On the taxation of durable goods

Sebastian Koehne*

May 11, 2016

Abstract

This paper proposes a Mirrleesian theory of commodity taxation in the presence of durable goods. Nondurable goods should be taxed uniformly provided that the preferences over nondurable consumption are weakly separable from labor effort. A uniform taxation across all goods is optimal if the utility from durable consumption is linear and the preferences are additively separable between durable consumption, nondurable consumption and labor effort. If those conditions are not met, diminishing marginal utilities and substitution effects justify the use of differential commodity taxes. To determine the sign of the tax differential, the paper combines the well-known Inverse Euler Equation with a novel *Substitution Euler Equation* that characterizes the marginal rate of substitution between durable and nondurable consumption across time. An application of this theory suggests that housing investment should face higher tax rates than nondurable consumption.

Keywords: optimal taxation; commodity taxation; durable goods; Atkinson–Stiglitz result; pre-committed goods; housing

JEL Classification: D82, H21

*Institute for International Economic Studies (IIES), Stockholm University, Sweden, Phone: +468163564, sebastian.koehne@iies.su.se, and CESifo, Germany. The paper has benefited from very insightful comments by Harold Cole, four anonymous referees, Johannes Becker, Pierre Boyer, Tobias Broer, Helmuth Cremer, Moritz Kuhn, Rick van der Ploeg and Nicolas Werquin. Sebastian Koehne gratefully acknowledges the financial support from the Knut och Alice Wallenbergs Stiftelse (grant 2012.0315).

1 Introduction

Approximately 40% of a typical household's consumption expenditure are spent on durable goods (housing, cars, furniture, consumer electronics, etc.).[1] How should durable goods be taxed? Despite the significance of durable goods for household consumption baskets, very little is known on their tax implications. Durable goods are not easily represented in standard taxation models because they combine aspects of nondurable goods and savings technologies—investing in a durable good yields a contemporaneous consumption flow as well as an entitlement to future flows. Describing this duality is impossible in static environments (e.g., Atkinson and Stiglitz, 1976) that, by construction, cannot distinguish between stock and flow variables. Moreover, common approaches to savings taxation are informative on tax differentials across time, but provide little guidance for intratemporal taxation because they consider environments with a single consumption good.[2]

The present paper addresses the consequences of durable goods for optimal taxation. The paper sets up an explicit model of durable goods and proposes a dynamic Mirrleesian theory of commodity taxation when durable and nondurable goods coexist. The main findings can be summarized as follows.

First, I show that *nondurable* goods should be taxed uniformly provided that the preferences over nondurable consumption are weakly separable from labor effort (Proposition 1). Stated differently, the Atkinson–Stiglitz result (Atkinson and Stiglitz, 1976) on uniform commodity taxation holds true for nondurable goods in dynamic frameworks.[3]

Second, as a theoretical benchmark, I derive a maximal case in which *all* goods should be taxed uniformly (Proposition 2). If the utility from durable consumption is linear and the preferences are additively separable between nondurable consumption, durable consumption and labor effort, a uniform taxation across all goods (nondurable goods and investment in durable goods) is optimal. This result is sharp. If any of its assumptions is relaxed, a uniform commodity taxation is no longer optimal in general (Proposition 3).

[1]The average annual expenditure on durable goods (shelter, household furnishings and equipment, apparel, vehicles, entertainment equipment) in the Consumer Expenditure Survey (CEX) 2011 is $25,390. The average annual total expenditure amounts to $63,972.

[2]For example, see Diamond and Mirrlees (1978), Rogerson (1985), Kocherlakota (2005), Albanesi and Sleet (2006), Farhi and Werning (2013), Golosov, Troshkin, and Tsyvinski (2015) and Abraham, Koehne, and Pavoni (2016).

[3]This finding extends a result by Golosov, Kocherlakota, and Tsyvinski (2003) to a model where durable and nondurable goods coexist.

Third, I study the properties of optimal differential commodity taxation when consumption utility is strictly concave. I derive a *Substitution Euler Equation* that characterizes the marginal rate of substitution between durable and nondurable consumption across time (Proposition 4). Building on the Substitution Euler Equation equation and the Inverse Euler Equation (Rogerson, 1985), I characterize the optimal tax wedge between durable investment and nondurable consumption (Proposition 5). I show that differential commodity taxation is beneficial even when the consumption preferences are additively separable from labor, because durable investment interferes with incentive problems later in life through diminishing marginal utilities and substitution effects. Diminishing marginal utilities imply that investing in a durable good modifies the valuation of future investment. Substitution effects stem from nonseparabilities across goods and capture how an investment in a durable good affects the future valuation of other consumption goods.

Specifically, if there is one durable and one nondurable good and these are (Edgeworth) substitutes,[4] I show that investment in the durable good should be implicitly taxed at a higher rate than the nondurable good provided that the consumption of these goods is sufficiently monotonically related. I demonstrate that the monotonicity property can be taken for granted if preferences are homothetic. Moreover, I construct an explicit tax system that implements optimal allocations as a competitive equilibrium with taxes.

Finally, I apply the model to the case of housing. Based on recent estimations of the preferences for housing, I find evidence that housing and nondurable consumption are Edgeworth substitutes. Thus, the theory in this paper suggests that housing investment should face higher tax rates than nondurable goods. I provide a stylized numerical example to illustrate the role of housing taxation in the present framework.

The paper proceeds as follows. The remainder of this section surveys the related literature. Section 2 sets up a multi-period optimal tax problem with durable and nondurable goods. Section 3 explores a special case in which all commodity wedges are zero. Section 4 studies differential commodity taxation in a setting with one durable and one nondurable good. Section 5 applies the model to the case of housing. Section 6 presents concluding remarks and discusses some extensions of the model. Appendix A collects the proofs of all theoretical results.

[4]Two goods are Edgeworth substitutes if the utility function has a negative cross derivative with respect to these goods. Since von Neumann–Morgenstern utility functions are unique up to positive affine transformations, the notion of Edgeworth substitutability does not depend on the representation of preferences.

1.1 Related literature

This paper relates to the vast literature on commodity taxation that emerges from the analysis by Atkinson and Stiglitz (1976). The Atkinson–Stiglitz theorem on uniform commodity taxation considers a static environment that, by construction, cannot distinguish between durable and nondurable goods. Golosov et al. (2003) extend the Atkinson–Stiglitz theorem to a dynamic environment under the assumption that all goods in every period provide utility in the given period only. Therefore, neither the original Atkinson–Stiglitz result nor the dynamic extension shed any light on the case of durable goods.

The present paper is closely related to the study of pre-committed goods by Cremer and Gahvari (1995a,b). By proposing an explicit model of durability in a dynamic framework, I show that the tax implications of durable goods differ considerably from those of pre-committed goods. Note that pre-committed goods are goods that are chosen before the resolution of uncertainty. Thus, pre-commitment relates more to the timing of consumption decisions than to the durability of a good. The discussion in Section 6.1 compares durable goods and pre-committed goods in more detail.

Further arguments for differential commodity taxes are proposed in particular by Christiansen (1984), Naito (1999) and Saez (2002). Christiansen (1984) analyzes differential taxation when some goods are positively or negatively related to leisure.[5] Naito (1999) studies an economy with two production sectors that differ in their skill intensity. If sector-specific income taxation is impossible, differential commodity taxes can help to redistribute toward low-skilled workers by affecting the wage distribution. Saez (2002) shows that differential commodity taxation can be desirable when preferences are heterogeneous. To uncover the novel effects of durable goods, the present paper abstracts from the previous mechanisms and studies a one-sector economy in which agents have homogenous preferences that are (weakly or additively) separable between consumption and labor.[6]

As shown by da Costa and Werning (2002), the role of commodity taxation is typically very similar for hidden action models and adverse selection models. In fact, the analysis in the present paper rests on variational arguments that change the consumption allocation but leave

[5]Relatedly, Jacobs and Boadway (2014) study optimal *linear* commodity taxation when the preferences are not weakly separable between consumption and labor.

[6]The production side of the economy corresponds to a single sector in which effective labor inputs are perfectly substitutable irrespective of skill.

consumption utility and labor effort unaffected. Therefore, the present results hold under very general specifications of informational frictions, including frameworks that combine hidden actions and adverse selection.

Grochulski and Kocherlakota (2010) and Koehne and Kuhn (2015) analyze labor and savings taxation when the consumption preferences are time-nonseparable because of habit formation. Durable goods also generate a particular form of a time-nonseparability. Thus, decentralizations of constrained efficient allocations typically rely on retrospective tax instruments in both frameworks. However, despite this similarity, the implications of durable goods for optimal commodity taxation differ crucially from those of habit formation models. First, habit formation alone does not justify differential commodity taxes. Differential taxes can only be helpful if the habit formation process differs across goods. However, to date, there is little empirical evidence to support that view. In contrast, the durability of goods can be clearly distinguished and measured. Second, habits are exogenous functions of past individual or aggregate consumption. Hence, although habits enter the utility function, they cannot be traded off against other goods because they are outcomes rather than decision variables. In contrast, there is generally some degree of substitutability between durable and nondurable goods. The variational arguments in the present paper exploit precisely this substitution margin.

2 Model

This section introduces durable goods into a dynamic Mirrleesian taxation problem similar to that of Golosov et al. (2003).

2.1 Preferences

There is a continuum of agents with identical, time-separable von Neumann–Morgenstern preferences. The agents live for $T \geq 2$ periods and discount the future at the rate $\beta \in (0,1)$. Their period utility depends on a vector $c_t \in \mathbb{R}_+^N$ of nondurable consumption goods, a vector $s_t \in \mathbb{R}_+^M$ of service flows from durable consumption goods, and labor effort $e_t \in \mathbb{R}_+$. The utility function is $U : \mathbb{R}_+^{N+M+1} \to \mathbb{R}$, $(c_t, s_t, e_t) \mapsto U(c_t, s_t, e_t)$, with $N, M \geq 1$. The utility function is strictly increasing in the first $N + M$ arguments, strictly decreasing in the last argument, and continuously differentiable in the first $N + M$ arguments on $\mathbb{R}_{++}^{N+M} \times \mathbb{R}_+$. For $k = 1, \ldots, N +$

M, the subscript notation U_k represents the partial derivative of U with respect to the k-th argument.

2.2 Durable goods

Durable goods generate service flows s_t proportional to the (individual-specific) stocks of durable goods $d_t \in \mathbb{R}^M_+$. More specifically, $s_t = \rho d_t := (\rho_1 d_{t,1}, \ldots, \rho_M d_{t,M})$, where $\rho \in \mathbb{R}^M_{++}$ is a vector of proportionality coefficients and ρd_t denotes the point-wise product of vectors ρ and d_t. The stocks of durable goods depreciate over time and can be adjusted by investment: $d_t = \delta d_{t-1} + i_t$, where $\delta d_{t-1} := (\delta_1 d_{t-1,1}, \ldots, \delta_M d_{t-1,M})$ is the point-wise vector product, $\delta \in (0,1)^M$ represents depreciation and $i_t \in \mathbb{R}^M$ denotes investment. Negative investment in durable goods is feasible but the stocks are required to remain nonnegative at all times. By allowing for negative investment, the model in particular includes the possibility that stocks of durable goods are reallocated across agents. The initial stocks of durable goods are identical for all agents and normalized to $d_0 = 0$. After period T, there is no activity and the stocks of durable goods vanish.

2.3 Uncertainty

Agents face idiosyncratic uncertainty regarding their productivity (or skill) $\theta_t \in \Theta \subset \mathbb{R}_{++}$. To sidestep some formalities on the measurability and integrability of random variables, I assume that the productivity set Θ is a *finite* subset of $\mathbb{R}_{++}$.[7] For $t = 1$, productivity θ_1 is distributed with probability weights $\pi_1(\theta_1) > 0$, with $\sum_{\theta_1 \in \Theta} \pi_1(\theta_1) = 1$. For $t > 1$, the productivity has the conditional probability weights $\pi_t\left(\theta_t | \theta^{t-1}\right) > 0$, where $\theta^{t-1} = (\theta_1, \ldots, \theta_{t-1}) \in \Theta^{t-1}$ denotes the history of productivities before period t, and $\sum_{\theta_t \in \Theta} \pi_t\left(\theta_t | \theta^{t-1}\right) = 1$ for all θ^{t-1}. The unconditional probability of a history θ^t is given by $\Pi^t(\theta^t) := \pi_1\left(\theta_1\right) \pi_2\left(\theta_2 | \theta^1\right) \cdots \pi_t\left(\theta_t | \theta^{t-1}\right)$. The distribution Π^t has full support for all t.

As usual in this class of models, I assume that a law of large numbers applies. The individual distribution of uncertainty is thus identical with the realized cross-sectional distribution. The expectation operator with respect to the unconditional distribution of skill histories θ^T is

[7] All results in this paper can be extended to productivity sets that are continuous intervals or infinite, countable sets. However, such extensions would complicate the exposition and introduce some technical issues without adding economic insight.

denoted by $\mathbb{E}[\cdot]$. The notation $\mathbb{E}_t[\cdot] := \mathbb{E}\left[\cdot|\theta^t\right]$ represents expectations conditional on the period-t history θ^t. Similarly, $\text{cov}_t(\cdot,\cdot)$ represents covariances conditional on the period-t history.

An agent with productivity θ_t and labor effort e_t generates $y_t = \theta_t e_t$ efficiency units of labor. Productivity and labor effort are private information, whereas effective labor y_t is publicly observable. A natural interpretation of this framework is that θ_t represents the wage rate and labor effort e_t represents the intensive margin of labor supply. The social planner (tax authority) only observes annual income y_t but not how productive a worker is nor how much labor the worker supplied.

2.4 Allocations

In addition to consumption goods and labor, there is a capital good. The social planner owns the capital stock and has a given initial capital endowment $\bar{K}_1 > 0$.

An *allocation* is a collection $(\mathbf{c}, \mathbf{d}, \mathbf{y}, \mathbf{K}) = (c_t, d_t, y_t, K_t)_{t=1}^T$ of the following objects for each t:

$$K_t \in \mathbb{R}_+, \qquad c_t : \Theta^t \to \mathbb{R}_+^N, \qquad d_t : \Theta^t \to \mathbb{R}_+^M, \qquad y_t : \Theta^t \to \mathbb{R}_+.$$

Here, K_t represents the aggregate capital stock, c_t denotes the bundle of nondurable consumption goods, d_t denotes the stocks of durable goods, and y_t represents effective labor. The last three objects are functions of the time-t history θ^t. Each allocation of durable stocks generates a unique sequence of individual-specific investments $i_t : \Theta^t \to \mathbb{R}^M$, $t = 1, \ldots, T$, and service flows $s_t : \Theta^t \to \mathbb{R}_+^M$, $t = 1, \ldots, T$, according to the identities $i_t = d_t - \delta d_{t-1}$ and $s_t = \rho d_t$ from above.

Under standard assumptions on preferences, consumption will be nonzero. This motivates the following definition.

Definition 1. An allocation $(\mathbf{c}, \mathbf{d}, \mathbf{y}, \mathbf{K})$ has *interior consumption* if there exists a scalar $\epsilon > 0$ with $c_t(\theta) \geq \epsilon$ and $d_t(\theta) \geq \epsilon$ for all t and all θ^t.

The social planner operates a general production technology that takes capital K_t and aggregate labor $\mathbb{E}[y_t]$ as inputs and produces nondurable consumption goods, investment in durable consumption goods, and future capital K_{t+1} as outputs. An allocation is *feasible* if

$$G\left(\mathbb{E}\left[c_t\right], \mathbb{E}\left[i_t\right], K_{t+1}, K_t, \mathbb{E}\left[y_t\right]\right) \leq 0 \quad \text{for all } t,$$

with the convention $K_{T+1} = 0$. The function $G : \mathbb{R}^{N+M+3} \to \mathbb{R}$ is continuously differentiable, strictly increasing in the first $N + M + 1$ arguments and strictly decreasing in the remaining two arguments. As usual, for $k = 1, \ldots, N + M + 3$, the subscript notation G_k represents the partial derivatives of G.

For example, the technology may be defined by a production function F that combines capital and labor to produce a final good and the final good is converted one-to-one into capital or any of the consumption goods:

$$G\left(C, I, K', K, Y\right) = \sum_{n=1}^{N} C_n + \sum_{m=1}^{M} I_m + K' - (1 - \delta^K)K - F(K, Y).$$

In particular, if $\delta^K = 1$ and $F(K, Y) = (1 + r)K + Y$, capital corresponds to a savings technology with exogenous return r and the production technology is linear in labor.

2.5 Optimal allocation problem

Because labor effort and productivity are private information, allocations need to satisfy incentive compatibility conditions. By the revelation principle, one can restrict the attention to direct mechanisms where the agents report their productivities to the planner, who then allocates consumption and labor. A *reporting strategy* is a sequence $\sigma = (\sigma_t)_{t=1,\ldots,T}$ of mappings $\sigma_t :$ $\Theta^t \to \Theta$. Denote the set of all reporting strategies by Σ and set $\sigma^t(\theta^t) := \left(\sigma_1(\theta^1), \ldots, \sigma_t(\theta^t)\right)$. A reporting strategy $\sigma \in \Sigma$ yields ex ante expected utility according to

$$w\left(\mathbf{c} \circ \sigma, \mathbf{d} \circ \sigma, \mathbf{y} \circ \sigma\right) := \sum_{t=1}^{T} \beta^{t-1} \mathbb{E}\left[U\left(c_t\left(\sigma^t\left(\theta^t\right)\right), \rho d_t\left(\sigma^t\left(\theta^t\right)\right), \frac{y_t\left(\sigma^t\left(\theta^t\right)\right)}{\theta_t}\right)\right].$$

An allocation is *incentive compatible* if no agent has an incentive to misreport the productivity, i.e., if

$$w\left(\mathbf{c}, \mathbf{d}, \mathbf{y}\right) \geq w\left(\mathbf{c} \circ \sigma, \mathbf{d} \circ \sigma, \mathbf{y} \circ \sigma\right) \quad \text{for all } \sigma \in \Sigma.$$

An allocation is *incentive feasible* if it is incentive compatible and feasible.

The planner has the ability to commit ex ante to an allocation. The function $\chi : \Theta \to \mathbb{R}_+$, with $\chi(\theta_1) > 0$ for at least one θ_1, defines the planner's Pareto weights based on the initial productivities. Given the capital endowment $\bar{K}_1$, the planner solves the following problem:

$$V\left(\bar{K}_1\right) = \sup_{\mathbf{c},\mathbf{d},\mathbf{y},\mathbf{K}} \sum_{t=1}^{T} \beta^{t-1} \mathbb{E}\left[\chi(\theta_1) U\left(c_t\left(\theta^t\right), \rho d_t\left(\theta^t\right), \frac{y_t\left(\theta^t\right)}{\theta_t}\right)\right] \tag{1}$$
$$\text{s.t. } (\mathbf{c}, \mathbf{d}, \mathbf{y}, \mathbf{K}) \text{ is incentive feasible; } K_1 \leq \bar{K}_1.$$

An allocation $(\mathbf{c}^*, \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}^*)$ is called *optimal* if the allocation is incentive feasible, satisfies $K_1^* \leq \bar{K}_1$ and solves

$$V\left(\bar{K}_1\right) = \sum_{t=1}^{T} \beta^{t-1} \mathbb{E}\left[\chi(\theta_1) U\left(c_t^*\left(\theta^t\right), \rho d_t^*\left(\theta^t\right), \frac{y_t^*\left(\theta^t\right)}{\theta_t}\right)\right]. \tag{2}$$

Throughout the paper, a maintained assumption is that V is finite.

2.6 Monotonicity with respect to initial capital

Some results in this paper will rely on the assumption that V, the optimized value of social welfare, is strictly increasing in the initial capital endowment $\bar{K}_1$.[8] By construction, V is weakly increasing in $\bar{K}_1$. As shown by the next result, V is strictly increasing in $\bar{K}_1$ under a common assumption on preferences.

Lemma 1. *Suppose that $U(c,s,e) = u(c,s) - v(e)$, where u is strictly increasing and continuous. Then, $V\left(\bar{K}_1\right) < V\left(\bar{K}_1'\right)$ for all $\bar{K}_1 < \bar{K}_1'$.*

Lemma 1 follows from the same logic as in the case without durable goods (Golosov et al., 2003) and a formal proof is thus omitted. The main idea of the proof is as follows. Suppose that, contrary to the claim, $V\left(\bar{K}_1\right) = V\left(\bar{K}_1'\right)$ for some $\bar{K}_1 < \bar{K}_1'$. Then, an allocation that solves $V\left(\bar{K}_1\right)$ is also optimal for the problem $V\left(\bar{K}_1'\right)$ but does not use all initial capital. One can therefore use the spare resources and slightly increase the consumption of one nondurable good in the first period. The increase can be done in such a way that the consumption utilities of all agents raise by the same amount in the first period. All consumption utilities in later periods remain fixed and all differences in lifetime consumption utilities across realizations

[8]This issue can be sidestepped if the planner problem is set up as a cost minimization problem rather than a welfare maximization problem.

remain fixed, too. Hence, by the additive separability of preferences, the modified allocation is still incentive compatible. Because the modified allocation yields more social welfare than the original allocation, the assumption $V(\bar{K}_1) = V(\bar{K}_1')$ must be false and hence $V(\bar{K}_1) < V(\bar{K}_1')$ must hold.

3 Benchmark analysis: Extension of the Atkinson–Stiglitz theorem

Atkinson and Stiglitz (1976) show that, if the preferences over consumption goods are weakly separable from labor, optimal static allocations are associated with a uniform taxation of commodities. Intuitively, this result can be explained as follows. Under weak separability, consumption decisions for a given level of income do not depend on the type of the agent. Hence, differential commodity taxes do not help to ease the labor-leisure distortion because they cannot relax incentive constraints that require different types to generate different incomes. However, differential commodity taxes clearly impose a distortion on the consumption choice. Therefore, provided that income taxes are set optimally, differential commodity taxes will create a distortion with no redistributive benefit.

In subsequent work, Konishi (1995) and Kaplow (2006) demonstrate that differential commodity taxes are dispensable even when income taxes are not optimal. Laroque (2005) provides an elementary proof of the Atkinson–Stiglitz theorem. Deaton (1979) shows that, when only linear income taxes are available, linear Engel curves (homothetic consumption preferences) are additionally required for a uniform taxation result. Closely related to the present paper, Golosov et al. (2003) extend the Atkinson–Stiglitz theorem to a dynamic model with *nondurable* consumption goods.

In this section, I explore whether the Atkinson–Stiglitz theorem generalizes to dynamic frameworks with investment in durable consumption goods. I show that significantly stronger conditions are required to guarantee the absence of commodity wedges in these frameworks.

3.1 Definition of commodity wedges

As usual in the literature on optimal dynamic taxation, the decentralization of allocations in dynamic Mirrleesian models is not unique. Therefore, the robust predictions from these models are about wedges (implicit tax distortions), not about explicit tax instruments.[9] Generally, these wedges measure the magnitude and sign of a distortion to an individual decision margin relative to an allocation without government intervention. Throughout the paper, I study commodity wedges that measure the extent to which the commodity choice (across goods within a given period) at opimal allocations is distorted. When these wedges are zero, it is suboptimal to impose commodity taxes (or subsidies) that differ across goods.

Defining the wedge for the choice between two nondurable goods is straightforward. Formally, the *commodity wedge* $\tau^t_{n,n'}$ between two nondurable goods $n, n' \in \{1, \ldots, N\}$ in period t measures the gap between the marginal rate of substitution and the marginal rate of transformation:

$$\frac{U_n\left(c_t, s_t, \frac{y_t}{\theta_t}\right)}{U_{n'}\left(c_t, s_t, \frac{y_t}{\theta_t}\right)} = \left(1 + \tau^t_{n,n'}\right) \frac{G_n\left(\mathbb{E}\left[c_t\right], \mathbb{E}\left[i_t\right], K_{t+1}, K_t, \mathbb{E}\left[y_t\right]\right)}{G_{n'}\left(\mathbb{E}\left[c_t\right], \mathbb{E}\left[i_t\right], K_{t+1}, K_t, \mathbb{E}\left[y_t\right]\right)}. \tag{3}$$

In a laissez faire scenario, the marginal rate of substitution $U_n/U_{n'}$ coincides with the marginal rate of transformation $G_n/G_{n'}$ and the commodity wedge is zero. When the wedge differs from zero, the corresponding allocation cannot be decentralized without manipulating the individual transformation possibilities between the goods. The definition in Eq. (3) implies more specifically that, to make the allocation compatible with (first-order) optimal individual substitution decisions, the marginal opportunity cost of good n in terms of n' needs to be distorted by a factor of $1 + \tau^t_{n,n'}$. Therefore, the wedge $\tau^t_{n,n'}$ represents an implicit tax rate on the relative marginal price of good n in terms of good n'. In particular, if $\tau^t_{n,n'} > 0$, good n is implicitly taxed at a higher rate than good n', and the opposite happens if $\tau^t_{n,n'} < 0$.[10]

Note that the commodity wedges between two goods may differ across agents at any given point in time. To see this possibility, recall that the allocation objects (c_t, s_t, y_t) depend on the history of realizations. Moreover, the skill level θ_t is also a random variable. Therefore, commodity wedges are generally stochastic objects in the present environment.

The same concept of a commodity wedge also extends to the investment in durable goods. However, expressing the marginal rate of substitution becomes slightly more complicated because an investment in a durable good generates service flows in all remaining periods. Holding all other variables fixed, the marginal utility from investing in durable good m at time t is

[9] For instance, see Golosov et al. (2003) for a more detailed discussion of this point.

[10] Note that, by construction, $\tau^t_{n',n} = 1/(1 + \tau^t_{n,n'}) - 1$. Hence, $\tau^t_{n,n'}$ is positive if $\tau^t_{n',n}$ is negative, and vice versa.

the sum of an immediate flow and an expected discounted future flow:

$$\rho_m U_{N+m}\left(c_t, s_t, \frac{y_t}{\theta_t}\right) + \rho_m \mathbb{E}_t\left[\sum_{k=t+1}^{T} (\beta\delta_m)^{k-t}\, U_{N+m}\left(c_k, s_k, \frac{y_k}{\theta_k}\right)\right].$$

Here, recall that U_{N+m} is the partial derivative of U with respect to the service flow from the m-th durable good, δ_m represents the depreciation of that good, and ρ_m maps stocks to service flows. With these preparations in mind, I define the *commodity wedge* $\tau^t_{N+m,n}$ between investment in the durable good $m \in \{1, \ldots, M\}$ and the nondurable good $n \in \{1, \ldots, N\}$ in period t using the following equation:

$$\begin{aligned}&\frac{\rho_m U_{N+m}\left(c_t, s_t, \frac{y_t}{\theta_t}\right) + \rho_m \mathbb{E}_t\left[\sum_{k=t+1}^{T} (\beta\delta_m)^{k-t}\, U_{N+m}\left(c_k, s_k, \frac{y_k}{\theta_k}\right)\right]}{U_n\left(c_t, s_t, \frac{y_t}{\theta_t}\right)} \\ &= \left(1+\tau^t_{N+m,n}\right)\frac{G_{N+m}\left(\mathbb{E}\left[c_t\right], \mathbb{E}\left[i_t\right], K_{t+1}, K_t, \mathbb{E}\left[y_t\right]\right)}{G_n\left(\mathbb{E}\left[c_t\right], \mathbb{E}\left[i_t\right], K_{t+1}, K_t, \mathbb{E}\left[y_t\right]\right)}.\end{aligned} \tag{4}$$

Analogous to the interpretation of the wedge between two nondurable goods, the commodity wedge $\tau^t_{N+m,n}$ represents an implicit tax rate on the relative marginal price of investment in the durable good m in terms of the nondurable good n. For example, if $\tau^t_{N+m,n} > 0$, investment in good m is implicitly taxed at a higher rate than the purchase of good n.

Consistent with the previous terminology, the *commodity wedge* $\tau^t_{N+m,N+m'}$ between the investment in two durable goods $m, m' \in \{1, \ldots, M\}$ can be inferred from the definition in Eq. (4) using the formula

$$\tau^t_{N+m,N+m'} = \frac{1+\tau^t_{N+m,n}}{1+\tau^t_{N+m',n}} - 1 \tag{5}$$

for any nondurable good $n \in \{1, \ldots, N\}$.

Importantly, the definitions in Eqs. (4) and (5) consider an *investment* in the stock of a durable good, not the purchase of a one-time service flow from that good. This distinction is crucial for the following analysis. In fact, if service flows from durable goods could be purchased directly and adjusted on a period-by-period basis without any friction, durable goods and nondurable goods would become equivalent for the agents. However, a framework with frictionless spot markets for durable services seems to be a poor description of reality. Even when durable goods are rented or leased, the underlying contracts typically bind the agents for some time. Moreover, consumers often face transaction costs when they change the provider of a rental service and/or the details of the service. Therefore, the rental of a durable good resembles an intermediate case between the formal concepts of investing in the stock of a durable good and purchasing a nondurable good. Following this interpretation, although the present model primarily studies the possibility to invest in durable goods, the findings may also be suggestive for frameworks where durable services can be rented or leased.

For durable goods, the commodity wedges are neither pure intratemporal wedges, nor pure intertemporal wedges. Although the commodity wedges measure the distortion of a decision margin between two goods at a given point in time, the definition involves allocation variables in future periods because the service flow from investing in durable goods is generally dynamic. In the limiting case where durable goods fully depreciate from one period to the next, investing in a durable good will generate only one instantaneous service flow and the commodity wedge reduces to a standard intratemporal commodity wedge familiar from static environments. Beyond that special case, however, investing yields a dynamic service flow and, hence, the commodity wedge obtains an intertemporal aspect. In this sense, the commodity wedge resembles the well-known concept of an intertemporal (savings) wedge. Note that the intertemporal wedge τ_s in models with a single, nondurable consumption good is commonly defined as

$$\frac{u'(c_t)}{\beta\mathbb{E}\left[u'(c_{t+1})\right]} = (1-\tau_s)R_{t+1}$$

where u' is the marginal utility of nondurable consumption, and R_{t+1} the interest rate between periods t and $t+1$ (the marginal rate of transformation across periods). As a comparison with Eq. (4) reveals, the commodity wedge for durable goods generalizes the standard concept of an intertemporal wedge, because it considers an investment in a stock (or "real asset") that yields service flows (in a specific consumption good) directly upon investment and in all subsequent periods. By contrast, saving in a one-period asset (as is commonly assumed) yields a flow only in the period immediately following.

The similarity to intertemporal wedges suggests that commodity wedges on durable goods do not in general translate into simple intratemporal transaction taxes. In particular, as demonstrated in Section 4.3 below, retrospective tax systems can be useful to affect the decision to invest in durable goods. These tax systems resemble formulations of savings taxes that depend on the information available at the time when assets pay off, not only on the information at the time when agents save (e.g., Kocherlakota, 2005; Albanesi and Sleet, 2006; Grochulski and Kocherlakota, 2010).

3.2 Results on the absence of commodity wedges

By construction, the commodity wedges are zero if and only if the marginal rates of substitution coincide with the marginal rates of transformation. First, I demonstrate that the wedges between *nondurable* goods are zero if the preferences over those goods are weakly separable from labor.

Definition 2. The preferences over nondurable goods are *weakly separable from labor* if there exists a function $u : \mathbb{R}_+^{N+M} \to \mathbb{R}$, strictly increasing and continuously differentiable (on the interior of its domain) in the first N arguments , and a function $\tilde{U} : \mathbb{R} \times \mathbb{R}_+^{M+1} \to \mathbb{R}$, strictly increasing and continuously differentiable in the first argument, such that $U(c, s, e) = \tilde{U}\left(u(c, s), s, e\right)$ for all $(c, s, e) \in \mathbb{R}_+^{N+M+1}$.

Proposition 1 (Nondurable goods)**.** *Suppose that $V\left(\bar{K}_1\right) < V\left(\bar{K}_1'\right)$ for all $\bar{K}_1 < \bar{K}_1'$. Suppose that the preferences over nondurable goods are weakly separable from labor. Then, for any optimal allocation with interior consumption, the commodity wedge between any two nondurable goods is zero in all periods.*

The proof of Proposition 1 and all further proofs are relegated to the appendix. Unlike earlier findings on uniform commodity taxation, Proposition 1 allows nondurable goods to coexist with durable goods. Following a logic similar to that of Atkinson and Stiglitz (1976) and Golosov et al. (2003), the proposition shows that a uniform taxation across *nondurable* goods remains optimal when the preferences over nondurable goods are weakly separable from labor.

In the last period, the distinction between durable and nondurable goods vanishes. Hence, akin to Proposition 1, I obtain the following result.

Remark 1 (Final period)**.** *Suppose that $V\left(\bar{K}_1\right) < V\left(\bar{K}_1'\right)$ for all $\bar{K}_1 < \bar{K}_1'$. Suppose that the preferences over durable and nondurable goods are weakly separable from labor: $U(c, s, e) = \tilde{U}\left(u(c, s), e\right)$. Then, for any optimal allocation with interior consumption, all commodity wedges are zero in period T.*

Next, I establish the main result of this section. The proposition implies a uniform taxation across *all* goods (nondurable goods and investment in durable goods) within all periods.

Proposition 2 (Uniform taxation)**.** *Let* $\alpha \in \mathbb{R}^M_{++}$. *Let* $u : \mathbb{R}^N_+ \to \mathbb{R}$ *be strictly increasing and continuously differentiable on the interior of its domain. Suppose that*

$$U(c,s,e) = u(c) + \alpha \cdot s - v(e) \quad \text{for all } (c,s,e) \in \mathbb{R}^{N+M+1}_+ \tag{6}$$

where $\alpha \cdot s := \sum_{m=1}^{M} \alpha_m s_m$ *denotes the scalar product. Then, for any optimal allocation with interior consumption, all commodity wedges are zero in all periods.*

Heuristically, the proof of Proposition 2 works as follows. Because of the additive separability and linearity of the preferences over durable consumption, the utility flow from investing in a durable good does not depend on the consumption of other (durable or nondurable) goods, nor on past or future investment. Put differently, investing in a durable good yields a deterministic flow of utility. Following this reasoning, durable goods and nondurable goods become equivalent and Proposition 1 suggests that all commodity wedges should be zero.

To ensure that all commodity wedges are zero, Proposition 2 relies on assumptions that are significantly stronger than those required in models without durable goods. The additive separability and linearity of the preferences over durable consumption are, in fact, violated for many common environments with durable goods; see the discussion of housing in Section 5, for instance. Therefore, Proposition 2 is not widely applicable. However, the proposition establishes an important theoretical benchmark because it identifies a maximal case where all commodity wedges are zero. As shown by Proposition 3 below, the result breaks down if the preference specification in Proposition 2 is relaxed.

Proposition 3. *Let* $u : \mathbb{R}^N_+ \to \mathbb{R}$ *be strictly increasing and continuously differentiable on the interior of its domain. (a) Let* $U(c,s,e) = u(c) + \hat{u}(s) - v(e)$. *If* $\hat{u}$ *is nonlinear, optimal allocations do not in general imply zero commodity wedges. (b) Let* $U(c,s,e) = \tilde{U}(u(c),s) - v(e)$, *and* $\tilde{U}(u(c),s)$ *be linear in* s. *If* $\tilde{U}$ *is not additively separable, optimal allocations do not in general imply zero commodity wedges. (c) Let* $\alpha \in \mathbb{R}^M_{++}$. *Let* $U(c,s,e) = \hat{U}(u(c) + \alpha \cdot s, v(e))$. *If* $\hat{U}$ *is not additively separable, optimal allocations do not in general imply zero commodity wedges.*

Parts (a) and (b) of Proposition 3 show that a uniform taxation is not generally optimal if the utility from durable consumption is nonlinear or the preferences over durable and nondurable consumption are not additively separable. Both of these channels are explored in detail in Section 4. Part (c) of Proposition 3 shows that a uniform taxation is not generally optimal if the additive separability between consumption and labor is relaxed. Below, Example 1 demonstrates this finding using a multiplicative preference specification where all consumption goods are complementary with leisure. Although durable and nondurable consumption goods are both complementary with present leisure, there is a differential motive for taxing durable goods, because durable investment is also complementary with *future* leisure, whereas nondurable consumption is not. This motivates an implicit tax on the investment in durable goods in order to make leisure less attractive in the future.

Example 1. Suppose that the utility from consumption is additively separable and linear in durable consumption as in Eq. (6) but that the utility is only weakly separable between consumption and labor:

$$U(c,s,e) := (u(c) + \alpha \cdot s)\, v(1-e)$$

where u and v are strictly positive, strictly increasing and continuously differentiable, $\alpha \in \mathbb{R}^M_{++}$ and $e \in [0,1)$. For simplicity, consider a two-period problem with no uncertainty in the first period. Productivity in the first period equals $\theta > 0$. In the second period, productivity is an element of the binary set $\{\theta_L, \theta_H\}$, with $\theta_L = 0$ and $\theta_H > 0$. The probability weights are given by $\pi(\theta_k) \in (0,1)$ for $k = L, H$. Given that the productivity in the second period can be zero in this example, it is convenient to replace effective labor with effort in the setup of the allocation problem.[11] By construction, in the second period the unproductive agent will choose effort $e_L = 0$.

The planner chooses an allocation in order to maximize social welfare

$$\max \quad (u(c) + \alpha \cdot \rho i)\, v\,(1-e) + \beta \sum_{k=L,H} \pi(\theta_k)\, (u(c_k) + \alpha \cdot \rho(i_k + \delta i))\, v\,(1-e_k)$$

subject to resource feasibility and the (downward) incentive compatibility constraint,

$$(u(c_H) + \alpha \cdot \rho(i_H + \delta i))\, v\,(1-e_H) \geq (u(c_L) + \alpha \cdot \rho(i_L + \delta i))\, v\,(1-e_L)\,.$$

[11]Note that y_L/θ_L is not well defined when $y_L = \theta_L = 0$.

Assuming an interior solution for consumption, the first-order conditions with respect to consumption in the first period imply

$$\begin{aligned}&\frac{\alpha_m \rho_m v\left(1-e\right) + \beta \alpha_m \rho_m \delta_m \sum_{k=L,H} \pi(\theta_k) v\left(1-e_k\right)}{u_n(c) v\left(1-e\right)} \\ &= \frac{G_{N+m}\left(c, i, K_2, K_1, \theta e\right)}{G_n\left(c, i, K_2, K_1, \theta e\right)} + \frac{\mu \alpha_m \rho_m \delta_m \Delta v}{\lambda G_n\left(c, i, K_2, K_1, \theta e\right)}\end{aligned} \tag{7}$$

where $\lambda > 0$ is the Lagrange multiplier for the feasibility constraint in the first period, μ the multiplier for the incentive constraint, and $\Delta v := v\left(1-e_L\right) - v\left(1-e_H\right) \geq 0$. Suppose that the productivity realization θ_H in the second period is sufficiently large. Then, the incentive constraint is binding, which implies $\mu > 0$ and $\Delta v > 0$. By the first-order conditions, the marginal rate of substitution between the investment in durable good m and the consumption of nondurable good n in the first period (the left-hand side of Eq. (7)) exceeds the marginal rate of transformation (the first term on the right-hand side of Eq. (7)). This means that investments in durable goods are implicitly taxed at a higher rate than nondurable goods.

4 Differential commodity taxation

This section studies optimal commodity taxation when the consumption preferences are strictly convex. To make the analysis of differential taxation more tractable, I focus on a setting with one durable and one nondurable good: $M = N = 1$. Moreover, in line with most of the literature on dynamic Mirrleesian taxation, I consider a utility function that is additively separable between consumption and labor effort:[12]

$$U(c, s, e) = u(c, s) - v(e), \tag{8}$$

where u is strictly increasing, *strictly* concave (unlike the specification in Proposition 2) and twice continuously differentiable. In particular, this specification gives rise to a version of the Inverse Euler Equation, which constitutes an important input for the following analysis.

The technology in this section is described by a strictly increasing, continuously differentiable production function F that produces a final good. The final good can be used for

[12]As suggested by the general reasoning of Example 1, in the present framework there is no sharp result for weakly separable preferences, because weak nonseparabilities have the potential to motivate taxes as well as subsidies to durable goods.

nondurable consumption, investment in the durable good and investment in capital:

$$G\left(C, I, K', K, Y\right) = C + I + K' - (1 - \delta^K)K - F(K, Y). \tag{9}$$

It is convenient to denote the gross interest rate in period t at an optimal allocation by

$$R_t^* := 1 - \delta^K + F_K\left(K_t^*, \mathbb{E}\left[y_t^*\right]\right).$$

Moreover, for $T \geq k > t \geq 1$, define the intertemporal discount factor

$$q_{t,k}^* := \frac{1}{\prod_{i=t+1}^{k} R_i^*}$$

and set $q_{t,t}^* = 1$.

4.1 Necessary conditions for intertemporal optimality

The following two results characterize the evolution of marginal consumption utilities and marginal rates of substitution between nondurable and durable consumption over time.

Lemma 2 (Inverse Euler Equation)**.** *Let* $(\mathbf{c}^*, \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}^*)$ *be an optimal allocation with interior consumption. Then, for any* $t < T$*,*

$$\frac{\beta R_{t+1}^*}{u_c\left(c_t^*, \rho d_t^*\right)} = \mathbb{E}_t\left[\frac{1}{u_c\left(c_{t+1}^*, \rho d_{t+1}^*\right)}\right]. \tag{10}$$

The Inverse Euler Equation is well known (Rogerson, 1985; Golosov et al., 2003) and stems from an intertemporal trade-off in the provision of utility from nondurable consumption. Suppose that, at a given point in time and for a given history, the planner increases the agent's instantaneous utility by raising nondurable consumption. In the next period, for all possible continuations, the planner lowers the level of nondurable consumption such that the agent's total welfare remains unaffected. Clearly, such a variation does not affect the incentive-compatibility constraint. Hence, if the original allocation was optimal, the variation cannot by cheaper for the planner than the original allocation. Based on this reasoning, Eq. (10) states that the marginal effect of such an intertemporal variation on aggregate resources must be zero.

The Inverse Euler Equation is a basic intertemporal result that applies whenever there is a nondurable consumption good. The introduction of durable goods does not at all affect the underlying logic. However, the coexistence of durable and nondurable goods give rise to a new class of incentive-feasible variations. Minimizing the resources within this class of variations, I obtain the following novel intertemporal optimality condition.

Proposition 4 (Substitution Euler Equation)**.** *Let* $(\mathbf{c}^*, \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}^*)$ *be an optimal allocation with interior consumption. Then, for any* $t < T$*,*

$$1 = \rho \sum_{k=t}^{T} q^*_{t,k} \delta^{k-t} \mathbb{E}_t \left[\frac{u_s\left(c^*_k, \rho d^*_k\right)}{u_c\left(c^*_k, \rho d^*_k\right)} \right]. \tag{11}$$

The Substitution Euler Equation follows from the following argument. Suppose that the investment in the durable good is reduced by one marginal unit in period t. In turn, nondurable consumption is increased in periods $t, \ldots, T$ such that the agent remains as well off as before (in every period and for every realization). The reduced investment in the durable good saves one unit of resources, which gives the left-hand side of Eq. (11). The right-hand side of Eq. (11) captures the date-t value of the resources needed to increase nondurable consumption accordingly. Note that the stock of the durable good in period k falls by δ^{k-t} units under this variation, and hence the durable service flow diminishes by $\rho\delta^{k-t}$ units in period k. To compensate the agent, nondurable consumption has to be increased by $\rho\delta^{k-t} u_s / u_c$ units for each possible realization in period k. Summing up over all periods k from t to T and discounting across time by the factor $q^*_{t,k}$, the date-t value of these resources is given by the right-hand side of the equation. A necessary condition for optimality is that no resources are freed up if durable investment is exchanged for nondurable consumption in such an incentive-neutral way. Therefore, the first-order condition Eq. (11) needs to hold at any optimal allocation.

4.2 Optimal commodity wedges

When trading off durable investment against contemporaneous nondurable consumption, the agent explores a variation fundamentally different from the one that underlies the Substitution Euler Equation. Rather than considering state-contingent adjustments of future nondurable consumptions, the relevant comparison is between a sequence of durable service flows and an instantaneous one-time flow from nondurable consumption. Therefore, the Substitution Euler Equation does not immediately indicate how the individual margin to invest in the durable good may be distorted. However, as this section demonstrates, by combining the Substitution Euler Equation with the Inverse Euler Equation and using general mathematical arguments on stochastic processes, it becomes possible to characterize the wedge between durable investment and nondurable consumption.

First, note that the utility value of investing in the durable good generally depends on the dynamic realization of the two-dimensional stochastic process for durable and nondurable consumption levels. If the two goods are monotonically related, this process can be reduced to a one-dimensional "sufficient statistic".

Definition 3. The durable and the nondurable good are *perfectly rank correlated after period t* if the following equivalence holds for all periods $\tau > t$ and all histories $\theta^\tau, \tilde{\theta}^\tau \in \Theta^\tau$:

$$c_\tau\left(\theta^\tau\right) \geq c_\tau\left(\tilde{\theta}^\tau\right) \iff d_\tau\left(\theta^\tau\right) \geq d_\tau\left(\tilde{\theta}^\tau\right). \tag{12}$$

The two goods are *perfectly rank correlated* if Eq. (12) holds for all $\tau \geq 1$ and all $\theta^\tau, \tilde{\theta}^\tau \in \Theta^\tau$.

Note that, in a static environment, two goods are perfectly rank correlated if the consumption of both goods increases with the realization of uncertainty. Thus, Definition 3 establishes a dynamic concept of the normality of goods. Although perfect rank correlation helps keep the mathematical analysis tractable, the underlying economic argument suggests that the results are robust as long as there is a sufficiently positive relationship between durable and nondurable consumption.

The following result provides a sufficient condition for perfect rank correlation in terms of model primitives.

Lemma 3 (Perfect rank correlation for homothetic preferences)**.** *Suppose that u is a monotonic transformation of a homogeneous function and strictly concave. Then, for any optimal allocation with interior consumption, the durable and the nondurable good are perfectly rank correlated.*

Next, I provide the main result on differential commodity taxation. Mathematically, the result is a nontrivial combination of the Substitution Euler Equation, the Inverse Euler Equation and the property that the covariance of two increasing functions of a random variable is positive.[13]

[13] See Schmidt (2003) for an elementary proof of this property. Note that the property can also be used to verify

Proposition 5 (Differential taxation). *Let $t < T$ and consider preferences and a production technology as specified in Eqs.* (8) *and* (9)*. If the durable and the nondurable good are perfectly rank correlated after period t, any optimal allocation with interior consumption has the following implications: If $u_{cs} \leq 0$, then investment in the durable good is implicitly taxed at a higher rate than nondurable consumption, i.e., $\tau^t_{2,1} \geq 0$. If consumption is not fully insured in periods $t+1, \ldots, T$, the previous inequality becomes strict. If $u_{cs} > 0$, the sign of the commodity wedge is ambiguous.*

Proposition 5 shows that investment in durable goods should be taxed differently than nondurable goods. Hence, although the preferences are additively separable between consumption and labor effort, the Atkinson–Stiglitz result on uniform (intra-period) commodity taxation does not apply. The key difference between durable and nondurable goods is that investment in durables affects the incentive problem in the following periods. This dynamic incentive effect is precisely the reason why the Atkinson–Stiglitz result fails.

Intuitively, suppose that the investment in the durable good in period t is increased by a small amount Δi. In period $\tau > t$, consider two candidate realizations $\tilde{\theta}^\tau, \hat{\theta}^\tau$ and suppose that the associated consumption levels are ranked as $\tilde{c}_\tau > \hat{c}_\tau$ and $\tilde{d}_\tau > \hat{d}_\tau$. An incremental investment in period t raises the stock of the durable good in period τ by $\delta^{\tau-t}\Delta i$ units. Hence, in response to a marginal increment at time t, the utility difference between the states $\tilde{\theta}^\tau, \hat{\theta}^\tau$ in period τ changes by Δu, where Δu is given by

$$\frac{\Delta u}{\rho \delta^{\tau-t}} = u_s(\tilde{c}_\tau, \rho \tilde{d}_\tau) - u_s(\hat{c}_\tau, \rho \hat{d}_\tau) = \int_{\hat{d}_\tau}^{\tilde{d}_\tau} u_{ss}(\tilde{c}_\tau, \rho \xi) d\xi + \int_{\hat{c}_\tau}^{\tilde{c}_\tau} u_{cs}(\kappa, \rho \hat{d}_\tau) d\kappa.$$

By concavity, the second derivative u_{ss} is negative. Hence, when durable and nondurable consumption are Edgeworth substitutes ($u_{cs} \leq 0$), the utility difference Δu in the above example consists of two negative terms. In that case, investment in the durable good unambiguously dampens the variation of future utility. This effect is socially harmful because it makes incentive provision in the remaining periods more difficult. To account for this negative externality, the durable good should be taxed more than the nondurable good in order to relax the incentive compatibility constraint. Note that this result is a combination of substitution effects (captured by the cross derivative u_{cs}) and diminishing marginal utility effects (captured by the

that the intertemporal wedge is positive. A positive covariance $\mathrm{cov}_t\left(u_c\left(c^*_{t+1}, \rho d^*_{t+1}\right), -1/u_c\left(c^*_{t+1}, \rho d^*_{t+1}\right)\right)$ is by definition equivalent to the condition $\mathbb{E}_t\left[u_c\left(c^*_{t+1}, \rho d^*_{t+1}\right)\right]\mathbb{E}_t\left[1/u_c\left(c^*_{t+1}, \rho d^*_{t+1}\right)\right] > 1$. Using the Inverse Euler Equation, we obtain $\beta R^*_{t+1}\mathbb{E}_t\left[u_c\left(c^*_{t+1}, \rho d^*_{t+1}\right)\right] > u_c\left(c^*_t, \rho d^*_t\right)$.

second derivative u_{ss}).[14] If durable and nondurable consumption are Edgeworth complements ($u_{cs} > 0$), those two effects oppose each other and the sign of the commodity wedge becomes ambiguous.

Note that substitution effects arise only when nondurable consumption in future periods is uncertain and the cross derivative of the utility function nonzero. Similarly, diminishing marginal utility effects emerge only when investment in the durable good varies across realizations in the future (and the utility function is strictly concave in durable consumption). Hence, in the extreme case where investment in the durable good occurs only once, the wedge on the durable good depends exclusively on the cross derivative of the utility function. Ceteris paribus, this insight suggests that adjustment frictions for durable goods may reduce the motive to tax the investment in those goods.

Finally, consider the consequences of having several durable or nondurable consumption goods: $M, N \geq 1$. Then, the substitution effects and diminishing marginal utility effects from the two-goods model are complemented by non-separabilities with other consumption goods. Similar to Eq. (19) in the proof of Proposition 5, it can be shown that the durable good m is implicitly taxed at a higher rate than the nondurable good n in period t if and only if

$$\sum_{\tau=t+1}^{T} q_{t,\tau}^{*} \delta_m^{\tau-t} \operatorname{cov}_t \left(-u_{N+m}\left(c_\tau^*, \rho d_\tau^*\right), \frac{1}{u_n\left(c_\tau^*, \rho d_\tau^*\right)} \right) \geq 0.$$

By arguments akin to the proof of Proposition 5, this inequality is satisfied if all goods are Edgeworth substitutes and perfectly rank correlated.[15]

4.3 Decentralization of optimal allocations

As anticipated by the discussion in Section 3.1, commodity wedges on durable goods are typically not equivalent to standard intratemporal transaction taxes. Because durable goods generate long-lasting service flows, labor decisions in subsequent periods may be affected by the contemporaneous investment in a durable good. Given that subsequent labor supplies can be adjusted after the investment has been made, explicit tax systems that decentralize optimal

[14]Diminishing marginal utility effects also explain why savings should be taxed in dynamic Mirrlees models (Diamond and Mirrlees, 1978; Golosov et al., 2003).

[15]For example, Edgeworth substitutability holds if durable consumption is weakly separable from nondurable consumption, $u(c, \rho d) = \hat{u}(\tilde{u}(c), \rho d)$, with an aggregator function $\hat{u}$ that is submodular.

allocations need to break the link between durable investment and subsequent labor supplies not just in expectation (at the time of investment) but also ex post.

Following this idea, I present a decentralization of optimal allocations through a tax system with a retrospective taxation of durable investment. The tax on durable investment is levied in the final period and depends on the full history of investments and incomes. In addition, the tax system includes an income tax and a capital tax similar to the construction by Kocherlakota (2005).

Let $(\mathbf{c}^*, \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}^*)$ be an optimal allocation with interior consumption. In line with earlier contributions, I maintain the following assumption throughout this subsection.[16]

Assumption. Set $DOM_t := \left\{y^t \in \mathbb{R}_+^t : y^t = \left(y_1^*(\theta_1), y_2^*(\theta^2), \ldots, y_t^*(\theta^t)\right) \text{ for some } \theta^t \in \Theta^t\right\}$. For every t, there exist functions $\hat{c}_t^* : DOM_t \to \mathbb{R}_+$ and $\hat{d}_t^* : DOM_t \to \mathbb{R}_+$ such that

$$\begin{aligned}
\hat{c}_t^*\left(y_1^*(\theta_1), \ldots, y_t^*(\theta^t)\right) &= c_t^*(\theta^t) \quad \text{for all } \theta^t \in \Theta^t, \\
\hat{d}_t^*\left(y_1^*(\theta_1), \ldots, y_t^*(\theta^t)\right) &= d_t^*(\theta^t) \quad \text{for all } \theta^t \in \Theta^t.
\end{aligned}$$

Under this assumption, the optimal allocation treats agents with identical histories of effective labor (but possibly different skill histories) symmetrically in terms of consumption. This condition makes it possible to decentralize the allocation with a tax system defined in terms of effective labor. For the investment plan $\mathbf{i}^*$ associated with the optimal allocation, the assumption straightforwardly implies that there exists a function $\hat{\imath}_t^* : DOM_t \to \mathbb{R}_+$ such that $\hat{\imath}_t^*\left(y_1^*(\theta^1), \ldots, y_t^*(\theta^t)\right) = i_t^*(\theta^t)$ for all $\theta^t \in \Theta^t$.

For the decentralization, I consider an economy with a representative firm that owns the production technology and employs capital and labor taking as given the wage rate $w_t^* := F_Y(K_t^*, \mathbb{E}[y_t^*])$ and the gross interest rate $R_t^* = 1 - \delta^K + F_K(K_t^*, \mathbb{E}[y_t^*])$. The agents take prices and the tax system as given. They supply labor and trade capital, nondurable consumption and investment in the durable good in a sequence of competitive markets. All agents begin with an initial capital endowment of K_1^*.

I define taxes on capital, labor and durable investment as follows. By setting a sufficiently large tax for sequences $y^t \notin DOM_t$, I can restrict attention to effective labor sequences in

[16] This assumption is satisfied for any incentive-compatible allocation when the skill process is independent over time. For a discussion of the assumption in models with persistent skills, see Kocherlakota (2005) or Grochulski and Kocherlakota (2010).

DOM_t. For $t > 1$ and sequences $y^t \in DOM_t$, I define a tax on capital holdings $k_t \in \mathbb{R}_+$ as

$$\mathcal{T}_t^k \left(y^t, k_t\right) := k_t \left(R_t^* - \frac{u_c\left(\hat{c}_{t-1}^*, \rho \hat{d}_{t-1}^*\right)}{\beta u_c\left(\hat{c}_t^*, \rho \hat{d}_t^*\right)} \right).$$

To simplify the notation, I have omitted the argument y^t in the functions $\hat{c}_t^*$ and $\hat{d}_t^*$ in the previous definition.[17] I maintain this simplification throughout the following analysis. The sequence of labor income taxes is defined as follows:

$$\begin{aligned}
\mathcal{T}_1^y \left(y_1\right) &:= w_1^* y_1 + R_1^* K_1^* - \hat{c}_1^* - \hat{\imath}_1^* - K_2^*, \\
\mathcal{T}_t^y \left(y^t\right) &:= w_t^* y_t + R_t^* K_t^* - \mathcal{T}_t^k \left(y^t, K_t^*\right) - \hat{c}_t^* - \hat{\imath}_t^* - K_{t+1}^* \quad \text{for all } 1 < t < T, \\
\mathcal{T}_T^y \left(y^T\right) &:= w_T^* y_T + R_T^* K_T^* - \mathcal{T}_T^k \left(y^T, K_T^*\right) - \hat{c}_T^* - \hat{\imath}_T^*.
\end{aligned}$$

Finally, for $y^T \in DOM_T$, I define a tax $\mathcal{T}_T^d$ on the sequence of durable investments $i^T \in \mathbb{R}^T$ (levied in the final period) as

$$\mathcal{T}_T^d \left(y^T, i^T\right) := \sum_{\tau=1}^{T} \kappa_\tau \left(i_\tau - \hat{\imath}_\tau^*\right),$$

where the marginal tax rates on investment are given by

$$\kappa_t := \frac{\rho \sum_{\tau=t}^{T} \left(\beta\delta\right)^{\tau-t} u_s\left(\hat{c}_\tau^*, \rho \hat{d}_\tau^*\right) - u_c\left(\hat{c}_t^*, \rho \hat{d}_t^*\right)}{\beta^{T-t} u_c\left(\hat{c}_T^*, \rho \hat{d}_T^*\right)}.$$

Note that the marginal tax rate κ_t on durable investment measures the ex post gap between the marginal rates of substitution and transformation between durable investment and non-durable consumption. In this construction, the denominator accounts for the fact that the tax is collected in the final period.

The main result of this subsection is as follows.

Proposition 6 (Decentralization). *Given prices $(w_t^*, R_t^*)_t$ and the tax system $\left(\mathcal{T}^y, \mathcal{T}^k, \mathcal{T}_T^d\right)$, the allocation $(\mathbf{c}^*, \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}^*)$ solves the individual maximization problem in the decentralized economy.*

Proposition 6 shows that the tax system $\left(\mathcal{T}^y, \mathcal{T}^k, \mathcal{T}_T^d\right)$ implements the optimal allocation as

[17] Similarly, I have omitted y^{t-1} in the functions $\hat{c}_{t-1}^*$ and $\hat{d}_{t-1}^*$.

a competitive equilibrium. In this result, the main novelty is the introduction of a retrospective tax on investment in the durable good. This tax is based on a similar idea as the approach to capital taxation by Kocherlakota (2005). Note that, because of future uncertainty, the marginal utility flow from a durable investment will be higher for some continuation paths than for others. Therefore, the tax on durable investment is set such that the marginal tax rate κ_t on investment is high precisely when the (before-tax) marginal utility flow from investment is high. This construction neutralizes the effect of durable investment on the preferences over continuation paths. Put differently, agents have no incentive to engage in "joint deviations" that change their current durable investment and mimic the future labor choice associated with a specific continuation path. Note that this construction of taxes works irrespective of the actual sign of the commodity wedge.

5 Application to housing taxation

Housing is a prime example of a durable good. Housing is particularly interesting from an optimal tax perspective because tax advantages for housing are widespread in many countries.[18] For instance, payments of mortgage interest are (partly or fully) tax-deductible in the United States, the Netherlands, Switzerland, Belgium, Ireland, Norway and Sweden. In the UK, there is a reduced value added tax on the construction of new houses and renovations.

Given the large body of research that estimates the preferences over housing and other consumption, the present analysis can be readily applied to housing taxation. In the applied economic literature on housing, the preferences are commonly specified by a utility function with a constant elasticity of substitution,

$$u(c,s) = \frac{\left[(1-\omega)c^{1-\frac{1}{\epsilon}} + \omega s^{1-\frac{1}{\epsilon}}\right]^{\frac{1-\frac{1}{\sigma}}{1-\frac{1}{\epsilon}}}}{1-\frac{1}{\sigma}}, \tag{13}$$

where c denotes nondurable consumption, s denotes housing services, the parameter $\epsilon > 0$ measures the intratemporal substitutability between housing and nondurable consumption, $\omega \in (0,1)$ controls the expenditure share on housing, and $\sigma > 0$ governs the intertempo-

[18]Different from many critiques of such tax advantages, the present approach is based on pure efficiency reasoning and is independent of the redistributional objective.

ral substitutability of the consumption-housing composite. [19] This specification implies that housing and nondurable consumption are substitutes in the Edgeworth sense $\left(u''_{ch} \leq 0\right)$ if and only if the parameters satisfy $\epsilon \geq \sigma$, i.e., if and only if the intratemporal elasticity of substitution exceeds the intertemporal elasticity. Moreover, this preference specification is homothetic. Therefore, Proposition 5 and Lemma 3 have the following consequence.

Corollary to Proposition 5. *If $\epsilon \geq \sigma$, housing investment should be (implicitly) taxed at a higher rate than nondurable consumption.*

Many papers have estimated the above CES specification in housing models. Several approaches rely on macroeconomic evidence and calibration strategies. For example, based on macro-level consumption data, Piazzesi, Schneider, and Tuzel (2007) provide a calibration for low risk aversion with $(\epsilon, \sigma) = (1.05, 0.2)$ and one for high risk aversion with $(\epsilon, \sigma) = (1.25, 0.0625)$. Their paper refers to several further calibrations of the CES function for housing where the intratemporal elasticity of substitution exceeds the intertemporal one. More recently, two papers estimate the CES function by matching cross-sectional and time series moments of wealth and housing profiles from PSID micro data. Li, Liu, Yang, and Yao (2015) estimate parameter values of $(\epsilon, \sigma) = (0.487, 0.140)$. Bajari, Chan, Krueger, and Miller (2013) follow a similar approach for logarithmic utility functions and estimate $(\epsilon, \sigma) = (4.550, 1)$. In sum, the available empirical evidence suggests $\epsilon > \sigma$, which implies that housing and nondurable consumption are Edgeworth substitutes.[20] Thus, according to the theory in this paper, housing investment should be taxed at a higher rate than nondurable consumption.[21]

5.1 Numerical illustration

To illustrate the role of housing taxation in the present framework, I explore a simple parametrized example. There are two periods with a duration of 20 years each. Thus, the example roughly

[19] For $\sigma = 1$, preferences take a logarithmic form.

[20] This pattern is also documented for more comprehensive measures of durable goods (Ogaki and Reinhart, 1998).

[21] Admittedly, the present analysis abstracts from several alternative motives for housing policy. For instance, capital market imperfections such as borrowing constraints may justify subsidies to housing. Moreover, political economy considerations may lead to outcomes that differ from the solution of a social planning problem. Such imperfections warrant an independent investigation because they have many consequences beyond the taxation of housing. In principle, capital market imperfections and the government's role can be affected by political processes, whereas the diminishing marginal utility and substitution effects highlighted in this paper are primitives that follow directly from individual preferences.

(ϵ, σ)	(0.487, 0.140)	(1.05, 0.2)	(1.25, 0.0625)	(4.550, 1)
housing wedge (%)	28.0	19.0	64.6	2.6
intertemp wedge (%)	29.0	23.7	39.7	5.1
welfare gain (%)	0.11	0.14	0.24	0.01

Table 1: Expected housing wedges, expected intertemporal wedges and welfare gains of differential taxation (measured in terms of consumption equivalent variation) for different preference parameters.

covers the working life of a typical employee. Following Kocherlakota (2010), I consider a binary skill process with a unit root for log skills.[22] Despite its stylized nature, this process represents some key empirical regularities regarding the life-cycle growth, the persistence and the cross-sectional variance of wages in the U.S. economy. The production function is linear in effective labor: $F(K,Y) = RK + Y$, with $R = 1/\beta$, and initial wealth in the economy is zero. Furthermore, I set $\rho = 1$ and assume that the housing stock depreciates at a rate of 1.7% per year. This rate corresponds to the annual maintenance cost of housing as estimated by Li et al. (2015).

The agents discount the future with a factor of 0.98 per annum and have CES consumption preferences as specified in Eq. (13). I consider a range of empirically plausible substitution elasticities (ϵ, σ) based on the external estimation/calibration results described above. The disutility of labor effort is $v(e) = \alpha e^{1+1/\eta}/(1+1/\eta)$, with a Frisch elasticity of $\eta = 0.5$. For the sake of comparability across allocations, I recalibrate the preference weights (α, ω) for each preference scenario such that the expenditure share on housing and the present-discounted value of lifetime income remain fixed.[23]

Numerical results. Table 1 presents the expected wedge between housing investment and nondurable consumption ("housing wedge") implied by the optimal allocations of different preference scenarios.[24] The housing wedge is sensitive to the preference parameters and ranges from 3% to 65%. Table 1 also documents the welfare gains of housing taxation. To this end, I solve an auxiliary model that constrains the planner to equalize the marginal rate

[22]Skills in the first period are random draws from the set $\Theta_1 = \{\exp(-0.5), \exp(0.5)\}$. From the first period to the second, skills grow at a stochastic rate randomly drawn from the set $\left\{\exp\left(0.4 - \sqrt{0.12}\right), \exp\left(0.4 + \sqrt{0.12}\right)\right\}$.

[23]For the parameter ω, I target an expenditure share on housing of 0.23 based on CEX 2011 data. Moreover, I set $\alpha = 1$ for the scenario with substitution elasticities as estimated by Li et al. (2015) and adjust this parameter for the other scenarios to maintain the same present-discounted value of lifetime income.

[24]The housing wedge is evaluated in the first period. By Remark 1, the wedge in the second period is zero.

of substitution between housing investment and nondurable consumption with the marginal rate of transformation. Then, I compare the welfare in the constrained model to the welfare in the baseline model and express the welfare change in consumption equivalent terms. As shown by the last row of Table 1, the welfare gains of housing taxation in this example are moderate (below 0.3% of lifetime consumption). For logarithmic utility ($\sigma = 1$), the welfare gains are negligible. Note that the housing wedge and the welfare gain of differential taxation are largest for the preference parameters $(\epsilon, \sigma) = (1.25, 0.0625)$. For this specification, the discrepancy between the intratemporal and the intertemporal elasticity of substitution is particularly pronounced, which results in a strong degree of Edgeworth substitutability between housing and nondurable consumption. Moreover, this specification is associated with a high aversion to risk, which means that suboptimal social insurance causes larger welfare losses.[25]

Given the stylized setup of example, the quantitative findings are mainly illustrative. Yet, one particular message is likely to hold more generally: optimal housing policy in the present framework is quite sensitive to the preference parameters. Although the current range of empirically plausible parameters generates a unique sign of the housing wedge, the magnitude of optimal housing distortions varies considerably with the parameterization.

6 Discussion and conclusion

This paper shows that optimal commodity taxes are generically non-uniform in the presence of durable goods. Nonseparabilities between durable and nondurable consumption, as well as nonlinearities of the utility from durable consumption, imply that differential commodity taxes improve welfare. Applied to housing policy, these findings suggest that housing investment should be taxed at a higher rate than nondurable consumption.

To conclude, I contrast the present results with the analysis of optimal taxes on pre-committed goods. Moreover, I briefly discuss two model extensions.

[25] The difference between the two substitution elasticities is also pronounced for $(\epsilon, \sigma) = (4.550, 1)$. However, this specification has a significantly lower coefficient of risk aversion, which makes improvements to social insurance less valuable.

6.1 Durable goods versus pre-committed goods

This paper leads to a novel interpretation of the analysis of pre-committed goods by Cremer and Gahvari (1995a,b). Assuming a separability between pre-committed and post-uncertainty goods, their main finding is that pre-committed goods should be subsidized relative to post-uncertainty goods.

Consider a two-period version of the present model with one durable and one nondurable consumption good in each period. Moreover, suppose that there is no uncertainty in the first period. Then, the consumption goods are pre-committed in the first period (i.e., decided before the realization of uncertainty) but they are post-uncertainty goods in the second period. Hence, durable and nondurable goods become pre-committed goods or post-uncertainty goods depending on the timing. Stated differently, the notion of pre-commitment does not distinguish durable goods from nondurable goods.

For the two-period setting, Remark 1 implies that a uniform taxation of goods is optimal in the second period. This result is closely related to the finding that post-uncertainty goods should be taxed uniformly. In contrast, the tax wedge between durable and nondurable goods in the first period (or more generally in non-terminal periods) in Proposition 5 does not have a counterpart in the analysis of pre-committed goods, because that analysis focuses on differentials between pre-committed and post-uncertainty goods, not on differentials within pre-committed goods. However, the motive to subsidize pre-committed goods relates to a dynamic result in the present paper. Note that a nondurable good in the first period is separable from the consumption goods in the second period, and it is decided before the resolution of uncertainty. The motive to subsidize this good relative to a post-uncertainty good means that there is an intertemporal wedge on *nondurable* goods, as implied by the Inverse Euler Equation.

Cremer and Gahvari (1995b) also analyze optimal housing policy in a calibrated model similar to the exploration in Section 5.1. Because they formalize housing as a pre-committed good, their quantitative findings differ from the results in the present paper. In particular, they find *subsidies* on housing of approximately 25% to be optimal. Based on the interpretation above, subsidies to pre-committed goods are closely related to intertemporal wedges on nondurable goods. The intertemporal wedges found in Section 5.1 (third row of Table 1) indeed have a quantitative magnitude that is broadly comparable to the subsidies found by Cremer and Gahvari (1995b).

6.2 Durable goods and time use

In the present framework, durable goods generate service flows that are independent of labor/leisure choices. Yet, in practice several examples of durable goods are related to leisure (e.g., audio and video entertainment, sports equipment, electronic gadgets) or household production activities (e.g., home appliances), suggesting that the flows from durable goods may not always be separable from labor.

In a reduced form, such interactions can be modeled by letting the service flow from durable goods directly affect the disutility of labor effort. Assuming that durable services are complements to leisure and/or household production, it seems natural to consider labor disutilities that have a positive cross derivative with respect to durable services and (market) labor effort.[26] Atkinson and Stiglitz (1976) and Christiansen (1984) explore nonseparabilities of this form in static environments. Based on their results, a positive cross derivative of the labor disutility function implies an additional rationale for taxing durable goods. Therefore, the time use aspect of durable goods will most likely reinforce the present findings on positive commodity wedges when durable and nondurable consumption are Edgeworth substitutes (Proposition 5).

6.3 Adverse selection and moral hazard

The mathematical analysis in this paper rests on variational arguments that exploit incentive-neutral perturbations of optimal allocations. More precisely, the analysis modifies the allocation of consumption across time and goods, but keeps the assignment of labor effort and consumption utility fixed. Therefore, the results hold under very general specifications of uncertainty.

For example, the multiplicative specification of effective labor, $y_t = \theta_t e_t$, can be replaced by a general framework where θ_t is a preference shock that affects the (dis)utility of labor effort. None of the results in this paper would change if the disutility of labor were given by a function $\hat{v}(y_t; \theta_t)$ rather than the current specification $v(y_t/\theta_t)$. Furthermore, the results

[26] For example, consider labor disutilities of the form $v(e,s) = -\tilde{v}(1-e,s)$, where $\tilde{v}$ is an increasing and concave utility function defined over leisure (or household production) $1-e$ and durable services s. Then, $v_{e,s} = \tilde{v}_{1-e,s}$.

in this paper remain valid when adverse selection and moral hazard coexist. For instance, suppose that there are two sources of (idiosyncratic) uncertainty. First, individual skills θ_t follow a stochastic process as before. Second, given labor effort e_t and skill θ_t, effective labor y_t is a random variable described by a distribution $F(y_t|e_t, \theta_t)$. Suppose that the timing of events is as follows. At the beginning of the period, agents learn their skill θ_t. Next, they choose a labor effort vector e_t. Then, effective labor y_t is realized. Skill and effort are private information, whereas effective labor is publicly observable. In this framework, labor effort in period t is assigned based on the histories (y^{t-1}, θ^t) and consumption is allocated based on the same histories and the current realization y_t. All results in this paper extend to this framework because they exploit consumption variations for a given assignment of labor effort. Notice that the form of those consumption variations is, in fact, independent of the question why labor effort differs across agents.

A Appendix: Proofs of all theoretical results

Proof of Proposition 1. The proof adapts the argument from Theorem 2 in Golosov et al. (2003) to the framework with durable goods. Let $(\mathbf{c}^*, \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}^*)$ be an optimal allocation with interior consumption. Let $\mathbf{i}^*$ be the associated investment plan.

Step 1: I claim that c_t^* solves the following cost minization problem:

$$\begin{aligned}&\min_{c_t \geq 0} G\left(\mathbb{E}\left[c_t\right], \mathbb{E}\left[i_t^*\right], K_{t+1}^*, K_t^*, \mathbb{E}\left[y_t^*\right]\right)\\ &\text{s.t. } u\left(c_t\left(\theta^t\right), \rho d_t^*\left(\theta^t\right)\right) = u\left(c_t^*\left(\theta^t\right), \rho d_t^*\left(\theta^t\right)\right) \text{ for all } \theta^t.\end{aligned}$$

Suppose that, contrary to the claim, there exists a mapping $c_t' : \Theta^t \to \mathbb{R}_+$ with

$$u\left(c_t'\left(\theta^t\right), \rho d_t^*\left(\theta^t\right)\right) = u\left(c_t^*\left(\theta^t\right), \rho d_t^*\left(\theta^t\right)\right) \text{ for all } \theta^t$$

and

$$G\left(\mathbb{E}\left[c_t'\right], \mathbb{E}\left[i_t^*\right], K_{t+1}^*, K_t^*, \mathbb{E}\left[y_t^*\right]\right) < G\left(\mathbb{E}\left[c_t^*\right], \mathbb{E}\left[i_t^*\right], K_{t+1}^*, K_t^*, \mathbb{E}\left[y_t^*\right]\right) \leq 0. \tag{14}$$

Define $\mathbf{c}' = \left(c_t', c_{-t}^*\right)$ and consider the allocation $(\mathbf{c}', \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}^*)$. The allocation is feasible. The allocation is also incentive compatible: for all reporting strategies σ we have

$$\begin{aligned}
& w\left(\mathbf{c}', \mathbf{d}^*, \mathbf{y}^*\right) \\
= & \sum_{\tau=1}^{T} \beta^{\tau-1} \mathbb{E}\left[\tilde{U}\left(u\left(c'_\tau\left(\theta^\tau\right), \rho d^*_\tau\left(\theta^\tau\right)\right), \rho d^*_\tau\left(\theta^\tau\right), \frac{y^*_\tau\left(\theta^\tau\right)}{\theta_\tau}\right)\right] \\
= & \sum_{\tau=1}^{T} \beta^{\tau-1} \mathbb{E}\left[\tilde{U}\left(u\left(c^*_\tau\left(\theta^\tau\right), \rho d^*_\tau\left(\theta^\tau\right)\right), \rho d^*_\tau\left(\theta^\tau\right), \frac{y^*_\tau\left(\theta^\tau\right)}{\theta_\tau}\right)\right] \\
\geq & \sum_{\tau=1}^{T} \beta^{\tau-1} \mathbb{E}\left[\tilde{U}\left(u\left(c^*_\tau\left(\sigma^\tau\left(\theta^\tau\right)\right), \rho d^*_\tau\left(\sigma^\tau\left(\theta^\tau\right)\right)\right), \rho d^*_\tau\left(\sigma^\tau\left(\theta^\tau\right)\right), \frac{y^*_\tau\left(\sigma^\tau\left(\theta^\tau\right)\right)}{\theta_\tau}\right)\right] \\
= & \sum_{\tau=1}^{T} \beta^{\tau-1} \mathbb{E}\left[\tilde{U}\left(u\left(c'_\tau\left(\sigma^\tau\left(\theta^\tau\right)\right), \rho d^*_\tau\left(\sigma^\tau\left(\theta^\tau\right)\right)\right), \rho d^*_\tau\left(\sigma^\tau\left(\theta^\tau\right)\right), \frac{y^*_\tau\left(\sigma^\tau\left(\theta^\tau\right)\right)}{\theta_\tau}\right)\right] \\
= & w\left(\mathbf{c}' \circ \sigma, \mathbf{d}^* \circ \sigma, \mathbf{y}^* \circ \sigma\right)
\end{aligned}$$

where the inequality follows from the incentive compatibility of $\left(\mathbf{c}^*, \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}^*\right)$. Moreover, the allocation delivers the same level of social welfare as the optimal allocation $\left(\mathbf{c}^*, \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}^*\right)$. Therefore, $\left(\mathbf{c}', \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}^*\right)$ is also an optimal allocation. However, by Eq. (14), $\left(\mathbf{c}', \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}^*\right)$ does not use all capital in period t. By the strict monotonicity of the production technology G in capital, there exists a sequence of capital stocks $\mathbf{K}'$, with $K'_1 < K^*_1$, such that $\left(\mathbf{c}', \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}'\right)$ solves the planner problem for initial capital K'_1. This implies $V\left(K'_1\right) = V\left(K^*_1\right)$, which is a contradiction.

Step 2: Derive the necessary first-order conditions for the cost minimization problem. Let $n \in \{1, \ldots, N\}$. The first-order condition with respect to $c_{t,n}\left(\theta^t\right)$ is

$$\Pi^t\left(\theta^t\right) G_n\left(\mathbb{E}\left[c^*_t\right], \mathbb{E}\left[i^*_t\right], K^*_{t+1}, K^*_t, \mathbb{E}\left[y^*_t\right]\right) = \mu\left(\theta^t\right) u_n\left(c^*_t\left(\theta^t\right), \rho d^*_\tau\left(\theta^\tau\right)\right)$$

where $\mu\left(\theta^t\right)$ is the Lagrange multiplier associated with the utility constraint for history θ^t. By dividing the condition for n' by the one for n, we obtain $\tau^t_{n,n'} = 0$. □

Proof of Remark 1. Let $\left(\mathbf{c}^*, \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}^*\right)$ be an optimal allocation with interior consumption. Using the same type of argument as in the proof of Proposition 1, the allocation can only be optimal if (c_T^*, d_T^*) solves the following cost minization problem:

$$\min_{c_T, d_T} G\left(\mathbb{E}\left[c_T\right], \mathbb{E}\left[d_T - \delta d_{T-1}^*\right], K_{t+1}^*, K_t^*, \mathbb{E}\left[y_t^*\right]\right)$$
$$\text{s.t. } u\left(c_T\left(\theta^T\right), \rho d_T\left(\theta^T\right)\right) = u\left(c_T^*\left(\theta^T\right), \rho d_T^*\left(\theta^T\right)\right) \text{ for all } \theta^T.$$

The first-order conditions of this problem imply that, in period T, for any pair of consumption goods and any realization θ^T, the marginal rate of substitution equals the marginal rate of transformation. □

Proof of Proposition 2. By the linearity and additive separability of $U(c, s, e)$ in s, the utility flow from investing in a durable good is separable from all other goods, and from past and future investments. Specifically, the utility flow from investing $i_{t,m}$ in durable good m at time t is given by

$$\beta^{t-1}\left(\alpha_m \rho_m i_{t,m} + \beta \alpha_m \rho_m \delta_m i_{t,m} + \cdots + \beta^{T-t} \alpha_m \rho_m \delta_m^{T-t} i_{t,m}\right) = \beta^{t-1} \alpha_m \rho_m \frac{1 - (\beta \delta_m)^{T-t+1}}{1 - \beta \delta_m} i_{t,m}.$$

Define a function

$$\tilde{u}^t(i_t) := \sum_m \alpha_m \rho_m \frac{1 - (\beta \delta_m)^{T-t+1}}{1 - \beta \delta_m} i_{t,m}.$$

Using the above formula, the ex ante consumption utility of any deterministic plan $(c_t, d_t)_t$ is given by

$$\sum_{t=1}^{T} \beta^{t-1}\left(u(c_t) + \alpha \cdot \rho d_t\right) = \sum_{t=1}^{T} \beta^{t-1}\left(u(c_t) + \alpha \cdot \rho \sum_{k=1}^{t} \delta^{t-k} i_k\right)$$
$$= \sum_{t=1}^{T} \beta^{t-1}\left(u(c_t) + \tilde{u}^t(i_t)\right).$$

Therefore, the framework is equivalent to a model with nondurable goods (c, i) and time-dependent utility functions $U^t(c, i, e) = u(c) + \tilde{u}^t(i) - v(e)$.

Let $(\mathbf{c}^*, \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}^*)$ be an optimal allocation with interior consumption. Let $\mathbf{i}^*$ be the associated investment plan. The preferences are additively separable between consumption and labor and therefore Lemma 1 applies. Using the same argument as in the proof of Proposition 1, the allocation can only be optimal if (c_t^*, i_t^*) solves the following cost minimization problem:

$$\begin{aligned}
&\min_{c_t, i_t} G\left(\mathbb{E}\left[c_t\right], \mathbb{E}\left[i_t\right], K_{t+1}^*, K_t^*, \mathbb{E}\left[y_t^*\right]\right) \\
&\text{s.t. } u\left(c_t\left(\theta^t\right)\right) + \tilde{u}^t\left(i_t\left(\theta^t\right)\right) = u\left(c_t^*\left(\theta^t\right)\right) + \tilde{u}^t\left(i_t^*\left(\theta^t\right)\right) \text{ for all } \theta^t.
\end{aligned}$$

The first-order conditions with respect to $i_{t,m}\left(\theta^t\right)$ and $c_{t,n}\left(\theta^t\right)$ are

$$\begin{aligned}
\Pi^t\left(\theta^t\right) G_{N+m}\left(\mathbb{E}\left[c_t^*\right], \mathbb{E}\left[i_t^*\right], K_{t+1}^*, K_t^*, \mathbb{E}\left[y_t^*\right]\right) &= \mu\left(\theta^t\right) \tilde{u}_m^t\left(i_t^*\left(\theta^t\right)\right) \\
\Pi^t\left(\theta^t\right) G_n\left(\mathbb{E}\left[c_t^*\right], \mathbb{E}\left[i_t^*\right], K_{t+1}^*, K_t^*, \mathbb{E}\left[y_t^*\right]\right) &= \mu\left(\theta^t\right) u_n\left(c_t^*\left(\theta^t\right)\right)
\end{aligned}$$

where $\mu\left(\theta^t\right)$ is the Lagrange multiplier associated with the utility constraint for history θ^t. By the definition of preferences, $U_{N+m}\left(c,s,e\right) = \alpha_m$ and $U_n\left(c,s,e\right) = u_n(c)$ for all (c,s,e). Therefore,

$$\begin{aligned}
&\frac{\tilde{u}_m^t\left(i_t^*\right)}{u_n\left(c_t^*\right)} = \frac{\alpha_m \rho_m \frac{1-\left(\beta\delta_m\right)^{T-t+1}}{1-\beta\delta_m}}{u_n\left(c_t^*\right)} \\
&= \frac{\rho_m U_{N+m}\left(c_t^*, s_t^*, \frac{y_t^*}{\theta_t}\right) + \rho_m \mathbb{E}_t\left[\sum_{k=t+1}^{T}\left(\beta\delta_m\right)^{k-t} U_{N+m}\left(c_k^*, s_k^*, \frac{y_k^*}{\theta_k}\right)\right]}{U_n\left(c_t^*, s_t^*, \frac{y_t^*}{\theta_t}\right)}.
\end{aligned}$$

Hence, by dividing the first-order conditions of the cost minimization problem by each other, $\tau_{N+m,n}^t = 0$ follows. Because m and n were arbitrary, we conclude from the identities

$$\begin{aligned}
1 + \tau_{N+m,N+m'}^t &= \frac{1 + \tau_{N+m,n}^t}{1 + \tau_{N+m',n}^t} \\
1 + \tau_{n,n'}^t &= \frac{1 + \tau_{N+m,n'}^t}{1 + \tau_{N+m,n}^t}
\end{aligned}$$

that all other commodity wedges are zero as well. □

Proof of Proposition 3. Consider the binary specification of uncertainty from Example 1. Note that the arguments derived in this environment remain valid if the unproductive agent has a sufficiently small, positive productivity. Therefore, the following insights extend to the formal model with strictly positive productivities as introduced in Section 2.

(a) Let $N = M = 1$ and consider preferences as follows: $U(c,s,e) := u(c) + \hat{u}(s) - v(e)$, where u, $\hat{u}$ and v are strictly increasing and continuously differentiable, and u and $\hat{u}$ are strictly concave. Consider a two-period problem with the same specification of uncertainty as in Example 1. Assuming an interior solution for consumption, the first-order conditions with respect to consumption in the first period imply

$$\begin{aligned} &\frac{\rho\hat{u}'(\rho i)+\beta\rho\delta\sum_{k=L,H}\pi(\theta_k)\hat{u}'(\rho(i_k+\delta i))}{u'(c)} \\ &= \frac{G_2(c,i,K_2,K_1,\theta e)}{G_1(c,i,K_2,K_1,\theta e)}+\frac{\mu\rho\delta\Delta\hat{u}'}{\lambda G_1(c,i,K_2,K_1,\theta e)} \end{aligned} \tag{15}$$

where $\lambda > 0$ is the Lagrange multiplier for the feasibility constraint in the first period, μ the multiplier for the incentive constraint, and $\Delta\hat{u}' := \hat{u}'(\rho(i_L+\delta i)) - \hat{u}'(\rho(i_H+\delta i))$. Suppose that the productivity realization θ_H in the second period is sufficiently large. Then, the incentive constraint is binding, which implies $\mu > 0$ and

$$u(c_H)+\hat{u}(\rho(i_H+\delta i)) > u(c_L)+\hat{u}(\rho(i_L+\delta i)).$$

By Remark 1 we have

$$\frac{\hat{u}'(\rho(i_H+\delta i))}{u'(c_H)} = \frac{\hat{u}'(\rho(i_L+\delta i))}{u'(c_L)}.$$

Because u and $\hat{u}$ are strictly increasing and strictly concave, we conclude that $c_H > c_L$ and $i_H > i_L$. Hence, $\Delta\hat{u}' > 0$. Using $\mu\Delta\hat{u}' > 0$, Eq. (15) implies that the marginal rate of substitution differs from the marginal rate of transformation.

(b) Let $N = M = 1$ and consider preferences as follows: $U(c,s,e) := u(c)s - v(e)$, where u is strictly positive and strictly increasing. Following steps very similar to the proof of part (a), we find once more that in the first period the marginal rate of substitution differs from the marginal rate of transformation.

(c) See Example 1. □

Proof of Lemma 2. Let $(\mathbf{c},\mathbf{d},\mathbf{y},\mathbf{K})$ be an incentive-feasible allocation with interior consumption. Let $\hat{\theta}^t \in \Theta^t$. Consider the following perturbation of nondurable consumption:

$$\begin{aligned} u\left(c_t^\varepsilon\left(\hat{\theta}^t\right),\rho d_t\left(\hat{\theta}^t\right)\right) &= u\left(c_t\left(\hat{\theta}^t\right),\rho d_t\left(\hat{\theta}^t\right)\right)+\varepsilon \\ u\left(c_{t+1}^\varepsilon\left(\hat{\theta}^t,\theta_{t+1}\right),\rho d_{t+1}\left(\hat{\theta}^t,\theta_{t+1}\right)\right) &= u\left(c_{t+1}\left(\hat{\theta}^t,\theta_{t+1}\right),\rho d_{t+1}\left(\hat{\theta}^t,\theta_{t+1}\right)\right)-\frac{\varepsilon}{\beta}, \quad \theta_{t+1}\in\Theta. \end{aligned}$$

For histories (θ^t, θ_{t+1}) with $\theta^t \neq \hat{\theta}^t$, set $c^\varepsilon_\tau = c_\tau$ for $\tau \in \{t, t+1\}$. Moreover, for periods $\tau \notin \{t, t+1\}$ set $c^\varepsilon_\tau = c_\tau$ for all histories. Adjust the capital stock of period $t+1$ in response to the changed consumption levels. Formally, define $K^\varepsilon_{t+1} := K_{t+1} - \zeta^\varepsilon$ with the help of the equation

$$\begin{aligned} & \Pi^t\left(\hat{\theta}^t\right) \sum_{\theta_{t+1}} \pi_{t+1}\left(\theta_{t+1}|\hat{\theta}^t\right) \left[c_{t+1}\left(\hat{\theta}^t, \theta_{t+1}\right) - c^\varepsilon_{t+1}\left(\hat{\theta}^t, \theta_{t+1}\right)\right] \\ = \quad & F\left(K_{t+1}, \mathbb{E}\left[y_{t+1}\right]\right) - F\left(K_{t+1} - \zeta^\varepsilon, \mathbb{E}\left[y_{t+1}\right]\right) + \left(1 - \delta^K\right) \zeta^\varepsilon. \end{aligned}$$

By construction, for all histories $\theta^T \in \Theta^T$, the perturbed allocation delivers the same lifetime utility as the original allocation. Hence, the perturbed allocation is also incentive compatible and yields the same social welfare. If the perturbed allocation requires fewer resources than the original allocation, there exists an incentive-feasible allocation with identical social welfare for a strictly smaller capital endowment. Then, by Lemma 1, the original allocation cannot be optimal.

Hence, a necessary condition for the optimality of $(\mathbf{c}, \mathbf{d}, \mathbf{y}, \mathbf{K})$ is that $\varepsilon = 0$ solves the following cost minimization problem:

$$\min_\varepsilon \left\{\Pi^t\left(\hat{\theta}^t\right) c^\varepsilon_t\left(\hat{\theta}^t\right) - \zeta^\varepsilon\right\}$$

This implies the first-order condition

$$0 = \Pi^t\left(\hat{\theta}^t\right) \frac{dc^\varepsilon_t\left(\hat{\theta}^t\right)}{d\varepsilon}|_{\varepsilon=0} - \frac{d\zeta^\varepsilon}{d\varepsilon}|_{\varepsilon=0}$$

which is equivalent to

$$\frac{\Pi^t\left(\hat{\theta}^t\right)}{u_c\left(c_t\left(\hat{\theta}^t\right), \rho d_t\left(\hat{\theta}^t\right)\right)} = \frac{\Pi^t\left(\hat{\theta}^t\right)}{\beta R_{t+1}} \sum_{\theta_{t+1}} \frac{\pi_{t+1}\left(\theta_{t+1}|\hat{\theta}^t\right)}{u_c\left(c_{t+1}\left(\hat{\theta}^t, \theta_{t+1}\right), \rho d_{t+1}\left(\hat{\theta}^t, \theta_{t+1}\right)\right)}$$

with $R_{t+1} := 1 - \delta^K + F_K\left(K_{t+1}, \mathbb{E}\left[y_{t+1}\right]\right)$. Dividing by $\Pi^t\left(\hat{\theta}^t\right)$, Eq. (10) follows. □

Proof of Proposition 4. Let $(\mathbf{c}, \mathbf{d}, \mathbf{y}, \mathbf{K})$ be an incentive-feasible allocation with interior consumption. Let $\hat{\theta}^t \in \Theta^t$. Consider a one-time perturbation of investment in the durable good:

$i_t^\varepsilon\left(\hat{\theta}^t\right) = i_t\left(\hat{\theta}^t\right) - \varepsilon$. The corresponding change in the stock of the durable good is

$$\begin{aligned} d_t^\varepsilon\left(\hat{\theta}^t\right) &= d_t\left(\hat{\theta}^t\right) - \varepsilon \\ d_\tau^\varepsilon\left(\hat{\theta}^t, \theta_{t+1}^\tau\right) &= d_\tau\left(\hat{\theta}^t, \theta_{t+1}^\tau\right) - \delta^{\tau-t}\varepsilon, \quad \tau > t,\ \theta_{t+1}^\tau \in \Theta^{\tau-t}, \end{aligned}$$

where $\theta_{t+1}^\tau = (\theta_{t+1}, \ldots, \theta_\tau)$ represents the realization in periods $t+1$ to τ. Adjust the level of nondurable consumption such that, in every period and for every realization, the agent obtains the same consumption utility as before:

$$\begin{aligned} u\left(c_t^\varepsilon\left(\hat{\theta}^t\right), \rho d_t^\varepsilon\left(\hat{\theta}^t\right)\right) &= u\left(c_t\left(\hat{\theta}^t\right), \rho d_t\left(\hat{\theta}^t\right)\right) \\ u\left(c_\tau^\varepsilon\left(\hat{\theta}^t, \theta_{t+1}^\tau\right), \rho d_\tau^\varepsilon\left(\hat{\theta}^t, \theta_{t+1}^\tau\right)\right) &= u\left(c_\tau\left(\hat{\theta}^t, \theta_{t+1}^\tau\right), \rho d_\tau\left(\hat{\theta}^t, \theta_{t+1}^\tau\right)\right), \quad \tau > t,\ \theta_{t+1}^\tau \in \Theta^{\tau-t}. \end{aligned}$$

For periods $\tau < t$ set $c_\tau^\varepsilon = c_\tau$ and $d_\tau^\varepsilon = d_\tau$ for all histories. Moreover, for periods $\tau \geq t$ and histories $\left(\theta^t, \theta_{t+1}, \ldots, \theta_\tau\right)$ with $\theta^t \neq \hat{\theta}^t$, set $c_\tau^\varepsilon = c_\tau$ and $d_\tau^\varepsilon = d_\tau$.

Define the capital stock K_T^ε in the last period as follows:

$$\begin{aligned} &\sum_{\theta_{t+1}^T \in \Theta^{T-t}} \Pi^T\left(\hat{\theta}^t, \theta_{t+1}^T\right)\left[c_T^\varepsilon\left(\hat{\theta}^t, \theta_{t+1}^T\right) - c_T\left(\hat{\theta}^t, \theta_{t+1}^T\right)\right] \\ = \ &\left(1-\delta^K\right) K_T^\varepsilon + F\left(K_T^\varepsilon, \mathbb{E}\left[y_T\right]\right) - \left(1-\delta^K\right) K_T - F\left(K_T, \mathbb{E}\left[y_T\right]\right). \end{aligned}$$

For periods $t < \tau < T$, define the capital stocks K_τ^ε recursively. Given $K_{\tau+1}^\varepsilon$, set K_τ^ε such that the feasibility constraint in period τ is satisfied:

$$\begin{aligned} &\sum_{\theta_{t+1}^\tau \in \Theta^{\tau-t}} \Pi^\tau\left(\hat{\theta}^t, \theta_{t+1}^\tau\right)\left[c_\tau^\varepsilon\left(\hat{\theta}^t, \theta_{t+1}^\tau\right) - c_\tau\left(\hat{\theta}^t, \theta_{t+1}^\tau\right)\right] \\ = \ &\left(1-\delta^K\right) K_\tau^\varepsilon + F\left(K_\tau^\varepsilon, \mathbb{E}\left[y_\tau\right]\right) - K_{\tau+1}^\varepsilon - \left(1-\delta^K\right) K_\tau - F\left(K_\tau, \mathbb{E}\left[y_\tau\right]\right) + K_{\tau+1}. \end{aligned}$$

By construction, for all histories $\theta^T \in \Theta^T$, the perturbed allocation delivers the same lifetime utility as the original allocation. Hence, the perturbed allocation is also incentive compatible and yields the same social welfare. If the perturbed allocation requires fewer resources in period t than the original allocation, there exists an incentive-feasible allocation with identical social welfare for a strictly smaller capital endowment. Then, by Lemma 1, the original allocation cannot be optimal.

Hence, the allocation $(\mathbf{c}, \mathbf{d}, \mathbf{y}, \mathbf{K})$ can only be optimal if $\varepsilon = 0$ solves the following cost minimization problem:

$$\min_{\varepsilon} \left\{ \Pi^t \left(\hat{\theta}^t\right) \left[c_t^{\varepsilon} \left(\hat{\theta}^t\right) + i_t^{\varepsilon} \left(\hat{\theta}^t\right) \right] + K_{t+1}^{\varepsilon} \right\}$$

This implies the first-order condition

$$0 = \Pi^t \left(\hat{\theta}^t\right) \frac{dc_t^{\varepsilon} \left(\hat{\theta}^t\right)}{d\varepsilon}|_{\varepsilon=0} - \Pi^t \left(\hat{\theta}^t\right) + \frac{dK_{t+1}^{\varepsilon}}{d\varepsilon}|_{\varepsilon=0}.$$

By the construction of the perturbed allocation, we have

$$\begin{aligned} \frac{dc_t^{\varepsilon} \left(\hat{\theta}^t\right)}{d\varepsilon}|_{\varepsilon=0} &= \rho \frac{u_s \left(c_t \left(\hat{\theta}^t\right), \rho d_t \left(\hat{\theta}^t\right)\right)}{u_c \left(c_t \left(\hat{\theta}^t\right), \rho d_t \left(\hat{\theta}^t\right)\right)} \\ \frac{dc_\tau^{\varepsilon} \left(\hat{\theta}^t, \theta_{t+1}^\tau\right)}{d\varepsilon}|_{\varepsilon=0} &= \rho \delta^{\tau-t} \frac{u_s \left(c_\tau \left(\hat{\theta}^t, \theta_{t+1}^\tau\right), \rho d_\tau \left(\hat{\theta}^t, \theta_{t+1}^\tau\right)\right)}{u_c \left(c_\tau \left(\hat{\theta}^t, \theta_{t+1}^\tau\right), \rho d_\tau \left(\hat{\theta}^t, \theta_{t+1}^\tau\right)\right)}. \end{aligned}$$

Moreover, the derivatives of K_τ^{ε}, $t < \tau < T$, and K_T^{ε} are implicitly given by

$$\begin{aligned} R_\tau \frac{dK_\tau^{\varepsilon}}{d\varepsilon}|_{\varepsilon=0} &= \sum_{\theta_{t+1}^\tau \in \Theta^{\tau-t}} \Pi^\tau \left(\hat{\theta}^t, \theta_{t+1}^\tau\right) \frac{dc_\tau^{\varepsilon} \left(\hat{\theta}^t, \theta_{t+1}^\tau\right)}{d\varepsilon}|_{\varepsilon=0} + \frac{dK_{\tau+1}^{\varepsilon}}{d\varepsilon}|_{\varepsilon=0} \\ R_T \frac{dK_T^{\varepsilon}}{d\varepsilon}|_{\varepsilon=0} &= \sum_{\theta_{t+1}^T \in \Theta^{T-t}} \Pi^T \left(\hat{\theta}^t, \theta_{t+1}^T\right) \frac{dc_T^{\varepsilon} \left(\hat{\theta}^t, \theta_{t+1}^T\right)}{d\varepsilon}|_{\varepsilon=0} \end{aligned}$$

where $R_\tau = \left(1 - \delta^K + F_K \left(K_\tau, \mathbb{E}\left[y_\tau\right]\right)\right)$ for $t < \tau \leq T$. This implies

$$\frac{dK_{t+1}^{\varepsilon}}{d\varepsilon}|_{\varepsilon=0} = \Pi^t \left(\hat{\theta}^t\right) \rho \sum_{\tau=t+1}^{T} q_t^\tau \delta^{\tau-t} \mathbb{E} \left[\frac{u_s \left(c_\tau, \rho d_\tau\right)}{u_c \left(c_\tau, \rho d_\tau\right)} |\hat{\theta}^t \right]$$

where $q_{t,\tau} = \left(R_{t+1} \cdots R_\tau\right)^{-1}$. After combining these equations and dividing by $\Pi^t \left(\hat{\theta}^t\right)$, we can write the first-order condition of the cost minimization problem as

$$0 = \rho \frac{u_s \left(c_t \left(\hat{\theta}^t\right), \rho d_t \left(\hat{\theta}^t\right)\right)}{u_c \left(c_t \left(\hat{\theta}^t\right), \rho d_t \left(\hat{\theta}^t\right)\right)} - 1 + \rho \sum_{\tau=t+1}^{T} q_{t,\tau} \delta^{\tau-t} \mathbb{E} \left[\frac{u_s \left(c_\tau, \rho d_\tau\right)}{u_c \left(c_\tau, \rho d_\tau\right)} |\hat{\theta}^t \right].$$

Using the conventions $q_{t,t} = \delta^0 = 1$, we obtain Eq. (11). □

Proof of Lemma 3. Let $(\mathbf{c}, \mathbf{d}, \mathbf{y}, \mathbf{K})$ be an incentive-feasible allocation with interior consumption.

Let $t < T$ and $\hat{\theta}^t \in \Theta^t$. Consider a one-time perturbation of the stock of the durable good $d_t^\varepsilon\left(\hat{\theta}^t\right) = d_t\left(\hat{\theta}^t\right) + \varepsilon$. In terms of investment, the perturbation is defined by $i_t^\varepsilon\left(\hat{\theta}^t\right) = i_t\left(\hat{\theta}^t\right) + \varepsilon$ and $i_{t+1}^\varepsilon\left(\hat{\theta}^t, \theta_{t+1}\right) = i_{t+1}\left(\hat{\theta}^t, \theta_{t+1}\right) - \delta\varepsilon$. Adjust the level of nondurable consumption in period t to compensate the agent for the reduced durable consumption service: $u\left(c_t^\varepsilon\left(\hat{\theta}^t\right), \rho d_t^\varepsilon\left(\hat{\theta}^t\right)\right) = u\left(c_t\left(\hat{\theta}^t\right), \rho d_t\left(\hat{\theta}^t\right)\right)$. For histories $\theta^t \neq \hat{\theta}^t$ set $c_t^\varepsilon = c_t$ and $d_t^\varepsilon = d_t$. Moreover, for periods $\tau \neq t$ set $c_\tau^\varepsilon = c_\tau$ and $d_\tau^\varepsilon = d_\tau$. Adjust the capital stock of period $t+1$ in response to the changed durable investment. Formally, define $K_{t+1}^\varepsilon := K_{t+1} - \zeta^\varepsilon$ with the help of the equation

$$\Pi^t\left(\hat{\theta}^t\right)\delta\varepsilon = F\left(K_{t+1}, \mathbb{E}\left[y_{t+1}\right]\right) - F\left(K_{t+1} - \zeta^\varepsilon, \mathbb{E}\left[y_{t+1}\right]\right) + \left(1 - \delta^K\right)\zeta^\varepsilon.$$

By construction, the perturbed allocation delivers the same utility as the original allocation for all periods and all histories. Hence, the perturbed allocation is also incentive compatible and yields the same social welfare. If the perturbed allocation requires fewer resources than the original allocation, there exists an incentive-feasible allocation with identical social welfare for a strictly smaller capital endowment. Then, by Lemma 1, the original allocation cannot be optimal.

Hence, a necessary condition for the optimality of $(\mathbf{c}, \mathbf{d}, \mathbf{y}, \mathbf{K})$ is that $\varepsilon = 0$ solves the following cost minimization problem:

$$\min_{\varepsilon}\left\{\Pi^t\left(\hat{\theta}^t\right)\left[c_t^\varepsilon\left(\hat{\theta}^t\right) + i_t^\varepsilon\left(\hat{\theta}^t\right)\right] - \zeta^\varepsilon\right\}.$$

This implies the first-order condition

$$\frac{\rho u_s\left(c_t\left(\hat{\theta}^t\right), \rho d_t\left(\hat{\theta}^t\right)\right)}{u_c\left(c_t\left(\hat{\theta}^t\right), \rho d_t\left(\hat{\theta}^t\right)\right)} = 1 - \frac{\delta}{R_{t+1}} \tag{16}$$

where $R_{t+1} := 1 - \delta^K + F_K\left(K_{t+1}, \mathbb{E}\left[y_{t+1}\right]\right)$. Moreover, by a very similar argument (compare Remark 1), the first-order condition for the final period is

$$\frac{\rho u_s\left(c_T\left(\hat{\theta}^T\right), \rho d_T\left(\hat{\theta}^T\right)\right)}{u_c\left(c_T\left(\hat{\theta}^T\right), \rho d_T\left(\hat{\theta}^T\right)\right)} = 1 \tag{17}$$

By Eqs. (16) and (17), the marginal rates of substitution between durable and nondurable consumption are equalized across realizations within every period. Because u is a monotonic transformation of a homogeneous function and strictly concave, we have

$$\frac{u_s(c,s)}{u_c(c,s)}=\frac{u_s(c',s')}{u_c(c',s')}\iff\frac{c}{s}=\frac{c'}{s'}.$$

Therefore, Eqs. (16) and (17) imply that the durable good and the nondurable good are consumed in fixed proportions within every period. In particular, they are perfectly rank correlated. $\square$

Proof of Proposition 5. Let $(\mathbf{c}^*,\mathbf{d}^*,\mathbf{y}^*,\mathbf{K}^*)$ be an optimal allocation with interior consumption. The Substitution Euler Equation implies

$$\rho\frac{u_s(c_t^*,\rho d_t^*)}{u_c(c_t^*,\rho d_t^*)}=1-\rho\sum_{\tau=t+1}^{T}q_{t,\tau}^*\delta^{\tau-t}\mathbb{E}_t\left[\frac{u_s(c_\tau^*,\rho d_\tau^*)}{u_c(c_\tau^*,\rho d_\tau^*)}\right].$$

Equivalently,

$$\begin{aligned}&\rho\frac{u_s(c_t^*,\rho d_t^*)+\sum_{\tau=t+1}^{T}(\beta\delta)^{\tau-t}\,\mathbb{E}_t\left[u_s(c_\tau^*,\rho d_\tau^*)\right]}{u_c(c_t^*,\rho d_t^*)}\\&=1-\rho\sum_{\tau=t+1}^{T}q_{t,\tau}^*\delta^{\tau-t}\mathbb{E}_t\left[u_s(c_\tau^*,\rho d_\tau^*)\left(\frac{1}{u_c(c_\tau^*,\rho d_\tau^*)}-\frac{\beta^{\tau-t}}{q_{t,\tau}^*u_c(c_\tau^*,\rho d_\tau^*)}\right)\right].\end{aligned}\tag{18}$$

By the Inverse Euler Equation, for all $\tau>t$ we have

$$\mathbb{E}_t\left[\frac{1}{u_c(c_\tau^*,\rho d_\tau^*)}-\frac{\beta^{\tau-t}}{q_{t,\tau}^*u_c(c_t^*,\rho d_t^*)}\right]=0.$$

Therefore, the expectation of the following product coincides with its covariance,

$$\begin{aligned}&\mathbb{E}_t\left[u_s(c_\tau^*,\rho d_\tau^*)\left(\frac{1}{u_c(c_\tau^*,\rho d_\tau^*)}-\frac{\beta^{\tau-t}}{q_{t,\tau}^*u_c(c_t^*,\rho d_t^*)}\right)\right]\\=\;&\mathrm{cov}_t\left(u_s(c_\tau^*,\rho d_\tau^*),\frac{1}{u_c(c_\tau^*,\rho d_\tau^*)}-\frac{\beta^{\tau-t}}{q_{t,\tau}^*u_c(c_t^*,\rho d_t^*)}\right)\\=\;&\mathrm{cov}_t\left(u_s(c_\tau^*,\rho d_\tau^*),\frac{1}{u_c(c_\tau^*,\rho d_\tau^*)}\right).\end{aligned}$$

Consequently, Eq. (18) shows that in period t the marginal rate of substitution between the service flow from investing in the durable good and the consumption of the nondurable good exceeds unity (the marginal rate of transformation) if and only if

$$\sum_{\tau=t+1}^{T} q_{t,\tau}^{*}\delta^{\tau-t}\operatorname{cov}_t\left(-u_s\left(c_\tau^*,\rho d_\tau^*\right),\frac{1}{u_c\left(c_\tau^*,\rho d_\tau^*\right)}\right)\geq 0. \tag{19}$$

Given that the consumption goods are perfectly rank correlated after period t, for all $\tau > t$ there exist a "sufficient statistic" $\lambda_\tau : \Theta^\tau \to \mathbb{R}$ and strictly increasing functions $C_\tau : \mathbb{R} \to \mathbb{R}_+$, $D_\tau : \mathbb{R} \to \mathbb{R}_+$ such that

$$c_\tau^* = C_\tau \circ \lambda_\tau \quad , \quad d_\tau^* = D_\tau \circ \lambda_\tau$$

The marginal utilities u_s and u_c depend on the realization of uncertainty only through the statistic λ_τ. Consider two realizations $\lambda = \lambda_\tau(\theta^\tau)$ and $\hat{\lambda} = \lambda_\tau(\hat{\theta}^\tau)$ with $\lambda \geq \hat{\lambda}$. Then, by the mean value theorem, there exists a number $\xi \in \left[D_\tau(\hat{\lambda}), D_\tau(\lambda)\right]$ and a number $\kappa \in \left[C_\tau(\hat{\lambda}), C_\tau(\lambda)\right]$ such that

$$\begin{aligned}
& u_s\left(C_\tau(\lambda),\rho D_\tau(\lambda)\right) - u_s\left(C_\tau(\hat{\lambda}),\rho D_\tau(\hat{\lambda})\right) \\
= \;& u_s\left(C_\tau(\lambda),\rho D_\tau(\lambda)\right) - u_s\left(C_\tau(\lambda),\rho D_\tau(\hat{\lambda})\right) + u_s\left(C_\tau(\lambda),\rho D_\tau(\hat{\lambda})\right) - u_s\left(C_\tau(\hat{\lambda}),\rho D_\tau(\hat{\lambda})\right) \\
= \;& \rho u_{ss}\left(C_\tau(\lambda),\rho\xi\right)\left[D_\tau(\lambda) - D_\tau(\hat{\lambda})\right] + u_{cs}\left(\kappa,\rho D_\tau(\hat{\lambda})\right)\left[C_\tau(\lambda) - C_\tau(\hat{\lambda})\right].
\end{aligned}$$

Hence, if $u_{cs} \leq 0$, the marginal utility $u_s\left(C_\tau(\lambda),\rho D_\tau(\lambda)\right)$ is strictly decreasing in λ. Similarly, for the marginal utility of nondurable consumption, there exists a number $\xi' \in \left[D_\tau(\hat{\lambda}), D_\tau(\lambda)\right]$ and a number $\kappa' \in \left[C_\tau(\hat{\lambda}), C_\tau(\lambda)\right]$ such that

$$\begin{aligned}
& u_c\left(C_\tau(\lambda),\rho D_\tau(\lambda)\right) - u_c\left(C_\tau(\hat{\lambda}),\rho D_\tau(\hat{\lambda})\right) \\
= \;& u_c\left(C_\tau(\lambda),\rho D_\tau(\lambda)\right) - u_c\left(C_\tau(\lambda),\rho D_\tau(\hat{\lambda})\right) + u_c\left(C_\tau(\lambda),\rho D_\tau(\hat{\lambda})\right) - u_c\left(C_\tau(\hat{\lambda}),\rho D_\tau(\hat{\lambda})\right) \\
= \;& \rho u_{cs}\left(C_\tau(\lambda),\rho\xi'\right)\left[D_\tau(\lambda) - D_\tau(\hat{\lambda})\right] + u_{cc}\left(\kappa',\rho D_\tau(\hat{\lambda})\right)\left[C_\tau(\lambda) - C_\tau(\hat{\lambda})\right].
\end{aligned}$$

Once more, if $u_{cs} \leq 0$, the marginal utility $u_c\left(C_\tau(\lambda),\rho D_\tau(\lambda)\right)$ is strictly decreasing in λ. This implies that the random variables $-u_s$ and $1/u_c$ in Eq. (19) are positively monotonically related. Hence, their covariance is nonnegative and Eq. (19) is satisfied. Moreover, because $u_{cc} < 0$ and $u_{ss} < 0$, the covariance is strictly positive unless λ_τ is constant. Thus, unless consumption is fully insured in periods $t+1, \ldots, T$, the result becomes strict. $\square$

Proof of Proposition 6. Fix some effective labor plan $y = (y_1, \ldots, y_T)$. Given this plan, the agent's intertemporal consumption problem is to solve

$$\max_{c,i,k} \sum_{t=1}^{T} \beta^{t-1} \mathbb{E}\left[u\left(c_t, \rho d_t\right)\right]$$

subject to $k_1 \leq K_1^*$ and the following period-by-period budget constraints:

$$\begin{aligned}
c_1 + i_1 + k_2 &\leq w_1^* y_1 + R_1^* k_1 - \mathcal{T}_1^y\left(y_1\right), \\
c_t + i_t + k_{t+1} &\leq w_t^* y_t + R_t^* k_t - \mathcal{T}_t^y\left(y^t\right) - \mathcal{T}_t^k\left(y^t, k_t\right) \quad \text{for } 1 < t < T, \\
c_T + i_T &\leq w_T^* y_T + R_T^* k_T - \mathcal{T}_T^y\left(y^T\right) - \mathcal{T}_T^k\left(y^T, k_T\right) - \mathcal{T}_T^d\left(y^T, i^T\right).
\end{aligned}$$

For any given labor plan y, this is a strictly concave problem with linear constraints. By monotonicity, we have $k_1 = K_1^*$ and all budget constraints are satisfied with inequality. The necessary and sufficient first-order conditions of the problem are

$$\begin{aligned}
0 &= -u_c\left(c_t, \rho d_t\right) + \rho \mathbb{E}_t\left[\sum_{\tau=t}^{T} (\beta\delta)^{\tau-t} u_s\left(c_\tau, \rho d_\tau\right)\right] - \beta^{T-t}\mathbb{E}_t\left[\kappa_t u_c\left(c_T, \rho d_T\right)\right] \quad \text{for } 1 \leq t \leq T, \\
0 &= -u_c\left(c_t, \rho d_t\right) + \beta \mathbb{E}_t\left[\left(R_{t+1}^* - \frac{d\mathcal{T}_{t+1}^k\left(y^{t+1}, k_{t+1}\right)}{dk_{t+1}}\right) u_c\left(c_{t+1}, \rho d_{t+1}\right)\right] \quad \text{for } 1 \leq t < T.
\end{aligned}$$

I claim that the solution of this problem is $(c_t, d_t, k_t) = \left(\hat{c}^*\left(y^t\right), \hat{d}^*\left(y^t\right), K_t^*\right)$ for all t. Clearly, by the construction of the tax system, the plan $\left(\hat{c}^*\left(y^t\right), \hat{d}^*\left(y^t\right), K_t^*\right)$ satisfies the period-by-period budget constraints. Moreover, it is easy to verify that the plan satisfies the necessary and sufficient first-order conditions. Hence, for a fixed labor plan y and given taxes and prices, it is optimal for the agent to choose the intertemporal consumption plan $\left(\hat{c}^*\left(y^t\right), \hat{d}^*\left(y^t\right), K_t^*\right)$ for all t.

It remains to show that the effective labor plan $y = \mathbf{y}^*$ is optimal for the agent. Yet, because the plan $\left(\hat{c}^*\left(y^t\right), \hat{d}^*\left(y^t\right), K_t^*\right)$ solves the intertemporal consumption problem for given y, choosing the labor plan is equivalent to the reporting problem in the social planner setup. Hence, because the allocation $(\mathbf{c}^*, \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}^*)$ is incentive compatible, the labor plan $\mathbf{y}^*$ solves the agent's decision problem in the decentralized economy. This step completes the proof. □

References

ABRAHAM, A., S. KOEHNE, AND N. PAVONI (2016): "Optimal income taxation when asset taxation is limited," *Journal of Public Economics*, 136, 14–29.

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

ATKINSON, A. AND J. STIGLITZ (1976): "The design of tax structure: Direct versus indirect taxation," *Journal of Public Economics*, 6, 55 – 75.

BAJARI, P., P. CHAN, D. KRUEGER, AND D. MILLER (2013): "A Dynamic Model of Housing Demand: Estimation and Policy Implications," *International Economic Review*, 54, 409–442.

CHRISTIANSEN, V. (1984): "Which commodity taxes should supplement the income tax?" *Journal of Public Economics*, 24, 195 – 220.

CREMER, H. AND F. GAHVARI (1995a): "Uncertainty and optimal taxation: In defense of commodity taxes," *Journal of Public Economics*, 56, 291–310.

——— (1995b): "Uncertainty, optimal taxation and the direct versus indirect tax controversy," *Economic Journal*, 1165–1179.

DA COSTA, C. AND I. WERNING (2002): "Commodity taxation and social insurance," MIT, Mimeo.

DEATON, A. (1979): "Optimally uniform commodity taxes," *Economics Letters*, 2, 357 – 361.

DIAMOND, P. A. AND J. A. MIRRLEES (1978): "A model of social insurance with variable retirement," *Journal of Public Economics*, 10, 295 – 336.

FARHI, E. AND I. WERNING (2013): "Insurance and taxation over the life cycle," *Review of Economic Studies*, 80, 596–635.

GOLOSOV, M., N. KOCHERLAKOTA, AND A. TSYVINSKI (2003): "Optimal Indirect and Capital Taxation," *Review of Economic Studies*, 70, 569–587.

GOLOSOV, M., M. TROSHKIN, AND A. TSYVINSKI (2015): "Redistribution and Social Insurance," *American Economic Review*, forthcoming.

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

JACOBS, B. AND R. BOADWAY (2014): "Optimal linear commodity taxation under optimal non-linear income taxation," *Journal of Public Economics*, 117, 201 – 210.

KAPLOW, L. (2006): "On the undesirability of commodity taxation even when income taxation is not optimal," *Journal of Public Economics*, 90, 1235 – 1250.

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, 1587–1621.

KOEHNE, S. AND M. KUHN (2015): "Optimal taxation in a habit formation economy," *Journal of Public Economics*, 122, 31–39.

KONISHI, H. (1995): "A Pareto-improving commodity tax reform under a smooth nonlinear income tax," *Journal of Public Economics*, 56, 413 – 446.

LAROQUE, G. R. (2005): "Indirect taxation is superfluous under separability and taste homogeneity: a simple proof," *Economics Letters*, 87, 141 – 144.

LI, W., H. LIU, F. YANG, AND R. YAO (2015): "Housing over time and over the life cycle: a structural estimation," Working Paper 15-4, Federal Reserve Bank of Philadelphia.

NAITO, H. (1999): "Re-examination of uniform commodity taxes under a non-linear income tax system and its implication for production efficiency," *Journal of Public Economics*, 71, 165 – 188.

OGAKI, M. AND C. M. REINHART (1998): "Measuring Intertemporal Substitution: The Role of Durable Goods," *Journal of Political Economy*, 106, 1078–1098.

PIAZZESI, M., M. SCHNEIDER, AND S. TUZEL (2007): "Housing, consumption and asset pricing," *Journal of Financial Economics*, 83, 531–569.

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

SAEZ, E. (2002): "The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes," *Journal of Public Economics*, 83, 217 – 230.

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