Optimal capital taxation for time-nonseparable preferences

Sebastian Koehne*  Moritz Kuhn†

February 11, 2013

Abstract

We study optimal capital taxation in a dynamic Mirrleesian model with time-nonseparable preferences. The model covers the widely used cases of habit formation and durable consumption. Time-nonseparable preferences change labor supply incentives across time and thereby generate novel motives to distort capital accumulation decisions. We decompose intertemporal wedges (implicit capital taxes) into three components and provide conditions under which intertemporal wedges are positive. We derive a recursive formulation of constrained efficient allocations and evaluate the quantitative importance of habit formation for intertemporal wedges. In our baseline parameterization, habit formation reduces average intertemporal wedges by about 40 percent compared to the time-separable case.

Keywords: optimal taxation; intertemporal wedge; habit formation; local substitution; 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 62996, mokuhn@uni-bonn.de
1 Introduction

The decision to accumulate assets today affects the incentive to supply labor in the future. As noted by Diamond and Mirrlees (1978), this link across time generates a rationale for taxing capital. This insight has been extended by Rogerson (1985) to general private information problems, and has become a cornerstone of the recent literature on ‘New Dynamic Public Finance’ (surveyed by Kocherlakota, 2010). The present paper argues that—in addition to assets—several further links across time may influence labor supply incentives. For instance, a worker’s preferences may depend on consumption experiences made in the past. Similarly, consumption goods can have a durable nature and act as a substitute for future consumption expenditures. Such intertemporal links have found broad support in the empirical literature (discussed below). As simple examples with different forms of time-nonseparability in Grochulski and Kocherlakota (2010) and in the present paper show, this generalization of preferences can reverse the standard taxation logic and call for negative or zero taxes on capital.\(^1\)

Motivated by these considerations, the present paper provides a general theoretical and quantitative analysis of optimal intertemporal wedges (implicit capital taxes) when individual preferences are time-nonseparable. Our contribution is threefold. First, we provide a decomposition of intertemporal wedges for 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 importance of habit formation for intertemporal wedges. In the class of time-nonseparable preferences, the case of habit formation seems particularly relevant for capital tax policy, as habit formation has reconciled theory and evidence for other important aspects of savings behavior under uncertainty such as the equity premium puzzle.\(^2\)

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

\(^1\) The two examples find nonpositive wedges even though consumption is substitutable across time in the example presented in the present paper, while it is complementary across time in the case of Grochulski and Kocherlakota (2010). Our decomposition result shows that different taxation motives are responsible for the negative intertemporal wedges in the two cases.

\(^2\) See Messinis (1999) for a review of habit formation in the macroeconomic literature.
information. To focus on the impact of time-nonseparabilities, we consider preferences that are additively separable between consumption and labor.\textsuperscript{3} The only restrictions imposed on lifetime consumption utility are monotonicity and concavity. This setup allows for a large range of special cases, including models of habit formation and models where the nonseparability arises from durable consumption. 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 \textit{implicit} taxes on capital and indicate that decentralizations of the social planning allocation must correct individual capital returns downward in one way or another.\textsuperscript{4}

By decomposing intertemporal wedges, we show that optimal capital taxes are determined by three forces when preferences are time-nonseparable. First of all, capital 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, capital should be taxed when consumption is complementary over time (as in habit formation models), because enhancing present consumption makes high consumption in the next period more attractive, which reinforces next period’s labor supply incentives (complementarity effect). By contrast, capital should be subsidized when consumption is substitutable over time (as in models with durable goods). Finally, capital should be taxed or subsidized, because the allocation of consumption between two periods affects labor supply incentives also in later periods (future incentive effect). The future incentive effect tends to counteract the complementarity effect. However, the future incentive effect depends in an intricate way on the details of preference nonseparabilities and the nature of the skill process, and is difficult to evaluate analytically, because it is not amenable to the perturbation approach by Rogerson (1985).

We identify special cases in which intertemporal wedges can be linked to economic fundamentals in a very simple way. In those cases, monotonicity of the marginal rates of intertemporal substitution suffices to guarantee that intertemporal wedges are positive. Our result shows in what sense the logic from the time-separable benchmark model extends to the present setup.

\textsuperscript{3}Nonseparable preferences between consumption and labor are explored by Golosov, Tsyvinski, and Werning (2006).

but it also highlights in what sense it does not. First, as long as the future incentive effect is negligible, intertemporal wedges are positive under assumptions significantly weaker than time-separability. Technically, this finding replaces the well-known Jensen’s inequality argument by a more general result on the positivity of the covariance of two monotonic functions. A particular case in which the future incentive effect is zero is the penultimate period of the model. On the other hand, however, if the future incentive effect is nonzero, the perturbation approach becomes infeasible, and the sign of the intertemporal wedge remains theoretically ambiguous.

The theoretical indeterminacy motivates our quantitative exercise. We evaluate the effect of time-nonseparable preferences on intertemporal wedges in a model parameterized to the U.S. economy. In line with empirical evidence on nonseparability at time-horizons of at least one year (Heaton, 1995), we explore preferences where time-nonseparability arises from habit formation. To the best of our knowledge, this is 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 even though the sufficient conditions for the emergence of positive intertemporal wedges are not satisfied, intertemporal wedges do remain positive for a wide range of habit parameters. As indicated by the decomposition result, we find a non-negligible negative future incentive effect. Compared to the time-separable case, average intertemporal wedges therefore drop by about 40 percent.

The paper proceeds as follows. The next section discusses the related literature. By means of two stylized examples, Section 3 shows that nonpositive intertemporal wedges can emerge both from intertemporal complementarity as well as from intertemporal substitutability of consumption. Section 4 sets up our general theoretical model. Section 5 decomposes intertemporal wedges into three parts and provides sufficient conditions for intertemporal wedges to be positive. Section 6 studies the quantitative importance of habit formation for intertemporal wedges, while Section 7 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, the preferences explored by Farhi and Werning (2008) do not have an expected utility representation, so their setup cannot be nested inside the present one (and vice versa). 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. Moreover, models with habit formation have contributed to the solution of macroeconomic puzzles related to consumption behavior (excess sensitivity puzzle, excess smoothness puzzle), asset pricing (equity premium puzzle) and the relationship between savings and growth; see the seminal papers by Ryder and Heal (1973), Abel (1990), Constantinides (1990), Campbell and Cochrane (1999) and the survey by 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 Two examples with nonpositive intertemporal wedges

By means of two simple examples, this section highlights how optimal savings distortions for time-nonseparable preferences differ from the time-separable case. Two key messages emerge. First, saving interacts with next period’s incentives not only via a wealth effect, but also via intertemporal substitution or complementarity effects. Second, saving may not only modify
incentive provision in the next period, but in fact in the entire future, which implies that savings distortions cannot be understood by considering two subsequent periods in isolation.

Example 1 (Intertemporal substitutability). Consider an agent who lives for two periods. He consumes in both periods and works only in the second period. The agent’s skill $\theta$ is private information and is distributed in the set $\{\theta_L, \theta_H\}$, where $\theta_L$ and $\theta_H > \theta_L$ both have nonzero probability. The agent learns his skill type at the beginning of the second period. Preferences over consumption $(c_1, c_2)$ and labor supply $l$ are given by

$$U(c_1, c_2) - V(l),$$

where $U$ is a concave function that is increasing in each argument, and $V$ is increasing and weakly convex. An agent with skill $\theta$ and labor supply $l$ produces $y = \theta l$ units of output.

A risk-neutral principal designs a consumption-output allocation $(c_1, c_2, y_i)_{i=L,H}$, where the index $i \in \{L, H\}$ refers to the agent’s skill type. The principal minimizes the costs of providing the agent with a given level of ex-ante utility subject to the constraint that skill is private information. To make the problem nontrivial, we assume $y_H > y_L$.

Suppose that the principal can save and borrow at an interest rate of zero. Moreover, suppose for a moment that preferences are time-separable, so that $U(c_1, c_2) = u(c_1) + u(c_2)$, 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 to save:

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

(2)

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, we say that consumption is substitutable over time. It is straightforward that at any (not necessarily optimal) allocation, we have the following identity:

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

(3)

Hence, in contrast to the time-separable case, the agent has no residual motive to save/borrow.
at an optimal allocation.

What is the difference between the two setups that is responsible for this result? The standard intuition for positive intertemporal wedges in the time-separable model is a wealth effect on labor supply incentives (Golosov, Kocherlakota, and Tsyvinski, 2003): if an agent brings more wealth into the period, then due to declining marginal utility he will work less in response to any given incentive scheme. Since this effect on the incentive problem is not internalized by the agent, undistorted saving cannot be optimal.

In principle, the same wealth effect is also at work in the nonseparable model. However, saving in period 1 now matters for incentives in period 2 not only via the level of wealth brought into that period, but also because \( c_1 \) has a direct impact on the preferences over consumption in period 2. In particular, when consumption is substitutable over time, there is a force that counteracts the wealth effect: higher saving means lower consumption in the first period; this raises the marginal utility of consuming in the second period, and so the agent becomes more responsive to the incentive scheme. In the above example, since the term \( U_{c_2}(c_1, c_2) - U_{c_1}(c_1, c_2) = 0 \) is completely independent of \( c_2 \), the wealth effect and the substitutability effect exactly offset each other, so that saving has a neutral impact on the preferences over date-2 consumption. 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

\[ U(c_0, c_1, c_2) = u(c_0) + u(c_1) + u(c_2 - c_1), \]

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, we say that consumption is *complementary* between dates 1 and 2.

As in the previous example, the agent works only in the final period. Allocations thus take the form \( (c_0, c_1, c_{2i}, y_i)_{i=L,H} \). Grochulski and Kocherlakota (2010) show that, under the
condition \( y_H > y_L \), any optimal allocation satisfies

\[
\frac{\mathbb{E}[U_{c_0}(c_0, c_1, c_2)]}{\mathbb{E}[U_{c_1}(c_0, c_1, c_2)]} = \frac{u'(c_0)}{u'(c_1) - \mathbb{E}[u'(c_2) - c_1]} > 1,
\]

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.

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 the agent is idle 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 General setup

This section sets up a general 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 and skill levels. Time is discrete and indexed by \( t = 1, 2, \ldots, T \), with \( T \in \mathbb{N} \cup \{\infty\} \).

### 4.1 Preferences

The agents have identical von-Neumann-Morgenstern preferences and maximize the expected value of

\[
U(c_1, \ldots, c_T) - V(l_1, \ldots, l_T)
\]

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 the partial derivatives of \( U \).
Notice that this setup allows for consumption complementarities across time (habit formation), \( U_{c_t,c_s} > 0 \) for \( t \neq s \), as well as consumption substitutabilities (durable consumption), \( U_{c_t,c_s} < 0 \) for \( t \neq s \), and combinations thereof. The setup also includes the time-separable case, \( U_{c_t,c_s} = 0 \) for \( t \neq s \), of course. Besides, the model allows for time-nonseparabilities in the preferences over hours, but this will be irrelevant for the questions addressed in this paper.\(^5\)

### 4.2 Skills

Agents differ with respect to their skills. An agent with hours \( l_t \) and skill \( \theta_t \) 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 \( \mu \) on the set \( \Theta^T := \Theta \times \cdots \times \Theta \), where \( \Theta \) is a finite subset of the positive reals. Denote the expectation operator with respect to \( \mu \) by \( \mathbb{E}[\cdot] \). At the beginning of period \( t \), every agent learns his current skill \( \theta_t \). 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 \( \text{cov}_t(\cdot, \cdot) \).

### 4.3 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.

### 4.4 Allocations

An allocation is a sequence \( (c, y) = (c_t, y_t)_{t=1}^T \) of consumption plans \( c_t : \Theta^T \rightarrow \mathbb{R}_+ \) and output plans \( y_t : \Theta^T \rightarrow \mathbb{R}_+ \) such that, for any period \( t \), \( c_t \) and \( y_t \) are functions of period-\( t \)

\(^5\)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.
information. That is, \(c_t\) and \(y_t\) are \(\theta^t\)-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 \(\sigma_t\) is \(\theta^t\)-measurable for all \(t\). Denote the truth-telling strategy by \(\sigma^*\), with \(\sigma^*(\theta^T) = \theta^T\) for all \(\theta^T\), and denote the set of all reporting strategies by \(\Sigma\). Since skills are privately observed, the planner must ensure that all agents reveal their information truthfully. Hence, allocations must satisfy the incentive compatibility constraint

\[
E \left[ U(c_1(\sigma^*), \ldots, c_T(\sigma^*)) - V \left( \frac{y_1(\sigma^*)}{\theta_1}, \ldots, \frac{y_T(\sigma^*)}{\theta_T} \right) \right] 
\geq E \left[ U(c_1(\sigma), \ldots, c_T(\sigma)) - V \left( \frac{y_1(\sigma)}{\theta_1}, \ldots, \frac{y_T(\sigma)}{\theta_T} \right) \right].
\]  

(7)

An allocation that satisfies (7) for all reporting strategies \(\sigma \in \Sigma\) is called incentive compatible.

4.5 Optimal allocations

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

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

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

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

(9)

5 Characterization 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 $(c_1(\theta^t), \ldots, c_T(\theta^T))$ is defined as the interest rate at which the agent is indifferent between saving and not saving. Formally, the shadow rate of return is
\[
E_t[U_c(c_1, \ldots, c_T)]
\]
\[
E_t[U_c(c_{t+1}, c_1, \ldots, c_T)].
\]

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
\[
\tau^K_t := 1 - \frac{E_t[U_c(c_1, \ldots, c_T)]}{R_t E_t[U_{c_{t+1}}(c_1, \ldots, c_T)]}.
\]

Note that $\tau^K_t$ is a random variable that depends on the date-$t$ history $\theta^t$ as indicated by the conditional expectations operator $E_t[\cdot]$. If the intertemporal wedge $\tau^K_t$ 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 $\tau^K_t$ is negative, we have an implicit subsidy.

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 $(c, y)$ be an optimal allocation. For any $t \in \{1, \ldots, T-1\}$, the intertemporal wedge equals $\tau^K_t = \gamma_t (A_t + B_t + C_t)$, where $\gamma_t$ is positive and where
\[
A_t = -\text{cov}_t \left( R_t U_{c_{t+1}}, \frac{1}{U_{c_{t+1}}} \right) \geq 0,
\]
\[
B_t = \text{cov}_t \left( U_{c_t}, \frac{1}{U_{c_{t+1}}} \right) \geq 0,
\]
\[
C_t = R_t - E_t \left[ \frac{U_{c_t}}{U_{c_{t+1}}} \right] \geq 0.
\]

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 savings 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

$$A_t = -\text{cov}_t \left( R_t U_{c_{t+1}}, \frac{1}{U_{c_{t+1}}} \right)$$  \hspace{1cm} (15)$$
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-nonseparabilities 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 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 (as in habit formation models), then $U_{c_t}$ is increasing in $c_{t+1}$, which implies that the term

$$B_t = \text{cov}_t \left( U_{c_t}, \frac{1}{U_{c_{t+1}}} \right)$$  \hspace{1cm} (16)$$
contributes positively to the intertemporal wedge. Intuitively, reducing consumption in period $t$ increases the relative attractiveness of low consumption in period $t + 1$ when consumption is complementary across time. 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 as in Example 1.\footnote{In Example 1, we have $A_1 > 0$, $B_1 = -A_1 < 0$ and $C_1 = 0$, which leads to $\tau^K = A_1 - A_1 = 0$.}
Finally, the intertemporal wedge has the component

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

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.\(^7\) More generally, we have the following result.

**Lemma 1.** Let \((c,y)\) 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 \( \theta^{t+1} \)-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 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).\(^8\) 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.\(^9\) Similar motives arise if consumption at date \( t + 1 \) is complementary with consumption at other future dates.\(^10\) 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.

\(^{7}\)If \( U(c_1, \ldots, c_T) = \sum \beta^{t-1} u(c_t) \), we can write \( C_t = u'(c_t) R_t (1/u'(c_t)) - \mathbb{E}_t [1/(\beta R_t u'(c_{t+1}))] \). The term in parentheses is equal to zero by the Inverse Euler equation from Rogerson (1985).

\(^{8}\)For the same reason, the Inverse Euler equation does not generalize to the time-nonseparable case; see Grochulski and Kocherlakota (2010).

\(^{9}\)This is precisely the reason why the intertemporal wedge \( \pi_0^f \) is negative in Example 2. Note that, since there is no uncertainty at date 1, both \( A_0 \) and \( B_0 \) are equal to zero, and thus \( C_0 = \pi_0^f < 0 \).

\(^{10}\)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 \( t + 1 \) will be more crucial for future incentives.
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.

In summary, intertemporal wedges for time-nonseparable preferences consist of three, partly opposing, components. This makes it difficult to determine the sign of intertemporal wedges for general specifications of time-nonseparability. Yet, as long as the future incentive effect is sufficiently small, the sign of intertemporal wedges can be linked to a simple property of the agent’s preferences as the next result shows.

**Proposition 2** (Positive intertemporal wedge). Let \((c, y)\) be an optimal allocation and \(t \in \{1, \ldots, T - 1\}\). In addition to the conditions of Lemma 1, suppose that the marginal rate of intertemporal substitution \(U_{c_{t+1}}/U_{c_t}\) is a decreasing function of \(c_{t+1}\). Then \(\tau^K_t \geq 0\) almost everywhere. Moreover, \(\tau^K_t > 0\) except when the associated history \(\theta^t\) has probability zero or when \(c_{t+1}(\theta^t, \theta_{t+1})\) is constant for almost all \(\theta_{t+1}\). 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 satisfied in the penultimate period in any setup with a finite time-horizon. Hence, intertemporal wedges are positive in the penultimate period for very general models of time-nonseparable preferences. The measurability assumption 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 $\theta_{t+1}$-measurable for time-nonseparable preferences.

6 Numerical analysis of a habit formation economy

As the previous discussion 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 the time-nonseparability arises from 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 when it comes to optimal taxation.

Our quantitative model captures some key features of the U.S. economy. In particular, the skill process is in line with 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 suppose 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 6.4. Even though skill fluctuations are still 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).
6.1 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 $\theta$ has the time-dependent probability weight $\pi_t(\theta)$, with $\sum_{\theta \in \Theta} \pi_t(\theta) = 1$, where $\Theta$ 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

$$U \left( \{ c_t \}_{t=1}^T \right) - V \left( \{ l_t \}_{t=1}^T \right) = \sum_{t=1}^T \beta^{t-1} \left( u(c_t, h_t) - v(l_t) \right), \quad h_t = H(c_{t-1}, h_{t-1}), \quad h_1 \text{ given}, \quad (18)$$

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 $[c, \bar{c}]$ and an output space of the type $[y, \bar{y}]$. We define $\text{dom}_t(h)$ 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 $\bar{c} \geq c_s \geq c$ and $\bar{y} \geq y_s \geq y$ for all $T \geq s \geq t$ and generates continuation utility

$$E \left[ \sum_{s=t}^T \beta^{t-1} \left( u(c_s, h_s) - v(y_s/\theta_s) \right) \right] = U, \quad \text{where } h_t = h, \ h_s = H(c_{s-1}, h_{s-1}) \text{ for } s > t. \quad (19)$$

Given the structure of our problem, we can express $\text{dom}_t(h)$ in closed form.
Lemma 2. For any \( h \in \mathbb{R}_+ \) and \( 1 \leq t \leq T \), the set \( \text{dom}_t(h) \) is a compact interval with bounds

\[
\max(\text{dom}_t(h)) = \sum_{s=t}^{T} \beta^{t-1} \left( u(c, h_s) - \sum_{\theta \in \Theta} \pi_s(\theta) v(y/\theta) \right), \quad \text{with} \quad h_t = h, \ h_s = H(\tau, h_{s-1}) \text{ for } s > t,
\]

\[
\min(\text{dom}_t(h)) = \sum_{s=t}^{T} \beta^{t-1} \left( u(c, h_s) - \sum_{\theta \in \Theta} \pi_s(\theta) v(y/\theta) \right), \quad \text{with} \quad h_t = h, \ h_s = H(\zeta, h_{s-1}) \text{ for } s > t.
\]

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

\[
C_1(U_1, h_1) := \min_{c,y} \mathbb{E} \left[ \sum_{t=1}^{T} \frac{c_t - y_t}{R^{t-1}} \right] \\
\text{s.t.} \quad \\
\mathbb{E} \left[ \sum_{t=1}^{T} \beta^{t-1} (u(c_t, h_t) - v(y_t/\theta_t)) \right] \geq \mathbb{E} \left[ \sum_{t=1}^{T} \beta^{t-1} (u(c_t(\sigma), h_t(\sigma)) - v(y_t(\sigma)/\theta_t)) \right] \quad \forall \sigma \in \Sigma \\
\mathbb{E} \left[ \sum_{t=1}^{T} \beta^{t-1} (u(c_t, h_t) - v(y_t/\theta_t)) \right] = U_1 \\
\tau \geq c_t \geq \xi \quad \forall t, \ \underline{y} \geq y_t \geq \overline{y} \quad \forall t, \quad h_t = H(c_{t-1}, h_{t-1}) \quad \forall t > 1, \ h_1 \text{ given.}
\]

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 \( C_{T+1} = 0 \)):

\[
C_t(U, h) = \min_{c,y,U'} \sum_{\theta \in \Theta} \pi_t(\theta) \left[ c(\theta) - v(y(\theta)) + \frac{1}{R} C_{t+1}(U'(\theta), H(c(\theta), h)) \right] \\
\text{s.t.} \quad u(c(\theta), h) - v(y(\theta)/\theta) + \beta U'(\theta) \geq u(c(\theta'), h) - v(y(\theta)/\theta) + \beta U'(\theta') \quad \forall \theta, \theta' \\
\sum_{\theta \in \Theta} \pi_t(\theta) \left[ u(c(\theta), h) - v(y(\theta)/\theta) + \beta U'(\theta) \right] = U \\
U'(\theta) \in \text{dom}_t(H(c(\theta), h)) \quad \forall \theta \\
\tau \geq c(\theta) \geq \xi, \ \underline{y} \geq y(\theta) \geq \overline{y} \quad \forall \theta
\]

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 (21) reaches beyond the case of Cobb-Douglas habit formation studied below. For instance, it includes the case of linear habit formation, $u(c_t, h_t) = \tilde{u}(c_t - \gamma h_t)$, 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 $u(c_t, h_t) = \tilde{u}(c_t + \delta h_t)$ 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 $(\text{dom}_t(h))_{h \in \mathbb{R}^+} \forall t = 1, \ldots, T$ following Lemma 2. We then exploit the functional equation (21) 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.

### 6.2 Parameters

There are $T = 10$ periods with a duration of 5 years each. Agents enter the model at age 18 and leave the model at the end of period 10 (aged 68). In the first period, agents have a deterministic skill level $\theta_1$. In period $t \geq 2$, skill $\theta_t$ 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.\footnote{Hansen (1993) uses average hourly earnings of a given age-subgroup divided by average hourly earnings of all subgroups as the relative efficiency measure.} 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 for agents in periods $t \geq 2$.

Table 1: Life-cycle profile of expected skills

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.560</td>
<td>0.754</td>
<td>0.912</td>
<td>1.034</td>
<td>1.119</td>
<td>1.168</td>
<td>1.180</td>
<td>1.156</td>
<td>1.095</td>
<td>0.999</td>
</tr>
</tbody>
</table>

Our preference specification follows Abel (1990), Carroll, Overland, and Weil (2000), and Diaz, Pijoan-Mas, and Rios-Rull (2003), and sets up habit formation in a Cobb-Douglas form:

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

(22)

where $\gamma$ 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 reference level, $c_t/h_t$, and the parameter $\gamma$ controls the relative importance of these two components.\(^{12}\) Period utility is of the CRRA type: $\tilde{u}(x) = x^{1-\sigma}/(1-\sigma)$. In line with Diaz, Pijoan-Mas, and Rios-Rull (2003), we choose parameters of $\sigma = 2$ and $\gamma = 0.75$. The value $\gamma = 0.75$ 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. The discount factor is $\beta = 0.965$ and the interest rate equals $R = 1/\beta$. The labor disutility function is $v(l) = \alpha l^{1+\frac{\psi}{\psi}}/(1 + \frac{1}{\psi})$, 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 $\lambda$ and is given by:

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

(23)

\(^{12}\) 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.
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.\(^{13}\)

We set the initial habit level to $h_1 = 0.7$. As we verify ex-post, this number coincides roughly 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 $c_l, \bar{c}, \underline{y}$, and $\bar{y}$ are never binding for the optimal allocation starting from this initial state.

### 6.3 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 ($\gamma = 0$).\(^{14}\) 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.

---

\(^{13}\)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 ($\lambda = 0.75$) as robustness checks.

\(^{14}\)To make the allocations comparable, we choose a scaling parameter of $\alpha = 2.75$ for the time-separable case, such that the discounted value of life-time output (and consumption) coincides with the habit formation case.
Figure 2a decomposes expected intertemporal wedges into the wealth effect, complementarity effect, and future incentive effect. The expected intertemporal wedge is positive for all periods. The wealth effect and complementarity effect are both positive, as shown by the formulas 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 longer than one period into the future then.

Habit formation has a quantitatively significant impact on intertemporal wedges (Figure 2b). The life-cycle average of the intertemporal wedge with habit formation is 0.019 (corresponding to an 11 percent reduction of the net interest rate). In the time-separable case it is 0.032. When we consider the sum of the wealth and complementarity effect, we find that it is almost as big as the intertemporal wedge in the time-separable case, and so the negative future incentive effect accounts for the observed difference.

The heterogeneity of intertemporal wedges is analyzed in Figure 3. For each period, we show minimum and maximum intertemporal wedges. We see that intertemporal wedges are positive not just in expectation, but for all possible histories at each point in time.

For habit formation as well as time-separable preferences, expected intertemporal wedges increase steeply towards the end of the agent’s life. The reason is that 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 the incentive problem becomes stronger over time, and this increases the intertemporal wedge; see also Golosov, Troshkin, and
Tsyvinski (2011).

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).

Farhi and Werning (2013) study optimal dynamic taxes in a life-cycle framework with time-separable preferences and a persistent skill process. They find intertemporal wedges that are smaller than the ones in our time-separable case. Part of this difference is arguably due to the lower coefficient of risk aversion in their model. Besides, incentives can be smoothed over more periods due to a retirement phase in Farhi and Werning (2013). This also tends to reduce intertemporal wedges.

6.4 Sensitivity analysis

To study the robustness of our results, we explore alternative parameters for the persistence of the habit process (\( \lambda = 0, \lambda = 0.75 \)), the coefficient of risk aversion (\( \sigma = 1, \sigma = 3 \)), 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 robustness checks point out 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 robustness checks show 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 the present one.

7 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 provide sufficient conditions for positive intertemporal wedges for special cases of time-nonseparability. 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

\[
\tau_t^K = 1 - \frac{E_t[U_{t+1}]}{R_t E_t[U_{t+1}] E_t \left[ \frac{1}{U_{t+1}} \right]}.
\]

(24)

Equivalently, we have

\[
\tau_t^K = \frac{R_t E_t[U_{t+1}]}{R_t E_t[U_{t+1}] E_t \left[ \frac{1}{U_{t+1}} \right]} \left[ E_t \left[ \frac{1}{U_{t+1}} \right] - E_t[U_{t+1}] E_t \left[ \frac{1}{U_{t+1}} \right] \right].
\]

(25)
After adding and subtracting a few terms, we obtain

$$\tau^K_t = \frac{R_t E_t[U_{ct+1}] E_t \left[ \frac{1}{U_{ct+1}} \right] - R_t + E_t \left[ \frac{U_{ct}}{U_{ct+1}} \right] - E_t[U_{ct}] E_t \left[ \frac{1}{U_{ct+1}} \right] + R_t - E_t \left[ \frac{U_{ct}}{U_{ct+1}} \right]}{R_t E_t[U_{ct+1}] E_t \left[ \frac{1}{U_{ct+1}} \right]}.$$

(26)

By the definition of a covariance, we have

$$\text{cov}_t \left( U_{ct+1}, \frac{1}{U_{ct+1}} \right) = 1 - E_t[U_{ct+1}] E_t \left[ \frac{1}{U_{ct+1}} \right],$$

(27)

$$\text{cov}_t \left( U_{ct}, \frac{1}{U_{ct+1}} \right) = E_t \left[ \frac{U_{ct}}{U_{ct+1}} \right] - E_t[U_{ct}] E_t \left[ \frac{1}{U_{ct+1}} \right].$$

(28)

This implies

$$\tau^K_t = \frac{-R_t \text{cov}_t \left( U_{ct+1}, \frac{1}{U_{ct+1}} \right)}{R_t E_t[U_{ct+1}] E_t \left[ \frac{1}{U_{ct+1}} \right]} + \frac{\text{cov}_t \left( U_{ct}, \frac{1}{U_{ct+1}} \right)}{R_t E_t[U_{ct+1}] E_t \left[ \frac{1}{U_{ct+1}} \right]} + \frac{R_t - E_t \left[ \frac{U_{ct}}{U_{ct+1}} \right]}{R_t E_t[U_{ct+1}] E_t \left[ \frac{1}{U_{ct+1}} \right]}$$

$$= \gamma_t (A_t + B_t + C_t),$$

(29)

(30)

where $$\gamma_t = \left( R_t E_t[U_{ct+1}] E_t \left[ \frac{1}{U_{ct+1}} \right] \right)^{-1}$$ is positive, since $$U_{ct+1} > 0$$.

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 $$1/U_{ct+1}$$ is an increasing function of $$c_{t+1}$$. Marginal utility $$U_{ct}$$ 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 $$\theta^t$$ that occurs with positive probability. Consider a perturbation that increases consumption at time $$t$$ by $$\xi$$ and reduces consumption at time $$t+1$$ by

$$\Box$$

26
ϕ. Then, depending on future shocks \((\theta_{t+1}, \ldots, \theta_T)\), the agent’s consumption utility changes by

\[
\Delta U = U\left(c_1(\theta^1), \ldots, c_T(\theta^T)\right) - U\left(c_1(\theta^1), \ldots, c_t(\theta^t) - \xi, c_{t+1}(\theta^{t+1}) + \phi, \ldots, c_T(\theta^T)\right) \tag{31}
\]

\[
= \int_{c_t(\theta^t) + \xi}^{c_t(\theta^t)} U_{c_t}(c_1(\theta^1), \ldots, \gamma, c_{t+1}(\theta^{t+1}), \ldots, c_T(\theta^T)) \, d\gamma \tag{32}
\]

\[
+ \int_{c_{t+1}(\theta^{t+1}) + \phi}^{c_{t+1}(\theta^{t+1})} U_{c_{t+1}}(c_1(\theta^1), \ldots, c_t(\theta^t) + \xi, \delta, \ldots, c_T(\theta^T)) \, d\delta. \tag{33}
\]

Since \(U_{c_t}\) and \(U_{c_{t+1}}\) are by assumption \(\theta^{t+1}\)-measurable, the above formula shows that \(\Delta U\) depends on \((\theta_{t+1}, \ldots, \theta_T)\) only through the variable \(\theta_{t+1}\). Hence we can find numbers \(\phi = \phi(\xi, \theta_{t+1})\) depending only on \(\xi\) and \(\theta_{t+1}\) 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 \((c, y)\) 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

\[
\xi - \frac{\mathbb{E}_t[\phi(\xi, \theta_{t+1})]}{R_t}, \tag{34}
\]

which yields the first-order condition

\[
1 - \frac{\mathbb{E}_t[U_{c_t}(c_1, \ldots, c_T)]}{R_t} = 0. \tag{35}
\]

Using the notation from Proposition 1, this implies \(C_t = 0\).

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

\[
-R_t \text{cov}_t\left(U_{c_{t+1}}, \frac{1}{U_{c_{t+1}}} \right) + \text{cov}_t\left(U_{c_t}, \frac{1}{U_{c_{t+1}}} \right) \geq 0. \tag{36}
\]

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

\[
-\mathbb{E}_t[U_{c_t}] + R_t\mathbb{E}_t[U_{c_{t+1}}] \geq 0. \tag{37}
\]
Since $R_t = E_t \left[ U_{c_t} / U_{c_{t+1}} \right]$, we can rewrite the previous line as

$$\text{cov}_t \left( -U_{c_{t+1}}, \frac{U_{c_t}}{U_{c_{t+1}}} \right) \geq 0.$$  \hfill (38)

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 $t, \ldots, T$ irrespective of his skill. Clearly all such allocations are incentive compatible. Setting $h_s = H(\tau, h_{s-1})$ for $s > t$, we have

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

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

$$\tilde{U}(\varsigma, \eta) = \sum_{s=t}^{T} \beta^{s-1} \left( u(\varsigma, h_s) - \sum_{\theta \in \Theta} \pi_s(\theta)v(y/\theta) \right)$$ \hfill (40)

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

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

$$\text{dom}_t(h) \supseteq [\tilde{U}(\varsigma, \eta), \tilde{U}(\tau, \eta)].$$  \hfill (41)

On the other hand, we clearly have

$$\text{dom}_t(h) \subseteq [\min(\text{dom}_t(h)), \max(\text{dom}_t(h))].$$  \hfill (42)
Using the results $\min(\text{dom}_t(h)) = \bar{U}(\underline{c}, \underline{y})$ and $\max(\text{dom}_t(h)) = \bar{U}(\bar{c}, \bar{y})$ from the first step, the two set inequalities taken together imply

$$\text{dom}_t(h) = [\bar{U}(\underline{c}, \underline{y}), \bar{U}(\bar{c}, \bar{y})].$$

(43)

This completes the proof.

\[\square\]

B Robustness checks

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 ($\lambda = 0.75$). 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 ($\sigma = 1$) and high ($\sigma = 3$) 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.
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.

with Asset Accumulation,” European University Institute. Mimeo. Available online at:
http://mpra.ub.uni-muenchen.de/38629/.


Explanation of Aggregate Stock Market Behavior,” Journal of Political Economy, 107(2),
205–251.


DIAZ, A., J. PIJOAN-MAS, AND J. RIOS-RULL (2003): “Precautionary savings and wealth dis-
tribution 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.


