Optimal taxation with dual atmospheric-consumption externalities

Enda Patrick Hargaden
Matthew C. Harris

February 2021

Abstract

This paper generalizes optimal taxation results to study goods that yield both atmospheric and consumption externalities, where the number of users affects the benefits of consuming the good. Commodity taxes therefore affect utility through both price effects and by affecting the number of users. We evaluate the importance of these consumption externalities to optimal tax policy in a dynamic setting using Monte Carlo simulations. Time-varying tax schedules can improve total surplus and increase total tax revenues, compared to a time-invariant sequence. This stands in contrast to a strand of the public finance literature de-emphasizing commodity taxation as a means for efficiently raising revenue.
1 Introduction

Despite a near consensus that commodity taxation is generally less efficient or redundant to an income tax (Atkinson and Stiglitz, 1976), commodity taxation is still a crucial source of revenue for many levels of government. In the United States, state and local governments raise over $580 billion in sales taxes. This constitutes a larger share of revenue than that from property, income, or corporate taxes (Tax Policy Center, 2020a,b). The IMF reports that 160 countries levy a Value Added Tax (IMF, 2018). Because of the distortionary effects of taxes and the prevalence of commodity taxation, setting commodity tax rates close to optimal can substantively improve welfare. As taxation affects individuals’ decisions, concerns about the distortionary effects of commodity taxation may be especially pronounced when taxing goods that yield externalities.

This paper studies goods that yield both atmospheric and consumption-based externalities. We follow the literature in delineating consumption externalities (e.g., network effects) as those for which the individual’s utility from consuming a good is affected by the total number of users of that good (see e.g., Eckerstorfer and Wendner, 2013). At least four of the five largest U.S. companies sell products that yield substantial consumption externalities.\(^1\) We define atmospheric externalities as those for which total consumption of the good affects the overall utility of every agent (Meade, 1952; Micheletto, 2008). Most atmospheric externalities, like pollution, do not directly affect demand. If a widget factory pollutes a lake, the demand for widgets is likely unchanged.

Goods that yield both consumption and atmospheric externalities are increasingly common. For example, car exhausts have negative health effects (Knittel, Miller and Sanders, 2016; Currie and Walker, 2011), and higher levels of pollution may affect the willingness to be a pedestrian (Neidell, 2009). Cell phones are network goods by nature, but also have atmospheric effects: cell phones’ ubiquity may affect you even if you do not own a phone.\(^2\) The total number of cell phone users is thus a dual atmospheric-consumption externality, and the question of this paper is how this duality affects the optimal tax rate for such goods. Even goods that have a nominal price of zero such as Twitter may have negative atmospheric effects (e.g. polarized political discourse, see Allcott, Gentzkow and Yu, 2019b; Allcott, Braghieri, Eichmeyer and Gentzkow, 2020), but positive consumption externalities if utility is increasing in the number of users on the platform. However, most prior work on optimal taxation of externality-generating goods focuses solely on atmospheric or positional externalities.

In this paper, we generalize Sandmo (1975) to allow for the fact that externality producing aspects of a good may yield both atmospheric and consumption externalities, and derive the optimal tax rate. Sandmo (1975) found the optimal tax on an externality-generating good was an additive and separable weighted average of Pigouvian taxation and a form of the Ramsey rule.

\(^1\)At the time of writing, the five largest U.S. corporations by market capitalization are Microsoft, Apple, Amazon, Alphabet (i.e. Google), and Facebook.

\(^2\)Through increased use of video recording, or in an emergency where your phone is stolen, for example.
for commodity taxation. This feature is important as it permits a taxation strategy that directly “targets” the externality-generating good under relatively general conditions (Kopczuk, 2003; Micheletto, 2008). We show the optimal tax rate of a dual atmospheric-consumption externality good has two components in common with Sandmo, and one new component. The common components are the Ramsey- and Pigou-like factors; and the new component captures the effects of the consumption externality on demand.

Like Sandmo (1975), the results from our model also have an additive and separable interpretation. For example, for goods that yield negative atmospheric externalities, the optimal tax rate increases if the good also yields a negative consumption externality. When atmospheric and consumption externalities are of opposing directions, the sign of the optimal tax rate may be counterintuitive. Positive consumption externalities reduce the optimal tax rate on a good, mitigating Pigouvian taxation of negative atmospheric externalities. When those consumption externalities are strong enough, we show that optimal tax policy may be a subsidy, even when the good generates negative atmospheric externalities such as pollution.

For commodities with consumption externalities, the dynamics are of particular interest. As contemporaneous adoption rates affect demand in subsequent periods, our introduction of consumption externalities raises the possibility of dynamic effects. A government with an intertemporal revenue requirement faces a tradeoff, where greater revenues today implies lower adoption rates and therefore a smaller tax base for that commodity in future periods. Without restrictions on how consumers form beliefs, multiple equilibria can exist. Thus, deriving a closed-form time path for optimal commodity taxes with consumption externalities is infeasible at the level of generality of the baseline model. We thus assess the dynamic implications of consumption externalities for optimal taxation using Monte Carlo simulations.

We find that when goods yield consumption externalities, time-varying commodity tax schedules can improve total surplus and increase total tax revenues, compared to a time-invariant sequence. This fact is somewhat surprising. The excess burden of taxation generally grows in the square of the rate. This fact encourages low rates over wide bases and, in the intertemporal context, penalizes deviations from a constant time-path. Our result that the optimal policy in respect to dual atmospheric-consumption externalities is time-varying is therefore somewhat unusual. Further, we show time-varying tax sequences can be welfare-improving whether the government does or does not face an exogenous revenue constraint. That is, dynamic setting of rates can improve welfare over a zero tax default. This stands in contrast to a strand of the public finance literature de-emphasizing commodity taxation as a means for efficiently raising revenue. These results do not contradict this literature, but highlight the potential gains from

---

3When we refer to ‘total tax revenues,’ we mean cumulative tax revenues from all considered periods.
4One of the most important results in public finance comes from Atkinson and Stiglitz (1976), which shows that under certain conditions (e.g., optimal non-linear income taxation) commodity taxation is redundant. In the presence of consumption externalities, income tax policy is unlikely to induce socially optimal consumption choices. We find that if the government has a revenue threshold it must meet over several periods, that a commodity tax policy of initial subsidization (relative to a time-invariant baseline) can be a revenue-neutral, welfare-improving
commodity taxation in this specific context.

This paper contributes to the literature in several ways. We add to the literature on optimal taxation by deriving the optimal tax rate for a commodity with dual atmospheric-consumption externalities. The externalities we consider differ from positional externalities previously considered in the public finance literature (e.g., Aronsson and Johansson-Stenman, 2010), and are closer in nature to network externalities. Our contribution builds directly on the work of Micheletto (2008), which generalizes Kopczuk (2003) on the principle of targeting. Micheletto finds the Sandmo additive result is not true for consumption externalities in general, but that it does hold under reasonable conditions. Similarly, we show that dual atmospheric-consumption externalities can be interpreted as a generalization of Sandmo’s results. Our second contribution relates to the optimal dynamic path of commodity taxes, and has substantial policy implications. We examine whether optimal tax policy in relation to these goods depends on how established these goods are. Simulation results indicate that intertemporal variation in tax rates can lead to considerable welfare gains, particularly if the government can borrow. The numerical simulations are not intended to suggest fully general first-best tax rates, but do illustrate the results under reasonable calibrations.

More generally, we view our work as a complement to the recent research extending canonical optimal tax results. Perhaps the most fruitful source of this literature has been incorporating behavioral agents into optimal tax frameworks, prominently Farhi and Gabaix (2020), Alcott, Lockwood and Taubinsky (2019a), Rees-Jones (2018), and Lockwood (2020). As a matter of theoretical interest, this paper is a contribution to the literature on optimal taxation. We argue that this model is applicable to a large class of goods. With technological progress, the relevance and prominence of network goods in particular will almost certainly increase. Many will become part of the tax base. For this reason, this paper is relevant to policy as well as to public economic theory.

The rest of the paper proceeds as follows: Section 2.1 lays out the baseline closed-form model, and derives the optimal tax rate when dual externalities are present. Sections 2.2 and 2.3 clarify the restrictions imposed, the methodology of, and results from the dynamic Monte Carlo simulations. Section 3 concludes.

2 Analysis

2.1 Baseline Model

We start by solving for the optimal tax rates in a static framework. This allows us to derive the principal results of the model at a level of generality not feasible in a dynamic setting with unrestricted expectations. We model a utilitarian planner maximizing the sum of utilities for strategy. As a practical matter, it seems unlikely that many governments impose an optimal non-linear income tax. Nine U.S. states, for example, have no income tax at all.
n identical consumers subject to a government revenue requirement $T$. Each consumer chooses labor effort $x_0$, the wages of which act as a numeraire, and the complement of labor is leisure. This is a second-best world where lump sum taxes are infeasible and leisure is untaxable. In addition to labor, there are $m$ taxable commodities in the economy. Consumers purchase goods based on tax-inclusive prices $P_i, i = 1, \ldots, m$. Commodity $m$ generates an externality $\alpha$, which for simplicity we will think of as total consumption of $x_m$. The consumer’s problem is thus to maximize

$$L = u(1 - x_0, x_1, \ldots, x_m, \alpha) + \lambda \left( x_0 - \sum_{i=1}^{m} P_i x_i \right)$$

We denote $u_i$ as the derivative of the utility function with respect to $x_i$, and therefore denote the derivative of utility with respect to $\alpha$ as $u_{m+1}$. For a negative externality, $u_{m+1} < 0$. Consumers do not consider their own effect on the externality, and we assume that the usual conditions for an interior maximum hold.

The solution of the consumer’s problem requires the following two first-order conditions:

$$FOC_1 : u_0 = \lambda$$

$$FOC_2 : u_i = \lambda P_i, \ i = 1, \ldots, m$$

Note also from the budget constraint that $-\frac{\partial x_0}{\partial P_k} + \sum_{i=1}^{m} P_i \frac{\partial x_i}{\partial P_k} + x_k = 0 \Rightarrow x_k = \frac{\partial x_0}{\partial P_k} - \sum_{i=1}^{m} P_i \frac{\partial x_i}{\partial P_k}$. We will use this fact later.

We follow convention by permitting governments to adjust the price vector $P$ to maximize society’s indirect utility $V(P)$:

$$V(P) = u [1 - x_0(P), x_1(P), \ldots, x_m(P, \alpha(P)), \alpha(P)]$$

Note that this formulation permits both that demand for $x_m$ be a function of $\alpha$, and that $\alpha$ directly affects utility. The case of cars and air pollution may be an intuitive example. The welfare effect of adjusting the price of good $k$ is:

$$\frac{\partial V}{\partial P_k} = -u_0 \frac{\partial x_0}{\partial P_k} + \sum_{i=1}^{m-1} u_i \frac{\partial x_i}{\partial P_k} + u_m \left( \frac{\partial x_m}{\partial P_k} + \frac{\partial \alpha}{\partial P_k} \frac{\partial x_m}{\partial \alpha} \right) + u_{m+1} \frac{\partial \alpha}{\partial P_k}$$

$$= -u_0 \frac{\partial x_0}{\partial P_k} + \sum_{i=1}^{m} u_i \frac{\partial x_i}{\partial P_k} + u_m \left( \frac{\partial \alpha}{\partial P_k} \frac{\partial x_m}{\partial \alpha} \right) + u_{m+1} \frac{\partial \alpha}{\partial P_k}$$
Substituting in (2), (3):

\[ = -\lambda \frac{\partial x_0}{\partial P_k} + \sum_{i=1}^{m} \lambda P_i \frac{\partial x_i}{\partial P_k} + u_m \left( \frac{\partial \alpha}{\partial P_k} \frac{\partial x_m}{\partial \alpha} \right) + u_{m+1} \frac{\partial \alpha}{\partial P_k} \]  

(7)

\[ = -\lambda \left( \frac{\partial x_0}{\partial P_k} - \sum_{i=1}^{m} P_i \frac{\partial x_i}{\partial P_k} \right) + u_m \left( \frac{\partial \alpha}{\partial P_k} \frac{\partial x_m}{\partial \alpha} \right) + u_{m+1} \frac{\partial \alpha}{\partial P_k} \]  

(8)

And using the fact \[ x_k = \frac{\partial x_0}{\partial P_k} - \sum_{i=1}^{m} P_i \frac{\partial x_i}{\partial P_k} \], we conclude that

\[ \frac{\partial V(P)}{\partial P_k} = -\lambda x_k + u_m \left( \frac{\partial \alpha}{\partial P_k} \frac{\partial x_m}{\partial \alpha} \right) + u_{m+1} \frac{\partial \alpha}{\partial P_k} \]  

(9)

Equation (9) will be useful shortly. Define the government’s problem as the maximization of \[ V(P) \] subject to raising a budget of at least \[ T \]. Define \[ t_i, \] the tax on good \[ i, \] as the difference between the final price and the producer price: \[ t_i = P_i - p_i. \] Implicitly this is assuming perfectly competitive production markets. This is an unreasonable assumption for many of the goods considered, but generalizing the optimal commodity tax to address market imperfections is not the focus of this paper. The government maximization problem can be summarized as

\[ \mathcal{L} = n V(P) - \beta \left[ n \sum_{i=1}^{m} (P_i - p_i)x_i - T \right] \]  

(10)

Now using Equation (9), we can see that a necessary condition for the optimal commodity tax rate is:

\[ \frac{\partial \mathcal{L}}{\partial P_k} = -\lambda x_k + u_m \left( \frac{\partial \alpha}{\partial P_k} \frac{\partial x_m}{\partial \alpha} \right) + u_{m+1} \frac{\partial \alpha}{\partial P_k} - \beta \left[ \sum_{i=1}^{m} t_i \frac{\partial x_i}{\partial P_k} + x_k \right] = 0 \]  

(11)

This can be simplified. Noting that \[ \frac{\partial \alpha}{\partial P_k} = n \frac{\partial x_m}{\partial P_k}, \]

\[ \sum_{i=1}^{m} t_i \frac{\partial x_i}{\partial P_k} = -\left( \frac{\lambda + \beta}{\beta} \right) x_k + \frac{n}{\beta} \left( u_{m+1} + u_m \frac{\partial x_m}{\partial \alpha} \right) \frac{\partial x_m}{\partial P_k} \]  

(12)

Let the coefficient matrix on \[ t_i \] (the transpose of the Jacobian of the taxable goods’ demand functions) be denoted \[ J^* \]. Further let \[ J \equiv \text{det}(J^*) \] and denote \[ J_{ik} \] as the cofactor of the element in row \[ i \], column \[ j \] of \[ J \]. Then, applying Cramer’s Rule:

\[ t_k = \left[ -\left( \frac{\lambda + \beta}{\beta} \right) \right] \left( \sum_{i=1}^{m} \frac{x_i J_{ik}}{J} \right) + \frac{n}{\beta} \left( u_{m+1} + u_m \frac{\partial x_m}{\partial \alpha} \right) \left( \frac{\sum_{i=1}^{m} \frac{\partial x_m}{\partial P_k} J_{ik}}{J} \right) \]  

(13)
As per Sandmo (1975), it can be shown that:

$$\sum_{i=1}^{m} \frac{\partial x_m}{\partial P_i} J_{ik} = \begin{cases} 0 & \text{for } k \neq m \\ J & \text{for } k = m \end{cases}$$

(14)

Consequently,

$$t_k = -\left(\frac{\lambda + \beta}{\beta}\right) \left(\frac{\sum_{i=1}^{m} x_i J_{ik}}{J}\right) + \frac{n}{\beta} \left(\frac{u_{m+1} + u_m \frac{\partial x_m}{\partial \alpha}}{J}\right) \times 1_{\{k=m\}}$$

(15)

$$\frac{t_k}{P_k} = \left(-\frac{1}{P_k}\right) \left(\frac{\lambda}{\beta} + 1\right) \left(\frac{\sum_{i=1}^{m} x_i J_{ik}}{J}\right) + \frac{n \lambda}{\beta} \frac{1}{P_k} \left(\frac{u_{m+1} + u_m \frac{\partial x_m}{\partial \alpha}}{J}\right) \times 1_{\{k=m\}}$$

(16)

Defining $\theta_i$ as the tax rate on good $i$, i.e. $\theta_i \equiv t_i/P_i$ and $\mu$ as the negative of the ratio of Lagrangian multipliers, i.e. $\mu \equiv -\lambda/\beta$,

$$\theta_k = \left(-\frac{1}{P_k}\right) (1 - \mu) \left(\frac{\sum_{i=1}^{m} x_i J_{ik}}{J}\right) - n \mu \left(\frac{1}{\lambda P_m}\right) \left(\frac{u_{m+1} + u_m \frac{\partial x_m}{\partial \alpha}}{J}\right) \times 1_{\{k=m\}}$$

(17)

Finally, substituting from (3) and rearranging, we have:

$$\theta_k = (1 - \mu) \left[\frac{-1}{P_k} \sum_{i=1}^{m} x_i J_{ik}\right], \quad k = 1, \ldots, (m - 1)$$

(18)

$$\theta_m = (1 - \mu) \left[\frac{-1}{P_m} \sum_{i=1}^{m} x_i J_{im}\right] - \mu \left[n \left(\frac{u_{m+1}}{u_m}\right)\right] - \mu \left[n \left(\frac{\partial x_m}{\partial \alpha}\right)\right]$$

(19)

This solves for the optimal tax rates. Equation (18) shows that the tax rate on the $m-1$ typical goods is a form of the Ramsey discouragement index, consistent with the findings of the previous literature.

Equation (19) defines the optimal rate for the $m^{th}$ good. It shows that the tax comprises three additively separable components: the first, a Ramsey-like factor which decreases in the sensitivity/elasticity of consumers to price the second, a Pigouvian factor increasing in the magnitude of the direct externality; and the third, an adjustment for how consumption responds to the externality.

The departure of Equations (18) and (19) with previous research is the third ‘consumption response’ component. If consumption does not depend on the externality, e.g. when the demand for widgets in unaffected by pollution in a lake, the consumption response component is zero and the optimal tax rate collapses to that found by Sandmo (1975).

This can be shown more clearly by grouping the final terms in Equation (19) together:

$$\theta_m = (1 - \mu) \left[\frac{-1}{P_m} \sum_{i=1}^{m} x_i J_{im}\right] + \mu \left[-n \left(\frac{u_{m+1}}{u_m}\right) + \mu \left(\frac{\partial x_m}{\partial \alpha}\right)\right]$$

(20)

With this formulation, we can interpret the result as a weighted average (with weight $\mu$) of
Ramsey taxation and adjusted-externality taxation. There may be disutility caused by $\alpha$, but the extent to which $\alpha$ increases consumption of $x_m$ can mitigate that negative effect. Indeed, if $\frac{\partial x_m}{\partial \alpha} > \frac{\partial u_m}{\partial \alpha} + 1$, then the optimal policy is to subsidize the “dirty” (i.e. negative atmospheric externality-generating) good.

This result shows that the optimal taxation of dual atmospheric-consumption externalities is more complex than the existing literature but retains several intuitive features. The tax remains an additive and separable weighted average of Ramsey commodity and Pigouvian externality taxation. The Pigouvian component is adjusted to account for the effect of the externality on demand. When the externality positively affects demand, then the optimal tax rate is lower. The next section of the paper analyzes how optimal taxation should respond in a dynamic context.

2.2 Intertemporal extension

Durable goods that are purchased today persist into future periods, and so too will the consumption externalities they generate. Thus for durable goods, which many network goods like cell phones are, the dynamics are of particular interest. If and when contemporaneous adoption rates affect demand in subsequent periods, consumption today has intertemporal or ‘dynamic’ external effects.

These intertemporal effects are relevant to the optimal tax problem. If expanding adoption today affects utility tomorrow, a government with an intertemporal revenue requirement faces a tradeoff: increased tax rates implies lower adoption today and therefore a smaller tax base for that good in future periods. Equivalently, decreased tax rates that induce market growth today may provide a larger tax base in future periods. In this situation, the durability of the good generates intertemporal revenue effects that the government must consider.

It is not obvious a priori what path optimal taxation would follow over time. The fact that deadweight loss generally increases exponentially in tax rates means a time-constant tax is intuitively appealing. However, this intuition need not be true with intertemporal effects. Time-varying tax schedules may, for example, be Pareto improving if initial subsidies induce earlier adoption. It is plausible that after a few periods’ growth the network effects become important enough that demand is quite insensitive to increases in taxes.

While not directly contradicting the conventional wisdom in public finance (e.g. Atkinson and Stiglitz, 1976), the notion that time-varying commodity tax schedules could be welfare-improving is an unusual result. However, in our context, any revenue-neutral increases in consumer surplus are possible because, insofar as tax rates influence adoption rates, tax rates affect utility from (and the demand for) the consumption externality-generating good in subsequent periods. Using the intuition of Equation (19), the Ramsey taxation component might dominate. Under these conditions, the optimal tax rate could be quite high even if there are large positive consumption externalities.\[5\]

\[5\]We explore one such case in the Appendix, where we calibrate the model to a setup with very large consumption
However, any analysis of intertemporal effects requires an assumption on how consumers form expectations. Consider the optimization problem specified in Equation (1). Consumers optimize given a level of $\alpha$, which in a dynamic context should be denoted as $\alpha_t$. For durable goods, the decision to purchase depends on the expectation of utility in future periods. For example, with standard Bellman-type notation,

$$V(x_t|\alpha_t) = u(x_t|\alpha_t) + \beta \mathbb{E}[V(x_{t+1}|\alpha_{t+1})]$$

and the decision today will depend on $\mathbb{E}[\alpha_{t+1}]$, leading to the well-known possibility of multiple equilibria. To overcome this problem, authors have generally imposed some structure on expectations or equilibrium selection. For example, Katz and Shapiro (1986) assume agents can coordinate to choose whichever equilibrium is Pareto optimal. Jackson and Watts (2002) assume expectations that are “somewhat myopic in that individuals do not forecast how their [personal decision] might affect future decisions of other individuals or more generally how it might influence the future evolution of [consumption].” We believe this latter assumption best matches the notion of atomistic agents ignoring their contribution to an externality, and it avoids the potential for expectations-induced multiple equilibria because the current information set (prices, quantities, taxes) is sufficient to determine purchasing behavior during the current period.

2.3 Monte Carlo Simulation

The derivation of optimal time paths of taxation in the theoretical model is mathematically infeasible at the level of generality in the tradition of Sandmo (1975), Kopczuk (2003), Micheletto (2008), and Aronsson and Johansson-Stenman (2018). There have been some attempts to derive closed-form solutions to related problems. For example, Greaker and Midttømme (2016) models optimal tax policy with two competing types of good, of which consumers must possess one. Their model is a (repeated) three-period game, and governments do not have an exogenous revenue constraint, so it is not clear the extent to which their findings would transfer to our setting. We therefore employ Monte Carlo simulations to demonstrate how the presence of network effects implies that a time-varying sequences of taxes can substantially improve welfare over a constant tax rate while remaining at least revenue neutral over the considered period. Alternatively, we empirically show that under the assumptions of the theoretical model (and conditional on certain parameter values) static tax sequences are not Pareto optimal.

2.3.1 Consumer utility and government objectives

To conduct these simulations, we impose some specific features for tractability (e.g., functional form of utility) and some adjustments for simplicity. In our simulations, income is exogenous externalities. We demonstrate that when consumption externalities are strong enough, the benefits of initial subsidization are increased. In this calibration, low initial tax rates induce such large network effects that the rate of adoption remains elevated even though taxes are substantially higher than baseline rates.

9
as we do not model the choice of labor supply. Doing so would greatly complicate the model, and evidence suggests labor supply is relatively unresponsive to taxes on specific goods (see e.g. Madden, 1995). We also make the consumer’s choice binary rather than continuous. While exploring the effects of commodity taxation on consumption behavior at the intensive margin may be of some interest, modeling both margins for individual consumers, would entail considerable complications without affecting the fundamental economic implications.

Acknowledging those adjustments, we take great care to preserve the important foundational assumptions of the baseline model. Specifically, we consider consumers’ decision to allocate their resources between one good that yields consumption and/or atmospheric externalities, and one private numeraire good. Consumers are rational and use all information available at time $t$ in their purchasing decisions, and do not consider their own effect on the externality. Following Goyal (2012), we call consumers ‘myopic’ in that their best guess of time $t+1$ values is their level today, i.e. $E[x_{t+1}] = x_t$, and the utility of consumers is additively separable. Mechanically, the government conducts a grid search over sequences of tax rates to maximize total surplus subject to satisfying a budget requirement.

We simulate purchasing decisions for a durable good for $n = 10,000$ individuals over six periods. When an individual purchases the good, there is a one-time purchase price of amount $p$. However, for each period in which the individual owns the good, the individual pays a tax of amount $\tau_t$. We impose a CES functional form on the individuals’ utility function. The individual’s utility from purchasing the good can be expressed as:

$$U_{it1} = (\gamma + \alpha \cdot s_t)^r + (y - p - \tau_t + \delta \cdot s_t)^r)^{1/r} + \epsilon_{it1}$$

Where $\gamma$ is the flow utility from owning the good, even if no one else does, $s_t$ is the share of the population who owns the good, $\alpha$ is the parameter that captures how much the purchase decisions of $j \in -i$ affect person $i$’s utility of consuming the good, and $\epsilon_{it1}$ is an idiosyncratic preference shock. The second term captures the utility of consuming the numeraire good. To the extent that the durable good of interest yields atmospheric externalities, those effects are parameterized by $\delta$. Substitution preferences are measured by $r$, income by $y$, purchase price of the good by $p$, and the tax rate by $\tau$.

If the individual has purchased the good in a previous period, their flow utility is the same as the above, except that they do not pay the purchase price $p$ again. They do, however, pay the usage taxes for every period they own the good. If the individual does not purchase the good, their utility can be expressed as:

---

6The number of periods chosen is somewhat ad hoc but is not pivotal for our results. If we evaluate $K$ possible tax rates for each of $T$ periods, we must calculate welfare for $K^T$ sequences of taxes. Therefore, increasing the number of periods increases computation time by a factor of $K$.

7Like all parameters in Monte Carlo simulations, the robustness of this assumption can be tested in future work.
\[ U_{it0} = ((y + \delta \cdot s_t)^{1/r} + \epsilon_{it0})^{1/r} = y + \delta \cdot s_t + \epsilon_{it0} \]

which is derived from the expression above where the individual gets none of the utility in the first subset of parentheses, but also does not pay the tax nor the purchase price in the second. Due to the CES form of the utility function, the expression collapses to be linear in income and the atmospheric externality. As consumers use the time-\(t\) information set in their decision, the current utility from purchasing the good becomes a sufficient statistic for the present discounted flow of utility. Valuing future flows would be an affine transformation of utility and could change the specific parameter values for which dynamic taxation improves efficiency, but not the overall pattern of results. While we explore adoption as a permanent state in some specifications, consumers in our model generally have the option of free disposal and in this sense our results offer a lower-bound as the potential gains from time-varying taxes.

Mechanically, the simulation proceeds as follows: in the first period, no one has the good, implying \(s_t = 0\). For each individual, we assume the idiosyncratic preference shocks are i.i.d. Type 1 Extreme Value, meaning that individuals purchase the good with probability:

\[ p(\text{Purchase}) = \frac{e^{U_{it}}} {e^{U_{it0}} + e^{U_{it1}}} \]

We take a pseudo-random draw, \(\eta_{it}\), from the \(U[0,1]\) distribution. If \(\eta_{it} > p(\text{purchase})\), the individual purchases the good. At the start of the each successive period, individuals observe the share of the population who currently own the good as a result of choices in previous periods, and that information affects contemporaneous purchase probabilities.

In these simulations, the government sector serves as a constraint. We assume a revenue requirement \(\tilde{w}\) is necessary to finance government activities or provide a public good, \(\Sigma_t(\tau_t > \tilde{w})\). We do not model the provision of that good. Rather, the government’s objective is to satisfy the revenue constraint with minimal distortions.

For each considered vector of parameters, \(\theta\), we simulate this 10,000 agent, six period model for each possible sequence of taxes \(\tau_t \in \{0, 0.1, 0.2, \ldots, 0.9\}\), \(\forall t \in \{1, \ldots, 6\}\). For each vector of parameters, then, we compare tax revenues and total consumer utility for the 1,000,000 considered sequences of tax rates. It should be remembered we are not necessarily looking for \(\max \sum_t \sum_i (U_{it}|\theta, \sum_t \tau_t > \tilde{w})\). Rather, we demonstrate that for a given amount of aggregate revenue \(\tilde{w}\) generated by a time-invariant sequence of tax rates, \(\tau_t = \tilde{\tau} \forall t\), when consumption externalities are present \((\alpha \neq 0)\) there exists a time-varying sequence of tax rates that increases total consumer utility while raising at least as much revenue for the government as the static sequence. We are searching for optimal tax sequences defined over a discrete, bounded interval, and exploring the conditions under which potential gains from dynamic taxation are larger or smaller.

We simulate the model under two different assumptions about the permanence of consumer
choice. The first assumes that a purchase is an absorbing state; once an individual owns the good, they own the good for the remainder of the simulated periods. This assumption is appropriate for cases where goods come with contracts (e.g., mobile phones) or where disposal of the good would be costly. However, this enables the government to ‘bait and switch’ consumers in a sense with low tax rates in early periods and high tax rates in later periods. We do not believe this is realistic and do not want this effect to drive our results. We therefore also simulate the model under an alternative condition where in each period consumers may choose to discard the good at no cost.

To further discipline the model, every period a fraction of consumers also receive a negative preference shock to encourage disposal of the good. The probability of disposal increases if taxes are increased. Essentially, the absorbing state and free disposal state represent feasible endpoints of the spectrum about the ease of disposing of the good when the government increases tax rates. Under either set of assumptions about permanence or disposal cost, because consumers enjoy flow utilities from the good without having to ‘re-purchase’ the good each period, dynamic taxation may be welfare improving without consumption externalities. We therefore simulate a comparison case for each set of parameters \( r, p, \gamma, y \) where \( \alpha = 0 \) to establish a baseline. Any gains from dynamic taxation in the presence of consumption externalities, relative to this baseline, can be directly attributed to the dynamic effects of tax rates and consumption externalities.

### 2.3.2 Results

Tables 1 through 3 contain the results from conducting simulations with several sets of parameters, to quantify welfare gains from dynamic taxation under different assumptions about consumer utility. Because the dynamic implications of taxation and adoption rates are driven by consumption externalities, we focus specifically on that aspect in the first two tables. For each set of parameters, we choose two static tax rates, \( \tilde{\tau}_1 = 0.3; \tilde{\tau}_2 = 0.5 \), as baselines for the model. We sum the collected taxes and consumer surplus under those static rates. The collected revenue becomes the benchmark exogenous revenue requirement in the dynamic case. We then evaluate consumer surplus for all tax sequences that collect tax revenues greater than or equal to the amount generated by static taxation. Among the tax sequences considered in our discrete grid search, we focus our attention on the tax sequence \( \tau^* \) that yields maximum consumer utility, conditional on meeting or exceeding the revenue requirement.

We use two measures of change in welfare. First, and most obvious, we use percentage change in consumer surplus from the baseline \( \tilde{\tau}_t = \tilde{\tau} \forall t \in 1, \ldots, 6 \):

---

8We use the mean utility from each good to produce logit probabilities for keeping the good, similar to the way we introduced randomness in the purchase decision.

9This abstracts from the possibility of resale, which would create negative disposal costs.

10These rates are admittedly ad-hoc, but were chosen because they are in the middle of the considered range of values, allowing for us to empirically evaluate the theoretical predictions of the model.
\[
\%\Delta CS = \frac{(CS|\tau^*) - (CS|\tilde{\tau})}{(CS|\tilde{\tau})}
\]

However, this measure is likely to understate gains from dynamic taxation specifically attributable to consumption externalities, because that comparison requires evaluating differences in \%\Delta CS with different denominators. For example, setting \(\alpha = 1\) rather than \(\alpha = 0\) raises the denominator because it raises the utility for each and every tax sequence considered, but may not affect the range of consumer surplus affected by tax schemes. Thus we also map the range of consumer surpluses generated by the grid of tax sequences into the \([0, 1]\) interval conditional on the specific values for \(\alpha, p, \gamma, y,\) and \(r\). We then evaluate how dynamic taxation affects consumer surplus over that range, relative to the baseline case. We express this \textit{Normalized} measure as:

\[
\text{Normalized Gains} = \frac{(CS|\tau^*, \alpha, p, \gamma, y, r) - (CS|\tilde{\tau}, \alpha, p, \gamma, y, r)}{\max_{\tau}(CS|\alpha, p, \gamma, y, r) - \min_{\tau}(CS|\alpha, p, \gamma, y, r)}
\]

Panel A of Table 1 presents results for a baseline set of parameters. These parameters were chosen because of several properties: substitution between the good of interest and ‘other’ consumption is slightly inelastic \((r = 0.8)\), the purchase price slightly exceeds the private utility alone \((p = 0.9, \gamma = 0.6)\), and income is large enough to keep all arguments in each component of the utility function greater than one.

The results in Panel A provide two pieces of insight to support that dynamic taxation improves welfare when consumption externalities are present. First, when the purchase decision is binding (permanent) dynamic taxation improves welfare somewhat due to the structure of the simulation, even when externalities are not present. (When consumption externalities \((\alpha > 0)\) are introduced, the gains of dynamic taxation increase by over 40 percent \((\text{i.e. 6 percent to 8.8 percent})\) relative to the baseline.) Second, when \(\tilde{\tau} = 0.3\) and individuals are allowed to costlessly dispose of the good, in the absence of consumption externalities \((\alpha = 0)\), dynamic taxation does not substantially improve welfare. However, when \(\alpha = 1\), a scheme of \(\tau^* = \{0.0, 0.0, 0.1, 0.3, 0.5, 0.8\}\) yields more tax revenue and welfare gains equal to 1.21 percent improvement over the baseline consumer surplus, or 4.96 percent of the support of CS gains attributable to taxation.

We also see that when \(\tilde{\tau} = 0.5\), dynamic taxation has more potential to improve welfare than the case where \(\tilde{\tau} = 0.3\). The greater the revenue requirement and higher the rate, the more opportunity there is to pursue optimal taxation through dynamics when consumption externalities are present. Comparing rows 3 and 5 for the \(\tilde{\tau} = 0.5\) case, we see less dramatic acceleration of tax rates when disposal of the good is an option to consumers. As expected, the ability to discard the phone constrains the government’s ability to ‘bait and switch’.

However, even when consumers have the choice of free disposal, we still see substantial gains from low initial rates. Taking the fifth row (bottom-right of the panel) as an example, the optimal tax sequence is an elevation of tax rates from 0.1 at the start up to 0.9 at the end. This improves welfare by over 7 percent of the support of consumer surplus. This is strong evidence that the
desirability of time-varying taxation is not driven by our assumptions about consumer myopia.

Panel A shows that when consumption externalities are present, dynamic taxation can improve welfare relative to a static scheme. However, the extent to which dynamic taxation can improve welfare is determined by several aspects of the individual’s choice problem. Of special importance is how pivotal taxes are in the individuals’ decision to buy the good when the good is new, or when only a small share of the population owns the good. The parameters in Panel A are such that the choice probabilities of purchase are quite high when the price \( p \) is small and tax rates are between 0.3 and 0.5. Panel B, by contrast, shows results when purchase prices increase from 0.9 to 1.5. In this situation of a larger gap between prices and flow utilities, and therefore fewer inframarginal consumers, dynamic taxation has greater potential to generate welfare gains. With \( \tilde{\tau} = 0.3 \) and a permanent purchase decision, the gains from dynamic taxation when \( \alpha = 1 \) are twice as large as the baseline case in Panel A. When individuals are permitted to dispose of the good, the gains from dynamic taxation are equal to 3.14 percent of the baseline case, or 13.6 percent or the support of CS gains attributable to taxation. When \( \tilde{\tau} = 0.5 \), dynamic taxation improves welfare even more, a 21.4 percent increase when consumers are constrained by contract to retain the good. Introducing the option of free disposable lowers this number somewhat, to 18.6 percent, but the gain is still substantial.

Finally, note that most of the identified \( \tau^* \) in Panel A and Panel B are only maxima because they are constrained by the grid search over the \([0, 1]\) interval by deciles. The optimal tax sequences generally begin at the lower-bound and tend towards the upper-bound. Our \([0, 1]\) feasible tax interval is reasonable but admittedly ad hoc. It may be instruction to analyze whether expanding the range to include initial subsidies and even higher taxes later can yield additional gains to consumer surplus. We include Panel C as an acknowledgement of this.

In Panel C, we relax the range of taxes to span \([-0.5, 0.4]\) in the first three periods and \([0.3, 1.2]\) in the last three periods. Comparing the \( \alpha = 1 \) case to the equivalent rows of Panel B, we see \( \tau^* \) is a smooth escalation from maximum subsidy in the first period to the maximum tax rate in the last period. We see a substantial gain in welfare, equal to between 18 and 24 percent of the support of consumer surplus. In the presence of consumption externalities, permitting initial subsidies further increases potential welfare gains over the case where tax rates are bounded below by zero. The final row (\( \alpha = 2 \)) further confirms this with greater deviations from time-invariance further enhancing welfare.

Figures 1 and 2 provide graphical intuition to supplement the findings in Table 1. Figure 1 depicts cumulative tax revenue in the default time-invariant case (solid line) and in the optimal tax sequence (dashed line). Figure 2 shows the accompanying per period number of number of users of the good in the time-invariance tax case (solid line) and optimal policy case (dashed line). All panels depict outcomes for \( \alpha = 1 \) and where consumers have the option of free disposal. Note that consumers do respond to low and high tax rates. The high tax rates in the last few periods of the optimal sequences do decrease the number of buyers. This is not always the case, however. When \( \alpha > 2 \), the number of buyers is approximately the same as the static case. Even
Table 1: Summary of simulation results under different baseline prices of the good

Panel A: \((p = 0.9; \gamma = 0.6; r = 0.8; y = 4)\)

<table>
<thead>
<tr>
<th>(\alpha)</th>
<th>Discardable</th>
<th>(\bar{\tau}_t = 0.3\forall t)</th>
<th>(\bar{\tau}_t = 0.5\forall t)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\tau^*)</td>
<td>%ΔCS Normalized</td>
<td>%ΔCS Normalized</td>
</tr>
<tr>
<td>(\alpha = 0)</td>
<td>No</td>
<td>0.0; 0.0; 0.0; 0.8; 0.7</td>
<td>1.12; 6.00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(\alpha = 0)</td>
<td>0.0, 0.0, 0.0, 0.8, 0.8, 0.9, 0.8</td>
</tr>
<tr>
<td>(\alpha = 0.5)</td>
<td>No</td>
<td>0.0; 0.0; 0.0; 0.7; 0.9</td>
<td>1.35; 8.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(\alpha = 0.5)</td>
<td>0.0, 0.0, 0.0, 0.8, 0.8, 0.9, 0.8</td>
</tr>
<tr>
<td>(\alpha = 1.0)</td>
<td>No</td>
<td>0.0, 0.0, 0.1, 0.6, 0.9</td>
<td>1.57; 8.82</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(\alpha = 1.0)</td>
<td>0.0, 0.0, 0.0, 0.8, 0.9, 0.9</td>
</tr>
<tr>
<td>(\alpha = 0)</td>
<td>Yes</td>
<td>0.2, 0.1, 0.2, 0.3, 0.3, 0.8</td>
<td>0.19; 1.28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(\alpha = 0)</td>
<td>0.5, 0.5, 0.3, 0.4, 0.5, 0.9</td>
</tr>
<tr>
<td>(\alpha = 1.0)</td>
<td>Yes</td>
<td>0.0, 0.0, 0.1, 0.3, 0.5, 0.8</td>
<td>1.21; 4.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(\alpha = 1.0)</td>
<td>0.1, 0.2, 0.4, 0.5, 0.7, 0.9</td>
</tr>
</tbody>
</table>

Panel B: \((p = 1.5; \gamma = 0.6; r = 0.8; y = 4)\)

<table>
<thead>
<tr>
<th>(\alpha)</th>
<th>Discardable</th>
<th>(\bar{\tau}_t = 0.3\forall t)</th>
<th>(\bar{\tau}_t = 0.5\forall t)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\tau^*)</td>
<td>%ΔCS Normalized</td>
<td>%ΔCS Normalized</td>
</tr>
<tr>
<td>(\alpha = 0)</td>
<td>No</td>
<td>0.0; 0.0; 0.0; 0.7; 0.1; 0.6</td>
<td>1.51; 9.41</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(\alpha = 0)</td>
<td>0.0, 0.0, 0.0, 0.9, 0.9, 0.4</td>
</tr>
<tr>
<td>(\alpha = 0.5)</td>
<td>No</td>
<td>0.0; 0.0; 0.0; 0.5; 0.9</td>
<td>2.58; 13.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(\alpha = 0.5)</td>
<td>0.0, 0.0, 0.0, 0.9, 0.9, 0.9</td>
</tr>
<tr>
<td>(\alpha = 1.0)</td>
<td>No</td>
<td>0.0; 0.0; 0.0; 0.5; 0.9</td>
<td>3.14; 13.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(\alpha = 1.0)</td>
<td>0.0, 0.0, 0.0, 0.4, 0.9, 0.9</td>
</tr>
<tr>
<td>(\alpha = 0)</td>
<td>Yes</td>
<td>0.1; 0.1; 0.2; 0.1; 0.5; 0.9</td>
<td>0.53; 4.49</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(\alpha = 0)</td>
<td>0.5, 0.3, 0.4, 0.4, 0.5, 0.9</td>
</tr>
<tr>
<td>(\alpha = 1.0)</td>
<td>Yes</td>
<td>0.0; 0.0; 0.0; 0.7; 0.9</td>
<td>3.26; 12.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(\alpha = 1.0)</td>
<td>0.0, 0.0, 0.3, 0.5, 0.7, 0.9</td>
</tr>
</tbody>
</table>

Panel C: \((p = 1.5; \gamma = 0.6; r = 0.8; y = 4)\) — expanded range of taxes/subsidies

<table>
<thead>
<tr>
<th>(\alpha)</th>
<th>Discardable</th>
<th>(\bar{\tau}_t = 0.3\forall t)</th>
<th>(\bar{\tau}_t = 0.5\forall t)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\tau^*)</td>
<td>%ΔCS Normalized</td>
<td>%ΔCS Normalized</td>
</tr>
<tr>
<td>(\alpha = 0)</td>
<td>Yes</td>
<td>0.1; 0.0; 0.1; 0.4; 0.4; 0.9</td>
<td>0.52; 3.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(\alpha = 0)</td>
<td>0.6; 0.5; 0.6; 0.4; 0.5; 1.1</td>
</tr>
<tr>
<td>(\alpha = 1.0)</td>
<td>Yes</td>
<td>-0.5; -0.2; 0.1; 0.3; 0.6; 1.2</td>
<td>5.07; 18.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(\alpha = 1.0)</td>
<td>-0.5; 0.0; 0.4; 0.6; 0.7; 1.1</td>
</tr>
<tr>
<td>(\alpha = 2.0)</td>
<td>Yes</td>
<td>-0.5; -0.5; 0.0; 0.4; 0.7; 1.2</td>
<td>5.30; 25.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(\alpha = 2.0)</td>
<td>-0.5; -0.5; 0.6; 0.7; 0.9; 1.1</td>
</tr>
</tbody>
</table>

Notes: Each row within a Panel presents the welfare-maximizing tax sequence \(\tau^*\) for different parameter specifications of unit price \(p\), flow utility \(\gamma\), elasticity of substitution \(r\), and income \(y\). The \(\alpha\) parameter represents the intensity of the consumption externality. All simulations are run with \(n = 10,000\) consumers. Please refer to the text for how we normalize the gains in consumer surplus.
when $\alpha = 1$, high taxes have smaller effects on disposal rates than when $\alpha = 0$.

Generally, compared to static sequences, optimal tax sequences (when positive consumption externalities are present) feature lower tax rates in initial periods. These lower rates generate a compensating effect of greater numbers of users comprising a larger tax base in subsequent periods. Take the top-left panels as an example. The constrained welfare maximizing tax sequence is $\{0, 0, 0.1, 0.3, 0.5, 0.8\}$, and in Figure 1 the cumulative revenue (dashed line) is contrasted with the cumulative revenue from a policy of a constant tax rate of 0.3 (solid line). Though it is not perfectly straight, the solid line shows a relatively consistent increase in cumulative revenues over the entire period. Unsurprisingly, the dashed line shows no tax revenue is raised in the initial periods when tax rates are zero. However, Figure 2 shows this encourages early adoption, increasing the number of users of the good. As well as creating consumer welfare through the consumption externality’s positive effects on utility, this expands the tax base for the relatively high rates that follow. Revenues quickly catch up and by the final period marginally surpass those raised in the time-invariant case. The higher tax rates mean the number of users tails off in the final periods, but not so much as to result in lower overall welfare. Using the mechanism of reduced sensitivity to price increases, the planner maximizes total welfare by increasing tax rates in later periods.

Comparing with the top-left and top-right panels in Figure 1, we see more revenue is raised when the tax rate is increased to 0.5 rather than 0.3. However, the earlier takeaways persist. The pattern of low initial taxes followed by higher taxes on a relatively larger base persists, and accumulated revenue only catches up in the final period. The corresponding comparison in Figure 2 is quite similar. The patterns of consumption/use take the same shape as when the tax rate was 0.3. Unsurprisingly, the number of users is uniformly lower in the high-tax case.

Of particular interest are the results in Panel C, where we expanded the range of feasible tax rates to include very low (i.e. negative) early rates and higher subsequent rates. This of course assumes governments have the ability to borrow, and at zero interest rates. This is consistent with our assumption of zero discounting of future utility. The bottom-left panel shows the optimal policy does induce government debt, and that debt persists until the fourth period. Initial adoption is very high, with about twice as many users consuming the good in the first period compared to the time-invariant baseline. Indeed, the optimal tax schedule induces more users than the time-invariant schedule through the fifth period. Only in the final period, when taxes are very high, does the number of users drop below the baseline. By this late stage, the increased number of users (and the utility boost the large network externality generates) has considerably increased welfare. The high tax rates in the final periods ensure cumulative revenue in the optimal schedule surpasses the baseline in the final period.

11Discount rates and interest rates would not substantively affect our results. Consumers would benefit even more from low initial tax rates, and governments would need to increase rates by more on the back end pay the interest. So long as consumption externalities have the requisite strength, the qualitative implications are unchanged.
Figures show paths of cumulative tax revenue for the time-invariant case (solid line) and optimal policy case (dashed line). All panels depict outcomes for $\alpha = 1$ and where consumers have the option of free disposal.
Figure 2: Graphical depiction of consumption paths

Patterns of Adoption and Use

Figures show paths of consumption/usage for the time-invariant case (solid line) and optimal policy case (dashed line). All panels depict outcomes for $\alpha = 1$ and where consumers have the option of free disposal.
Table 2: Simulation results under different baseline prices, flow utility, and substitution elasticity

<table>
<thead>
<tr>
<th>Panel A: ( (p = 1.5; \gamma = 0.6; r = 2.0; y = 4) )</th>
<th>( \bar{\tau}_t = 0.3 \forall t )</th>
<th>( \bar{\tau}_t = 0.5 \forall t )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>Discardable</td>
<td>( \tau^* )</td>
</tr>
<tr>
<td>( \alpha = 0 )</td>
<td>No</td>
<td>0.9; 0.9; 0.6; 0.4; 0.3; 0.1</td>
</tr>
<tr>
<td>( \alpha = 2.0 )</td>
<td>No</td>
<td>0.0; 0.0; 0.0; 0.2; 0.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: ( (p = 0.9; \gamma = 0.0; r = 0.5; y = 4) )</th>
<th>( \bar{\tau}_t = 0.3 \forall t )</th>
<th>( \bar{\tau}_t = 0.5 \forall t )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>Discardable</td>
<td>( \tau^* )</td>
</tr>
<tr>
<td>( \alpha = 0 )</td>
<td>No</td>
<td>0.8; 0.9; 0.9; 0.6; 0.0; 0.0</td>
</tr>
<tr>
<td>( \alpha = 1.0 )</td>
<td>No</td>
<td>0.0; 0.0; 0.0; 0.3; 0.6; 0.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C: ( (p = 0.9; \gamma = 0.6; r = 0.5; y = 4) ) — expanded range of taxes/subsidies</th>
<th>( \bar{\tau}_t = 0.3 \forall t )</th>
<th>( \bar{\tau}_t = 0.5 \forall t )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>Discardable</td>
<td>( \tau^* )</td>
</tr>
<tr>
<td>( \alpha = 0 )</td>
<td>No</td>
<td>0.0; 0.0; 0.3; 0.4; 0.5; 0.5</td>
</tr>
<tr>
<td>( \alpha = 1.0 )</td>
<td>No</td>
<td>0.0; 0.0; 0.4; 0.4; 0.5; 0.4</td>
</tr>
</tbody>
</table>

Notes: Each row within a Panel presents the welfare-maximizing tax sequence \( \tau^* \) for different parameter specifications of unit price \( p \), flow utility \( \gamma \), elasticity of substitution \( r \), and income \( y \). The \( \alpha \) parameter represents the intensity of the consumption externality. All simulations are run with \( n = 10,000 \) consumers. Please refer to the text for how we normalize the gains in consumer surplus.
Table 2 depicts the optimal tax rate under alternative specifications but holding the absorbing state assumption constant. Panel A shows that when substitution is highly elastic ($r = 2$) and there are no consumption externalities ($\alpha = 0$), in this finite horizon model, there are actually slight gains to be had from taxing at high rates early to keep individuals from buying in to the market, compared to a constant tax rate. When substitution between the two goods is highly elastic (e.g. $r \geq 2$), consumption externalities do lead to gains from dynamic taxation, but those gains are not substantial not until $\alpha > 0.0$.

Panel B shows that when the elasticity $r = 0.5$ but when the good yields utility only through the network (private flow utility $\gamma = 0$), dynamic taxation improves welfare over the static case. Clearly, those gains can only be attributed to the presence of those network effects. There are also cases in Table 2 where dynamic taxation is shown to not add substantial value. When substitution between the good of interest and other consumption is already sufficiently inelastic, consumption externalities do not substantively contribute to the gains from dynamic taxation. In the Panel C case where $r = 0.5$ and $\gamma = 0.6$, consumption externalities do not affect the welfare gains from dynamic taxation. Any/all gains are derived from incentivizing individuals to purchase the durable good (for which demand is inelastic) early on.

These simulations show that time-varying tax schedules improve welfare when consumption externalities are present. When the structure of the problem is such that dynamic taxation schedules improve welfare for private goods, those improvements are increasing in the strength of the consumption externality. Some parameterizations, particularly those where the distribution of purchase probabilities is nearly degenerate, leave little room for consumption externalities to yield gains through dynamic taxation. In many cases, consumption externalities can create room for dynamic taxation if the $\alpha$ parameter is large enough. Whether those large values are reasonable depends on the particular good or market being considered.

While this section implicitly treats the intertemporal development of the network as a relevant externality, Table 3 presents results with explicit incorporation of atmospheric externalities. We see that holding $\alpha$ constant, the incorporation of an atmospheric externality $\delta$ does not substantially change the pattern of taxation when $\tilde{\tau} = 0.3$. When $\tilde{\tau} = 0.5$ there is perhaps a tendency to have lower initial tax rates, but that tendency is not particularly large.

We do note that the normalized gains from the optimal tax sequence increases in $\delta$: when there are positive atmospheric externalities, the gains from dynamic taxation are larger; and the existence of negative atmospheric externalities diminish the benefits of dynamic taxation. For example, with $\tilde{\tau} = 0.3$ and $\alpha = 1$ the normalized gain from dynamic taxation is 0.14 when $\delta = 0.5$ but only 0.10 when $\delta = -0.5$. This is entirely consistent with our closed form results.

One limitation of our analysis is the discreteness of state space. The 0.1 increments of the tax space is quite stark. Consider a situation where the continuous-space optimal tax rate increases from 0.26 to 0.34. In our discrete space, this will round to a constant tax rate of $\{0.3, 0.3\}$, but an almost equivalent continuous-space optimum of 0.24 to 0.35 will round to $\{0.2, 0.4\}$ in our setting. While random shocks should even out with 10,000 consumers, the interaction of shocks...
Table 3: Simulation results incorporating explicitly dual atmospheric-consumption externalities

**Panel A:** \( p = 1.5; \gamma = 0.6; r = 0.8; y = 4 \)

<table>
<thead>
<tr>
<th>( \alpha; \delta )</th>
<th>Discardable</th>
<th>( \tilde{\tau}_t = 0.3 \forall t )</th>
<th>( % \Delta CS )</th>
<th>Normalized</th>
<th>( \tilde{\tau}_t = 0.5 \forall t )</th>
<th>( % \Delta CS )</th>
<th>Normalized</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha = 0; \delta = 0 )</td>
<td>Yes</td>
<td>0.1; 0.1; 0.2; 0.1; 0.5; 0.9</td>
<td>0.53</td>
<td>4.48</td>
<td>0.5, 0.3, 0.4, 0.5, 0.9</td>
<td>0.43</td>
<td>3.55</td>
</tr>
<tr>
<td>( \alpha = 0; \delta = 0.5 )</td>
<td>Yes</td>
<td>0.1; 0.1; 0.0; 0.3; 0.5; 0.9</td>
<td>0.97</td>
<td>6.84</td>
<td>0.1, 0.5, 0.3, 0.5, 0.6, 0.9</td>
<td>0.85</td>
<td>5.84</td>
</tr>
<tr>
<td>( \alpha = 0.5; \delta = -0.2 )</td>
<td>Yes</td>
<td>0.0; 0.0; 0.1; 0.2; 0.5; 0.9</td>
<td>1.58</td>
<td>8.82</td>
<td>0.0, 0.0, 0.5, 0.9, 0.9</td>
<td>1.41</td>
<td>7.59</td>
</tr>
<tr>
<td>( \alpha = 0.5; \delta = 0 )</td>
<td>Yes</td>
<td>0.0; 0.0; 0.1; 0.2; 0.5; 0.9</td>
<td>1.76</td>
<td>9.43</td>
<td>0.0, 0.2, 0.4, 0.5, 0.7, 0.9</td>
<td>1.71</td>
<td>8.75</td>
</tr>
<tr>
<td>( \alpha = 1.0; \delta = -0.5 )</td>
<td>Yes</td>
<td>0.0; 0.0; 0.1; 0.3; 0.4; 0.7</td>
<td>2.43</td>
<td>10.29</td>
<td>0.0; 0.1; 0.2; 0.5; 0.8; 0.9</td>
<td>3.70</td>
<td>14.25</td>
</tr>
<tr>
<td>( \alpha = 1.0; \delta = 0 )</td>
<td>Yes</td>
<td>0.0; 0.0; 0.0; 0.0; 0.7; 0.9</td>
<td>3.31</td>
<td>12.65</td>
<td>0.0; 0.0; 0.3; 0.5; 0.7; 0.9</td>
<td>5.26</td>
<td>18.78</td>
</tr>
<tr>
<td>( \alpha = 1.0; \delta = 0.5 )</td>
<td>Yes</td>
<td>0.0; 0.0; 0.0; 0.3; 0.4; 0.8</td>
<td>4.03</td>
<td>14.22</td>
<td>0.0; 0.0; 0.2; 0.5; 0.8; 0.9</td>
<td>6.57</td>
<td>21.56</td>
</tr>
</tbody>
</table>

Notes: Each row within presents the welfare-maximizing tax sequence \( \tau^* \) for given parameter values of unit price \( p \), flow utility \( \gamma \), elasticity of substitution \( r \), income \( y \), and purely atmospheric externality \( \delta \). The \( \alpha \) parameter represents the intensity of the consumption externality. All simulations are run with \( n = 10,000 \) consumers. Please refer to the text for how we normalize the gains in consumer surplus.
with discrete space can lead to imprecision. We thus encourage readers to focus interpretation on broad trends rather than details of any particular sequence.

3 Conclusion

Given that taxes drive a wedge between consumer and producer prices, this paper studies how a planner would choose that tax rate for goods with dual atmospheric-consumption externalities. Beyond theoretical interest, it is likely that such products will be a large source of government revenue in coming decades.\textsuperscript{12}

We develop a model of optimal taxation where the demand for an externality-generating good is affected by its own consumption. This setting is applicable quite broadly: to modern network goods like phones and operating systems, but also to a variety of other important non-network products like automobiles, social networks, and fashionable clothing. Indeed any component of the utility function which is affected the society’s total level of consumption can be investigated with this model.

The solution to the model generalizes previous results from the literature, including those of Pigou (1920), Ramsey (1927), and Sandmo (1975). The tax rate comprises three additively separable factors related to substitution elasticities, the magnitude of the direct externality, and the effect of the externality on consumption behavior. Negative atmospheric externalities should be taxed, and we show that the optimal tax rate is higher if the atmospheric externality lowers utility from private consumption. Equivalently, the optimal tax rate is lower if the atmospheric externality increases the utility from private consumption. If the consumption effects are strong enough, the optimal policy may even be to subsidize goods that generate negative atmospheric externalities.

Anticipating growth in the number of dual atmospheric-consumption goods in the economy, we investigate if the government should tax early-stage goods differently to well-established ones. Alternatively stated, we ask if the optimal taxation of these goods is static through time. We find that it is not. Simulating consumer choices for a spectrum of potential tax rates and finding the sequence that maximizes total surplus, we show that it can be optimal to subsidize these goods in early periods. This finding holds even when the goods come with free disposal. Incentivizing early adoption makes consumer less sensitive to subsequent tax increases, lowering excess burden in the long-run. Relative to a static baseline, initial subsidization can be revenue-neutral and welfare-enhancing.

We end on two points. Firstly, while our simulations have shown cases where time-varying taxes in general and initial subsidization in particular can delivery substantial gains, we also demonstrate several parameter values where this is not true. Secondly, we note a potential application of our work to the public policy of pandemics. One does not typically consider

\textsuperscript{12}For example, France began collecting a 3% digital services tax in December 2020.
indoor dining to have network properties. When the spread of disease is a significant feature of the world, activities like indoor dining gain characteristics of a network bad. Future work could analyze optimal taxation under these conditions.

References


A Appendix: Additional Calibrations

The results in the main paper present what we consider to be reasonable calibration values. While unitless and thus scalable, we intentionally chose flow utilities (represented by $\gamma$, typically equal to 0.6) to be of similar magnitude to the consumption externality (represented by $\alpha$, and typically set equal to either 0, 0.5, or 1).

One advantage of numerical simulations is the ease with which parameter values can be modified. This permits analysis under more extreme specifications that, while not necessarily in the main paper, may be of interest. In this section we perform two such specifications. The first calibration is with a very large consumption externality, setting $\alpha = 2.5$, a magnitude that is substantially larger than the default values. The second calibration we present is with a zero-tax default. The results in the main paper are benchmarked against revenues generated from time-invariant tax rates equal to either 0.3 or 0.5. In the second specification below, we demonstrate that a time-varying sequence of taxes can improve welfare even when the exogenous revenue constraint is zero.

What happens when we make the externalities very large? Figure 3 shows cumulative revenue and adoption rates when the consumption externality $\alpha = 2.5$, a magnitude larger than in any of the tables in the main paper. The baseline tax rate is constant at 0.3 for periods 1 through 6. The optimal sequence is 0 through the initial four periods, before sharply increasing to 0.7 and 0.8 in the final two periods, i.e. $\{0, 0, 0, 0, 0.7, 0.8\}$. The other parameters are consistent with those in Table 1, namely flow utility $\gamma$ of 0.6, a pre-tax price of 1.5, an income of 4, and a baseline substitution elasticity $r$ of 0.8.

Tax revenues steadily increase in the baseline case. There is some evidence of convexity in revenues as the number of users increases, but these effects are mild. For the optimal tax sequence, tax revenues are zero for the first four periods. This is due to the tax rate being zero, and is suggestive that subsidies (which are outside the feasible set in this simulation as tax rates are constrained between 0 and 1) may be optimal. It seems likely that more gains could be achieved were this less constrained, as in Panel C in Table 1. Revenues increase sharply in the final two periods, comfortably exceeding those raised in the baseline case.

The drivers of these results are consumers/the numbers of users of the good. The number of users in the baseline case is approximately linear, starting at about 1,000 users in the first period, and growing by about 1,500 users each period. In the optimal case, the number of users at the beginning of the sequence is about twice that of the baseline case. This is due to the lower tax rate inducing more consumers to enter the market. The growth rate in the optimal case is large, with an additional 2,000 users every period while taxes are zero. Taxes affect behavior, as noted by the sharp kink when taxes increase at period five. Increasing the tax rate from 0 to 0.7 greatly slows the growth in users. Despite the increase in taxes, however, he number of users is monotonically increasing. The large userbase is somewhat self-sustaining, and initial subsidization is profitable in a revenue-raising sense.
Figure 3: Revenue and consumption with very large consumption externalities

Figures show paths of tax revenues and number of users for the time-invariant case (solid line) and optimal policy case (dashed line). Both panels depict outcomes when consumers have free disposal and where $\alpha = 2.5$, a value high than we usually consider.
The results in Figure 3 are a product of the large consumption externality. They show an interesting case where the number of users is uniformly higher under the optimal tax sequence than in the baseline case, suggesting a Pareto improvement. A related question is whether time-varying tax rates can be welfare-improving relative to a zero-tax default. The main focus of the paper is optimal commodity taxation under a binding non-zero revenue requirement. Thus while this simulation is supplementary, it is instructive to investigate optimal policy without a revenue requirement.

Figure 4 provides evidence that initial subsidization can improve welfare even if the government has no revenue requirement. One sequence that generates positive tax revenue and more consumer surplus relative to the baseline is \{-1, -0.9, 0, 0.2, 0.5, 1\}. The paths of cumulative tax revenue and the numbers of users for these sequences are below.

Relative to a time-invariant zero tax baseline, the exogenous constraint for the optimal sequence is obviously to generate weakly positive cumulative revenues. In the baseline case, the number of users follows a concave path towards full adoption. The dynamics are slightly different in the optimal tax sequence. We see that the optimal sequence achieves positive revenues through heavy subsidization in the initial two periods to generate near-full adoption of the good. These early gains improve welfare for consumers. Thereafter the number of users is relatively flat. While not infinitely inelastic, demand does not respond much to the subsequent taxes. In particular, the number of users is quite insensitive to the dramatic escalation of taxes in the final three periods. The number of users in the optimal sequence is lower in the final period compared to baseline, but by an insufficient amount to cancel out the gains in the early periods. In particular, consumer surplus is 4.6 percent higher in the optimal case than in the baseline case.
Figures show paths of tax revenues and number of users for the time-invariant tax-free case (solid line) and optimal policy case (dashed line). Both panels depict outcomes when consumers have free disposal and where $\alpha = 2$. 