On the Taxation of Durable Goods

Sebastian Koehne*

June 18, 2015

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 goods, nondurable goods and labor effort. If those conditions are not met, wealth effects and substitution effects justify the use of differential commodity taxes. To characterize the sign of the tax differential, the paper exploits a "Substitution Euler Equation" that links the marginal rate of substitution between durable and nondurable consumption across time. Finally, 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, SE-10691 Stockholm, Sweden, Phone: +46 8 16 35 64, sebastian.koehne@iies.su.se, and CESifo, Munich, Germany. Sebastian Koehne gratefully acknowledges financial support from Knut och Alice Wallenbergs Stiftelse (grant 2012.0315).

1 Introduction

The analysis of commodity taxation by Atkinson and Stiglitz (1976) is arguably one of the most fundamental contributions to public finance theory. The analysis has influenced academic discussions and policy circles alike.[1] However, as argued by the present paper, the Atkinson–Stiglitz framework fails to capture a major share of goods. In fact, approximately 40% of a typical household's consumption expenditure are spent on durable goods (housing, cars, furniture, consumer electronics, etc.).[2] Durable goods are not easily represented in the static Atkinson–Stiglitz framework because they are stock (rather than flow) variables.

The present 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 on uniform commodity taxation holds true for nondurable goods in dynamic frameworks. This finding extends a result by Golosov, Kocherlakota, and Tsyvinski (2003) to a model that includes durable goods in addition to nondurable goods.

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

Third, I study the properties of optimal differential commodity taxation. I derive a "Substitution Euler Equation" that links the marginal rate of substitution between durable and nondurable consumption across time (Proposition 3). Then, I exploit this equation to characterize the optimal tax differential between durable and nondurable goods. I show that wealth effects and substitution effects justify the use of differential commodity taxes in order to facilitate the

[1] For example, the insights by Atkinson and Stiglitz (1976) have motivated some of the policy recommendations of the recent 'Mirrlees Review' on tax policy in the United Kingdom.

[2] 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.

provision of incentives in the future (Propositions 4 to 7). Wealth effects emerge because investment in a durable good modifies the valuation of future investment. Substitution effects stem from nonseparabilities across goods and describe how an investment in the durable good affects the future valuation of other consumption goods. Specifically, if a durable and a nondurable good are (Edgeworth) substitutes, the durable good should be taxed at a higher rate than the nondurable good.[3]

Fourth, I show that optimal commodity taxation depends on the frequency at which durable goods are adjusted. In the polar case where investment in a durable good occurs only once, wealth effects (captured by the curvature of the utility from durable consumption) play no role and commodity taxation is entirely determined by the nonseparability of the utility function across goods. This finding suggests that adjustment frictions may moderate the tax differential between durable and nondurable goods.

Fifth, I explore the consequences of uncertain depreciation rates. I show that, when the depreciation of durable goods is stochastic and unobservable, subsidies to durable goods may help to provide insurance against this source of uncertainty. Sixth, 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.

This paper is structured 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 optimal commodity taxes are uniform. Section 4 studies differential commodity taxes in a two-period setting with one durable and one nondurable good. Section 5 documents the robustness of the basic results with respect to the time horizon and the number of goods. Moreover, some alternative specifications of uncertainty, e.g., stochastic depreciation rates, are considered. Section 6 provides concluding remarks and applies the model to the case of housing. Appendix A collects the proofs of all theoretical results.

[3]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 is closely related to the study of pre-committed goods by Cremer and Gahvari (1995a,b). I extend their approach by proposing an explicit model of durability in a dynamic framework. I show that the tax implications of durable goods differ notably from those of pre-committed goods. Thus, pre-commitment does not easily translate into durability. At the same time, I suggest a novel interpretation of their findings: subsidies to pre-committed goods can be seen as a rationale for intertemporal wedges on *nondurable* goods in dynamic frameworks. The discussion in Section 6.2 compares durable goods and pre-committed goods in more detail.

More generally, the present 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. More recently, Jacobs and Boadway (2014) study optimal *linear* commodity taxation in a similar framework. 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.

In an unpublished paper, da Costa and Werning (2002) show that the role of commodity taxation coincides for hidden action models and adverse selection models. The analysis in the present paper rests on variational arguments that change the consumption allocation but leave 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 (2014) 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. However, the implications of durable goods for optimal taxation, and especially for commodity taxation, are quite distinct. First, habit formation alone does not justify differential commodity taxes. Differential taxes may 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, durable goods give rise to different applications and relate to the taxation of housing and pre-committed goods. Moreover, durable goods have several aspects that matter for optimal taxation but lack a counterpart in models of habit formation, such as the frequency of adjustment or uncertainty with respect to depreciation. Finally, unlike earlier works on time-nonseparable preferences, the present paper identifies the sign of the commodity tax wedge *analytically* in an environment with a very general specification of uncertainty. In contrast, Koehne and Kuhn (2014) provide a numerical quantification of optimal labor and savings wedges for a habit formation model with a single consumption good and a special case of transitory binary skill shocks. Grochulski and Kocherlakota (2010) explore the properties of efficient decentralizations (with respect to labor supply and saving) in a similar habit formation model with one consumption good.

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 periods (with $2 \leq T \leq \infty$) 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 stock 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 stock of durable goods depreciates 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 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 stock is required to remain nonnegative at all times. The initial stock of durable goods is identical across agents and normalized to $d_0 = 0$. After period T, there is no activity and the stock of durable goods vanishes.

2.3 Uncertainty

Agents face idiosyncratic uncertainty regarding their productivity $\theta_t \in \Theta \subset \mathbb{R}_{++}$. To sidestep some formalities on the measurability and integrability of random variables, I assume that the productivity set Θ is a *finite* subset of $\mathbb{R}_{++}$. 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.[4]

As usual in this class of models, I assume that a law of large numbers applies. The individual distribution of uncertainty is thus identical to the realized cross-sectional distribution. The expectation operator with respect to the unconditional distribution of skill histories θ^T is denoted by $\mathbb{E}[\,\cdot\,]$. The notation $\mathbb{E}_t\left[\,\cdot\,\right] := \mathbb{E}\left[\,\cdot\,|\theta^t\right]$ represents expectations conditional on the period-t history θ^t. Similarly, $\mathrm{cov}_t\left(\,\cdot\,,\,\cdot\,\right)$ 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

[4]All results in this paper can be extended to productivity sets that are intervals or countable sets, and distributions without full support. However, such extensions will complicate the exposition and introduce some technical issues without adding economic insight.

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 capital stock, c_t denotes the bundle of nondurable consumption goods, d_t denotes the stock 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 investment $i_t : \Theta^t \to \mathbb{R}^M$, $t = 1, \ldots, T$, 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 $\delta^K = -r$ and $F(K,Y) = Y$, 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 fully 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:

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

An allocation $\left(\mathbf{c}^*, \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}^*\right)$ 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. For this property, it is sufficient to have a utility function that is bounded above.

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 $\bar{K}_1$.[5] 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 grow 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 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 $V\left(\bar{K}_1\right) = V\left(\bar{K}_1'\right)$ must be false and hence $V\left(\bar{K}_1\right) < V\left(\bar{K}_1'\right)$ must hold.

[5] This issue does not arise if the planner problem is set up as a cost minimization problem rather than a welfare maximization problem.

3 A maximal case of uniform commodity taxation

Optimal allocations impose tax distortions ("wedges") that any decentralizing tax system must generate. In the special case analyzed in this section, the consumption choice will be undistorted. This implies that optimal allocations can be decentralized by labor and capital taxes combined with *uniform* commodity taxes.

Note that an investment in the stock of durable goods affects the service flows in all remaining periods. The marginal utility from investing in durable good m at time t is 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 depreciation for that good, and ρ_m maps stocks to service flows.

An allocation is associated with uniform commodity taxes if, for any pair of goods within any period, the marginal rate of substitution coincides with the marginal rate of transformation. Formally, the definition is as follows.

Definition 2. An allocation $(\mathbf{c}, \mathbf{d}, \mathbf{y}, \mathbf{K})$ implies a *uniform taxation across all goods (within any period)* if, for all t and all $n \in \{1, \ldots, N\}$, $m \in \{1, \ldots, M\}$,

$$\begin{aligned} &\frac{\rho_m U_{N+m}\left(c_t, s_t, \frac{y_t}{\theta_t}\right) + \rho_m \mathbb{E}_t\left[\sum_{k=t+1}^{T} (\beta\delta_m)^{k-t}\, U_{N+m}\left(c_k, s_k, \frac{y_k}{\theta_k}\right)\right]}{U_n\left(c_t, s_t, \frac{y_t}{\theta_t}\right)} \\ &= \frac{G_{N+m}\left(\mathbb{E}\left[c_t\right], \mathbb{E}\left[i_t\right], K_{t+1}, K_t, \mathbb{E}\left[y_t\right]\right)}{G_n\left(\mathbb{E}\left[c_t\right], \mathbb{E}\left[i_t\right], K_{t+1}, K_t, \mathbb{E}\left[y_t\right]\right)}. \end{aligned} \tag{3}$$

The allocation implies a *uniform taxation across all nondurable goods (within any period)* if, for all t and all $n, n' \in \{1, \ldots, N\}$,

$$\frac{U_{n'}\left(c_t, s_t, \frac{y_t}{\theta_t}\right)}{U_n\left(c_t, s_t, \frac{y_t}{\theta_t}\right)} = \frac{G_{n'}\left(\mathbb{E}\left[c_t\right], \mathbb{E}\left[i_t\right], K_{t+1}, K_t, \mathbb{E}\left[y_t\right]\right)}{G_n\left(\mathbb{E}\left[c_t\right], \mathbb{E}\left[i_t\right], K_{t+1}, K_t, \mathbb{E}\left[y_t\right]\right)}. \tag{4}$$

Definition 2 requires some explanation. First, note that Eq. (3) only considers pairs that consist of one durable good m and one nondurable good n. However, if Eq. (3) is satisfied for all such pairs, the marginal rates of substitution and transformation will also coincide for any pair of durables (m, m') or any pair of nondurables (n, n'). Hence, the first part of Definition 2 does indeed refer to situations where *all* goods are taxed the same. Second, Definition 2 restricts the tax differentials across goods within any period to be zero but allows for tax differentials across time. Given that dynamic private information models generally lead to positive intertemporal wedges, tax differentials across time are a common feature of optimal allocations. Third, recall that the allocation objects (c_t, s_t, y_t) are functions of the period-t history. Therefore, Eqs. (3) and (4) are satisfied if and only if they hold for all possible histories.

Next, I show that a uniform taxation of *nondurable* goods is optimal if the preferences over those goods are weakly separable from labor.

Definition 3. 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, any optimal allocation with interior consumption implies a uniform taxation across all nondurable goods.*

The proof of Proposition 1 and all further proofs are relegated to the appendix. In models without durable goods, it is well known that a uniform taxation of goods is optimal when the preferences over consumption are weakly separable from labor (Atkinson and Stiglitz, 1976; Golosov et al., 2003). Proposition 1 allows nondurable goods to coexist with durable goods and shows, following a similar logic, that a uniform taxation across *nondurable* goods remains optimal when the preferences over nondurable goods are weakly separable from labor.

Now, I establish the main result of this section. The proposition guarantees a uniform taxation across *all* goods (durables and nondurables).

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*

$$U(c, s, e) = u(c) + \alpha \cdot s - v(e) \quad \textit{for all } (c, s, e) \in \mathbb{R}_+^{N+M+1} \tag{5}$$

where $\alpha \cdot s := \sum_{m=1}^{M} \alpha_m s_m$ *denotes the scalar product. Then, any optimal allocation with interior consumption implies a uniform taxation across all goods.*

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 and 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 goods should be taxed the same.

To ensure a uniform taxation of all goods, 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 frameworks with durable goods; see the discussion of housing in Section 6.1, for instance. Therefore, Proposition 2 is not widely applicable. However, the proposition establishes an important theoretical benchmark because it identifies a maximal case where the taxation of all goods is uniform. If any of the assumptions is relaxed, a uniform taxation will be no longer optimal.

Example. Suppose that the utility from consumption is additively separable and linear in durable consumption as in Eq. (5) but that period 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 second period is an element of the binary set $\{\theta_L, \theta_H\}$, with $\theta_L < \theta_H$ and probability weights $0 < \pi(\theta_L), \pi(\theta_H) < 1$. The production technology is linear in labor and the planner chooses an allocation to maximize social welfare

$$\max\ (u(c)+\alpha\cdot\rho i)\, v\left(1-\frac{y}{\theta}\right)+\beta\sum_{k=L,H}\pi(\theta_k)\left(u(c_k)+\alpha\cdot\rho(i_k+\delta i)\right) v\left(1-\frac{y_k}{\theta_k}\right)$$

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

$$\left(u(c_H)+\alpha\cdot\rho(i_H+\delta i)\right) v\left(1-\frac{y_H}{\theta_H}\right)\geq\left(u(c_L)+\alpha\cdot\rho(i_L+\delta i)\right) v\left(1-\frac{y_L}{\theta_H}\right).$$

The first-order conditions with respect to consumption in the first period imply

$$\begin{aligned}&\frac{\alpha_m\rho_m v\left(1-\frac{y}{\theta}\right)+\beta\alpha_m\rho_m\delta_m\sum_{k=L,H}\pi(\theta_k)v\left(1-\frac{y_k}{\theta_k}\right)}{u_n(c)v\left(1-\frac{y}{\theta}\right)}\\&=\frac{G_{N+m}\left(c,i,K_2,K_1,y\right)}{G_n\left(c,i,K_2,K_1,y\right)}+\frac{\mu\alpha_m\rho_m\delta_m\Delta v}{\lambda G_n\left(c,i,K_2,K_1,y\right)}\end{aligned}\tag{6}$$

where λ is the Lagrange multiplier for the feasibility constraint, μ the multiplier for the incentive constraint, and $\Delta v := v\left(1-\frac{y_L}{\theta_H}\right)-v\left(1-\frac{y_H}{\theta_H}\right)\geq 0$. Suppose that optimal consumption is interior and the incentive constraint in the second period is binding. Then, $\mu>0$ and $\Delta v>0$. By the first-order conditions, the marginal rate of substitution of durable good m for nondurable good n in the first period (the left-hand side of Eq. (6)) exceeds the marginal rate of transformation (the first term on the right-hand side of Eq. (6)). This means that durable goods are implicitly taxed at a higher rate than nondurables. Intuitively, by the multiplicative specification of utility, all consumption goods are complementary with leisure in this example. However, durable investment is also complementary with *future* leisure, whereas nondurable consumption is not. This motivates a tax on durable goods in order to make leisure less attractive in the second period.

As shown by the above example, the uniform taxation result breaks down if the additive separability between consumption and labor in Eq. (5) is relaxed. By constructing similar examples, it can be verified that the additive separability between durable consumption (or nondurable consumption) and labor cannot be relaxed individually either. Moreover, in the following section, it will become clear that nonseparability between durable and nondurable consumption, or nonlinearity of the utility from durable consumption, will also give rise to differential commodity taxes. Hence, the uniform taxation of goods in Proposition 2 remains a knife-edge case.

4 Differential commodity taxation for two goods in two periods

This section studies optimal commodity taxation in a two-period setting with one durable and one nondurable good. Rather than focusing on explicit tax instruments, I analyze the distortion of the consumption choice at optimal allocations. This level of generality is useful for at least three reasons. First, dynamic private information problems are much more tractable for theoretical analysis if no constraints other than incentive compatibility and resource feasibility are present. Second, the distortions imposed by optimal allocations often provide guidelines for tax rates in settings with simple, linear instruments (Farhi and Werning, 2013). Third, complex forms of commodity taxation constitute a feature of many existing tax systems. For example, investment in owner-occupied housing is often treated differently from other goods because interest payments for mortgages can be deducted from the income tax. Such provisions make the actual taxation of housing nonlinear and dependent on income.

Define the (implicit) tax differential between a durable and a nondurable good as follows.

Definition 4. Let $n \in \{1, \ldots, N\}$, $m \in \{1, \ldots, M\}$ and $t \in \{1, \ldots, T\}$. An allocation $(\mathbf{c}, \mathbf{d}, \mathbf{y}, \mathbf{K})$ implies that *the durable good m should be taxed at a higher rate than the nondurable good n (in period t)* if

$$\begin{aligned} &\frac{\rho_m U_{N+m}\left(c_t, s_t, \frac{y_t}{\theta_t}\right) + \rho_m \mathbb{E}_t\left[\sum_{k=t+1}^{T} \left(\beta\delta_m\right)^{k-t} U_{N+m}\left(c_k, s_k, \frac{y_k}{\theta_k}\right)\right]}{U_n\left(c_t, s_t, \frac{y_t}{\theta_t}\right)} \\ &\geq \frac{G_{N+m}\left(\mathbb{E}\left[c_t\right], \mathbb{E}\left[i_t\right], K_{t+1}, K_t, \mathbb{E}\left[y_t\right]\right)}{G_n\left(\mathbb{E}\left[c_t\right], \mathbb{E}\left[i_t\right], K_{t+1}, K_t, \mathbb{E}\left[y_t\right]\right)}. \end{aligned} \tag{7}$$

If the inequality is strict, the durable good m should be taxed at a *strictly* higher rate than the nondurable good n. If the reverse inequality holds, the durable good m should be taxed at a *(strictly) lower* rate than the nondurable good n.

Once more, recall that the allocation variables depend on the period-t history. Thus, Definition 4 only applies when the respective inequality is satisfied for all histories.

If Eq. (7) holds, the marginal rate of substitution of the durable good for the nondurable good exceeds the marginal rate of transformation. In this case, any decentralization of the allocation must distort the price of the durable good upward. Such a distortion can, for instance, be achieved by a commodity tax at the time of transaction or by an income tax that depends on the stock of the durable good.

4.1 Necessary conditions for intertemporal optimality

To make the analysis of differential commodity taxation more tractable, I consider a two-period setting with one durable and one nondurable good: $T = 2$ and $M = N = 1$. (These assumptions will be relaxed in Section 5.1.) The utility function is additively separable between consumption and labor effort:

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

where u is strictly increasing, *strictly* concave (unlike the specification in Proposition 2) and twice continuously differentiable. The technology is described by a strictly increasing, continuously differentiable production function F that produces a final good. The final good can be used for 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 at an optimal allocation by $R^* := 1 - \delta^K + F_K\left(K_2^*, \mathbb{E}\left[y_2^*\right]\right)$.

The following two equations characterize 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,*

$$\frac{\beta R^*}{u_c\left(c_1^*, \rho d_1^*\right)} = \mathbb{E}_1\left[\frac{1}{u_c\left(c_2^*, \rho d_2^*\right)}\right]. \tag{10}$$

Proposition 3 (Substitution Euler Equation)**.** *Let* $(\mathbf{c}^*, \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}^*)$ *be an optimal allocation with interior consumption. Then,*

$$1 = \rho \frac{u_s\left(c_1^*, \rho d_1^*\right)}{u_c\left(c_1^*, \rho d_1^*\right)} + \frac{\rho\delta}{R^*}\mathbb{E}_1\left[\frac{u_s\left(c_2^*, \rho d_2^*\right)}{u_c\left(c_2^*, \rho d_2^*\right)}\right]. \tag{11}$$

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. The introduction of durable goods does not affect the underlying logic. However, the Substitution Euler Equation is novel. This equation is based on the following reasoning. Suppose that investment in the durable good is reduced by one marginal unit at $t = 1$. In turn, nondurable consumption is increased in both periods so 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 explains the left-hand side of Eq. (11). The right-hand side of Eq. (11) captures the present-value of the resources needed to increase nondurable consumption accordingly. The Substitution Euler Equation states that no resources are freed up if durable consumption is exchanged for nondurable consumption in an incentive-neutral way.

4.2 Optimal commodity taxation with one-time investment or repeated investment in durables

For a moment, I will assume that investment in the durable good is only possible in the first period. This assumption allows me to isolate one specific channel of optimal commodity taxation. Moreover, this setup relates to models with explicit adjustment frictions (transaction costs, search costs, nondivisibilities, etc.). By comparing the tax implications when the durable good cannot be adjusted to those when investment occurs in every period, one can draw tentative conclusions on how the optimal taxes on durable goods with large adjustment frictions differ from durable goods whose adjustment is frictionless.

Proposition 4 (Differential taxation with one-time investment)**.** *Let $T = 2$ and $M = N = 1$. Suppose that investment in the durable good is only possible in the first period. Then, any optimal allocation with interior consumption has the following implications: If $u_{cs} \leq 0$, the durable good should be taxed at a higher rate than the nondurable good (in period 1). If $u_{cs} < 0$ and consumption is not fully insured in the second period, the result becomes strict. If $u_{cs} \geq 0$, the result is reversed.*

For the setup with one-time investment, Proposition 4 shows that durable goods should be taxed differently from nondurable goods if the preferences are nonseparable across 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 period. This dynamic incentive effect is precisely the reason why the Atkinson–Stiglitz result fails.

Intuitively, if durable consumption is an (Edgeworth) substitute with nondurable consumption ($u''_{cs} \leq 0$), investment in the durable good reduces the marginal value of nondurable consumption in the following period. In that case, variations of nondurable consumption translate into smaller variations of utility. This effect is socially harmful because it tightens the incentive compatibility constraint. To account for this negative externality, the durable good should be taxed more than the nondurable good. Conversely, if durable consumption is complementary with nondurable consumption, the argument is reversed and the durable good should be taxed at a lower rate than the nondurable good.

The setup with one-time investment in the durable good identifies one rationale for differential taxation but abstracts from the role of future investment. Now, I return to the standard framework with investment in every period. Then, the case for differential commodity taxation becomes even stronger.

Proposition 5 (Differential taxation with repeated investment)**.** *Let $T = 2$ and $M = N = 1$. Suppose that investment in the durable good is possible in both periods. Then, any optimal allocation with interior consumption has the following implications: (i) The durable good should be taxed at a higher rate than the nondurable good in period 1. If consumption is not fully insured in the second period, this result becomes strict. (ii) The durable good should be taxed at the same rate as the nondurable good in period 2.*

The second part of Proposition 5 makes intuitive sense because the distinction between durable and nondurable goods vanishes in the last period of the two-period setting. Thus, Proposition 1 suggests that all goods should be taxed uniformly in period 2.

The first part of Proposition 5 establishes a stark result: in the first period, given any strictly concave utility function, durable goods should be taxed more than nondurable goods. To understand this result, note that the value of investment in the durable good depends on the stock of the durable good. Because of concavity, any current investment in the durable good lowers the marginal value of future investment. Consequently, future variations of investment become a less effective tool for incentive provision. This reasoning motivates a tax on the durable good in the first period to relax the incentive compatibility constraint. Note that this argument is similar to the rationale for savings distortions in models with nondurable goods (Diamond and Mirrlees, 1978).

In summary, the two-period model shows that concavities in durable consumption ("wealth effects") and interactions between durable and nondurable consumption justify the use of differential commodity taxes. These two insights apply in the same way also to models with longer time horizons (Section 5.1). Yet, the implications of frameworks with repeated investment will become slightly more nuanced for longer horizons. More precisely, the tax differential will in general depend on the sum of wealth effects and interaction effects, and concavity of the utility function alone will no longer guarantee a positive tax wedge on durable goods.[6]

5 Extensions

Next, I extend the results from Section 4 to settings with many periods and a large number of consumption goods. I also show that the results hold under very general specifications of uncertainty. Moreover, I study the consequences of stochastic depreciation rates.

5.1 Many periods and many goods

First, suppose that $T \geq 2$ and $M = N = 1$. Let the preferences and production technology be specified as in Eqs. (8) and (9) in Section 4. Proposition 4 extends to this more general framework without any difficulty. More precisely, the result is as follows.

Proposition 6. *Let $T \geq 2$ and $M = N = 1$. Let $t < T$ and suppose that investment in the durable good is only possible in period t. Then, any optimal allocation with interior consumption has the following implications: If $u_{cs} \leq 0$, the durable good should be taxed at a higher rate than the nondurable good (in period t). If $u_{cs} < 0$ and consumption is not fully insured in periods $t+1, \ldots, T$, the result becomes strict. If $u_{cs} \geq 0$, the result is reversed.*

Next, consider the case of repeated investment in the durable good. Proposition 5 extends in a straightforward way to the last two periods of the T-period framework. Proposition 5 can also be adapted to earlier periods. However, some further structure is helpful to establish that

[6]The stark result in part (i) of Proposition 5 relies on the finding that the marginal rate of substitution between durable and nondurable consumption is undistorted in the second period. In models with more than two periods, this finding only applies to the last period.

durable goods are taxed at a higher rate than nondurable goods. With repeated investment in the durable good, its service flow becomes a random process. Hence, the utility value of investing in the durable good depends on the dynamic realization of the two-dimensional process for durable and nondurable goods. If the two goods are monotonically related, this process can be reduced to a one-dimensional "sufficient statistic".

Definition 5. The nondurable good n and the durable good m are *perfectly rank correlated* after period t if

$$c_{\tau,n}\left(\theta^{\tau}\right) \geq c_{\tau,n}\left(\tilde{\theta}^{\tau}\right) \iff d_{\tau,m}\left(\theta^{\tau}\right) \geq d_{\tau,m}\left(\tilde{\theta}^{\tau}\right)$$

for all periods $\tau > t$ and all histories $\theta^{\tau}, \tilde{\theta}^{\tau} \in \Theta^{\tau}$.

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 5 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 tax differential in the model with repeated investment satisfies the following result.

Proposition 7. *Let $T \geq 2$ and $M = N = 1$. Let $t < T$ and suppose that investment in the durable good is possible in every period. 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 $u_{cs} \leq 0$, the durable good should be taxed at a higher rate than the nondurable good in period t. If consumption is not fully insured in periods $t+1, \ldots, T$, the result becomes strict.*

To rationalize Proposition 7, note that the planner uses consumption variations to incentivize the agents to supply labor. Suppose that 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

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

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 dampens the variation of future utility. This effect makes the incentive provision in the remaining periods more difficult and justifies a tax on the durable good to relax the incentive compatibility constraint. Note that this result is a combination of substitution effects (captured by the cross derivative u_{cs}) and wealth effects (captured by the second derivative u_{ss}) that were already visible in the two-period model; see Proposition 4 and Proposition 5.

Finally, consider the consequences of having several durable or nondurable consumption goods: $M, N \geq 1$. Then, the substitution and wealth effects from the two-goods model are complemented by non-separabilities with other consumption goods. Similar to Eq. (21) in the proof of Proposition 7, it can be shown that the durable good m should be 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 7, this inequality is satisfied if all goods are Edgeworth substitutes (provided that the consumption goods are sufficiently monotonic). [7]

5.2 Adverse selection and moral hazard

The mathematical analysis in this paper rests on 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,

[7] 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.

the results easily extend to setups with more general processes of effective labor.

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 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 $F(y_t|e_t, \theta_t)$. 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. The form of these consumption variations is independent of the question why labor effort differs across agents.

5.3 Stochastic depreciation

Uncertainty with respect to the depreciation of durable goods may be a further motivation for tax policy. When the depreciation shocks are observable, the planner can provide full insurance against these shocks without affecting any other decision margin. However, when the shocks are private information, insurance provision becomes less direct and commodity taxation may be justified.

I set up a tractable framework to explore the role of unobservable depreciation shocks. Consider a model with one durable and one nondurable good. For simplicity, suppose that the depreciation rate of the durable good can take two values: $\delta \in \{\delta_L, \delta_H\}$, with $\delta_L < \delta_H$ and probability weights $\pi_L, \pi_H > 0$. I assume that the utility function u is strictly concave and satisfies the single-crossing property:[8]

$$\frac{d\frac{u_c(c,\rho(i+\delta))}{u_s(c,\rho(i+\delta))}}{d\delta} > 0 \quad \text{for all } c, \rho, i, \delta > 0.$$

Here, c represents durable consumption, i denotes investment in the durable good, ρ maps durable stocks to service flows, and u_c and u_s represent the marginal utilities with respect to nondurable and durable consumption flows. Because of the single-crossing property, the insurance problem can be set up in a relaxed form, with only a downward incentive compatibility constraint:

$$\begin{aligned}
\max \quad & u(c_0, \rho i_0) + \beta \sum_{k=L,H} \pi_k u\left(c_k, \rho(i_k + \delta_k i_0)\right) \\
\text{s.t. } & Y - c_0 - i_0 - q \sum_{k=L,H} \pi_k \left[c_k + i_k\right] = 0 \\
& u\left(c_H, \rho(i_H + \delta_H i_0)\right) = u\left(c_L, \rho(i_L + \delta_H i_0)\right)
\end{aligned}$$

Here, Y denotes the (exogenous) level of aggregate resources and q represents the discounting of costs over time.

The necessary first-order conditions of this problem imply that the consumption choice is undistorted in the case of a positive depreciation shock:

$$\frac{\rho u_s\left(c_H, \rho(i_H + \delta_H i_0)\right)}{u_c\left(c_H, \rho(i_H + \delta_H i_0)\right)} = 1.$$

By the single-crossing property, the optimal plan for the negative depreciation shock satisfies $c_L < c_H$ and $i_L > i_H$. Since the indifference curves are strictly convex, this ordering implies

$$\frac{\rho u_s\left(c_L, \rho(i_L + \delta_H i_0)\right)}{u_c\left(c_L, \rho(i_L + \delta_H i_0)\right)} < \frac{\rho u_s\left(c_H, \rho(i_H + \delta_H i_0)\right)}{u_c\left(c_H, \rho(i_H + \delta_H i_0)\right)}. \tag{12}$$

Denote the Lagrange multipliers for the resource constraint and incentive constraint by λ and

[8] The single-crossing property holds if and only if $u_{cd} > u_c u_{dd}/u_d$. Intuitively, this condition is satisfied if durable and nondurable consumption are not too closely substitutable.

μ, respectively. The first-order conditions for c_L and i_L yield

$$\frac{\rho u_s\left(c_L, \rho(i_L+\delta_L i_0)\right)}{u_c\left(c_L, \rho(i_L+\delta_L i_0)\right)}=\frac{\lambda q \pi_L+\mu \rho u_s\left(c_L, \rho(i_L+\delta_H i_0)\right)}{\lambda q \pi_L+\mu u_c\left(c_L, \rho(i_L+\delta_H i_0)\right)}.$$

By combining the first-order condition with Eq. (12), it becomes clear that the solution to the insurance problem satisfies

$$\frac{\rho u_s\left(c_L, \rho(i_L+\delta_L i_0)\right)}{u_\rho\left(c_L, \rho(i_L+\delta_L i_0)\right)}<1.$$

Hence, for the consumption plan associated with negative depreciation shock, investment in the durable good is subsidized relative to nondurable consumption. Note that an agent with a positive depreciation shock values investment less than an agent with a negative shock. Thus, a subsidy to investment yields some slackness in the incentive compatibility constraint and facilitates the provision of insurance.[9]

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.

The findings in this paper have implications for the taxation of housing. They also shed some light on the interpretation of optimal taxes on pre-committed goods.

6.1 Application to housing taxes

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.[10]

[9]The implied tax wedge is once more affected by the interaction of durable and nondurable consumption. Note that the slope of the indifference curve depends less strongly on the depreciation shock δ when the two goods are close substitutes. In the limit case of perfect substitutes, $u(c, \rho(i+\delta)) = \hat{u}\left(c+\rho(i+\delta)\right)$, the marginal rate of substitution between durable and nondurable consumption in fact becomes independent of δ. In that case, the solution of the insurance problem is a pooling allocation where the marginal rate of substitution equals the marginal rate of transformation. This insight suggests that substitutability of durable and nondurable consumption leads to a lower subsidy on durable goods. Note that this finding is similar to the results in the main part of the paper: substitutability again calls for relatively higher taxes (lower subsidies) on durable goods.

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

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 economic literature on housing, the preferences are commonly specified by a utility function with a constant elasticity of substitution,

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

where c denotes nondurable consumption, h 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 intertemporal substitutability of the consumption-housing composite. This specification implies that housing and nondurable consumption are strict substitutes in the Edgeworth sense $\left(u''_{ch} < 0\right)$ if and only if the parameters satisfy $\epsilon > \sigma$, i.e., if and only if the intratemporal elasticity of substitution exceeds the intertemporal elasticity. Therefore, if $\epsilon > \sigma$, Propositions 4 to 7 suggest that housing should face higher tax rates than nondurable consumption.

Many papers have estimated the above CES specification. 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 that $\epsilon > \sigma$, which means that housing and nondurable consumption are Edgeworth substitutes. Thus, according to the theory in this paper, housing should face higher tax rates than nondurable consumption.[11]

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 decisions, whereas the wealth and substitution effects highlighted in this paper are primitives that follow directly from individual preferences.

6.2 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 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.

Part (ii) of Proposition 5 shows 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 part (i) of 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

[11] As pointed out in Section 5.3, housing subsidies may be helpful if the depreciation rates (and thus the house values) are stochastic and unobservable. For housing, the assumption of nonobservability seems restrictive because depreciation shocks may be correlated across space. Moreover, housing transactions are public information (through laws on property registration).

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 shown by the Inverse Euler Equation in Lemma 2.

Appendix

A 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:

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

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

$$G\left(\mathbb{E}\left[c_t'\right], \mathbb{E}\left[i_t^*\right], K_{t+1}^*, K_t^*, \mathbb{E}\left[y_t^*\right]\right) < G\left(\mathbb{E}\left[c_t^*\right], \mathbb{E}\left[i_t^*\right], K_{t+1}^*, K_t^*, \mathbb{E}\left[y_t^*\right]\right) \leq 0. \tag{13}$$

Define $\mathbf{c}' = \left(c_t', c_{-t}^*\right)$ and consider the allocation $\left(\mathbf{c}', \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}^*\right)$. 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 $(\mathbf{c}^*, \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}^*)$. Moreover, the allocation delivers the same level of social welfare as the optimal allocation $(\mathbf{c}^*, \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}^*)$. Therefore, $(\mathbf{c}', \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}^*)$ is also an optimal allocation. However, by Eq. (13), $(\mathbf{c}', \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}^*)$ 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 $(\mathbf{c}', \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}')$ solves the planner problem for initial capital K_1'. This implies $V(K_1') = V(K_1^*)$, 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 Eq. (4). □

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

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

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. By proceeding 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, $U_{N+m}(c, s, e) = \alpha_m$ and $U_n(c, s, e) = u_n(c)$ for all (c, s, e).

Therefore,

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

Hence, by dividing the first-order conditions of the cost minimization problem by each other, Eq. (3) follows. □

Proof of Lemma 2. Let $(\mathbf{c}, \mathbf{d}, \mathbf{y}, \mathbf{K})$ be an incentive-feasible allocation with interior consumption. Let $\theta_1' \in \Theta$. Consider the following perturbation of nondurable consumption:

$$\begin{aligned} u\left(c_1^{\varepsilon}\left(\theta_1'\right), \rho d_1\left(\theta_1'\right)\right) &= u\left(c_1\left(\theta_1'\right), \rho d_1\left(\theta_1'\right)\right) + \varepsilon \\ u\left(c_2^{\varepsilon}\left(\theta_1', \theta_2\right), \rho d_2\left(\theta_1', \theta_2\right)\right) &= u\left(c_2\left(\theta_1', \theta_2\right), \rho d_2\left(\theta_1', \theta_2\right)\right) - \frac{\varepsilon}{\beta}. \end{aligned}$$

For histories (θ_1, θ_2) with $\theta_1 \neq \theta_1'$, set $c_t^{\varepsilon} = c_t$ for $t = 1, 2$. Adjust the capital stock of the second period in response to the changed consumption levels. Formally, define $K_2^{\varepsilon} := K_2 - \zeta^{\varepsilon}$ with the help of the equation

$$\pi_1\left(\theta_1'\right) \sum_{\theta_2} \pi_2\left(\theta_2 | \theta_1'\right) \left[c_2\left(\theta_1', \theta_2\right) - c_2^{\varepsilon}\left(\theta_1', \theta_2\right)\right] = F\left(K_2, \mathbb{E}\left[y_2\right]\right) - F\left(K_2 - \zeta^{\varepsilon}, \mathbb{E}\left[y_2\right]\right) + \left(1 - \delta^K\right)\zeta^{\varepsilon}.$$

By construction, for all histories (θ_1, θ_2), 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 initial 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:

$$\min_{\varepsilon}\left\{\pi_1\left(\theta_1'\right) c_1^{\varepsilon}\left(\theta_1'\right) - \zeta^{\varepsilon}\right\}$$

This implies the first-order condition

$$0 = \pi_1\left(\theta_1'\right)\frac{dc_1^{\varepsilon}\left(\theta_1'\right)}{d\varepsilon}|_{\varepsilon=0} - \frac{d\zeta^{\varepsilon}}{d\varepsilon}|_{\varepsilon=0}$$

which is equivalent to

$$0 = \frac{\pi_1\left(\theta_1'\right)}{u_c\left(c_1\left(\theta_1'\right),\rho d_1\left(\theta_1'\right)\right)} - \frac{1}{\beta R}\sum_{\theta_2}\pi_2\left(\theta_2|\theta_1'\right)\frac{\pi_1\left(\theta_1'\right)}{u_c\left(c_2\left(\theta_1',\theta_2\right),\rho d_2\left(\theta_1',\theta_2\right)\right)}$$

with $R := 1-\delta^K + F_K\left(K_2,\mathbb{E}\left[y_2\right]\right)$. Dividing by $\pi_1\left(\theta_1'\right)$, Eq. (10) follows. $\square$

Proof of Proposition 3. Let $(\mathbf{c},\mathbf{d},\mathbf{y},\mathbf{K})$ be an incentive-feasible allocation with interior consumption. Let $\theta_1' \in \Theta$. Consider the following perturbation of investment in the durable good:

$$\begin{aligned}
i_1^{\varepsilon}\left(\theta_1'\right) &= i_1\left(\theta_1'\right) - \varepsilon \\
i_2^{\varepsilon}\left(\theta_1',\theta_2\right) &= i_2\left(\theta_1',\theta_2\right) \\
d_1^{\varepsilon}\left(\theta_1'\right) &= d_1\left(\theta_1'\right) - \varepsilon \\
d_2^{\varepsilon}\left(\theta_1',\theta_2\right) &= d_2\left(\theta_1',\theta_2\right) - \delta\varepsilon.
\end{aligned}$$

Adjust the level of nondurable consumption so that, in every period and for every realization, the agent obtains the same consumption utility as before:

$$\begin{aligned}
u\left(c_1^{\varepsilon}\left(\theta_1'\right),\rho d_1^{\varepsilon}\left(\theta_1'\right)\right) &= u\left(c_1\left(\theta_1'\right),\rho d_1\left(\theta_1'\right)\right) \\
u\left(c_2^{\varepsilon}\left(\theta_1',\theta_2\right),\rho d_2^{\varepsilon}\left(\theta_1',\theta_2\right)\right) &= u\left(c_2\left(\theta_1',\theta_2\right),\rho d_2\left(\theta_1',\theta_2\right)\right).
\end{aligned}$$

For histories (θ_1,θ_2) with $\theta_1 \neq \theta_1'$, set $d_t^{\varepsilon} = d_t$ and $c_t^{\varepsilon} = c_t$ for $t = 1,2$. Moreover, adjust the capital stock of the second period in response to the changed nondurable consumption levels. Formally, define $K_2^{\varepsilon} := K_2 + \zeta^{\varepsilon}$ with the help of the equation

$$\pi_1\left(\theta_1'\right)\sum_{\theta_2}\pi_2\left(\theta_2|\theta_1'\right)\left[c_2^{\varepsilon}\left(\theta_1',\theta_2\right) - c_2\left(\theta_1',\theta_2\right)\right] = F\left(K_2+\zeta^{\varepsilon},\mathbb{E}\left[y_2\right]\right) - F\left(K_2,\mathbb{E}\left[y_2\right]\right) + \left(1-\delta^K\right)\zeta^{\varepsilon}.$$

Similar to the proof of Lemma 2, the allocation $(\mathbf{c},\mathbf{d},\mathbf{y},\mathbf{K})$ can only be optimal if $\varepsilon = 0$ solves the following cost minimization problem:

$$\min_{\varepsilon}\left\{\pi_1\left(\theta_1'\right)\left[c_1^{\varepsilon}\left(\theta_1'\right)+i_1^{\varepsilon}\left(\theta_1'\right)\right]+\zeta^{\varepsilon}\right\}$$

This implies the first-order condition

$$0=\pi_1\left(\theta_1'\right)\frac{dc_1^{\varepsilon}\left(\theta_1'\right)}{d\varepsilon}|_{\varepsilon=0}-\pi_1\left(\theta_1'\right)+\frac{d\zeta^{\varepsilon}}{d\varepsilon}|_{\varepsilon=0}.$$

By the construction of the perturbed allocation, we have

$$\begin{aligned}\frac{dc_1^{\varepsilon}\left(\theta_1'\right)}{d\varepsilon}|_{\varepsilon=0} &= \rho\frac{u_s\left(c_1\left(\theta_1'\right),\rho d_1\left(\theta_1'\right)\right)}{u_c\left(c_1\left(\theta_1'\right),\rho d_1\left(\theta_1'\right)\right)}\\ \frac{dc_2^{\varepsilon}\left(\theta_1',\theta_2\right)}{d\varepsilon}|_{\varepsilon=0} &= \rho\delta\frac{u_s\left(c_2\left(\theta_1',\theta_2\right),\rho d_2\left(\theta_1',\theta_2\right)\right)}{u_c\left(c_2\left(\theta_1',\theta_2\right),\rho d_2\left(\theta_1',\theta_2\right)\right)}.\end{aligned}$$

Moreover, the derivative of ζ^{ε} is implicitly given by

$$\left(1-\delta^K+F_K\left(K_2,\mathbb{E}\left[y_2\right]\right)\right)\frac{d\zeta^{\varepsilon}}{d\varepsilon}|_{\varepsilon=0}=\pi_1\left(\theta_1'\right)\sum_{\theta_2}\pi_2\left(\theta_2|\theta_1'\right)\frac{dc_2^{\varepsilon}\left(\theta_1',\theta_2\right)}{d\varepsilon}|_{\varepsilon=0}.$$

After combining these equations and dividing by $\pi_1\left(\theta_1'\right)$, we can write the first-order condition of the cost minimization problem as

$$0=\rho\frac{u_s\left(c_1\left(\theta_1'\right),\rho d_1\left(\theta_1'\right)\right)}{u_c\left(c_1\left(\theta_1'\right),\rho d_1\left(\theta_1'\right)\right)}-1+\frac{\rho\delta}{R}\sum_{\theta_2}\pi_2\left(\theta_2|\theta_1'\right)\frac{u_s\left(c_2\left(\theta_1',\theta_2\right),\rho d_2\left(\theta_1',\theta_2\right)\right)}{u_c\left(c_2\left(\theta_1',\theta_2\right),\rho d_2\left(\theta_1',\theta_2\right)\right)},$$

where $R:=1-\delta^K+F_K\left(K_2,\mathbb{E}\left[y_2\right]\right)$. This establishes Eq. (11). □

Proof of Proposition 4. Let $\left(\mathbf{c}^*,\mathbf{d}^*,\mathbf{y}^*,\mathbf{K}^*\right)$ be an optimal allocation with interior consumption. Given that investment in the durable good is only possible in the first period, the stock of the durable good is independent of the realization of uncertainty in the second period, i.e., we have $d_2^*=d_2^*\left(\theta_1\right)$. Note that Lemma 2 and Proposition 3 remain valid in this modified framework because they do not rely on perturbations of durable investment after the first period.

The Substitution Euler Equation (Eq. (11)) implies

$$\rho\frac{u_s\left(c_1^*,\rho d_1^*\right)}{u_c\left(c_1^*,\rho d_1^*\right)}=1-\frac{\rho\delta}{R^*}\mathbb{E}_1\left[\frac{u_s\left(c_2^*,\rho d_2^*\right)}{u_c\left(c_2^*,\rho d_2^*\right)}\right].$$

Equivalently,

$$\begin{aligned}&\rho\frac{u_s\left(c_1^*,\rho d_1^*\right)}{u_c\left(c_1^*,\rho d_1^*\right)}+\rho\beta\delta\mathbb{E}_1\left[\frac{u_s\left(c_2^*,\rho d_2^*\right)}{u_c\left(c_1^*,\rho d_1^*\right)}\right]\\ =\ &1-\frac{\rho\delta}{R^*}\mathbb{E}_1\left[u_s\left(c_2^*,\rho d_2^*\right)\left(\frac{1}{u_c\left(c_2^*,\rho d_2^*\right)}-\frac{\beta R^*}{u_c\left(c_1^*,\rho d_1^*\right)}\right)\right].\end{aligned}$$

The Inverse Euler Equation (Eq. (10)) implies

$$\frac{\beta R^*}{u_c\left(c_1^*,\rho d_1^*\right)}=\mathbb{E}_1\left[\frac{1}{u_c\left(c_2^*,\rho d_2^*\right)}\right].$$

Hence, the Substitution Euler Equation can be written in terms of a covariance,

$$\begin{aligned}&\rho\frac{u_s\left(c_1^*,\rho d_1^*\right)}{u_c\left(c_1^*,\rho d_1^*\right)}+\rho\beta\delta\mathbb{E}_1\left[\frac{u_s\left(c_2^*,\rho d_2^*\right)}{u_c\left(c_1^*,\rho d_1^*\right)}\right]\\ =\ &1-\frac{\rho\delta}{R^*}\operatorname{cov}_1\left(u_s\left(c_2^*,\rho d_2^*\right),\frac{1}{u_c\left(c_2^*,\rho d_2^*\right)}-\frac{\beta R^*}{u_c\left(c_1^*,\rho d_1^*\right)}\right),\end{aligned}$$

where $\operatorname{cov}_1(\cdot,\cdot)$ represents the covariance with respect to the realization of uncertainty in period 2, conditional on the realization in period 1. Since c_1^* and d_1^* are certain after period 1, the equation can be simplified to

$$\rho\frac{u_s\left(c_1^*,\rho d_1^*\right)}{u_c\left(c_1^*,\rho d_1^*\right)}+\rho\beta\delta\mathbb{E}_1\left[\frac{u_s\left(c_2^*,\rho d_2^*\right)}{u_c\left(c_1^*,\rho d_1^*\right)}\right]=1-\frac{\rho\delta}{R^*}\operatorname{cov}_1\left(u_s\left(c_2^*,\rho d_2^*\right),\frac{1}{u_c\left(c_2^*,\rho d_2^*\right)}\right).$$

The marginal rate of substitution of the durable good for the nondurable good exceeds unity (the marginal rate of transformation) if and only if

$$\operatorname{cov}_1\left(-u_s\left(c_2^*,\rho d_2^*\right),\frac{1}{u_c\left(c_2^*,\rho d_2^*\right)}\right)\geq 0. \tag{14}$$

Recall that $d_2^*=d_2^*(\theta_1)$ does not depend on θ_2 in this framework. Hence, the marginal utilities $u_s\left(c_2^*,\rho d_2^*\right)$ and $u_c\left(c_2^*,\rho d_2^*\right)$ are uncertain only through their dependence on $c_2^*(\theta_1,\theta_2)$. By concavity, $u_c\left(c_2^*,\rho d_2^*\right)$ is strictly decreasing in c_2^*. If $u_{cs}\leq 0$, $u_s\left(c_2^*,\rho d_2^*\right)$ is weakly decreasing in c_2^*. In that case, Eq. (14) measures the covariance of two increasing functions of the same random variable. Therefore, the covariance is nonnegative. Moreover, if $u_{cs}<0$ and $c_2^*(\theta_1,\theta_2)$ is not constant across θ_2, the covariance is strictly positive. Finally, if $u_{cs}\geq 0$ or $u_{cs}>0$, the respective inequalities are reversed. $\square$

Proof of Proposition 5. Let $(\mathbf{c}^*, \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}^*)$ be an optimal allocation with interior consumption. Let $(\theta_1', \theta_2') \in \Theta^2$ and consider the following perturbation of second-period consumption:

$$\begin{aligned} i_2^{\varepsilon}\left(\theta_1', \theta_2'\right) &= i_2^*\left(\theta_1', \theta_2'\right) + \varepsilon \\ d_2^{\varepsilon}\left(\theta_1', \theta_2'\right) &= d_2^*\left(\theta_1', \theta_2'\right) + \varepsilon \\ u\left(c_2^{\varepsilon}\left(\theta_1', \theta_2'\right), \rho d_2^{\varepsilon}\left(\theta_1', \theta_2'\right)\right) &= u\left(c_2^*\left(\theta_1', \theta_2'\right), \rho d_2^*\left(\theta_1', \theta_2'\right)\right). \end{aligned}$$

Set $c_1^{\varepsilon} = c_1^*$, $d_1^{\varepsilon} = d_1^*$ and $c_2^{\varepsilon}\left(\theta^2\right) = c_2^*\left(\theta^2\right)$, $c_2^{\varepsilon}\left(\theta^2\right) = c_2^*\left(\theta^2\right)$ for all histories $\theta^2 \neq \left(\theta_1', \theta_2'\right)$. 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.

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

$$\min_{\varepsilon} \left\{ c_2^{\varepsilon}\left(\theta_1', \theta_2'\right) + i_2^{\varepsilon}\left(\theta_1', \theta_2'\right) \right\}$$

This implies the first-order condition

$$\rho \frac{u_s\left(c_2^*\left(\theta_1', \theta_2'\right), \rho d_2^*\left(\theta_1', \theta_2'\right)\right)}{u_c\left(c_2^*\left(\theta_1', \theta_2'\right), \rho d_2^*\left(\theta_1', \theta_2'\right)\right)} = 1. \tag{15}$$

Hence, in the second period, the marginal rate of substitution of the durable good for the nondurable good equals unity (the marginal rate of transformation). This establishes that the durable good should be taxed at the same rate as the nondurable good in period 2.

Now consider the tax distortion for period 1. By following the same steps as in the proof of Proposition 4, the marginal rate of substitution of the durable good for the nondurable good exceeds the marginal rate of transformation if and only if

$$\operatorname{cov}_1\left(u_s\left(c_2^*, \rho d_2^*\right), \frac{1}{u_c\left(c_2^*, \rho d_2^*\right)}\right) \leq 0.$$

By Eq. (15), this condition is equivalent to

$$\operatorname{cov}_1\left(u_s\left(c_2^*,\rho d_2^*\right),\frac{1}{u_s\left(c_2^*,\rho d_2^*\right)}\right)\leq 0. \tag{16}$$

Since the utility function is strictly concave, the covariance in Eq. (16) is nonpositive. Moreover, if c_2^* or d_2^* depends on the realization of uncertainty in the second period, the covariance is strictly negative. □

Proof of Proposition 6. Let $(\mathbf{c}^*,\mathbf{d}^*,\mathbf{y}^*,\mathbf{K}^*)$ be an optimal allocation with interior consumption. By proceeding as in Lemma 2, the following Inverse Euler Equation holds for all $t<T$:

$$\frac{\beta R_{t+1}^*}{u_c\left(c_t^*,\rho d_t^*\right)}=\mathbb{E}_t\left[\frac{1}{u_c\left(c_{t+1}^*,\rho d_{t+1}^*\right)}\right], \tag{17}$$

where the interest rate is defined as $R_{t+1}^*:=1-\delta^K+F_K\left(K_{t+1}^*,\mathbb{E}\left[y_{t+1}^*\right]\right)$. For $k>t$, define the intertemporal discount factor $q_t^k:=\left(\prod_{i=t+1}^k R_i^*\right)^{-1}$ and set $q_t^t=1$. Similar to Proposition 3, the generalized Substitution Euler Equation for all $t\leq T$ is given by

$$1=\rho\sum_{k=t}^{T}q_t^k\delta^{k-t}\mathbb{E}_t\left[\frac{u_s\left(c_k^*,\rho d_k^*\right)}{u_c\left(c_k^*,\rho d_k^*\right)}\right]. \tag{18}$$

The generalized Substitution Euler Equation implies

$$\rho\frac{u_s\left(c_t^*,\rho d_t^*\right)}{u_c\left(c_t^*,\rho d_t^*\right)}=1-\rho\sum_{k=t+1}^{T}q_t^k\delta^{k-t}\mathbb{E}_t\left[\frac{u_s\left(c_k^*,\rho d_k^*\right)}{u_c\left(c_k^*,\rho d_k^*\right)}\right].$$

Equivalently,

$$\begin{aligned}&\rho\frac{u_s\left(c_t^*,\rho d_t^*\right)+\sum_{k=t+1}^T(\beta\delta)^{k-t}\,\mathbb{E}_t\left[u_s\left(c_k^*,\rho d_k^*\right)\right]}{u_c\left(c_t^*,\rho d_t^*\right)}\\&=1-\rho\sum_{k=t+1}^{T}q_t^k\delta^{k-t}\mathbb{E}_t\left[u_s\left(c_k^*,\rho d_k^*\right)\left(\frac{1}{u_c\left(c_k^*,\rho d_k^*\right)}-\frac{\beta^{k-t}}{q_t^k u_c\left(c_t^*,\rho d_t^*\right)}\right)\right].\end{aligned} \tag{19}$$

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

$$\mathbb{E}_t\left[\frac{1}{u_c\left(c_k^*,\rho d_k^*\right)}-\frac{\beta^{k-t}}{q_t^k u_c\left(c_t^*,\rho d_t^*\right)}\right]=0.$$

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

$$\begin{aligned}
& \mathbb{E}_t \left[u_s\left(c_k^*, \rho d_k^*\right) \left(\frac{1}{u_c\left(c_k^*, \rho d_k^*\right)} - \frac{\beta^{k-t}}{q_t^k u_c\left(c_t^*, \rho d_t^*\right)} \right) \right] \\
= \; & \operatorname{cov}_t \left(u_s\left(c_k^*, \rho d_k^*\right), \frac{1}{u_c\left(c_k^*, \rho d_k^*\right)} - \frac{\beta^{k-t}}{q_t^k u_c\left(c_t^*, \rho d_t^*\right)} \right) \\
= \; & \operatorname{cov}_t \left(u_s\left(c_k^*, \rho d_k^*\right), \frac{1}{u_c\left(c_k^*, \rho d_k^*\right)} \right).
\end{aligned}$$

Consequently, Eq. (19) shows that the marginal rate of substitution of the durable good for the nondurable good in period t exceeds unity (the marginal rate of transformation) if and only if

$$\sum_{k=t+1}^{T} q_t^k \delta^{k-t} \operatorname{cov}_t \left(-u_s\left(c_k^*, \rho d_k^*\right), \frac{1}{u_c\left(c_k^*, \rho d_k^*\right)} \right) \geq 0. \tag{20}$$

Now the result follows from the same arguments as in the proof of Proposition 4. □

Proof of Proposition 7. Let $(\mathbf{c}^*, \mathbf{d}^*, \mathbf{y}^*, \mathbf{K}^*)$ be an optimal allocation with interior consumption. Based on Eq. (20) from the proof of Proposition 6, the marginal rate of substitution of the durable good for the nondurable good in period t exceeds the marginal rate of transformation if and only if

$$\sum_{\tau=t+1}^{T} q_t^\tau \delta^{\tau-t} \operatorname{cov}_t \left(-u_s\left(c_\tau^*, \rho d_\tau^*\right), \frac{1}{u_c\left(c_\tau^*, \rho d_\tau^*\right)} \right) \geq 0. \tag{21}$$

Given that the consumption goods are perfectly rank correlated, 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 $\lambda = \lambda_\tau\left(\theta^\tau\right)$ and $\hat{\lambda} = \lambda_\tau\left(\hat{\theta}^\tau\right)$with $\lambda \geq \hat{\lambda}$. Then, by the mean value theorem, there exists a number $\xi \in \left[D_\tau\left(\hat{\lambda}\right), D_\tau\left(\lambda\right)\right]$ and a number $\kappa \in$

$\left[C_\tau\left(\hat{\lambda}\right), C_\tau(\lambda)\right]$ such that

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

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 $\xi' \in \left[D_\tau\left(\hat{\lambda}\right), D_\tau(\lambda)\right]$ and a number $\kappa' \in \left[C_\tau\left(\hat{\lambda}\right), C_\tau(\lambda)\right]$ such that

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

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. (21) are monotonic in the same direction. Hence, their covariance is nonnegative and Eq. (21) 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. □

References

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.

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.

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.

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.

KOEHNE, S. AND M. KUHN (2014): "Optimal taxation in a habit formation economy," *Journal of Public Economics*, http://dx.doi.org/10.1016/j.jpubeco.2014.12.007.

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.

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.