for this result. Since there is no uncertainty in the first two periods, both A_0 and B_0 are equal to zero, and thus $C_0 = \tau_0^K < 0$.

Intuitively, subsidizing the agent's savings margin here helps the principal with the incentive problem in period 2. Notice that saving in period 0 makes the agent richer in period 1, which due to consumption complementarity increases the agent's marginal utility in period 2. Consequently, saving between periods 0 and 1 has a socially desirable effect on labor supply incentives in period 2 (and no effect on incentives in period 1 as private information is absent then). Since the agent does not internalize that effect, optimal allocations feature a negative intertemporal wedge in order to subsidize the agent's savings activity.

4.2 Sufficient conditions for positive intertemporal wedges

Since intertemporal wedges for time-nonseparable preferences consist of three partly opposing components, it is difficult to determine the sign of intertemporal wedges for general specifications of time-nonseparability. Some results can be obtained by extending the techniques that are familiar from time-separable models. However, this requires relatively strong assumptions as the next finding shows.

Proposition 2 (Positive intertemporal wedge)**.** *Let* $(\mathbf{c}, \mathbf{y})$ *be an optimal allocation and* $t \in \{1, \ldots, T-1\}$*. Suppose that marginal utilities* $U_{c_t}(c_1, \ldots, c_T)$ *and* $U_{c_{t+1}}(c_1, \ldots, c_T)$ *are* θ^{t+1}*-measurable and that the marginal rate of intertemporal substitution* $U_{c_{t+1}}/U_{c_t}$ *is a decreasing function of* c_{t+1}*. Then* $\tau_t^K \geq 0$ *almost everywhere. Moreover,* $\tau_t^K > 0$ *except when the associated history* θ^t *has probability zero or when* $c_{t+1}(\theta^t, \theta_{t+1})$ *is constant for almost all* θ_{t+1}*. By contrast, the last two inequalities are reversed when* $U_{c_{t+1}}/U_{c_t}$ *is an increasing function of* c_{t+1}*.*

Under the assumptions of Proposition 2, optimal allocations impose implicit taxes on capital (positive intertemporal wedges) if the agent's marginal rate of intertemporal substitution, $U_{c_{t+1}}/U_{c_t}$, is decreasing in consumption at date $t+1$. This condition seems hardly restrictive as it states that the agent's value of having one extra consumption unit at time $t+1$ relative to time t falls with the level of consumption at time $t+1$. Since time-separable preferences satisfy this property, our result contains the finding that intertemporal wedges are positive for time-separable preferences (Golosov, Kocherlakota, and Tsyvinski, 2003) as a special case. Proposition 2 is based on the insight that consumption substitutabilities between periods t and $t+1$ cannot be excessive when the marginal rate of intertemporal substitution $U_{c_{t+1}}/U_{c_t}$ is decreasing in c_{t+1}. Hence, the complementarity effect cannot dominate the wealth effect of saving. The proof replaces the well-known Jensen's inequality argument by a more general result on the positive covariance of two monotonic functions of a random variable; see the proof of Proposition 2 in Appendix A for details.

In addition to monotonicity of the agent's marginal rate of intertemporal substitution, Proposition 2 assumes that the contribution of c_t and c_{t+1} to the agent's life-time utility $U(c_1, \ldots, c_T)$ depends only on information known until period $t+1$. This assumption is a strong one. It is obviously satisfied in the penultimate period in any setup with a finite time-horizon. Hence we have the following result.

Corollary. *Consider a history θ^{T-1} that occurs with positive probability and suppose that $c_T(\theta^{T-1}, \theta_T)$ is not constant for almost all θ_T. If $U_{c_T}/U_{c_{T-1}}$ is decreasing in c_T, then $\tau^K_{T-1} > 0$.*

The corollary states that intertemporal wedges are positive at the end of the agent's working life for very general models of time-nonseparable preferences. Finally, we note that the measurability assumption of Proposition 2 is also satisfied when consumption is fully insured from period $t+2$ onwards. In other cases, however, the link between life-time utility and consumption at a given point in time depends potentially on the entire life-time consumption path, so that $U_{c_{t+1}}$ is typically not θ^{t+1}-measurable for time-nonseparable preferences.

5 Quantitative evaluation of a habit formation economy

As the previous analysis has shown, the sign of the optimal distortion on savings remains theoretically ambiguous for time-nonseparable preferences apart from special cases analyzed in Proposition 2. It is therefore a quantitative question whether savings should be taxed or subsidized, and how large those taxes or subsidies should be.

For the quantitative analysis, we focus on preferences where time-nonseparability is due to habit formation. This case seems particularly relevant for capital taxation, because habit formation has helped explain savings behavior under uncertainty in several macroeconomic models as the survey by Messinis (1999) shows. Moreover, Heaton (1995) provides empirical evidence for intertemporal substitutability of consumption at short time-horizons, with habit formation occurring over periods of one year or longer. Given that the period for personal taxation in the U.S. (and most other countries) is one year, these findings suggest that habit formation is the empirically most relevant case of time-nonseparability when it comes to optimal taxation.

Our quantitative model captures some key features of the U.S. economy. In particular, the skill process matches the empirical life-cycle profile and the cross-sectional variance of wages. Given that the time-nonseparability of preferences already introduces an additional continuous state variable, we assume for tractability reasons that the distribution of skills is independent across time. As a robustness check, we allow the cross-sectional variance of skills to depend on age in Section 5.4. Even though skill fluctuations remain transitory in that environment, we pick up one important aspect of persistent processes that is visible in the data: the cross-sectional variance of log-wages increases with age (Heathcote, Storesletten, and Violante, 2012).

5.1 Recursive formulation

We use a recursive approach to compute optimal allocations. In our setup optimal allocations can be written recursively using only two state variables: promised utility and the agent's habit level.

Let $T < \infty$ and suppose that the distribution of skills is independent (but not necessarily identical) across time. As usual, we suppose that the distribution is independent and identical across agents. In period $t = 1, \ldots, T$, skill θ has the time-dependent probability weight $\pi_t(\theta)$, with $\sum_{\theta \in \Theta} \pi_t(\theta) = 1$, where Θ is a finite set of positive real numbers. Set the gross interest rate to $R_t = R \geq 1$ for all t. Moreover, suppose agents' preferences are given as

$$U\left(\{c_t\}_{t=1}^{T}\right) - V\left(\{l_t\}_{t=1}^{T}\right) = \sum_{t=1}^{T} \beta^{t-1} \left(u(c_t, h_t) - v(l_t)\right), \quad h_t = H(c_{t-1}, h_{t-1}), \quad h_1 \text{ given}, \tag{17}$$

where $0 < \beta < 1$ is the agent's discount factor, $u : \mathbb{R}_+^2 \to \mathbb{R}$ is a continuous, concave function that is increasing in its first argument, and $v : \mathbb{R}_+ \to \mathbb{R}$ is continuous, increasing and weakly convex. Consumption utility $u(c_t, h_t)$ is a function of current consumption c_t and a history-dependent reference level h_t. The reference level h_t is obtained iteratively using last period's consumption and last period's reference level using the continuous function $H : \mathbb{R}_+^2 \to \mathbb{R}_+$. As usual, we suppose that the implied specification of life-time consumption utility increases in c_t for all t.

To obtain a problem that is amenable to numerical methods, we require compact spaces for consumption and output. We therefore work with a consumption space of the form $[\underline{c}, \overline{c}]$ and an output space of the type $[\underline{y}, \overline{y}]$. We define $\mathrm{dom}_t(h)$ to be the set of continuation utilities U with the property that, given time-t reference level $h_t = h$, there exists an incentive compatible continuation allocation $(c_s, y_s)_{s=t,\ldots,T}$ which satisfies $\overline{c} \geq c_s \geq \underline{c}$ and $\overline{y} \geq y_s \geq \underline{y}$ for all $T \geq s \geq t$ and generates continuation utility

$$\mathbb{E}\left[\sum_{s=t}^{T} \beta^{t-1} \left(u(c_s, h_s) - v(y_s/\theta_s)\right)\right] = U, \quad \text{where } h_t = h,\ h_s = H(c_{s-1}, h_{s-1}) \text{ for } s > t. \tag{18}$$

Given the structure of our problem, we can express $\mathrm{dom}_t(h)$ in closed form.

Lemma 2. *For any $h \in \mathbb{R}_+$ and $1 \leq t \leq T$, the set $\mathrm{dom}_t(h)$ is a compact interval with bounds*

$$\max(\mathrm{dom}_t(h)) = \sum_{s=t}^{T} \beta^{s-1} \left(u(\overline{c}, h_s) - \sum_{\theta\in\Theta} \pi_s(\theta) v(\underline{y}/\theta)\right), \ \textit{with } h_t = h,\ h_s = H(\overline{c}, h_{s-1}) \textit{ for } s > t,$$

$$\min(\mathrm{dom}_t(h)) = \sum_{s=t}^{T} \beta^{s-1} \left(u(\underline{c}, h_s) - \sum_{\theta\in\Theta} \pi_s(\theta) v(\overline{y}/\theta)\right), \ \textit{with } h_t = h,\ h_s = H(\underline{c}, h_{s-1}) \textit{ for } s > t.$$

Given an initial reference level h_1 and ex-ante utility $U_1 \in \mathrm{dom}_1(h_1)$, we define the value of the planner's cost minimization problem as:

$$C_1(U_1, h_1) := \min_{\mathbf{c},\mathbf{y}} \mathbb{E}\left[\sum_{t=1}^{T} \frac{c_t - y_t}{R^{t-1}}\right] \tag{19}$$

s.t.

$$\mathbb{E}\left[\sum_{t=1}^{T} \beta^{t-1}\left(u(c_t, h_t) - v(y_t/\theta_t)\right)\right] \geq \mathbb{E}\left[\sum_{t=1}^{T} \beta^{t-1}\left(u(c_t(\sigma), h_t(\sigma))) - v(y_t(\sigma)/\theta_t)\right)\right] \quad \forall \sigma \in \Sigma$$

$$\mathbb{E}\left[\sum_{t=1}^{T} \beta^{t-1}\left(u(c_t, h_t) - v(y_t/\theta_t)\right)\right] = U_1$$

$$\bar{c} \geq c_t \geq \underline{c} \quad \forall t, \quad \bar{y} \geq y_t \geq \underline{y} \quad \forall t, \quad h_t = H(c_{t-1}, h_{t-1}) \quad \forall t > 1, \quad h_1 \text{ given.}$$

By extending standard recursive techniques to the present problem, the planner's cost function C_1 can be obtained recursively using the following functional equation for all t (using the convention $C_{T+1} = 0$):

$$C_t(U, h) = \min_{c,y,U'} \sum_{\theta \in \Theta} \pi_t(\theta)\left[c(\theta) - y(\theta) + \frac{1}{R} C_{t+1}(U'(\theta), H(c(\theta), h))\right] \tag{20}$$

s.t.

$$u(c(\theta), h) - v(y(\theta)/\theta) + \beta U'(\theta) \geq u(c(\theta'), h) - v(y(\theta')/\theta) + \beta U'(\theta') \quad \forall \theta, \theta'$$

$$\sum_{\theta \in \Theta} \pi_t(\theta)\left[u(c(\theta), h) - v(y(\theta)/\theta) + \beta U'(\theta)\right] = U$$

$$U'(\theta) \in \mathrm{dom}_t(H(c(\theta), h)) \quad \forall \theta$$

$$\bar{c} \geq c(\theta) \geq \underline{c}, \;\; \bar{y} \geq y(\theta) \geq \underline{y} \quad \forall \theta$$

Compared to the recursive formulation of incentive problems with time-separable preferences by Spear and Srivastava (1987) and Phelan and Townsend (1991), time-nonseparability adds the agent's reference level h as a second state variable to the planner's problem. Clearly, the state variable in the time-separable case (promised utility) is no longer sufficient here, because the planner faces heterogeneous types of agents when preferences are time-nonseparable. On the other hand, no additional states other than the agent's current reference level h are needed, because the agent's type is fully determined by observable information, which allows to separate the incentive constraint into a sequence of temporary incentive constraints similar to the time-separable case.

The recursive formulation in equation (20) reaches beyond the case of Cobb-Douglas habit formation studied below. For instance, it includes the case of linear habit formation, $u(c_t, h_h) = \tilde{u}(c_t - \gamma h_t)$, explored by Constantinides (1990) and Campbell and Cochrane (1999). The formulation also includes the case of durable goods (intertemporal substitutability of consumption) if we set $u(c_t, h_h) = \tilde{u}(c_t + \delta h_t)$ and interpret c_t as the current expenditure on a durable good and h_t as the previous stock of the durable good.

The previous insights give rise to a simple computational approach. We first solve for the sequence of domain restrictions $(\mathrm{dom}_t(h))_{h\in\mathbb{R}_+, t=1,\ldots,T}$ following Lemma 2. We then exploit the functional equation (20) to obtain the sequence of cost functions $(C_t)_{t=1,\ldots,T}$ of the planner's problem using standard numerical optimization procedures. The associated policy functions are then iterated forward to generate the optimal allocation.

5.2 Parameters

There are $T = 10$ periods with a duration of 5 years each. Agents enter the model at age 18. In each period, skill θ_t is randomly drawn from the set $\{\theta_{tL}, \theta_{tH}\}$, where both realizations have equal probability and $\theta_{tL} < \theta_{tH}$. Draws are independent across agents and time. We choose the life-cycle profile of expected skills in line with Hansen (1993, Table II), who estimates relative efficiency profiles of workers in the United States over the years 1955 to 1988.[9] Fitting those profiles to 5-year intervals generates the numbers in Table 1. Skills are hump-shaped over the life-cycle and peak in period 7 (age 48). Regarding the variance of skills, we target the cross-sectional variance of log-wages in the United States in the period 1967–2006 (Heathcote, Storesletten, and Violante, 2012, Table 3), which leads to a variance of log-skills of 0.351.

Table 1: Life-cycle profile of expected skills

$\mathbb{E}[\theta_1]$	$\mathbb{E}[\theta_2]$	$\mathbb{E}[\theta_3]$	$\mathbb{E}[\theta_4]$	$\mathbb{E}[\theta_5]$	$\mathbb{E}[\theta_6]$	$\mathbb{E}[\theta_7]$	$\mathbb{E}[\theta_8]$	$\mathbb{E}[\theta_9]$	$\mathbb{E}[\theta_{10}]$
0.560	0.754	0.912	1.034	1.119	1.168	1.180	1.156	1.095	0.999

Our preference specification follows Abel (1990), Carroll, Overland, and Weil (2000), and

[9] Hansen (1993) uses average hourly earnings of a given age-subgroup divided by average hourly earnings of all subgroups as the relative efficiency measure.

Diaz, Pijoan-Mas, and Rios-Rull (2003), and sets up habit formation in a Cobb-Douglas form:

$$u(c_t, h_t) = \tilde{u}\left(c_t h_t^{-\gamma}\right) = \tilde{u}\left(c_t^{1-\gamma}\left(\frac{c_t}{h_t}\right)^{\gamma}\right), \tag{21}$$

where γ is a number between zero and one. Note that period utility depends on a Cobb-Douglas aggregate of absolute consumption, c_t, and absolute consumption relative to the habit level, c_t/h_t, and the parameter γ controls the relative importance of these two components.[10] In line with Diaz, Pijoan-Mas, and Rios-Rull (2003), we choose $\gamma = 0.75$. This value corresponds to the case of 'strong habits' explored by Carroll, Overland, and Weil (2000) and is reasonably close to empirical results by Fuhrer (2000), who estimates a value of 0.80 based on aggregate consumption data. Period utility is of the CRRA type: $\tilde{u}(x) = x^{1-\sigma}/(1-\sigma)$. We set $\sigma = 3$ in line with recent estimations by Paravisini, Rappoport, and Ravina (2010). The discount factor is $\beta = 0.98^5$ and the interest rate equals $R = 1/\beta$. The labor disutility function is $v(l) = \alpha l^{1+\frac{1}{\psi}}/(1+\frac{1}{\psi})$, with a Frisch elasticity of labor supply of $\psi = 0.5$, and $\alpha = 1$.

The habit process in our model has a persistence coefficient of λ and is given by:

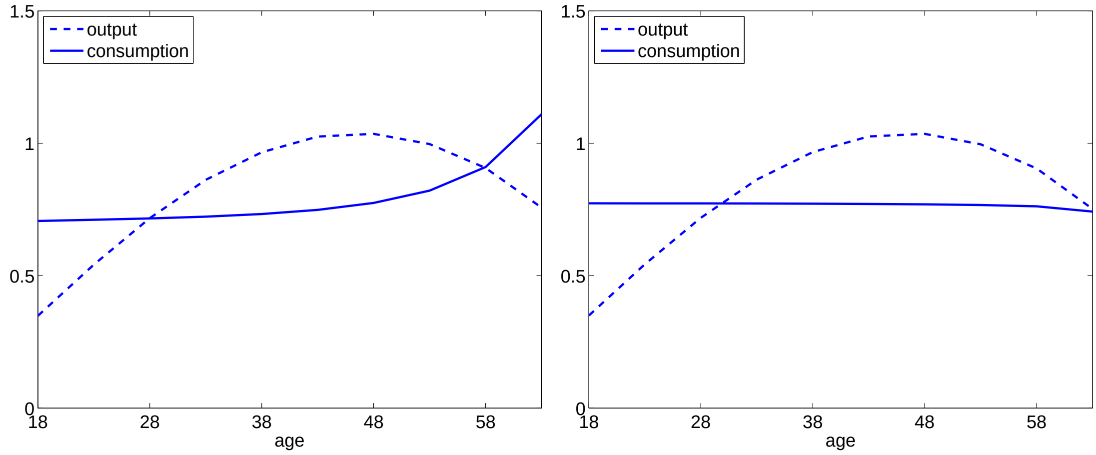
$$h_t = (1-\lambda)c_{t-1} + \lambda h_{t-1}. \tag{22}$$

Diaz, Pijoan-Mas, and Rios-Rull (2003) set λ to 0.75 for yearly periods. Carroll, Overland, and Weil (2000) explore a continuous time model in which habits adjust to permanent changes in consumption with a half-life of 3.5 years. Adapted to 5-year periods, both approaches imply very similar coefficients and we therefore choose $\lambda = 0.75^5$ as our baseline case.[11]

We set the initial habit level to $h_1 = 0.7$. As we verify ex-post, this number coincides approximately with the agent's consumption level in the first period. We set the initial utility promise U_1 such that the planner's budget is balanced, i.e., $C_1(U_1, h_1) = 0$. We verify that the bounds $\underline{c}$, $\overline{c}$, $\underline{y}$, and $\overline{y}$ are never binding for the optimal allocation starting from this initial state.

[10] Another common specification of habit formation is the linear one: $u(c_t, h_t) = \tilde{u}(c_t - \gamma h_t)$; see Constantinides (1990) and Campbell and Cochrane (1999). For our present purposes, the Cobb-Douglas formulation is more convenient, since period utilities are well defined whenever c_t and h_t are positive. The linear formulation has the drawback of ruling out all pairs (c_t, h_t) with $c_t < \gamma h_t$, which makes the computation of the domain restriction and of the optimal allocation somewhat more cumbersome.

[11] The available empirical evidence on the persistence of habits is mixed. Depending on the environment, some studies find persistence levels close to zero (Fuhrer, 2000), while other estimations point to significantly larger values (Heaton, 1995). We provide a case with non-persistent habits ($\lambda = 0$) and a case with high persistence ($\lambda = 0.5$) as robustness checks.



(a) Habit formation (b) Time-separable preferences

Figure 1: Expected consumption and output over the life-cycle

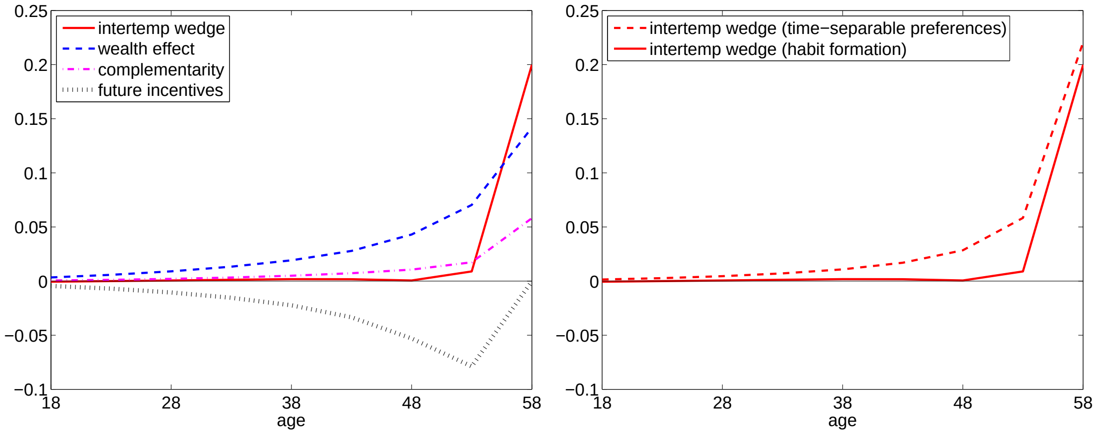
5.3 Results

Figure 1a presents the paths of expected output and consumption for our baseline case ($\gamma = 0.75$, $\lambda = 0.24$, $h_1 = 0.7$). Expected output follows the hump-shaped pattern of the skill process. Expected consumption increases over the life-cycle and grows by about 60 percent from age 18 to age 63. Figure 1b shows the corresponding paths for the case of time-separable preferences ($\gamma = 0$).[12] The path of expected output is very similar to the habit formation case. Expected consumption, however, is virtually flat (but slightly monotonically decreasing) for time-separable preferences. This shows that habit formation has a positive impact on the optimal growth rate of consumption.

Figure 2a decomposes expected intertemporal wedges into the wealth effect, complementarity effect, and future incentive effect. The expected intertemporal wedge is virtually zero—apart from the very end of the agent's working life. The wealth effect and complementarity effect are positive in all periods, as shown by equations (7) and (8) in Proposition 1. As expected, the future incentive effect is negative. In line with Lemma 1, the future incentive effect is zero in the penultimate period, because there is no incentive problem more than one period ahead.

The quantitative impact of habit formation is sizable (Figure 2b). The life-cycle average of the intertemporal wedge with habit formation is 0.024 (corresponding to a 24.8 percent tax on

[12]To make the allocations comparable, we choose a scaling parameter of $\alpha = 3.75$ for the time-separable case, such that the discounted value of life-time output (and consumption) coincides with the habit formation case. This adjustment has a negligible effect on intertemporal wedges: averaged over the life-cycle, wedges are 0.0393 for $\alpha = 1$ and 0.0390 for $\alpha = 3.75$.



(a) Habit formation (b) Habit formation and time-separable preferences

Figure 2: Expected intertemporal wedges over the life-cycle

net interest $R-1$). In the time-separable case it is 0.039 (corresponding to a 40.6 percent tax on net interest). The difference is even more pronounced if we focus on workers aged between 18 and 53. For those workers, the average wedge with habit formation is only one ninth of the average wedge with time-separable preferences. Habit formation also changes the qualitative features of intertemporal wedges: the increase of intertemporal wedges towards the end of the working life is much steeper with habit formation.

In our results, the future incentive effect dominates the complementarity effect by a large margin. Intuitively, the future incentive effect encourages saving (and thus next period's consumption) in order to relax incentive problems in the later future. The complementarity effect, by contrast, discourages saving in order to relax the incentive problem in the immediately following period. There are two reasons why the future incentive effect is more significant than the complementarity effect. First, towards the end of the working life, incentive provision must rely less on future promises and more on costly immediate consumption rewards. This makes relaxing future incentive problems important. Second, mean skills increase until age 48 and are relatively flat thereafter. Therefore, effort from younger workers contributes less to social output, and thus incentive provision in early periods is less crucial. For these two reasons, the future incentive effect dominates the complementarity effect (except at the very end of the working life).

The heterogeneity of intertemporal wedges is analyzed in Figure 3. For each period, we display minimum and maximum intertemporal wedges among all possible histories. We see

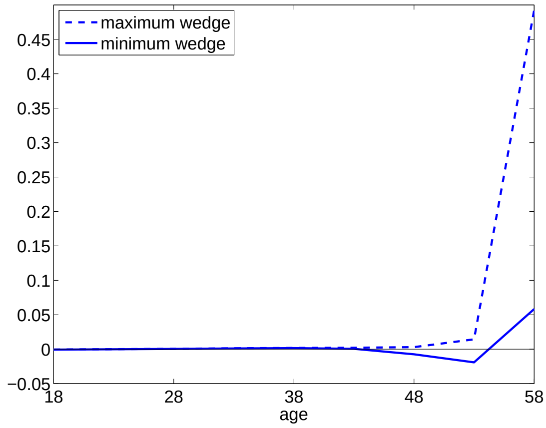


Figure 3: Minimum and maximum intertemporal wedges for habit formation model

that intertemporal wedges are indistinguishable from zero until age 48, ambiguous at age 53, and clearly positive at age 58.

For habit formation as well as time-separable preferences, expected intertemporal wedges increase steeply towards the end of the agent's life. As argued before, incentives are provided more by immediate consumption rewards and less by future promises as the agent approaches the final periods. Hence, the negative impact of saving on next period's incentive problem becomes stronger over time, and this increases the wealth effect; compare Golosov, Troshkin, and Tsyvinski (2011). For the habit formation model, this also increases the complementarity effect. Moreover, the future incentive effect reaches zero in the penultimate period, which additionally boosts the intertemporal wedge in that period.

For the case of time-separable preferences, the quantitative exploration closest to ours is by Albanesi and Sleet (2006). Based on an infinite horizon model with independently identically distributed skills, they find intertemporal wedges of typically less than one percent. The closest counterpart in our finite horizon setup is arguably the intertemporal wedge in the initial period. For the time-separable case, we obtain an intertemporal wedge of 0.002 in the initial period, which is in the same range as the wedges in Albanesi and Sleet (2006). The agent's time-discount factor is almost identical to our setup, so the implied tax rates on capital returns are also similar.

Farhi and Werning (2013) study optimal dynamic taxes in a life-cycle framework with time-separable preferences and a persistent skill process. Averaged over the life-cycle, they find intertemporal wedges corresponding to a tax on net interest of about 10 percent, which is smaller than the number in our time-separable case. Much of this difference is due to the lower coefficient of risk aversion in their framework. Indeed, for time-separable logarithmic utility, our model generates intertemporal wedges that correspond to an average tax on net interest of 11.4 percent.

5.4 Sensitivity analysis

To study the sensitivity of our results, we explore alternative parameters for the persistence of the habit process ($\lambda = 0$, $\lambda = 0.5$), the coefficient of risk aversion ($\sigma = 2$, $\sigma = 4$), and the initial habit level ($h_1 = 0.5$, $h_1 = 1.5$). We also explore a case in which the variance of the skill process increases over the life-cycle. All our results are qualitatively robust to these changes. For further details, we refer the reader to Appendix B.

The sensitivity analysis reveals how the future incentive effect depends on the preference specification and the details of the incentive problem. If habits are non-persistent ($\lambda = 0$), then next period's consumption fully determines the habit level in the period thereafter. The motive to encourage saving in order to increase consumption in the next period and thereby relax the incentive problem in the period thereafter is then strong. As a result, the future incentive effect becomes bigger in absolute value. When habits are persistent, the effect of next period's consumption on habits in the period thereafter is mitigated. At the same time, the impact of next period's consumption on habits in later periods becomes stronger. While the first effect decreases the future incentive effect, the second one increases it. In the extreme case of fully persistent habits, consumption in any period becomes irrelevant for future habits and Lemma 1 implies that the future incentive effect is zero. The sensitivity analysis shows more generally that the future incentive effect falls with the degree of persistence already at moderate ranges of persistence. Finally, we note that the future incentive effect becomes more pronounced if the variance of the skill process increases over time, because future incentive problems become more important relative to immediate ones.

6 Concluding remarks

This paper studies optimal capital taxation in a model with private information and time-nonseparable preferences. We characterize optimal allocations in terms of intertemporal wedges and decompose intertemporal wedges into three components. One component is the standard *wealth effect* known from the time-separable case. The two novel components are a *complementarity effect* and a *future incentive effect*. The former is due to complementarities (or substitutabilities) of consumption between adjacent periods, while the latter effect captures consequences on the more distant future. We discuss two examples where these additional effects reverse the standard optimal taxation logic and generate nonpositive intertemporal wedges (capital subsidies). We then evaluate intertemporal wedges quantitatively for a habit formation economy. In this case, the complementarity effect contributes positively to intertemporal wedges, whereas the future incentive effect is negative. However, the future incentive effect is quantitatively more significant, so that intertemporal wedges fall by about 40 percent compared to the time-separable case.

In many strands of the literature, habit formation has successfully bridged the gap between theory and evidence. Our quantitative results suggest that this might also be the case for models of optimal taxation with private information. With time-separable preferences, predicted optimal tax rates on capital income can exceed observed tax rates by a considerable margin (Golosov, Troshkin, and Tsyvinski, 2011; Abraham, Koehne, and Pavoni, 2012). Our quantitative results show that habit formation substantially reduces the optimal tax rates on capital.

Moreover, our quantitative framework can serve as a starting point for taxation problems with alternative forms of time-nonseparability. Housing seems a particularly interesting case, as housing is an asset and a durable consumption good at the same time. In addition, housing typically accounts for the largest fraction of households' wealth. Studying the optimal taxation of housing might shed light on the widely observed practice of treating housing wealth differently from other asset classes. Finally, the recursive approach proposed in this paper extends beyond optimal taxation and applies to arbitrary private information problems with time-nonseparable preferences.

Appendix

A Proofs

Proof of Proposition 1. The intertemporal wedge between periods t and $t+1$ is equal to

$$\tau_t^K = 1 - \frac{\mathbb{E}_t[U_{c_t}]\,\mathbb{E}_t\left[\frac{1}{U_{c_{t+1}}}\right]}{R_t\,\mathbb{E}_t[U_{c_{t+1}}]\,\mathbb{E}_t\left[\frac{1}{U_{c_{t+1}}}\right]}. \tag{23}$$

Equivalently, we have

$$\tau_t^K = \frac{R_t\,\mathbb{E}_t[U_{c_{t+1}}]\,\mathbb{E}_t\left[\frac{1}{U_{c_{t+1}}}\right] - \mathbb{E}_t[U_{c_t}]\,\mathbb{E}_t\left[\frac{1}{U_{c_{t+1}}}\right]}{R_t\,\mathbb{E}_t[U_{c_{t+1}}]\,\mathbb{E}_t\left[\frac{1}{U_{c_{t+1}}}\right]}. \tag{24}$$

After adding and subtracting a few terms, we obtain

$$\tau_t^K = \frac{R_t\,\mathbb{E}_t[U_{c_{t+1}}]\,\mathbb{E}_t\left[\frac{1}{U_{c_{t+1}}}\right] - R_t + \mathbb{E}_t\left[\frac{U_{c_t}}{U_{c_{t+1}}}\right] - \mathbb{E}_t[U_{c_t}]\,\mathbb{E}_t\left[\frac{1}{U_{c_{t+1}}}\right] + R_t - \mathbb{E}_t\left[\frac{U_{c_t}}{U_{c_{t+1}}}\right]}{R_t\,\mathbb{E}_t[U_{c_{t+1}}]\,\mathbb{E}_t\left[\frac{1}{U_{c_{t+1}}}\right]}. \tag{25}$$

By the definition of a covariance, we have

$$\operatorname{cov}_t\left(U_{c_{t+1}}\,,\,\frac{1}{U_{c_{t+1}}}\right) = 1 - \mathbb{E}_t[U_{c_{t+1}}]\,\mathbb{E}_t\left[\frac{1}{U_{c_{t+1}}}\right], \tag{26}$$

$$\operatorname{cov}_t\left(U_{c_t}\,,\,\frac{1}{U_{c_{t+1}}}\right) = \mathbb{E}_t\left[\frac{U_{c_t}}{U_{c_{t+1}}}\right] - \mathbb{E}_t[U_{c_t}]\,\mathbb{E}_t\left[\frac{1}{U_{c_{t+1}}}\right]. \tag{27}$$

This implies

$$\tau_t^K = \frac{-R_t\operatorname{cov}_t\left(U_{c_{t+1}}\,,\,\frac{1}{U_{c_{t+1}}}\right)}{R_t\,\mathbb{E}_t[U_{c_{t+1}}]\,\mathbb{E}_t\left[\frac{1}{U_{c_{t+1}}}\right]} + \frac{\operatorname{cov}_t\left(U_{c_t}\,,\,\frac{1}{U_{c_{t+1}}}\right)}{R_t\,\mathbb{E}_t[U_{c_{t+1}}]\,\mathbb{E}_t\left[\frac{1}{U_{c_{t+1}}}\right]} + \frac{R_t - \mathbb{E}_t\left[\frac{U_{c_t}}{U_{c_{t+1}}}\right]}{R_t\,\mathbb{E}_t[U_{c_{t+1}}]\,\mathbb{E}_t\left[\frac{1}{U_{c_{t+1}}}\right]} \tag{28}$$

$$= \gamma_t(A_t + B_t + C_t), \tag{29}$$

where $\gamma_t = \left(R_t\mathbb{E}_t[U_{c_{t+1}}]\,\mathbb{E}_t\left[\frac{1}{U_{c_{t+1}}}\right]\right)^{-1}$ is positive, since $U_{c_{t+1}} > 0$.

Since the function $x \mapsto 1/x$ is decreasing, the covariance term in A_t is nonpositive, which renders A_t nonnegative. To verify that B_t can be zero, positive, or negative, first note that by concavity $1/U_{c_{t+1}}$ is an increasing function of c_{t+1}. Marginal utility U_{c_t} can be increasing or decreasing in c_{t+1}, depending on whether consumption goods c_t and c_{t+1} are complements or substitutes. Hence, the covariance term in B_t can have a positive, negative or neutral sign. Finally, to analyze the sign of C_t, observe first that for time-separable preferences C_t is zero by the Inverse Euler equation from Rogerson (1985). A case with a negative C_t is presented in Example 2. To obtain a case in which C_t is positive, we simply change the preferences in that example to $U(c_0, c_1, c_2) = u(c_0) + u(c_1) + u(c_2 - c_0)$. □

Proof of Lemma 1. Fix some history θ^t that occurs with positive probability. Consider a perturbation that increases consumption at time t by ξ and reduces consumption at time $t+1$ by ϕ. Then, depending on future shocks $(\theta_{t+1}, \ldots, \theta_T)$, the agent's consumption utility changes by

$$\Delta U = U\left(c_1\left(\theta^1\right), \ldots, c_T\left(\theta^T\right)\right) - U\left(c_1\left(\theta^1\right), \ldots, c_t\left(\theta^t\right) - \xi, c_{t+1}\left(\theta^{t+1}\right) + \phi, \ldots, c_T\left(\theta^T\right)\right) \quad (30)$$

$$= \int_{c_t(\theta^t)+\xi}^{c_t(\theta^t)} U_{c_t}\left(c_1\left(\theta^1\right), \ldots, \gamma, c_{t+1}\left(\theta^{t+1}\right), \ldots, c_T\left(\theta^T\right)\right) d\gamma \quad (31)$$

$$+ \int_{c_{t+1}(\theta^{t+1})+\phi}^{c_{t+1}(\theta^{t+1})} U_{c_{t+1}}\left(c_1\left(\theta^1\right), \ldots, c_t\left(\theta^t\right) + \xi, \delta, \ldots, c_T\left(\theta^T\right)\right) d\delta. \quad (32)$$

Since U_{c_t} and $U_{c_{t+1}}$ are by assumption θ^{t+1}-measurable, the above formula shows that ΔU depends on $(\theta_{t+1}, \ldots, \theta_T)$ only through the variable θ_{t+1}. Hence we can find numbers $\phi = \phi(\xi, \theta_{t+1})$ depending only on ξ and θ_{t+1} so that we have $\Delta U = 0$ for all $(\theta_{t+1}, \ldots, \theta_T)$. The consumption perturbation is then neutral with respect to the incentive constraint.

The allocation $(\mathbf{c}, \mathbf{y})$ can only be optimal if the perturbed consumption scheme requires at least as many resources as the original scheme c. Hence, $\xi = 0$ must minimize

$$\xi - \frac{\mathbb{E}_t\left[\phi(\xi, \theta_{t+1})\right]}{R_t}, \quad (33)$$

which yields the first-order condition

$$1 - \frac{\mathbb{E}_t\left[\frac{U_{c_t}(c_1,\ldots,c_T)}{U_{c_{t+1}}(c_1,\ldots,c_T)}\right]}{R_t} = 0. \quad (34)$$

Using the notation from Proposition 1, this implies $C_t = 0$. □

Proof of Proposition 2. Fix some history θ^t that occurs with positive probability. Lemma 1 implies $C_t = 0$. By Proposition 1, the intertemporal wedge is therefore nonnegative if and only if

$$-R_t \operatorname{cov}_t\left(U_{c_{t+1}}\,,\,\frac{1}{U_{c_{t+1}}}\right)+\operatorname{cov}_t\left(U_{c_t}\,,\,\frac{1}{U_{c_{t+1}}}\right)\geq 0. \tag{35}$$

Using again the result $R_t = E_t\left[U_{c_t}/U_{c_{t+1}}\right]$ from Lemma 1 and dividing by $\mathbb{E}_t\left[1/U_{c_{t+1}}\right] > 0$, the previous line is equivalent to the condition

$$-\mathbb{E}_t\left[U_{c_t}\right]+R_t\mathbb{E}_t\left[U_{c_{t+1}}\right]\geq 0. \tag{36}$$

Since $R_t = E_t\left[U_{c_t}/U_{c_{t+1}}\right]$, we can rewrite the previous line as

$$\operatorname{cov}_t\left(-U_{c_{t+1}}\,,\,\frac{U_{c_t}}{U_{c_{t+1}}}\right)\geq 0. \tag{37}$$

Concavity of the utility function U implies that the negated marginal utility $-U_{c_{t+1}}$ is increasing in c_{t+1}. In addition, $U_{c_t}/U_{c_{t+1}}$ is increasing in c_{t+1} by assumption. Since the covariance of two increasing functions of a random variable is nonnegative (Schmidt, 2003), we have established the first part of the proposition. The second part follows from the fact that the covariance of two increasing functions of a random variable is positive unless the random variable is constant almost everywhere. □

Proof of Lemma 2. Let $\tilde{U}(c^*, y^*)$ be the time-t continuation utility of an agent with reference level $h_t = h$ who consumes a fixed level c^* and produces a fixed output y^* in periods $t, \dots, T$ irrespective of his skill. Clearly all such allocations are incentive compatible. Setting $h_s = H(\overline{c}, h_{s-1})$ for $s > t$, we have

$$\tilde{U}(\overline{c}, \underline{y}) = \sum_{s=t}^{T}\beta^{s-1}\left(u(\overline{c}, h_s) - \sum_{\theta\in\Theta}\pi_s(\theta)v(\underline{y}/\theta)\right) \tag{38}$$

and it is obvious that no other incentive compatible allocation can deliver a higher continuation utility. Similarly, using $h_s = H(\underline{c}, h_{s-1})$ for $s > t$, we have

$$\tilde{U}(\underline{c}, \overline{y}) = \sum_{s=t}^{T}\beta^{s-1}\left(u(\underline{c}, h_s) - \sum_{\theta\in\Theta}\pi_s(\theta)v(\overline{y}/\theta)\right) \tag{39}$$

and no other incentive compatible allocation can deliver a lower continuation utility.

To verify that $\mathrm{dom}_t(h)$ is an interval, note that $\mathrm{dom}_t(h)$ contains all numbers that can be written as $\tilde{U}(c^*, y^*)$ for some $c^* \in [\underline{c}, \overline{c}]$, $y^* \in [\underline{y}, \overline{y}]$. By the continuity of $\tilde{U}$ (ensured by the continuity of u, v, H) this covers all numbers in the interval $[\tilde{U}(\underline{c}, \overline{y}), \tilde{U}(\overline{c}, \underline{y})]$. Hence, we have

$$\mathrm{dom}_t(h) \supseteq [\tilde{U}(\underline{c}, \overline{y}),\, \tilde{U}(\overline{c}, \underline{y})]. \tag{40}$$

On the other hand, we clearly have

$$\mathrm{dom}_t(h) \subseteq [\min(\mathrm{dom}_t(h)),\, \max(\mathrm{dom}_t(h))]. \tag{41}$$

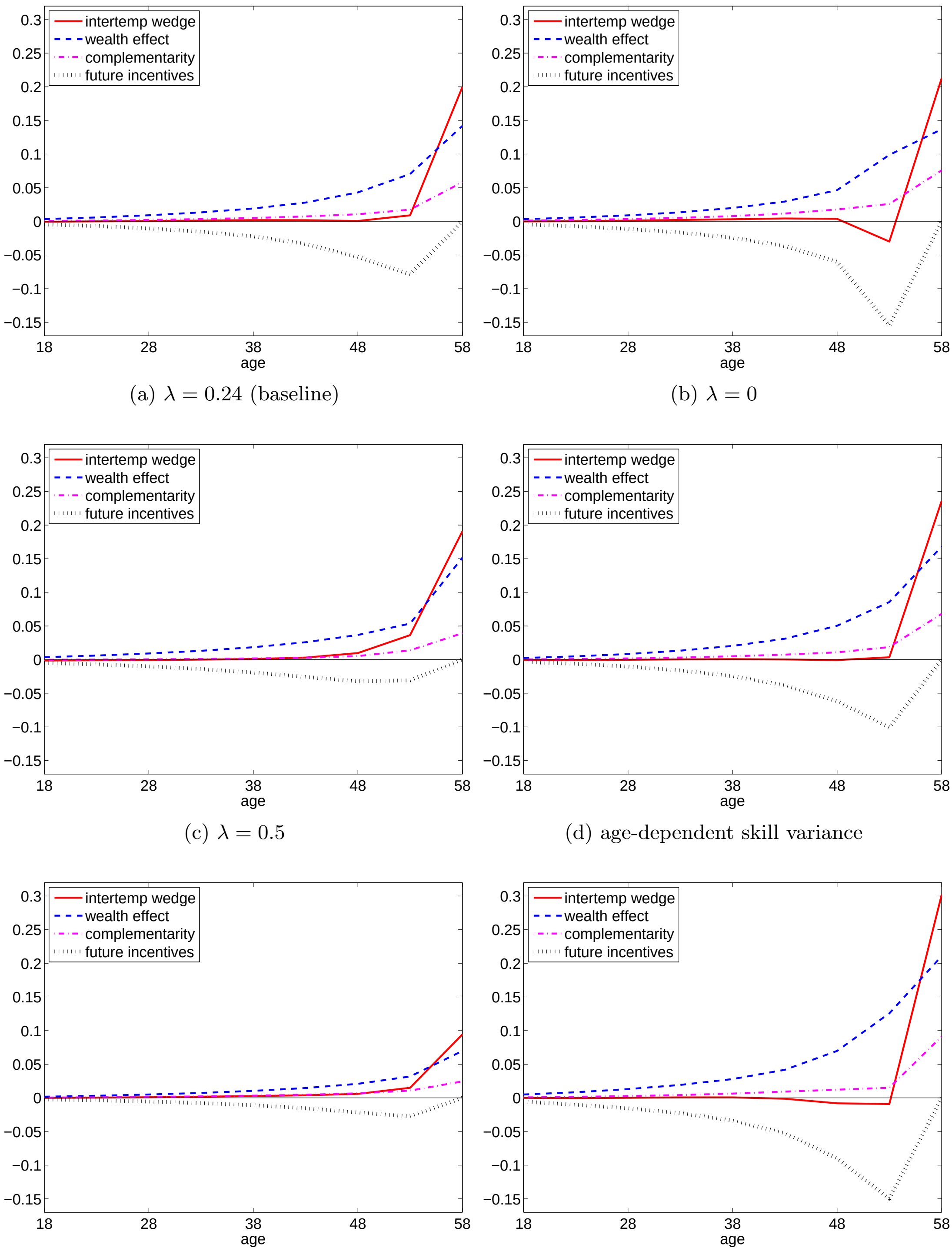
Using the results $\min(\mathrm{dom}_t(h)) = \tilde{U}(\underline{c}, \overline{y})$ and $\max(\mathrm{dom}_t(h)) = \tilde{U}(\overline{c}, \underline{y})$ from the first step, the two set inequalities taken together imply

$$\mathrm{dom}_t(h) = [\tilde{U}(\underline{c}, \overline{y}),\, \tilde{U}(\overline{c}, \underline{y})]. \tag{42}$$

This completes the proof. □

B Sensitivity analysis

We first study the role of the persistence parameter of the habit process. Figures 4a, 4b, and 4c compare expected intertemporal wedges for the baseline case ($\lambda = 0.24$), non-persistent habits ($\lambda = 0$) and highly persistent habits ($\lambda = 0.5$). We then consider a case in which the variance of the skill process increases over time (Figure 4d). Skills have the same expected values as in the baseline case, but their variance is age-dependent in line with findings by Heathcote, Storesletten, and Violante (2012, Figure 1A). Figures 4e and 4f explore low ($\sigma = 2$) and high ($\sigma = 4$) values for the coefficient of relative risk aversion. Finally, we explore alternative values for the initial habit level. The effect on intertemporal wedges is negligible and we therefore omit the results for the sake of brevity.



(a) $\lambda = 0.24$ (baseline)

(b) $\lambda = 0$

(c) $\lambda = 0.5$

(d) age-dependent skill variance

(e) $\sigma = 2$

(f) $\sigma = 4$

Figure 4: Decomposition of expected intertemporal wedges for alternative persistence parameters (a,b,c), skill process with age-dependent variance (d), and alternative coefficients of risk aversion (e,f)

References

Abel, A. B. (1990): "Asset Prices under Habit Formation and Catching up with the Joneses," *American Economic Review*, 80(2), 38–42.

Abraham, A., S. Koehne, and N. Pavoni (2012): "Optimal Income Taxation with Asset Accumulation," European University Institute. Mimeo. Available online at: `http://mpra.ub.uni-muenchen.de/38629/`.

Albanesi, S., and C. Sleet (2006): "Dynamic Optimal Taxation with Private Information," *Review of Economic Studies*, 73(1), 1–30.

Campbell, J. Y., and J. H. Cochrane (1999): "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107(2), 205–251.

Carroll, C. D., J. Overland, and D. N. Weil (2000): "Saving and Growth with Habit Formation," *American Economic Review*, 90(3), 341–355.

Clark, A. E. (1999): "Are wages habit-forming? evidence from micro data," *Journal of Economic Behavior & Organization*, 39(2), 179–200.

Constantinides, G. M. (1990): "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy*, 98(3), 519–43.

Diamond, P. A., and J. A. Mirrlees (1978): "A model of social insurance with variable retirement," *Journal of Public Economics*, 10(3), 295 – 336.

Diaz, A., J. Pijoan-Mas, and J. Rios-Rull (2003): "Precautionary savings and wealth distribution under habit formation preferences," *Journal of Monetary Economics*, 50(6), 1257–1291.

Dunn, K. B., and K. J. Singleton (1986): "Modeling the term structure of interest rates under non-separable utility and durability of goods," *Journal of Financial Economics*, 17(1), 27 – 55.

Eichenbaum, M., and L. P. Hansen (1990): "Estimating Models with Intertemporal Substitution Using Aggregate Time Series Data," *Journal of Business & Economic Statistics*, 8(1), 53–69.

Farhi, E., and I. Werning (2008): "Optimal savings distortions with recursive preferences," *Journal of Monetary Economics*, 55(1), 21–42.

Farhi, E., and I. Werning (2013): "Insurance and Taxation over the Life Cycle," *The Review of Economic Studies*.

Frederick, S., and G. Loewenstein (1999): "Hedonic Adaptation," in *Well-being: The foundations of hedonic psychology*, ed. by D. Kahneman, E. Diener, and N. Schwarz, pp. 302–329. Russell Sage Foundation Press.

Fuhrer, J. C. (2000): "Habit Formation in Consumption and Its Implications for Monetary-Policy Models," *The American Economic Review*, 90(3), pp. 367–390.

Golosov, M., N. Kocherlakota, and A. Tsyvinski (2003): "Optimal Indirect and Capital Taxation," *Review of Economic Studies*, 70(3), 569–587.

Golosov, M., M. Troshkin, and A. Tsyvinski (2011): "Optimal dynamic taxes," Discussion paper, National Bureau of Economic Research.

Golosov, M., and A. Tsyvinski (2006): "Designing Optimal Disability Insurance: A Case for Asset Testing," *Journal of Political Economy*, 114(2), 257–279.

Golosov, M., A. Tsyvinski, and I. Werning (2006): "New Dynamic Public Finance: A User's Guide," *NBER Macroeconomics Annual*, 21, 317–363.

Gottardi, P., and N. Pavoni (2011): "Ramsey Asset Taxation under Asymmetric Information," European University Institute. Mimeo.

Grochulski, B., and N. Kocherlakota (2010): "Nonseparable preferences and optimal social security systems," *Journal of Economic Theory*, 145(6), 2055 – 2077.

Hansen, G. (1993): "The Cyclical and Secular Behaviour of the Labour Input: Comparing Efficiency Units and Hours Worked," *Journal of Applied Econometrics*, 8(1), 71–80.

Heathcote, J., K. Storesletten, and G. Violante (2012): "Consumption and labor supply with partial insurance: An analytical framework," Federal Reserve Bank of Minneapolis and New York University, Mimeo.

Heaton, J. (1993): "The interaction between time-nonseparable preferences and time aggregation," *Econometrica*, 61(2), 353–385.

Heaton, J. (1995): "An Empirical Investigation of Asset Pricing with Temporally Dependent Preference Specifications," *Econometrica*, 63(3), 681–717.

Helson, H. (1964): *Adaptation-Level Theory*. Harper & Row New York.

Hopenhayn, H., and A. Jarque (2010): "Unobservable Persistent Productivity and Long Term Contracts," *Review of Economic Dynamics*, 13(2), 333–349.

Jarque, A. (2010): "Repeated moral hazard with effort persistence," *Journal of Economic Theory*, 145(6), 2412–2423.

Kocherlakota, N. (2010): *The New Dynamic Public Finance*. Princeton University Press.

Kocherlakota, N. R. (2005): "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation," *Econometrica*, 73(5), 1587–1621.

Kwon, I. (2006): "Incentives, Wages, and Promotions: Theory and Evidence," *RAND Journal of Economics*, 37(1), 100–120.

Messinis, G. (1999): "Habit Formation and the Theory of Addiction," *Journal of Economic Surveys*, 13(4), 417–42.

Mukoyama, T., and A. Sahin (2005): "Repeated moral hazard with persistence," *Economic Theory*, 25(4), 831–854.

Paravisini, D., V. Rappoport, and E. Ravina (2010): "Risk Aversion and Wealth: Evidence from Person-to-Person Lending Portfolios," Working Paper 16063, National Bureau of Economic Research.

Phelan, C., and R. M. Townsend (1991): "Computing Multi-Period, Information-Constrained Optima," *The Review of Economic Studies*, 58(5), 853–881.

Ravina, E. (2007): "Habit Persistence and Keeping Up with the Joneses: Evidence from Micro Data," New York University. Mimeo.

Rogerson, W. P. (1985): "Repeated Moral Hazard," *Econometrica*, 53(1), 69–76.

Ryder, Jr., H. E., and G. M. Heal (1973): "Optimum Growth with Intertemporally Dependent Preferences," *Review of Economic Studies*, 40(1), 1–33.

Schmidt, K. (2003): "On the covariance of monotone functions of a random variable," DSVM Working Paper 04/2003, University of Dresden.

Spear, S., and S. Srivastava (1987): "On repeated moral hazard with discounting," *Review of Economic Studies*, 54(4), 599–617.

Werning, I. (2011): "Nonlinear capital taxation," MIT. Mimeo.