Static and Dynamic Taxation of Network Goods

Andrew Hanson Enda Patrick Hargaden Matthew Harris

November 2025

Abstract

We derive optimal tax rates for goods with network externalities. Efficiency considerations are subtle in this setting because the externality directly shapes consumers' willingness-to-pay. In the baseline static case, we show the optimal linear consumption tax for a revenue-constrained government is additive and separable. We extend this to a dynamic setting where the network evolves as a stock. The optimal policy now balances current and future network externalities, with Metcalfe's Law emerging as a special case. We show that optimal taxation may require early subsidization when the network is small, followed by higher taxes as adoption grows.

JEL Codes: H21, H23, D85.
Keywords: Optimal taxation, network goods, dynamic taxation.

The authors are grateful for comments and suggestions from David Agrawal, Johannes Becker, Kat Bilicka, Don Bruce, Ron Davies, Alan de Bromhead, Naomi Feldman, Bill Fox, Jim Hines, Olof Johansson-Stenman, Wojciech Kopczuk, Jacob Lundberg, Ben Lockwood, David Madden, Ben Meiselman, Matt Murray, Bill Neilson, Cormac O'Dea, Daniel Reck, Nathan Seegert, Charles Sims, Joel Slemrod, Nora Strecker, and conference and seminar participants at the Central Bank of Ireland, the European Economic Association, the IIPF Annual Congress, the Irish Economic Association, the Irish Public Economics Workshop, UCD, and Umeå University. Hargaden: UCD School of Economics, Dublin, Ireland; enda.hargaden1@ucd.ie; +353 1 716 7777. Hanson: Department of Economics, University of Tennessee, Knoxville. Harris: Boyd Center for Business and Economic Research & Department of Economics, University of Tennessee, Knoxville.

1 Introduction

Network goods are goods or services for which the value for each user increases with the total number of users. Each new user therefore generates network externalities for other users. Examples include the telephone, the fax machine, social media platforms, digital currencies, autonomous vehicles, and buyer-seller marketplaces such as eBay and Airbnb. Growth rates in these markets have greatly outpaced that of traditional goods in recent decades: as of 2025, six of the ten largest U.S. companies by market capitalization sell products that yield substantial network externalities.[1] Network externalities of this sort are also found throughout the economy in less obvious places, such as eating in a restaurant during a pandemic.[2]

This paper studies the optimal taxation of such network goods. We rule out the use of lump-sum taxes to focus on the more policy-relevant second-best policies. We examine this problem through three lenses: a static model in a very general form; a two-period model with a specific utility form (CES); and, to allow us to incorporate additional features like consumer heterogeneity and market power, a multi-period quantitative model.

The first model examines the taxation of network goods through a static lens in the tradition of Sandmo (1975). We generalize Sandmo by including the number of network good users in the consumer demand function, and model a government levying a consumption tax to meet a revenue requirement. Even in static settings, network externalities affect optimal tax rates. For intuition, consider a network good which also generates a negative atmospheric externality affecting everyone, e.g. polarized political discourse (Allcott, Braghieri, Eichmeyer and Gentzkow, 2020). Any positive network externality tends to cancel out any negative atmospheric externality, and the optimal tax rate can be positive or negative depending on the relative magnitudes of these effects. This is the first contribution of our paper.

To analyze the dynamic relationship between consumption taxation and network goods, we also develop a two-period model. We treat the number of users as a stock variable, flow utility as directly proportional to that stock, and allow tax policy to affect the growth rate of that stock. The government commits at the beginning to a sequence of tax rates, and consumers optimize given this information (e.g., prices in the final period will affect consumption in the first period). We derive closed-form solutions for the optimal tax rates in this setting. We find that the government should in general set a schedule of time-varying tax rates for network goods. The reasoning follows from the dynamic pricing literature in industrial organization (e.g. Katz and Shapiro, 1985; Bhattacharya and Vogt, 2004; Fainmesser and Galeotti, 2016). The IO literature establishes entry pricing as profit-maximizing for network good firms, paralleling infant-industry arguments of e.g. Criscuolo, Martin, Overman and Van Reenen (2019). We characterize when the same principle applies for a revenue-raising government. Cultivating network growth can effectively lock consumers into

---

[1] Alphabet (Google), Amazon, Apple, Meta (Facebook), Microsoft, and Tesla. An argument could be made for the inclusion of Broadcom and Nvidia in this list, making it eight from ten.

[2] Bars and restaurants can act as a vector for respiratory viruses such as COVID-19, and the prevalence or stock of infections likely also influences the demand for indoor dining.

the market, reducing their sensitivity to subsequent tax increases, and this intertemporal effect on the tax base must be considered at the optimum. We characterize sufficient conditions for when tax rates should optimally vary through time, and document cases where dynamic consumption taxation is not welfare-improving.

Traditional optimal consumption tax models examine the tradeoffs between static efficiency costs and contemporaneous budget requirements. We consider the relationship between consumption taxes and *future* elasticities, i.e. an intertemporal fiscal externality. Network effects are an ideal setting to explore the dynamic effects of consumption taxation because they have strong intertemporal externalities: more users today lead to more users and value tomorrow. In this setting, governments may be willing to forgo some efficiency or revenue today to reduce price sensitivity tomorrow, thereby lowering the efficiency costs of future tax policy. The magnitude of these tradeoffs is captured by an element we coin the ”installed tax-base wedge”, which can be cleanly and analytically separated from traditional Ramsey and Pigouvian elements.

For robustness, we further explore the intertemporal externality with a quantitative model. The quantitative model simulates the purchasing decision for a durable network good for $n = 10,000$ individuals over six periods, and can account for different market structures. At any point, consumers are free to discard the good, limiting the government’s ability to extract rents from commitments made under lower tax rates.[3] Mechanically, the government conducts a grid search over a discrete sequence of tax rates to maximize total surplus subject to its budget requirement.[4]

The results from the quantitative model demonstrate that dynamic effects can be large in magnitude and are robust to important model specification changes. We highlight four results. First, stronger positive network externalities increase the gains from subsidizing the network good in early periods and subsequently raising taxes. This result holds even with a standard Constant Elasticity of Substitution (CES) utility function.[5] Second, when a good yields both positive network effects and negative atmospheric externalities, subsidizing the good in early periods to increase adoption can increase overall surplus; however, for the policy to be Pareto-improving, there must be higher compensatory transfers to non-users in later periods. Third, if negative atmospheric externalities are sufficiently strong, the optimal response is to heavily tax from the initial period onward, despite the potential long-run revenue gains to the government. Finally, we show, with some caveats, that time-varying taxation of network goods can be Pareto-improving when the industry structure is either perfectly competitive or when the firm can set prices as they would in a monopolistic setting.

In the United States, sales taxes raise over \$500bn of revenue every year, and more than €1,000bn is raised through the VAT in Europe. It seems likely that network goods will grow as a share of consumption in the coming years, will be a source of tax revenue in the future, and this

---

[3] Time inconsistency is a particular problem in the capital taxation literature, where governments retain the option of a once-off 100% tax rate. In the consumption tax setting, if consumers are locked into contracts or are otherwise constrained in their choices, the government may have a similar incentive to abruptly increase tax rates. Incorporating free disposal is a realistic way to add discipline to the model.

[4] The grid search approach has the added benefit of ensuring we locate discrete-grid approximations of *global* maxima.

[5] In other words, our model has constant intratemporal elasticities, but variable intertemporal elasticities.

paper attempts to model that optimal tax formally. In this sense, the paper is most closely related to Sandmo (1975), Kopczuk (2003), Eckerstorfer and Wendner (2013), Aronsson and Johansson-Stenman (2018), and others, taking an optimal taxation approach to externalities. Separately, a small literature exists on excess inertia for clean technology. For example, Greaker and Midttømme (2016) examine the role of taxes in encouraging adoption rates of new goods and Kind, Koethenbuerger and Schjelderup (2008) focuses on ensuring optimality in two-sided markets; neither quite reflect the optimal consumption tax/revenue-raising question of this paper. Additionally, our paper can be seen as in the same vein as the literature on dynamic taxation; e.g., Barrage (2020) and Akcigit, Hanley and Stantcheva (2022). The paper also addresses themes found in the optimal taxation literature incorporating "modern" realities like superstars (Scheuer and Werning, 2017) and robots/AI (Guerreiro, Rebelo and Teles, 2022; Thuemmel, 2022; Bastani and Waldenström, 2024; Jacquet and Lehmann, 2025). Finally, the topic of our paper is related to digital service taxation (Borders, Balladares, Barake and Baselgia, 2023). Digital Service Taxes have now been introduced in more than a dozen countries (KPMG, 2025), and elements of this paper may be useful to that emerging literature (e.g. Cui, 2019; Agrawal and Fox, 2021).

2 Closed-form Theory

2.1 Static 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 similar to Kopczuk (2003) and Micheletto (2008). We model a utilitarian planner maximizing the sum of utilities for 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 α, which for simplicity we will think of as total consumption of x_m. The consumer's problem is thus to maximize

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

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 α as u_{m+1}. For a negative externality, such as if autonomous vehicles decreased safety, $u_{m+1} < 0$. Consumers do not consider their own effect on the externality, and we assume that the usual conditions for an interior maximum hold.

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

$$V(P) = u\left[1 - x_0(P), x_1(P), \ldots, x_m\left(P, \alpha(P)\right), \alpha(P)\right] \tag{2}$$

Note that this formulation permits both that demand for x_m be a function of α (the network effect), and that α directly affects utility with α=0 a special case. The welfare effect of adjusting the price of good k is:

$$\frac{\partial V(P)}{\partial P_k} = -u_0 \frac{\partial x_0}{\partial P_k} + \sum_{i=1}^{m} u_i \frac{\partial x_i}{\partial P_k} + u_m \left(\frac{\partial \alpha}{\partial P_k}\frac{\partial x_m}{\partial \alpha}\right) + u_{m+1}\frac{\partial \alpha}{\partial P_k} \tag{3}$$

Using the fact (from the consumer budget constraint) that $x_k = \frac{\partial x_0}{\partial P_k} - \sum_{i=1}^{m} P_i \frac{\partial x_i}{\partial P_k}$, and substituting in the FOCs for the consumer problem, we conclude that:

$$\frac{\partial V(P)}{\partial P_k} = -\lambda x_k + u_m \left(\frac{\partial \alpha}{\partial P_k}\frac{\partial x_m}{\partial \alpha}\right) + u_{m+1}\frac{\partial \alpha}{\partial P_k} \tag{4}$$

In all sections of this paper, the government's objective is to maximize utility (indirect or direct) subject to some revenue requirement. Here, we 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. We recognize this is a strong assumption, and one which we will relax when we give firms price-setting power in Section 3 . However, we maintain the zero profits assumption to avoid discussions of a profit tax in this section and to facilitate comparison with canonical results.

Under these conditions, the government maximization problem can be summarized as

$$\mathscr{L} = nV(P) - \beta\left[n\sum_{i=1}^{m}(P_i - p_i)x_i - T\right] \tag{5}$$

Using (4), we can see that a necessary condition for the optimal commodity tax rate is:

$$\frac{\partial \mathscr{L}}{\partial P_k} = -\lambda x_k + u_m \left(\frac{\partial \alpha}{\partial P_k}\frac{\partial x_m}{\partial \alpha}\right) + u_{m+1}\frac{\partial \alpha}{\partial P_k} - \beta\left[\sum_{i=1}^{m} t_i \frac{\partial x_i}{\partial P_k} + x_k\right] = 0 \tag{6}$$

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

$$\sum_{i=1}^{m} t_i \frac{\partial x_i}{\partial P_k} = -\left(\frac{\lambda + \beta}{\beta}\right) x_k + \frac{n}{\beta}\left(u_{m+1} + u_m \frac{\partial x_m}{\partial \alpha}\right)\frac{\partial x_m}{\partial P_k} \tag{7}$$

Let the coefficient matrix on t_i (the transpose of the Jacobian of the taxable goods' demand functions) be denoted $J^\star$. Further let $J \equiv det(J^\star)$ and denote J_{ik} as the cofactor of the element in row i, column

j of J. Then, applying Cramer's Rule:

$$t_k = \left[-\left(\frac{\lambda+\beta}{\beta}\right)\right]\left(\frac{\sum_{i=1}^{m} x_i J_{ik}}{J}\right) + \frac{n}{\beta}\left(u_{m+1} + u_m \frac{\partial x_m}{\partial \alpha}\right)\left(\frac{\sum_{i=1}^{m} \frac{\partial x_m}{\partial P_i} J_{ik}}{J}\right) \tag{8}$$

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

$$\sum_{i=1}^{m} \frac{\partial x_m}{\partial P_i} J_{ik} = \begin{cases} 0 & \text{for } k \neq m \\ J & \text{for } k = m \end{cases} \tag{9}$$

Consequently,

$$t_k = -\left(\frac{\lambda+\beta}{\beta}\right)\left(\frac{\sum_{i=1}^{m} x_i J_{ik}}{J}\right) + \frac{n}{\beta}\left(u_{m+1} + u_m \frac{\partial x_m}{\partial \alpha}\right) \times \mathbb{1}_{\{k=m\}} \tag{10}$$

$$\frac{t_k}{P_k} = \left(\frac{-1}{P_k}\right)\left(\frac{\lambda}{\beta}+1\right)\left(\frac{\sum_{i=1}^{m} x_i J_{ik}}{J}\right) + \frac{n}{\beta}\frac{\lambda}{\lambda}\frac{1}{P_k}\left(u_{m+1} + u_m \frac{\partial x_m}{\partial \alpha}\right) \times \mathbb{1}_{\{k=m\}} \tag{11}$$

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

$$\theta_k = \left(\frac{-1}{P_k}\right)(1-\mu)\left(\frac{\sum_{i=1}^{m} x_i J_{ik}}{J}\right) - n\mu\left(\frac{1}{\lambda P_m}\right)\left(u_{m+1} + u_m \frac{\partial x_m}{\partial \alpha}\right) \times \mathbb{1}_{\{k=m\}} \tag{12}$$

Finally, substituting from the consumer FOC and rearranging, we have:

$$\theta_k = (1-\mu)\left[\frac{-1}{P_k}\frac{\sum_{i=1}^{m} x_i J_{ik}}{J}\right], k = 1, \ldots, (m-1) \tag{13}$$

$$\theta_m = (1-\mu)\left[\frac{-1}{P_m}\frac{\sum_{i=1}^{m} x_i J_{im}}{J}\right] - \mu\left[n\left(\frac{u_{m+1}}{u_m}\right)\right] - \mu\left[n\left(\frac{\partial x_m}{\partial \alpha}\right)\right] \tag{14}$$

This solves for the optimal tax rates. Equation (13) shows that the tax rate on the $m-1$ typical goods is a form of the Ramsey discouragement index which decreases in the sensitivity/elasticity of the consumers to price. The discouragement index is scaled by $1-\mu$, where $-\mu$ is the ratio of the Lagrangian multipliers.

Equation (14) defines the optimal rate for the m^{th} good. It shows that the tax comprises three additively separable components: the first, the Ramsey-like discouragement factor; the second, a Pigouvian factor increasing in the magnitude of the direct atmospheric externality; and the third, an adjustment for the network externality/how consumption responds to the externality.

The departure of Equations (13) and (14) with previous research is the third 'consumption response' component. If consumption does not depend on the externality, e.g. when the demand for widgets is 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 (14) together:

$$\theta_m = (1-\mu)\left[\frac{-1}{P_m}\frac{\sum_{i=1}^{m} x_i J_{im}}{J}\right] + \mu\left[-n\left(\frac{u_{m+1}}{u_m} + \frac{\partial x_m}{\partial \alpha}\right)\right] \tag{15}$$

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

This result shows that while the optimal taxation of network externalities is more complex than the existing literature, it can be seen as a generalization that retains an additive and separable form. Further, the optimal tax rule retains intuitive features, notably that the Pigouvian component is adjusted to account for the effect of the network externality. Consider a network good that generates a negative atmospheric externality like higher crime. Through its positive effects on demand, the network good has a lower Pigouvian correction than suggested by its negative atmospheric effects.

This model is quite general, but a clear shortcoming is that many network externalities are best considered as stock variables. If tax rates affect adoption rates, they will also alter the size of the network externality in subsequent periods. In this setting, the government should consider trading-off revenue not just with a contemporaneous increases in the network externality but all future flows of the increased stock. The government's decision thus becomes a dynamic problem.

Introducing dynamic considerations also means that optimal taxation of network goods is no longer as simple as an additively-separable adjustment. The second contribution of this paper is a model of optimal network taxation in a two-period setting.

2.2 Two-Period Model

In this section we provide a theoretical framework for optimal taxation in a two-period model with a durable consumption good featuring network externalities. We provide a number of conditions under which a time-varying tax sequence, "dynamic taxation", is welfare improving. We find that the presence of network externalities themselves is not a sufficient condition for dynamic taxation to be welfare improving. The strength of these externalities, cross-period demand elasticities, and intertemporal substitution also govern whether dynamic taxation is welfare improving.

Households Consider the optimization problem of a representative household which lives for two periods. As this is a more involved model than Section 2.1, some notational changes are necessary. In each period the household receives an exogenous endowment of real goods labeled y_1 and y_2 respectively. The household may then choose to purchase the network good (q_1 and q_2) which delivers flow services c_t or the outside option consumption good (z_1 and z_2), subject to budget constraints. To keep things tractable, we focus attention on a CES formulation with time discounting β. The network good is durable with depreciation rate δ and taxed on the consumption of flow services at per-unit net rates τ_1 and τ_2. We assume the households do not internalize (yet do benefit from) the consumption externalities afforded by the network good. The function $f(c, X)$ thus aggregates total consumption from the network good and its externalities; the latter of which is captured by X, which could be considered the aggregate consumption of the network good across all households. Households take X_t as given.

The household problem is thus,

$$\max_{z_1,z_2,q_1,q_2} \left[\theta z_1^{\rho} + (1-\theta) f(c_1, X_1)^{\rho}\right]^{1/\rho} + \beta \left[\theta z_2^{\rho} + (1-\theta) f(c_2, X_2)^{\rho}\right]^{1/\rho}$$

$$\begin{aligned}
&\text{subject to} \\
&c_1 = q_1 \\
&c_2 = (1-\delta) c_1 + q_2 \\
&y_1 \geq z_1 + q_1 + \tau_1 c_1 \\
&y_2 \geq z_2 + q_2 + \tau_2 c_2
\end{aligned}$$

The solution to the household problem are the following necessary first order conditions with respect to $\{z_1, z_2, q_1, q_2\}$. These combine into the following Euler equations:

$$\begin{aligned}
(1+\tau_1) &= \frac{1-\theta}{\theta}\left(\frac{z_1}{f(c_1, X_1)}\right)^{1-\rho} f_c(c_1, X_1) + (1-\delta)\Lambda_{priv} \\
(1+\tau_2) &= \frac{1-\theta}{\theta}\left(\frac{z_2}{f(c_2, X_2)}\right)^{1-\rho} f_c(c_2, X_2)
\end{aligned}$$

Where $\Lambda_{priv} \equiv \beta\left(\frac{u(z_2,c_2)}{u(z_1,c_1)}\right)^{1-\rho}\left(\frac{z_2}{z_1}\right)^{\rho-1}$ is the intertemporal discount factor or the marginal rate of substitution between z_1 and z_2. If we define the marginal rate of substitution between the network good and the outside good for the private consumer as $MRS_{q,t} \equiv \frac{1-\theta}{\theta}\left(\frac{z_t}{f(c_t,X_t)}\right)^{1-\rho} f_1(c_t, X_t)$ then this can be compactly written as

$$\begin{aligned}
(1+\tau_1) &= MRS_{q,1} + (1-\delta)\Lambda_{priv} \\
(1+\tau_2) &= MRS_{q,2} \qquad (16)
\end{aligned}$$

For the remainder of this section we choose to note equilibrium objects as functions of investment q_t instead of the flow utility services c_t. This is simply notational preference, and the reader should keep in mind the two equations from the household problem — $c_1 = q_1$ and $c_2 = (1-\delta) c_1 + q_2$ — which linearly maps one set to the other.

Government Consider now the problem of a government seeking to raise revenue R by levying per-unit taxes on the network good's flow of services. The government wishes to do this in such a way that maximizes social welfare of the representative household subject to the revenue constraint.

Furthermore, we assume the government is aware of the externalities inherent in the network good, which leads it to optimize under the condition that for the representative agent $f(c, X)\mid_{X=c} \equiv g(c)$. We thus follow a Ramsey primal approach with the government choosing τ_1 and τ_2 subject to the revenue constraint, resource constraints, and household implementability constraints.

The government's problem is

$$
\begin{aligned}
\max_{q_1,q_2,\tau_1,\tau_2} & \left[\theta (z_1)^\rho + (1-\theta) g (q_1)^\rho\right]^{1/\rho} + \beta \left[\theta (z_2)^\rho + (1-\theta) g \left((1-\delta) q_1 + q_2\right)^\rho\right]^{1/\rho} \\
& \text{subject to} \\
R &\leq \tau_1 q_1 + \tau_2 \left[(1-\delta) q_1 + q_2\right] \\
y_1 &= z_1 + q_1 \\
y_2 &= z_2 + q_2 \\
1+\tau_1 &= MRS_{q,1} + (1-\delta)\Lambda_{priv} \equiv \chi_1(q_1, q_2) \\
1+\tau_2 &= MRS_{q,2} \equiv \chi_2(q_2)
\end{aligned}
$$

where the revenue constraint uses the two household investment identities to change variables from taxation on $\{c_1, c_2\}$ to taxing a function of $\{q_1, q_2\}$. Note that the two implementability conditions imply the tax rates are complicated functions of $\{q_1, q_2\}$. For expediency purposes we define these functions as $\chi_1(q_1, q_2)$ and $\chi_2(q_2)$ respectively. We simplify this problem to one in which the implementability constraints are internalized into the revenue generation function and the resource constraints hold with equality:

$$
\begin{aligned}
\max_{q_1,q_2} & \left[\theta (y_1 - q_1)^\rho + (1-\theta) g (q_1)^\rho\right]^{1/\rho} + \beta \left[\theta (y_2 - q_2)^\rho + (1-\theta) g \left((1-\delta) q_1 + q_2\right)^\rho\right]^{1/\rho} \\
& \text{subject to} \\
R &\leq \tilde{R}(q_1, q_2) \equiv \left[\chi_1(q_1, q_2) - 1\right] q_1 + \left[\chi_2(q_2) - 1\right]\left[(1-\delta) q_1 + q_2\right]
\end{aligned}
$$

Define the marginal cost of raising public funds in units of the foregone utility from the outside good as $MC_t \equiv \frac{\mu_R}{\theta u_g(z_t, c_t)^{1-\rho} z_t^{\rho-1}}$, where μ_R is the shadow price of raising an additional unit of revenue.[6] Then the Ramsey primal solution follows the two necessary first order conditions:

$$
\begin{aligned}
MRS_{q,1}^{soc} + (1-\delta)\Lambda_{soc} MRS_{q,2}^{soc} &= 1 - MC_1 \frac{\partial \tilde{R}}{\partial q_1} \\
MRS_{q,2}^{soc} &= 1 - MC_2 \frac{\partial \tilde{R}}{\partial q_2}
\end{aligned}
\tag{17}
$$

$\Lambda_{soc} \equiv \beta \left(\frac{u_g(z_2, c_2)}{u_g(z_1, c_1)}\right)^{1-\rho} \left(\frac{z_2}{z_1}\right)^{\rho-1}$ is the socially optimal intertemporal discount factor, where the g subscript is meant to convey the social planner's network-aware utility function. The partial derivative $\frac{\partial \tilde{R}}{\partial q_t}$ represents the marginal ability to raise revenue from taxing the network good in

---

[6] The Lagrange multiplier for the revenue-raising constraint.

period t.

Finally, $MRS^{soc}_{q,t}$ is the socially optimal marginal rate of substitution in period $t \in \{1, 2\}$ and is related to the private $MRS_{c,t}$ by a Pigouvian factor:

$$\begin{aligned} MRS^{soc}_{q,t} &= \frac{1-\theta}{\theta}\left(\frac{z_t}{g\left(c_t\right)}\right)^{1-\rho} g_1\left(c_t\right) \\ &= MRS_{q,t} \times \left(\frac{f\left(c_t, X_t\right)}{g\left(c_t\right)}\right)^{1-\rho} \frac{g_1\left(c_t\right)}{f_1\left(c_t, X_t\right)} \end{aligned}$$

This factor $H\left(c_t, X_t\right) \equiv \left(\frac{f\left(c_t, X_t\right)}{g\left(c_t\right)}\right)^{1-\rho} \frac{g_1\left(c_t\right)}{f_1\left(c_t, X_t\right)}$ represents the wedge between the social valuation and the private valuation of the network good. Thus a Pigouvian tax should be set such that, in equilibrium, $H\left(c_t, X_t\right) = 1$. In which case the private and social valuations would be equal.

Equilibrium We now solve for the equilibrium optimal tax rates.

Definition 1. A Ramsey equilibrium, given taxes $\{\tau_1, \tau_2\}$, is a set of allocations $\{c_1, X_1, q_1, c_2, X_2, q_2$ such that:

- Taking taxes and quantities $\{X_1, X_2\}$ as given, households maximize utility given budget constraints
- Taking household optimality conditions as given, the Ramsey planner maximizes household utility using $g\left(c\right) = f\left(c, X\right)|_{X=c}$ subject to the revenue constraint
- Goods markets clear in each period
- $c_t = X_t$

This implies the following about the Pigouvian factor which we label $H\left(c_t\right)$:

$$\begin{aligned} MRS^{soc}_{q,t} &= MRS_{q,t} \times H\left(c_t, X_t\right)|_{X_t=c_t} \\ H\left(c_t\right) &\equiv 1 + \frac{f_X\left(c_t, X_t\right)}{f_c\left(c_t, X_t\right)}|_{X_t=c_t} \end{aligned} \tag{18}$$

This maps cleanly into our understanding of Pigouvian taxation. Under positive network externalities $f_X > 0$ and thus $H\left(c_t\right) > 1$ which implies the social valuation is higher than the private valuation. The opposite is true under negative externalities. Combining equation (18) with those in (17) yield the following optimality conditions in equilibrium:

$$\begin{aligned} MRS_{q,1}H\left(c_1\right) + \left(1-\delta\right)\Lambda_{soc}MRS_{q,2}H\left(c_2\right) &= 1 - MC_1\frac{\partial\tilde{R}}{\partial q_1} \\ MRS_{q,2}H\left(c_2\right) &= 1 - MC_2\frac{\partial\tilde{R}}{\partial q_2} \end{aligned} \tag{19}$$

Furthermore, recall the implementability constraints in equation (16) which relate the private $MRS_{q,t}$ to the optimal tax rates and the private intertemporal discount factor. Inserting these implementability constraints into equation (19) leads to the following proposition:

Proposition 1. *The optimal tax rates in the presence of a durable good which offers utility from flow services and network externalities are given by the Ramsey primal solution:*

$$1+\tau_1^* = \frac{1}{H(c_1)} - \underbrace{(1-\delta)\,MRS_{q,2}\left[\Lambda_{soc}\frac{H(c_2)}{H(c_1)} - \Lambda_{priv}\right]}_{\textit{installed tax-base wedge}} - \frac{MC_1}{H(c_1)}\frac{\partial\tilde{R}}{\partial q_1}$$

$$1+\tau_2^* = \frac{1}{H(c_2)} - \frac{MC_2}{H(c_2)}\frac{\partial\tilde{R}}{\partial q_2}$$

Each of the optimal tax rates contain a Pigouvian term related to the strength of the network externalities $1/H(c_t)$. When the marginal network externality at a given level of consumption is strong relative to the private valuation ($f_2(c_t, c_t) > f_1(c_t, c_t)$) the fiscal authority would optimally choose to reduce taxes in that period. Furthermore, each of the optimal tax rates has a revenue-based wedge in $\frac{MC_t}{H(c_t)}\frac{\partial\tilde{R}}{\partial q_t}$. The partial derivatives encode how effective taxing one more unit of the network good would be at raising revenue. This is valued in units of foregone utility of outside option consumption via the private shadow price MC_t. Multiplication by the Pigouvian term $1/H(c_t)$ transforms this private shadow price valuation into the socially optimal valuation.

The unique term in the first period's optimal tax rate governs the planner's use of relative tax or subsidy to manipulate the installed tax base of the network good. We have called this term the "installed tax-base wedge" for a number of reasons. The term in brackets represents a wedge in the way that the planner values intertemporal transfer of goods relative to the private valuation. Recall that Λ_{priv} represents the relative valuation of a unit of second period consumption to a unit of first period consumption for the private agent. Λ_{soc} represents the socially optimal relative valuation. However it must be adjusted by the ratio of Pigouvian terms to account for network externality differences across time. This relative intertemporal valuation is amplified or diminished by the intratemporal valuation of the network good through $MRS_{q,2}$ and discounted by the depreciation rate. In other words, the planner would choose to optimally reduce the tax in period one relative to period two in one or more of the following three situations:

1. Public intertemporal valuation is much stronger than private intertemporal valuation ($\Lambda_{soc}\frac{H(c_2)}{H(c_1)} > \Lambda_{priv}$). This is related to a notion of dynamic inefficiency in the competitive equilibrium. Private agents undervalue the benefits of "saving" in the network good, even given Pigouvian correction for the network externalities.

2. Private agents are network-good-poor in the second period relative to the outside good (large $MRS_{q,2}$), so consumption smoothing would prefer to defer consumption in the first period.

3. The "savings" technology of the network good is strong so depreciation is low. Thus

investment in the network good in the first period yields higher effective income in the second period.

An alternative method for understanding the installed tax-base wedge term is to consider the following comparative static thought experiment. Consider raising the tax rate in the first period by $\Delta\tau_1 \to 0$ such that relative allocations to z_1 and c_1 are unaffected. In otherwords we compare two hypothetical budget constraints for the household:

$$\begin{aligned} y_1 - z_1 &= (1+\tau_1)q_1 \\ y_1 - z_1 &= (1+\tau_1+\Delta\tau_1)q_1' \end{aligned}$$

This means that the affordable quantity of the network good has fallen by $q_1 \mapsto q_1' = \frac{1+\tau_1}{1+\tau_1+\Delta\tau_1}q_1$. This loss is valued in private utility from the second period–due to a lower initial stock–at $MRS_{q,2}$ discounted by the durable portion that would have been usable $1-\delta$. In private present-value terms that is valued at Λ_{priv}. If the social planner values that incremental loss more $\Lambda_{soc}\frac{H(c_2)}{H(c_1)} > \Lambda_{priv}$, then the planner would choose to lower the first period tax.

Although the depreciation and $MRS_{q,2}$ terms are important, the critical element is the dynamic term; the installed tax-base wedge. Ultimately it is this term which deviates from traditional Pigouvian and Ramsey taxation concerns. The following proposition highlights the fact that the taxes need not in general follow the constant commodity taxation Ramsey benchmark (even after Pigouvian correction):

Proposition 2. *The difference in per-period optimal taxes contains three constituent pieces:*

$$\tau_2^* - \tau_1^* = \left[\frac{1}{H(c_2)} - \frac{1}{H(c_1)}\right] + (1-\delta)\, MRS_{c,2}\left[\Lambda_{soc}\frac{H(c_2)}{H(c_1)} - \Lambda_{priv}\right] + \left[\frac{MC_1}{H(c_1)}\frac{\partial\tilde{R}}{\partial q_1} - \frac{MC_2}{H(c_2)}\frac{\partial\tilde{R}}{\partial q_2}\right]$$

which represent, roughly:

1. *Pigouvian differential*
2. *Installed tax-base wedge*
3. *Marginal cost of public funds differential*

We can understand the final two terms from the perspective of the benefit a higher installed tax base in the second period would bring. As already discussed, the planner would wish to reduce τ_1 relative to τ_2 if the network externality benefits are large in the second period. Furthermore, if raising revenue is marginally more costly in the first period, the planner would choose to more heavily tax in the second period to meet revenue requirements. In contrast, the first term describes an extra Pigouvian incentive to build the network early. If network effects are relatively strong in the first period, the planner would prefer to tax the second period more heavily and encourage early adoption.

Specific Example To visualize the optimality of dynamic taxation in this two period model, we make a set of modeling assumptions. In particular, assume that the externalities take the following form:

$$f\left(c, X\right) = c + X^{p}$$
$$g\left(c\right) = c + c^{p}$$

where we assume the parameter p is constant across time and represents the possibility of increasing returns to scale from the network good. The case of $p > 1$ is discussed in Theorem 3. *Metcalfe's Law* that the value of a network is proportional to the square of its size can be considered if we set $p = 2$. More generally, a natural question is when would the social planner, despite needing to raise revenue R, optimally choose to incentivize network build up in the first period with relatively lower taxes (or even subsidy)? The following highlights that the depreciation of the privately-held component of the network plays a critical role.

Assumption 1 (Feasibility, compactness)**.** Define the feasible set $\Gamma(\delta) := \{q \in [0, y_1] \times [0, y_2] : \tilde{R}(q, \delta) \geq R\}$ and the Ramsey argmax correspondence

$$\mathcal{A}(\delta) := \arg\max_{q \in \Gamma(\delta)} W(q, \delta), \qquad W(q, \delta) := u_{1,g}(c_1, z_1) + \beta\, u_{2,g}(c_2, z_2),$$

with $u_{t,g}(\cdot)$ computed using $g(\cdot) = c + c^{p}$.[7] Assume $\Gamma(\delta) \neq \emptyset$ in a neighborhood of $\delta = 1$ (e.g., R small enough).

Theorem 3 (Gross flow-tax ratio > 1 for high depreciation)**.** *Under Assumption 1, suppose $p > 1$ and, at $\delta = 1$, every Ramsey optimizer satisfies $c_2 < c_1$ (e.g., $\beta < 1$ with symmetric y_t and R small, or $y_2 < y_1$ with $\beta = 1$ and R small). Then there exists $\overline{\delta} \in (0, 1)$ such that for all $\delta \in (\overline{\delta}, 1)$ and all $q \in \mathcal{A}(\delta)$,*

$$\frac{1 + \tau_2}{1 + \tau_1} = \frac{\chi_2}{\chi_1} > 1.$$

Proof. See Appendix □

Theorem 3 proves that, at least in a local sense for revenue attainment, it is optimal for the Ramsey planner to set taxes low in the first period to build up the network base and then to raise taxes in the second period to meet revenue requirements. An immediate sub-result from this theorem is that the revenue-neutral optimum under symmetric endowments is the subsidize-then-tax strategy. Theorem 4 makes this more precise.

Theorem 4 (Revenue-neutral improvement: subsidize–then–tax under flow taxes)**.** *Assume $\delta = 1$ (no carryover), CES preferences with $\rho \in (0, 1)$, $\theta \in (0, 1)$, and externality aggregator $f(c, X) = c + X^{p}$ with $p \geq 0$. The social aggregator is $g(c) = c + c^{p}$, and in the representative-agent equilibrium $X_t = c_t$.*

---

[7] Because $X_t = c_t$ in equilibrium and $f(c, X) = c + X^{p}$, we have $f(\cdot) = g(\cdot)$ pointwise; we keep the "g" to stress social evaluation.

Let endowments be (y_1, y_2)*, and let* $\beta \in (0, 1]$ *be the planner's discount factor. Consider flow taxes* τ_t *that enter the period-t budget as* $z_t + (1 + \tau_t)c_t \leq y_t$ *and government revenue* $R(\tau) = \tau_1 c_1(\tau_1) + \tau_2 c_2(\tau_2)$.

Let (z_t^0, c_t^0) *denote the (unique) undistorted allocation at* $\tau \equiv 0$*, and define*

$$\lambda_t^0 := \frac{\partial u_t^{\text{priv}}}{\partial z}\Big|_{(z_t^0, c_t^0)}, \qquad \phi_1 := \lambda_1^0, \qquad \phi_2 := \beta\, \lambda_2^0.$$

Suppose $\phi_2 < \phi_1$*. Then there exists* $\varepsilon > 0$ *such that for any small* $|\eta| \in (0, \varepsilon)$*, the* revenue-neutral *perturbation*

$$d\tau_1 = -\eta, \qquad d\tau_2 = \eta\, \frac{c_1^0}{c_2^0}$$

satisfies $dR = 0$ *to first order and generates a strictly positive first-order welfare change* $dW > 0$*. Hence* $(\tau_1, \tau_2) = (0, 0)$ *is not locally optimal under the revenue-neutral constraint* $R(\tau) = 0$.

In particular, with symmetric primitives and endowments $(y_1 = y_2)$*, we have* $\phi_2 < \phi_1$ *whenever* $\beta < 1$*; and with* $\beta = 1$*, we have* $\phi_2 < \phi_1$ *whenever* $y_2 > y_1$.

Proof. See Appendix B. □

The combination of Theorems 3 and 4 assert that time-varying taxation is optimal in a revenue-neutral situation not just for a perfectly depreciating network good, but allows for some degree of durability $\delta \in (\bar{\delta}, 1)$. In a general revenue-raising case this is not a trivial result, because the Pigouvian term would prefer to subsidize in the second period due to the increasing returns from the network when the base is larger. In general this term strictly dominates the other two concerns; hence the necessary assumption that $c_2 < c_1$ holds at least at $\delta = 1$. We hypothesize, moving into the more general multi-period model of the following section, that this Pigouvian domination erodes as the length of time increases. Given a long enough horizon to present-discount the gains from the installed tax-base wedge, it is possible that the assumptions required for optimally increasing tax rates may weaken. However, an analytic examination of this hypothesis is intractable.

The durability of the network good plays a crucial role in determining whether this is the optimal strategy in general. The threshold value $\bar{\delta}$ is in general non-negligible. For instance, consider the following parameterization: $(\beta, \theta, \rho, y_1, y_2, R) = (0.99, 0.5, 0.9, 4, 4, 0.04)$. In other words, the two goods are close substitutes. The revenue requirements represent just one percent of "annual income" (per period endowment). Figure 1 shows the threshold value as a function of the externality exponent p. For the current parameterization this threshold hits a minimum value around $p = 2.75$, but still requires high depreciation $\delta \approx 0.59$ for increasing taxes to be welfare-improving. Clearly the Pigouvian concerns from network effects are crucial.

We can more concisely rationalize the importance of the depreciation term through the lens of the implementability constraints. Recall that the following must hold in the Ramsey equilibrium (equation 16):

$$\frac{1 + \tau_2}{1 + \tau_1} = \frac{MRS_{q,2}}{MRS_{q,1} + (1 - \delta)\, \Lambda_{priv}}$$

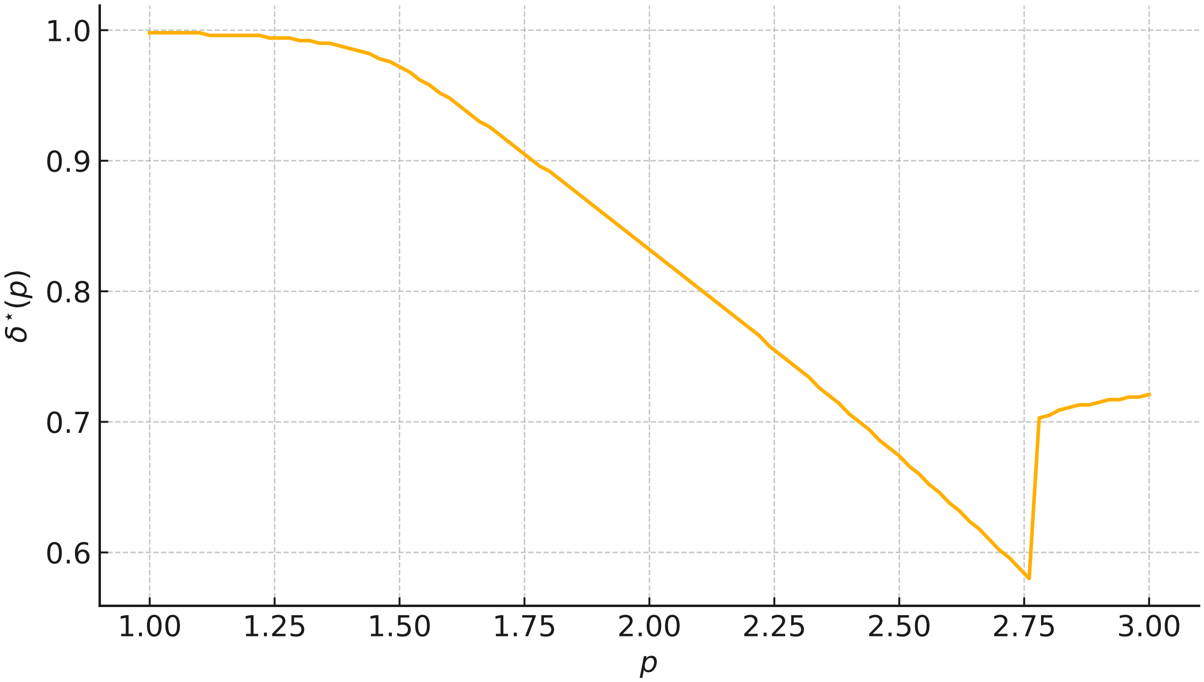


Figure 1: Threshold depreciation $\delta^\star(p)$ over $p \in [1,3]$ under the equality-constrained Ramsey problem with flow taxes. The two-period model features a durable network good with private aggregator $f(c,X) = c + X^p$ and planner aggregator $g(c) = c + c^p$, taking $X = c$ in equilibrium and a common exponent p across periods. Calibration: $(\beta, \theta, \rho, y_1, y_2, R) = (0.99, 0.50, 0.90, 4, 4, 0.04)$. For each p, we solve the welfare maximum subject to the exact revenue requirement via equality-constrained SLSQP with box constraints $0 \leq q_t \leq y_t$ (implying $z_t, c_t \geq 0$). The curve reports the smallest δ for which the gross flow-tax ratio satisfies $(1+\tau_2)/(1+\tau_1) > 1$ at the optimum. Algorithm: coarse search over $\delta \in [0, 0.99]$ (step 0.02) with neighbor warm starts, followed by local refinement (step 0.001) around the first hit; p is sampled in steps of 0.02. Points are accepted only if the revenue slack is $< 10^{-6}$ and the projected KKT stationarity norm is $< 10^{-4}$.

Consider the following comparative static thought experiment: hold allocations fixed and perturb δ slightly. Surely the numerator must increase given $\frac{\partial MRS_{q,2}}{\partial \delta} = \frac{\partial MRS_{q,2}}{\partial c_2}\frac{\partial c_2}{\partial \delta}$ and $\frac{\partial c_2}{\partial \delta} = -q_1$ by the budget constraint. CES utility preserves the diminishing marginal rate of substitution property. In other words, all else equal, increased depreciation makes the network good more scarce in the future so that households are willing to pay marginally higher taxes to obtain it. Again, holding allocations fixed, the denominator must fall given $\frac{\partial}{\partial \delta}\left[MRS_{q,1} + (1-\delta)\,\Lambda\right] = -\Lambda$. In this case a network good purchased in the present is worth less due to lower continuation value.

Certainly this thought experiment is incomplete, because households would rebalance their allocations thereby changing both marginal rates of substitutions and the intertemporal discount factor in complicated ways. Therefore, Theorem 4 provides sufficient conditions under which these reallocation effects do not dominate the aforementioned direct effects of raising δ. In order to account for all of the endogenous changes we turn instead to a quantitative experiment in the following section.

Remark 1. Although not a direct focus of the paper, congestion effects (negative network externalities) switch the signs of the Pigouvian and installed tax base wedge. Therefore in our analytic two-period model the restrictions for an optimally increasing tax schedule are relaxed; in particular the depreciation threshold is much lower.

3 Multi-period Quantitative Model

The previous section shows that intertemporal variation in tax rates can increase both consumer surplus and government revenue when network externalities are present.

However, the derivation of *optimal* time paths of taxation at is mathematically infeasible at the level of generality of Sandmo (1975). In this section, we therefore employ Monte Carlo simulations to further explore how the presence and strength of network and atmospheric externalities affect how intertemporal variation in tax rates can affect consumer surplus and government revenue. We compare the consumer surplus and government revenues to those of a benchmark case of a constant tax rate. The goal is to identify the sequences of taxes that maximize consumer surplus while remaining at least revenue neutral over the considered period. Alternatively, we empirically show that under the assumptions of the theoretical models (and conditional on certain parameter values) static tax sequences are not Pareto optimal when taxing network goods.

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. We treat income as exogenous 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).[8] 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 insights on how network externalities affect optimal taxation.

Acknowledging those adjustments, we preserve the important foundational assumptions of the static model in Sections 2.1 and 2.2. Specifically, we consider consumers' decision to allocate their resources between one good that yields network 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 the effects of their own actions on either externality.

Following Goyal (2012) and Jackson and Watts (2002), consumers' expectation of state variables (such as the size of the network) in time $t+1$ is their contemporaneous value, i.e. $\mathbb{E}\left[x_{t+1}\right] = x_t$. This is a standard assumption in the economics of networks, closely resembles our two-period model, and is reasonable as the distance between periods decrease. Mechanically, the government conducts a grid search over sequences of tax rates to maximize total surplus subject to satisfying a budget requirement.

---

[8] One avenue for future research is how mixed income-commodity taxation would respond if the network good is complementary/substitutable for labor.

We simulate purchasing decisions for a durable good for $n = 10,000$ individuals over six periods.[9] 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 τ_t. Consistent with Section 2.2, we select a CES functional form for the individuals' utility function. The individual's utility from purchasing the good can be expressed as:

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

The first term describes the utility from owning the externality-producing good, where γ is the private flow utility from owning the good (i.e., the utility from owning the good even if no one else does), s_t is the share of the population who owns the good, α is the parameter that captures how much the purchase decisions of $j \in -i$ affect person i's utility of consuming the good. If $\alpha = 0$, this is a regular private good. The second term captures the utility of consuming the numeraire good (which by default is income net of expenditures on the externality-inducing good as we are not modeling savings). To the extent that the durable good of interest yields atmospheric externalities, those effects are parameterized by δ. Substitution preferences are measured by r, income by y, purchase price of the good by p, and the tax rate by τ. Finally, ϵ_{it1} is an idiosyncratic preference shock.

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:

$$U_{it0} = ((y + \delta \cdot s_t)^r)^{1/r} + \epsilon_{it0} = y + \delta \cdot s_t + \epsilon_{it0}$$

which is derived from the expression above where the individual gets none of the utility in the first subset of parentheses, but also pays neither 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.[10] 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:

---

[9] The 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.

[10] 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.

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

To simulate purchase decisions, we take a pseudo-random draw, η_{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 revenue requirement of the government serves as a constraint. We assume a revenue requirement $\tilde{w}$ is necessary to finance government activities or provide a public good and that the sum of taxes collected in considered periods must exceed that threshold, $(\Sigma_t \tau_t > \tilde{w})$. We do not model the provision of the public good (i.e., the expenditure decision). Rather, the government's objective is to satisfy the revenue constraint with minimal distortions. We simulate this 10,000 agent, six period model for each possible sequence of taxes, conditional on a vector of parameters θ:

$$\tau_t \in \{0, 0.1, 0.2, \ldots, 0.9\}, \ \forall t \in \{1, \ldots, 6\}$$

We compare tax revenues and total consumer utility for the $K^T = 1{,}000{,}000$ considered sequences of tax rates for a given vector of parameters, θ, using a grid search to find a discrete approximation to the optimal tax sequence:

$$\max_{\tau_1, \ldots, \tau_6} \sum_t \sum_i \left(U_{it} | \theta, \sum_t \tau_t > \tilde{w} \right)$$

over a finite, bounded set of discrete values. As the objective is to explore the conditions under which potential gains from dynamic taxation are greater or lesser, solving for a precise maximum over a continuous set with infinite values adds minimal insight.

We simulate the model under two different assumptions about disposal costs that represent feasible endpoints of the continuum when the government increases tax rates. 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 prohibitively 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 also simulate the model under an alternative condition that represents the other end of the disposal cost continuum — where in each period consumers may choose to discard the good at no cost.[11] The probability of disposal increases if taxes are increased.[12] Under either set of assumptions about permanence or disposal

---

[11] The true other end of the continuum is that customers could costlessly resell the good on the secondary market at the full price paid. However, market price would be an endogenous state variable determined by the number of individuals who had purchased the good in previous periods and the disposal incentives induced by changes in taxes. Allowing consumers to simply dispose of the good at a price of zero is a cleaner option that still permits insight on how disposal costs factor into the gains from intertemporal variation in taxes.

[12] We use the mean utility from each good to produce logit probabilities for keeping the good, similar to the way we

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/network externalities. We therefore simulate a comparison case for each set of parameters (r, p, γ, 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.

3.2 Quantifying Gains from Intertemporal Variation in Taxes

Tables 1 through 2 contain the results from conducting simulations with several sets of parameters, to quantify potential welfare gains from dynamic taxation under different sets of parameters and different strengths of the atmospheric and consumption externalities.

We choose two static tax rates, $(\tilde{\tau}_1 = 0.3; \tilde{\tau}_2 = 0.5)$, as baselines for the purpose of comparison. These rates are admittedly *ad hoc*, but were chosen because they are in the middle of the considered range of values, allowing us to explore the effects of lower taxes in earlier periods and higher taxes in later periods. For each sequence of taxes, we consider the sum of collected taxes and consumer surplus over all six time periods. 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 $\tau^\star$ that yields maximum consumer utility, conditional on meeting or exceeding the revenue requirement.

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

$$\%\Delta CS = \frac{(CS|\tau^\star, \theta) - (CS|\tilde{\tau}, \theta)}{(CS|\tilde{\tau}, \theta)}$$

However, dynamic taxation may lead to welfare gains in the absence of any externalities due to the durable nature of the good. Therefore, this measure is likely to understate gains from dynamic taxation that are specifically attributable to consumption externalities. When consumption/network externalities are introduced, the parameter vector changes from θ to θ'. If consumption externalities are positive (e.g., setting $\alpha = 1$ rather than $\alpha = 0$), then $(CS|\tilde{\tau}, \theta') > (CS|\tilde{\tau}, \theta)$. Positive network externalities will therefore increase not just the numerator, but the utility for each and every tax sequence considered. Because the denominator is increasing, this leads to understatement of the importance of dynamic taxation when interpreting of these results.

As an alternative way of evaluating the magnitude of the gains from dynamic taxation, we 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. Intuitively, this measure can be thought of as "insofar as taxes affect consumer surplus, how much better off can time varying

---

introduced randomness in the purchase decision.

taxation make *consumers* when network externalities are present?" We express this measure as:

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

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 with discrete space can lead to imprecision. We thus encourage readers to focus interpretation on broad trends rather than details of any particular sequence.

3.3 Principal Results from the Multi-period Model

Results from these simulations provide insight on where dynamic taxation can increase consumer utility, but also where these schemes can do unintended harm. The potential for dynamic taxation to increase consumer utility while remaining at least revenue neutral is chiefly determined by the strength of the network effect and the extent to which taxes are pivotal in purchase decisions.

While many of the results explore the extent to which intertemporal variation in tax rates *can* increase both consumer surplus and government revenues, there will obviously be parts of the parameter space where intertemporal variation in taxes is *not* effective. This is an important caveat, as the presence of network externalities does not necessarily imply that a dynamic taxation scheme should be implemented. When atmospheric and network externalities have opposing signs, some caution is recommended in implementing dynamic taxation. However, this section shows that when certain conditions are met, dynamic taxation can Pareto-dominate a static tax scheme baseline.

Panel A of Table 1 presents results from a set of parameters for which taxes are *not* that pivotal. This vector of parameters creates a setting where intertemporal taxation plausibly could affect consumer surplus, but has little empirical impact: 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. Because the purchase price is greater than the private valuation of the good, there is potential for network effects (and adoption rates induced by tax policy) to matter.

There are several takeaways from Panel A. First, when disposal costs are prohibitive (or the purchase decision is binding/permanent) dynamic taxation improves welfare with or without network externalities. This is quite intuitive: prohibitive disposal costs make consumers less responsive to taxes and thus make those taxes less distortionary. The Normalized gains in consumer surplus are substantially lower when consumers have free disposal. As expected, the ability to discard the good constrains the efficacy of suddenly raising taxes on consumers. Consequently it is

Table 1: Summary of simulation results under different baseline prices of the good

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

		$\tilde{\tau}_t = 0.3 \forall t$			$\tilde{\tau}_t = 0.5 \forall t$		
α	Discardable	$\tau^\star$	%ΔCS	Normalized	$\tau^\star$	%ΔCS	Normalized
$\alpha = 0$	Permanent	0.0; 0.0; 0.0; 0.0; 0.8; 0.7	1.12	0.0600	0.0, 0.0, 0.0, 0.8, 0.9, 0.8	1.75	0.0915
$\alpha = 0.5$	Permanent	0.0; 0.0; 0.0; 0.0; 0.7; 0.9	1.35	0.0827	0.0, 0.0, 0.0, 0.8, 0.9, 0.8	2.87	0.1426
$\alpha = 1.0$	Permanent	0.0, 0.0, 0.0, 0.1, 0.6, 0.9	1.57	0.0882	0.0, 0.0, 0.0, 0.8, 0.9, 0.9	2.99	0.1542
$\alpha = 0$	Discard	0.2, 0.1, 0.2, 0.3, 0.3, 0.8	0.19	0.0128	0.5, 0.5, 0.3, 0.4, 0.5, 0.9	0.16	0.0125
$\alpha = 1.0$	Discard	0.0, 0.0, 0.1, 0.3, 0.5, 0.8	1.21	0.0496	0.1, 0.2, 0.4, 0.5, 0.7, 0.9	1.95	0.0724

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

		$\tilde{\tau}_t = 0.3 \forall t$			$\tilde{\tau}_t = 0.5 \forall t$		
α	Discardable	$\tau^\star$	%ΔCS	Normalized	$\tau^\star$	%ΔCS	Normalized
$\alpha = 0$	Permanent	0.0; 0.0; 0.0; 0.7; 0.1; 0.6	1.51	0.0941	0.0, 0.0, 0.0, 0.9, 0.9, 0.4	2.61	0.1574
$\alpha = 0.5$	Permanent	0.0; 0.0; 0.0; 0.0; 0.5; 0.9	2.58	0.1340	0.0, 0.0, 0.0, 0.4, 0.9, 0.9	4.57	0.2219
$\alpha = 1.0$	Permanent	0.0; 0.0; 0.0; 0.0; 0.5; 0.9	3.14	0.1362	0.0, 0.0, 0.0, 0.4, 0.9, 0.9	5.95	0.2142
$\alpha = 0$	Discard	0.1; 0.1; 0.2; 0.1; 0.5; 0.9	0.53	0.0449	0.5, 0.3, 0.4, 0.4, 0.5, 0.9	0.43	0.0355
$\alpha = 1.0$	Discard	0.0; 0.0; 0.0; 0.0; 0.7; 0.9	3.26	0.1294	0.0, 0.0, 0.3, 0.5, 0.7, 0.9	5.17	0.1860

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

		$\tilde{\tau}_t = 0.3 \forall t$		
α	Discardable	$\tau^\star$	%ΔCS	Normalized
$\alpha = 0$	Discard	0.1; 0.0; 0.1; 0.4; 0.4; 0.9	0.52	0.0393
$\alpha = 1.0$	Discard	-0.5; -0.2; 0.1; 0.3; 0.6; 1.2	5.07	0.1835

Notes: Each row within a Panel presents the consumer surplus-maximizing tax sequence $\tau^\star$ for different parameter specifications of unit price p, flow utility γ, elasticity of substitution r, and income y. The α 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.

important to allow consumers to discard the good to discipline the results. Panel A makes it clear that the optimal tax sequence exhibits less dramatic acceleration of tax rates when disposal of the good is costless.

Focusing on those cases where consumers may discard the good, the gains of dynamic taxation increase when network externalities are present ($\alpha > 0$). In the $\tilde{\tau}_t = 0.3$ case, the gains to consumers increase from 1.3% to 5.0% of the Normalized space (while satisfying the government revenue constraint). We see similar but quantitatively larger effects when $\tilde{\tau}_t = 0.5$, with Normalized consumer gains increasing from 1.3% to 7.2%.

This latter point reflects a general tendency. Note that when $\tilde{\tau} = 0.5$, dynamic taxation has more potential to improve welfare than when $\tilde{\tau} = 0.3$. This is driven by some intuitive properties. First, when $\tilde{\tau} = 0.5$, there is more space to cut taxes in the [0.0, 0.9] interval in earlier periods. Second, one reason that intertemporal variation in taxes did not have as large an effect when $\tilde{\tau} = 0.3$ was that taxes of that magnitude were simply not that pivotal: a large share of the agents were buying the good anyway. When $\tilde{\tau} = 0.5$, conditional on the other parameters in the model, taxes have a greater effect at the baseline. More broadly, the greater the revenue requirement and the higher the tax rate, the more opportunity there is to pursue optimal taxation through dynamics when network externalities are present.

Panel B shows results from a simulation with almost identical parameters as Panel A, but where purchase prices have increase from 0.9 to 1.5. Under the conditions in Panel A, a large share of consumers were willing to buy the good irrespective of taxes. With a purchase price of 1.5, taxes are more pivotal. From an initial purchase price of 1.5, a time invariant tax of $\tilde{\tau} = 0.3$ is enough to inhibit network formation. Therefore when there is a larger gap between prices and private valuations, dynamic subsidy/taxation has greater potential to increase consumer surplus.

With $\tilde{\tau} = 0.3$ and individuals are permitted to dispose of the good, the gains from dynamic taxation are equal to 3.3 percent of the baseline case when $\alpha = 1$, or 12.9 percent of the support of CS gains attributable to taxation. When $\tilde{\tau} = 0.5$, time varying tax rates have the potential for even larger gains in consumer surplus. Even with free disposal, a time-varying sequence of taxes can increase consumer surplus by 5.2 percent over the baseline, or 18.6 percent of the considered range of tax-related variation in consumer surplus.

Finally, note that most of the identified $\tau^{\star}$ sequences in Panels A and 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, increase in periods 2–5 and reach the the upper-bound in the last period.

Because the optimal sequences in Panels A and B are bound by corner solutions, we relax the range of taxes in Panel C to span $[-0.5, 0.4]$ in the first three periods and $[0.3, 1.2]$ in the last three periods.[13] The bottom-left panel of Table 1 shows the optimal policy does induce government debt,

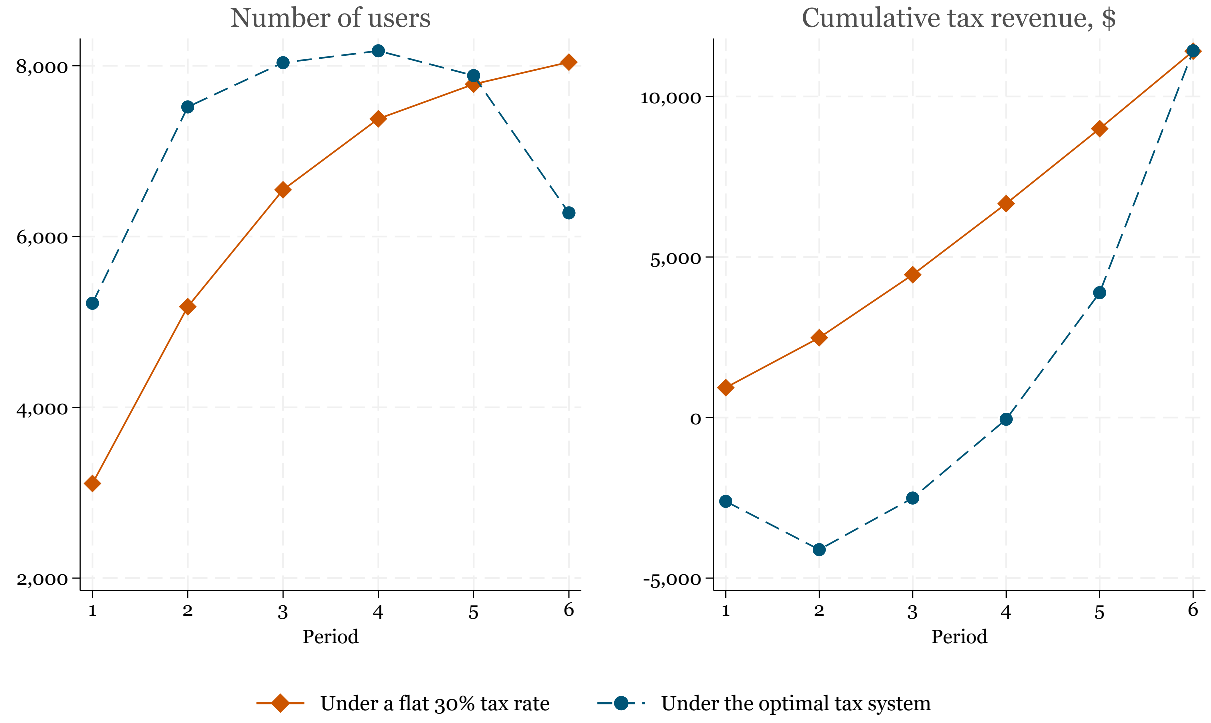
---

[13] This of course assumes governments have the ability to borrow, and at zero interest rates. Discount 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 to pay the interest. As long as consumption externalities have the requisite strength, the qualitative implications are unchanged.

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.

Comparing Panel C's results when $\alpha = 1$ to the last row of Panel B, we see $\tau^{\star}$ is still a smooth escalation from maximum subsidy in the first period to the maximum tax rate in the last period. We see an even greater gain in consumer surplus compared to the baseline case. When the set of tax values is expanded to include subsidies in the early periods, dynamic taxation can increase consumer surplus by 5.1 percent of the baseline, approximately 1.5 times larger than when the tax sequences are constrained to [0.0,0.9]. In the presence of consumption externalities, expanding the set of potential tax values to include subsidies further increases potential gains in consumer surplus, compared to the case where tax rates are bounded below by zero.

Figure 2: Graphical contrast of consumption and cumulative tax revenue under baseline and optimal tax schemes



Figures show paths of the number of users cumulative tax revenue for the time-invariant case (solid line) and optimal policy case (dashed line). We see initial subsidization of the network good encourages early adoption, and this facilitates recouping foregone tax revenues in the final period. Panels depict outcomes for $\alpha = 1$ and where consumers have the option of free disposal. Similar patterns emerge with higher levels of α and/or a more binding revenue requirement.

Figure 2 provides a graphical summary to supplement the findings in Table 1's Panel C. Figure 2 includes two charts, depicting behaviour in the default time-invariant case (solid line, a policy of a constant tax rate of 0.3) and in the optimal tax sequence (dashed line). The charts depict outcomes for $\alpha = 1$ where consumers have the option of free disposal. The left panel of Figure 2 shows the patterns of adoption under the two tax schemes. The optimal sequence encourages early adoption, increasing the number of users of the good. We see substantially higher usage in the first five periods, suggesting higher welfare. As well as creating consumer surplus through the network externality's positive effects on utility, this expands the tax base for the relatively high rates that follow. However the strategy is not unambiguously better, as number of users is lower in the last period when taxes are particularly high. The right panel of Figure 2 plots the cumulative tax revenue over the six periods. The solid line shows a relatively consistent (and certainly monotonic) increase in cumulative revenues over the entire period. In contrast, the initial subsidization from the optimal tax system means the government runs a deficit for the first two periods. However, the escalation of tax rates from the third period on means cumulative revenue quickly catches up and by the final period marginally surpasses those raised in the time-invariant case. The higher tax rates means 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.

The level of differentiation of the good in question can also affect the potential for intertemporal variation in taxes to increase consumer surplus under network externalities. While we do not explicitly model product differentiation, the CES utility function allows us to adjust the extent to which the composite numeraire good is a substitute for the externality generating good. These results are available in Appendix D.

3.4 Dynamic Taxation under Atmospheric and Network Externalities

While the prior section implicitly treats the intertemporal development of the network as the relevant externality, we now analyze a scenario that also includes atmospheric externalities.

In the absence of atmospheric externalities, we have seen that intertemporal variation in tax rates can improve overall consumer surplus while increasing government revenues. When both atmospheric and network externalities are of the same sign, the revenue-neutral gains to consumer surplus from subsidizing/taxing the good in earlier periods increase in magnitude. However, when the signs of the atmospheric and network externalities are opposing, the welfare implications for dynamic taxation are a bit more nuanced.

Table 2 presents results with explicit incorporation of negative atmospheric externalities concurrent with positive network externalities. In particular, Table 2 contains results from a set of simulations where taxes are pivotal, the good is discardable, the parameter on the network externalities is weakly positive ($\alpha \geq 0$), and the parameter on the atmospheric externalities are weakly negative ($\delta \leq 0$).

When the good does not produce atmospheric externalities, the implications of time varying

Table 2: Simulation results incorporating explicitly dual atmospheric-consumption externalities

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

		$\tilde{\tau}_t = 0.3 \forall t$			$\tilde{\tau}_t = 0.5 \forall t$		
α	δ	$\tau^\star$	%ΔCS	Normalized	$\tau^\star$	%ΔCS	Normalized
0.0	0.0	0.0; 0.1; 0.2; 0.3; 0.4; 0.8	0.46	0.042	0.4, 0.2, 0.4, 0.5, 0.6, 0.9	0.39	0.031
0.5	0.0	0.0; 0.0; 0.1; 0.2; 0.5; 0.9	1.80	0.096	0.0, 0.2, 0.2, 0.7, 0.7, 0.9	1.99	0.101
1.0	0.0	0.0; 0.0; 0.0; 0.0; 0.7; 0.9	3.41	0.130	0.0, 0.0, 0.3, 0.5, 0.7, 0.9	5.15	0.184
1.5	0.0	0.0; 0.0; 0.0; 0.0; 0.6; 0.9	3.94	0.128	0.0, 0.0, 0.0, 0.6, 0.8, 0.9	7.56	0.227
2.0	0.0	0.0; 0.0; 0.0; 0.2; 0.4; 0.9	3.87	0.122	0.0, 0.0, 0.1, 0.6, 0.7, 0.9	7.87	0.237
0.0	-1.0	0.9; 0.9; 0.8; 0.4; 0.0; 0.0	0.64	0.103	0.9, 0.9, 0.9, 0.9, 0.5, 0.1	0.54	0.085
0.5	-1.0	0.3; 0.1; 0.1; 0.2; 0.3; 0.8	0.30	0.023	0.6, 0.2, 0.2, 0.5, 0.6, 0.9	0.29	0.022
1.0	-1.0	0.0; 0.0; 0.2; 0.2; 0.4; 0.7	1.92	0.091	0.0, 0.1, 0.2, 0.5, 0.8, 0.9	2.58	0.116
1.5	-1.0	0.0; 0.0; 0.0; 0.2; 0.4; 0.9	2.95	0.107	0.0, 0.0, 0.0, 0.6, 0.9, 0.9	5.01	0.169
2.0	-1.0	0.0; 0.0; 0.0; 0.3; 0.3; 0.9	3.10	0.102	0.0, 0.0, 0.0, 0.6, 0.9, 0.9	5.92	0.182
0.0	-1.5	0.9; 0.9; 0.9; 0.4; 0.0; 0.0	2.04	0.408	0.9, 0.9, 0.9, 0.9, 0.8, 0.0	1.90	0.377
0.5	-1.5	0.9; 0.9; 0.5; 0.1; 0.1; 0.3	0.67	0.062	0.9, 0.9, 0.9, 0.8, 0.6, 0.2	0.57	0.053
1.0	-1.5	0.0; 0.0; 0.0; 0.2; 0.5; 0.9	0.86	0.046	0.0, 0.3, 0.3, 0.4, 0.6, 0.9	1.03	0.052
1.5	-1.5	0.0; 0.0; 0.0; 0.2; 0.5; 0.8	2.18	0.086	0.0, 0.0, 0.0, 0.6, 0.9, 0.9	3.68	0.136
2.0	-1.5	0.0; 0.0; 0.0; 0.2; 0.4; 0.9	2.32	0.079	0.0, 0.0, 0.1, 0.5, 0.9, 0.9	4.47	0.143
0.0	-2.0	0.9; 0.9; 0.9; 0.9; 0.0; 0.0	3.58	0.684	0.9, 0.9, 0.9, 0.9, 0.9, 0.0	3.23	0.618
0.5	-2.0	0.9; 0.9; 0.9; 0.6; 0.0; 0.0	2.36	0.280	0.9, 0.9, 0.9, 0.9, 0.7, 0.1	2.21	0.258
1.0	-2.0	0.7; 0.2; 0.2; 0.2; 0.1; 0.7	0.20	0.013	0.9, 0.2, 0.4, 0.3, 0.6, 0.8	0.25	0.016
1.5	-2.0	0.0; 0.0; 0.0; 0.2; 0.5; 0.8	1.36	0.064	0.0, 0.1, 0.1, 0.5, 0.8, 0.9	2.06	0.092
2.0	-2.0	0.0; 0.0; 0.0; 0.1; 0.5; 0.9	1.87	0.099	0.0, 0.1, 0.1, 0.7, 0.7, 0.8	3.34	0.167

Notes: Each row within presents the welfare-maximizing tax sequence $\tau^\star$ for given parameter values of unit price p, flow utility γ, elasticity of substitution r, income y, and purely atmospheric externality δ. The α 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.

taxation are the same as Table 1 Panel B: a weakly monotonically increasing sequence of taxes that increases from the lowest value to the highest value improves consumer surplus by 4–8% (depending on the baseline) relative to a static sequence.

When $\delta = -1.0$, however, the case for time varying taxation improving consumer surplus is murkier. The case where $\alpha = 0$ and $\delta = -1.0$ is analogous to a classic polluting good. Taxing the good at the highest rates in earlier periods reduces the number of consumers adopting the good, thereby improving total consumer surplus. When $\alpha = 0.5$ — when there are some positive network externalities — the potential gains from time varying tax rates are simultaneously convoluted and negligible. Finally, in the cases where $\alpha = 1.5$ and higher, the welfare gains from network externalities finally outweigh the negative atmospheric externalities in a substantive way. The tax strategy that yields the largest gains in consumer surplus is therefore to not tax in earlier periods, increase adoption rates, and tax more heavily once the network externalities are entrenched.

The last two sections of Table 2 contain results from scenarios with even larger negative atmospheric externalities ($\delta = -1.5$ and $\delta = -2.0$). Here, we see that when network externalities are relatively small, there are gains to consumer surplus from taxing the good heavily and early, compared to a static tax rate. When there are strong positive network externalities $\alpha \geq 1.5$, taxing at lower rates in earlier periods improves consumer surplus while remaining at least revenue neutral. However, there are also parts of the parameter space (e.g. $0 < \alpha < 1$) where the potential welfare gains from dynamic taxation are minimal and the optimal strategy is unclear. Thus for goods with economically significant atmospheric and consumption externalities of opposite signs, there is certainly the potential for poorly considered dynamic taxation to do more harm than good.

While we view these results as broadly applicable from a qualitative perspective, we recommend caution when interpreting the numeric specifics of these findings. All results from quantitative models are conditional on assumptions about functional form and other parameters. In cases like this where there are clear trade-offs, it is tempting to solve for an 'exchange rate' between α and δ. If we consider a good that produces atmospheric and consumption externalities of opposite signs, what is the threshold ratio value of $\frac{|\alpha|}{|\delta|}$ that gives clear guidance on when to subsidize the good early on versus tax it out of existence? There are multiple correct answers for that value, all of which depend on the context-specific assumptions about consumer utility.

3.5 Heterogeneous Agents and Pareto Optimality

In our quantitative model, we assume that agents are homogeneous except for taste shocks. However, extensions for heterogeneous agents are relatively straight forward. For example, suppose the good generates only a positive network externality, but that some share of the population s_N is never going to buy the good. If interpret this as due to differences in preferences, and some fraction of society just do not enjoy the network good, the potential gains in utility from dynamic taxation ΔCS_R are simply:

$$\Delta CS_R = \Delta CS_U \cdot (1 - s_N)$$

where ΔCS_U is the potential gains in consumer utility if everyone was a prospective adopter and s_N is the share that will never buy the good. As long as no one is made *worse* off by the network good, dynamic taxation schemes are at least weakly Pareto improving. To the extent that the efficiently generated revenue from the network good can crowd-out inefficient commodity taxes (which are a common empirical reality) these schemes may be strictly Pareto improving.

However, there should be concerns about potential regressiveness of dynamic taxation when some consumers are *constrained* (by income or other accessibility concerns) from purchasing the good, and the good in has the properties like those in Table 2. In other words, if the good generates positive network externalities, the temptation to encourage adoption through dynamic taxation may exist. However, if the good produces negative atmospheric externalities, and (without loss of generality) low income consumers cannot afford to join the network, a tax scheme that encourages adoption will make those who cannot reap the benefits of network effects worse off.

Let us consider what conditions would need to be true to sufficiently compensate the non-users. Consider a population of measure 1, with population share s who own the externality-generating good and $(1-s)$ who do not own the good. Denoting the parameter on the network externality α^* and the atmospheric externality δ^*, the total amount of disutility attributable to negative atmospheric externalities experienced by non-users is equal to:

$$(1-s)\left(U(\cdot|\delta=\delta^*)-U(\cdot|\delta=0)\right)$$

or the share of the population that are non-users times the per-individual disutility from the negative externality. In the CES functional form above, the per-person disutility from the negative externality, $(U\cdot|\delta=\delta^*)-U(\cdot|\delta=0))=s\cdot\delta^*$. The total disutility imposed by the negative atmospheric externality on the population of non-users is therefore $(1-s)\cdot s\cdot\delta^*$.

From the positive network externality, the general form for the total amount of utility attributable to the positive network externality experienced by users is equal to:

$$s\cdot\left(U(\cdot|\alpha=\alpha^*)-U(\cdot|\alpha=0)\right)$$

In the CES utility function considered here, $(U(\cdot|\alpha=\alpha^*)-U(\cdot|\alpha=0))\geq s\cdot\alpha^*$. For the next step, we briefly treat that weak inequality as an equality, and claim that the total utility experienced by users of the network good to network externalities is equal to $s\cdot s\cdot\alpha$.

Putting these two expressions together under the specification in this paper, the total utility generated from the network externality for users is greater than the disutility from the atmospheric externality imposed on non-users if:

$$s\cdot\alpha>(1-s)\cdot\delta \quad \text{or} \quad \frac{\alpha}{\delta}>\frac{(1-s)}{s}$$

If this is true, the total surplus accrued by users from the network is sufficient to where each non-user can be effectively held harmless from the negative atmospheric externality with a transfer at least equal to $s\delta$ pooled from the users of the externality generating good. Because in our specification, $(U(\cdot|\alpha=\alpha^*)-U(\cdot|\alpha=0))\geq s\cdot\alpha^*$, this condition is excessively strict. Interpreting the above inequality, assume there are a large share of users, $s\approx 0.8$. In this case, there are approximately four users for every non-user. Even if the parameter on the atmospheric externality is larger than the parameter on the network externality, if the government can arrange transfers from users to non-users, the transfers per-user will not be large enough to negate the network externalities enjoyed by users. However, we emphasize that this is a *necessary* condition for the existence of transfers that can off set any regressive effects of dynamic taxation, but by no means a *sufficient* condition for transfers to be feasible.

3.6 Dynamic Taxation and Network Externalities with a Monopolist Firm

In this paper, we have so far assumed that the industry of the network good is perfectly competitive, because it allows us to focus on the government's optimization problem. In reality, however, firms in emerging markets (with or without network externalities) have considerable market power, and will set prices to maximize profits.

In this section we run an additional set of simulations to verify that dynamic taxation can improve the efficiency of tax collection when the firm is a monopolist and the good yields network externalities. We use a similar framework as in previous simulations: 10,000 representative agents with an income endowment of 4, choose whether or not to buy a durable good. At the time consumers make their decision on whether to purchase the good, they have the same utility function, preference shocks, and information set as before (share of the population owning the good, current prices and tax rates). As in previous examples, the purchase price is incurred by the consumer once, but any ownership taxes are levied each period in which the customer owns the good. In each period, the consumer has the ability to discard the good at no cost.

While in previous simulations we searched over a grid of sequences of tax rates for the highest value of total consumer utility, now we are searching over a grid of sequences of *prices* looking for the sequence that maximizes *profit*. In these simulations, the government is the first to act in preannouncing a sequence of tax rates. We consider two cases:

$$\begin{aligned} \tau_t &= 0 \quad \forall t \in \{1, \ldots, 6\} \\ \tau_t &= \{-0.5, -0.5, 0, 0, 0.5, 0.5\} \end{aligned} \tag{20}$$

In words, the two cases are: *(i)* zero taxes for six periods; and *(ii)* initial subsidization ($\tau = -0.5$) for two periods, two periods of zero taxes, and two periods of strictly positive taxes ($\tau = 0.5$). Once the government specifies these tax rates, the monopolist then searches over an array of price sequences, $p_t \in [1.0, 3.0] \quad \forall t \in \{1, \ldots, 6\}$, for the sequence that maximizes profit. Panel A of Table 3 shows results from these simulations, comparing the total consumer utility, profit, and government revenue collected under the two tax sequences for values of $\alpha \in \{0, 1, 2\}$.

The qualitative implications for dynamic taxation for efficient revenue generation are similar to those under perfect competition. When the good does not yield network externalities, the time varying tax sequence is less than revenue neutral. However, when $\alpha = 1$, a sequence of taxes where the good is subsidized in the first two periods and taxed in the last two periods by an equal rate raises strictly positive revenue and increases producer surplus without sacrificing consumer utility. In short, when the firm has market power and network effects are strong enough, dynamic taxation can efficiently generate revenue while increasing total surplus. If we consider the 'effective rate' as the amount of revenue collected divided by the revenue from the sales of the durable good, dynamic taxation enables the government to have a 7.9% tax rate under these conditions.

When $\alpha = 2$, the network effects are sufficiently strong so that the ad hoc dynamic taxation increases consumer utility by almost 3% compared to a tax rate of zero in all periods. Profits also increase by 2.5% and the government raises revenue with an effective rate of 16%. Note also that in

Table 3: Simulation results for dynamic taxation when the firm is a monopolist

Panel A: $(\gamma = 0.2; r = 0.8; y = 4)$ $\tilde{\tau} = 0.0$

α	tax scheme	price sequence	utility	profit	tax revenue	effective rate
0	$\{0.0, 0.0, 0.0, 0.0, 0.0, 0.0\}$	$\{1.8, 1.8, 1.8, 1.6, 1.6, 1.4\}$	231968.4	16504.6	0.0	0.0
0	$\{-0.5, -0.5, 0.0, 0.0, 0.5, 0.5\}$	$\{2.0, 1.8, 1.8, 1.8, 1.4, 1.2\}$	232000.9	17024.2	-415.5	-2.44%
1	$\{0.0, 0.0, 0.0, 0.0, 0.0, 0.0\}$	$\{1.8, 1.8, 2.0, 2.0, 1.8, 1.6\}$	247329.4	16504.6	0.0	0.0
1	$\{-0.5, -0.5, 0.0, 0.0, 0.5, 0.5\}$	$\{2.0, 2.2, 2.0, 1.8, 1.6, 1.4\}$	247490.6	20570.0	1629.0	7.92%
2	$\{0.0, 0.0, 0.0, 0.0, 0.0, 0.0\}$	$\{1.8, 2.2, 2.4, 2.4, 2.2, 2.0\}$	270931.5	23573.4	0.0	0.0
2	$\{-0.5, -0.5, 0.0, 0.0, 0.5, 0.5\}$	$\{2.0, 2.4, 2.4, 2.4, 2.0, 1.8\}$	278454.5	24177.4	3855.5	15.94%

Panel B: $(\gamma = 0.2; r = 0.8; y = 4)$ $\tilde{\tau} = 0.3$

α	tax scheme	price sequence	utility	profit	tax revenue	effective rate
0	$\{0.3, 0.3, 0.3, 0.3, 0.3, 0.3\}$	$\{1.8, 1.6, 1.6, 1.6, 1.4, 1.2\}$	228230.1	13789.0	4717.5	34.21%
0	$\{-0.2, -0.2, 0.3, 0.3, 0.8, 0.8\}$	$\{1.8, 1.8, 1.6, 1.4, 1.4, 1.2\}$	228345.6	14350.4	4201.1	29.27%
1	$\{0.3, 0.3, 0.3, 0.3, 0.3, 0.3\}$	$\{1.6, 1.6, 1.6, 1.6, 1.6, 1.4\}$	238474.4	17114.8	6923.1	40.45%
1	$\{-0.2, -0.2, 0.3, 0.3, 0.8, 0.8\}$	$\{1.8, 1.8, 1.6, 1.6, 1.4, 1.4\}$	238639.7	17753.6	8044.7	45.31%
2	$\{0.3, 0.3, 0.3, 0.3, 0.3, 0.3\}$	$\{1.4, 1.8, 2.0, 2.0, 2.0, 1.8\}$	262919.3	20813.6	9012.9	43.30%
2	$\{-0.2, -0.2, 0.3, 0.3, 0.8, 0.8\}$	$\{1.8, 2.0, 2.0, 2.0, 1.6, 1.6\}$	265508.5	21252.2	13202.1	62.12%

the sequences of profit maximizing prices, the firm with market power increases (decreases) its prices by \$0.20 in periods when the \$0.50 subsidy (tax) is in place.

Panel B reports results from simulations similar to Panel A, but comparing outcomes from profit maximizing behavior under a static tax rate of $\tilde{\tau} = 0.3$ to a time-varying sequence $\tau = \{-0.2, -0.2, 0.3, 0.3, 0.8, 0.8\}$. This sequence represents an ad hoc reduction in taxes by 0.5 in the first two periods, but raises taxes by 0.5 in the last two periods. Similar to results in Panel A, when there are no network effects, intertemporal variation in taxes reduces collected revenue. When there are moderate network effects ($\alpha = 1$), the sequence of taxes with the initial subsidy and raised rates in later periods leads to greater revenue collections and slight increases in total consumer utility and profit. Finally, when network effects are sufficiently strong ($\alpha = 2$), dynamic taxation is Pareto improving. All parties (consumers, firms, and government) better off under the initial subsidy and subsequent elevated rates compared to a time-invariant sequence of tax rates.

4 Conclusion

This paper studies how a planner would choose the tax rate for goods with potentially both atmospheric and network externalities. Beyond theoretical interest, it is likely that such products will be a large source of government revenue in coming decades.[14]

---

[14] For one example, in October 2025 the National Assembly of France voted to increase that country's Digital Services Tax rate from 3% to 6%.

We first develop a static model of optimal taxation where the marginal utility of, and thus demand for, a good is affected by the total consumption of that good. This setting is applicable quite broadly: to traditional network goods like phones and operating systems, but also to a variety of other more recent phenomena like autonomous vehicles or indoor-dining during a pandemic. Indeed any component of the utility function which is affected the society's total level of consumption can be investigated with this model. We preclude the use of lump-sum taxes to focus on second-best policies.

The solution to the static model generalizes previous results from the literature, including those of Pigou (1920) and Sandmo (1975). The tax rate comprises three additively separable factors related to substitution elasticities, the magnitude of any atmospheric externalities, and the effect the network externality has on consumption behavior. To the extent that many networks goods (such as social media) may create negative effects on society at large, those negative atmospheric externalities should be taxed, and we show that the optimal tax rate is higher if the network externality lowers utility from private consumption. Equivalently, the optimal tax rate is lower if the network externality is positive. If the network effects are strong enough, the optimal policy may even be to subsidize goods that generate negative atmospheric externalities.

Anticipating growth in the number of network goods in the economy, we investigate if the government should tax early-stage network 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, at least under reasonable parameterizations. In both a two-period theoretical model and a six-period quantitative model, 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, and under varying market structures. 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 believe this result, on the optimality of time-varying consumption tax rates, is novel to the literature.

We end on three points. Firstly, while our results show cases where time-varying taxes in general and initial subsidization in particular can deliver substantial gains, we also demonstrate several parameter values where this is not true. Secondly, the real-world pattern of internet goods only becoming part of the tax base after their widespread use can be interpreted as a politically-constrained approximation to potentially optimality initial subsidization. Thirdly, we note a potential application of our work to the public policy of pandemics. One does not typically consider 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. We encourage further analysis of the topic in this area.

References

Agrawal, David R and William F Fox, "Taxing goods and services in a digital era," *National Tax Journal*, 2021, *74* (1), 257–301.

Akcigit, Ufuk, Douglas Hanley, and Stefanie Stantcheva, "Optimal taxation and R&D policies," *Econometrica*, 2022, *90* (2), 645–684.

Allcott, Hunt, Luca Braghieri, Sarah Eichmeyer, and Matthew Gentzkow, "The welfare effects of social media," *American Economic Review*, March 2020, *110* (3), 629–76.

Aronsson, Thomas and Olof Johansson-Stenman, "Paternalism against Veblen: Optimal taxation and non-respected preferences for social comparisons," *American Economic Journal: Economic Policy*, 2018, *10* (1), 39–76.

Barrage, Lint, "Optimal dynamic carbon taxes in a climate–economy model with distortionary fiscal policy," *The Review of Economic Studies*, 2020, *87* (1), 1–39.

Bastani, Spencer and Daniel Waldenström, "AI, Automation and Taxation," in "Handbook on Labour Markets in Transition," Edward Elgar Publishing, 2024, pp. 354–370.

Bhattacharya, Jay and William B. Vogt, "A Simple Model of Pharmaceutical Price Dynamics," *Journal of Law and Economics*, 2004, pp. 599–626.

Borders, Kane, Sofía Balladares, Mona Barake, and Enea Baselgia, "Digital Service Taxes," *EU Tax Observatory*, 2023.

Criscuolo, Chiara, Ralf Martin, Henry G. Overman, and John Van Reenen, "Some Causal Effects of an Industrial Policy," *American Economic Review*, January 2019, *109* (1), 48–85.

Cui, Wei, "The superiority of the digital services tax over significant digital presence proposals," *National Tax Journal*, 2019, *72* (4), 839–856.

Eckerstorfer, Paul and Ronald Wendner, "Asymmetric and non-atmospheric consumption externalities, and efficient consumption taxation," *Journal of Public Economics*, 2013, *106*, 42–56.

Fainmesser, Itay P. and Andrea Galeotti, "Pricing Network Effects," *The Review of Economic Studies*, 2016, *83* (1), 165–198.

Goyal, Sanjeev, *Connections: An Introduction to the Economics of Networks*, Princeton University Press, 2012.

Greaker, Mads and Kristoffer Midttømme, "Network effects and environmental externalities: Do clean technologies suffer from excess inertia?," *Journal of Public Economics*, 2016, *143*, 27–38.

Guerreiro, Joao, Sergio Rebelo, and Pedro Teles, “Should robots be taxed?,” *The Review of Economic Studies*, 2022, *89* (1), 279–311.

Jackson, Matthew O and Alison Watts, “The evolution of social and economic networks,” *Journal of Economic Theory*, 2002, *106* (2), 265–295.

Jacquet, Laurence and Etienne Lehmann, “Production Regulation Principles and Tax Reforms,” February 2025. CESifo Working Paper No. 11705.

Katz, Michael L and Carl Shapiro, “Network externalities, competition, and compatibility,” *American Economic Review*, 1985, *75* (3), 424–440.

Kind, Hans Jarle, Marko Koethenbuerger, and Guttorm Schjelderup, “Efficiency enhancing taxation in two-sided markets,” *Journal of Public Economics*, 2008, *92* (5-6), 1531–1539.

Kopczuk, Wojciech, “A note on optimal taxation in the presence of externalities,” *Economics Letters*, 2003, *80* (1), 81–86.

KPMG, “Taxation of the Digitalized Economy: Developments Summary,” Technical Report, KPMG LLP October 31, 2025. Accessed at https://kpmg.com/us on November 12, 2025.

Madden, David, “Labour supply, commodity demand and marginal tax reform,” *Economic Journal*, 1995, *105* (429), 485–497.

Micheletto, Luca, “Redistribution and optimal mixed taxation in the presence of consumption externalities,” *Journal of Public Economics*, 2008, *92* (10-11), 2262–2274.

Pigou, Arthur Cecil, *The Economics of Welfare*, London: Macmillan and Co., 1920.

Sandmo, Agnar, “Optimal taxation in the presence of externalities,” *Swedish Journal of Economics*, 1975, pp. 86–98.

Scheuer, Florian and Iván Werning, “The taxation of superstars,” *The Quarterly Journal of Economics*, 2017, *132* (1), 211–270.

Thuemmel, Uwe, “Optimal taxation of robots,” *Journal of the European Economic Association*, 2022, pp. 1–37.

A Appendix: Proof of Theorem 3

We begin the proof by with the following definition and lemma:

Definition 2 (Difference)**.** For $q = (q_1, q_2)$ define

$$D(\delta; q) := \chi_2(\delta; q) - \chi_1(\delta; q) = \left[\mathrm{MRS}_{q,2} - \mathrm{MRS}_{q,1}\right](\delta; q) - (1-\delta)\Lambda_{\mathrm{priv}}(\delta; q),$$

and the Ramsey-relevant set $D^*(\delta) := \{D(\delta; q) : q \in \mathcal{A}(\delta)\}$. At $\delta = 1$ set

$$\Delta(q) := D(1; q) = \mathrm{MRS}_{q,2}(1; q) - \mathrm{MRS}_{q,1}(1; q), \qquad \Delta_{\min} := \inf_{q \in \mathcal{A}(1)} \Delta(q).$$

Note that $\frac{1+\gamma\tau_2}{1+\gamma\tau_1} = \frac{\chi_2}{\chi_1} > 1 \iff D(\delta; q) > 0$.

Lemma 1 (Existence and upper hemicontinuity)**.** *Under Assumption 1: (i) $\mathcal{A}(\delta)$ is nonempty and compact for each $\delta \in [0, 1)$; (ii) $\mathcal{A}(\delta)$ is upper hemicontinuous at $\delta = 1$.*

Proof. Continuity of $W(q, \delta)$ and of $\tilde{R}(q, \delta)$ in (q, δ) implies $\Gamma(\delta)$ is closed in the compact box $[0, y_1] \times [0, y_2]$, hence compact and nonempty by Assumption 1. By Weierstrass, $\mathcal{A}(\delta)$ is nonempty and compact. The feasible-set correspondence has closed graph and nonempty compact values; Berge's Maximum Theorem yields upper hemicontinuity of $\mathcal{A}(\delta)$ at $\delta = 1$. □

The definition shows that in order to prove the optimal tax in the second period is larger than that in the first period, it suffices to show that the difference $D(\delta; q)$ is positive. The Lemma suggests that the optimal allocations (and hence tax rates) will exist in the feasible set $\Gamma(\delta)$ for a small neighborhood around $\delta = 1$. And hence the Ramsey-relevant set $D^*(\delta)$ is defined and exists at least locally to $\delta = 1$.

The next two Lemmas help to establish that the difference is continuous and converges to $\Delta(q)$ in the limit as $\delta \uparrow 1$.

Lemma 2 (Uniform interiority and bounds)**.** *There exists $\varepsilon > 0$ and a neighborhood U of $\delta = 1$ such that for all $\delta \in U$ and $q \in \mathcal{A}(\delta)$,*

$$z_t(\delta; q) \geq \varepsilon, \qquad f(c_t(\delta; q)) \geq \varepsilon, \qquad t = 1, 2.$$

Consequently, $\mathrm{MRS}_{q,t}$ *and* Λ_{priv} *are continuous on $\mathcal{A}(\delta)$ and uniformly bounded over $\delta \in U$.*

Proof. For $\rho \in (0, 1)$ the CES aggregator has Inada derivatives in each argument: holding the other argument fixed, $\partial u / \partial z \to \infty$ as $z \downarrow 0$, and similarly for $f(c)$ since $f_1 \equiv 1$. With finite prices/taxes, an optimum cannot put z_t or $f(c_t)$ at zero; otherwise a small reallocation increases W without violating the revenue constraint. By compactness and upper hemicontinuity (Lemma 1), these lower bounds can be chosen uniformly in a neighborhood of $\delta = 1$. Continuity and uniform boundedness of $\mathrm{MRS}_{q,t}$ and Λ_{priv} follow. □

Lemma 3 (Limit along optimal selections)**.** *Let* $\delta_n \uparrow 1$ *and choose* $q^n \in \mathcal{A}(\delta_n)$ *with* $q^n \to \bar{q}$*. Then* $\bar{q} \in \mathcal{A}(1)$ *and*

$$D(\delta_n; q^n) \longrightarrow D(1; \bar{q}) = \Delta(\bar{q}).$$

Proof. Upper hemicontinuity (Lemma 1) implies $\bar{q} \in \mathcal{A}(1)$. By Lemma 2, $\mathrm{MRS}_{q,t}$ and Λ_{priv} are continuous and uniformly bounded, hence

$$D(\delta_n; q^n) = \left[\mathrm{MRS}_{q,2} - \mathrm{MRS}_{q,1}\right](\delta_n; q^n) - (1 - \delta_n)\Lambda_{\mathrm{priv}}(\delta_n; q^n) \to \left[\mathrm{MRS}_{q,2} - \mathrm{MRS}_{q,1}\right](1; \bar{q}) = \Delta(\bar{q}).$$

□

With these in hand, we are finally able to prove continuity of the set $\{\delta_n : D(\delta_n, q^n) > 0\}$ away from $\delta = 1$.

Proposition 5 (Neighborhood positivity from a positive $\Delta_{\min}$)**.** *If* $\Delta_{\min} > 0$*, then there exists* $\overline{\delta} \in (0, 1)$ *such that for all* $\delta \in (\overline{\delta}, 1)$ *and all* $q \in \mathcal{A}(\delta)$*,*

$$D(\delta; q) > 0 \quad \Longleftrightarrow \quad \frac{1 + \gamma\tau_2}{1 + \gamma\tau_1} > 1.$$

Proof. By Lemma 2 there exists $\overline{\Lambda} < \infty$ with $\Lambda_{\mathrm{priv}}(\delta; q) \leq \overline{\Lambda}$ for all δ near 1 and all $q \in \mathcal{A}(\delta)$. Choose $\overline{\delta} \in (0, 1)$ with $(1 - \delta)\overline{\Lambda} < \Delta_{\min}/2$ for all $\delta \in (\overline{\delta}, 1)$. Let $\delta \in (\overline{\delta}, 1)$ and $q \in \mathcal{A}(\delta)$. Take any sequence $\delta_n \uparrow 1$ and $q^n \in \mathcal{A}(\delta_n)$ with $q^n \to \bar{q} \in \mathcal{A}(1)$ (Lemma 1). Lemma 3 gives $\left[\mathrm{MRS}_{q,2} - \mathrm{MRS}_{q,1}\right](\delta_n; q^n) \to \Delta(\bar{q}) \geq \Delta_{\min}$. For n large, $\left[\mathrm{MRS}_{q,2} - \mathrm{MRS}_{q,1}\right](\delta_n; q^n) \geq \frac{3}{4}\Delta_{\min}$; then

$$D(\delta_n; q^n) \geq \tfrac{3}{4}\Delta_{\min} - (1 - \delta_n)\overline{\Lambda} > \tfrac{3}{4}\Delta_{\min} - \tfrac{1}{2}\Delta_{\min} = \tfrac{1}{4}\Delta_{\min} > 0.$$

By continuity and compactness of $\mathcal{A}(\delta)$, the same positivity holds for all maximizers at that δ. □

Proposition 5 shows that it is possible to find a depreciation rate of the private network asset away from one which implies optimal allocations in which the tax is higher in the second period. This is akin to the notion (although not isomorphic) that the planner would relatively subsidize network creation in order to increase the size of the tax base in the second period. However, the proposition relies on the assumption that the optimum at $\delta = 1$ implies the same. The following Lemma describes situations in which that is the case.

Lemma 4 (When $\Delta_{\min} > 0$ under identical primitives)**.** *Let* $f = g = c + c^p$*. At* $\delta = 1$*,* $\chi_t = \mathrm{MRS}_{q,t}$*. Moreover,*

$$\mathrm{MRS}^{\mathrm{soc}}_{q,t} = H(c_t)\,\mathrm{MRS}_{q,t}, \qquad H(c) := 1 + \frac{f_2}{f_1} = 1 + p\,c^{p-1}.$$

If $p > 1$ *and every Ramsey optimizer at* $\delta = 1$ *satisfies* $c_2 < c_1$*, then there exists* $\eta > 0$ *such that* $\Delta(q) \geq \eta$ *for all* $q \in \mathcal{A}(1)$*; hence* $\Delta_{\min} > 0$*.*

Proof. Since $f_1 \equiv 1$ and $f_2(c, c) = p\,c^{p-1}$, the displayed identity holds. If $p > 1$, then H is strictly increasing and $1/H$ strictly decreasing. At $\delta = 1$ and for small enough R, the social MRS across periods is bounded and arbitrarily close to a common level (e.g., 1 in the first best). Thus there exists $\varepsilon > 0$ with

$$\left|\mathrm{MRS}^{\mathrm{soc}}_{q,t} - 1\right| \le \varepsilon \qquad (t = 1, 2,\ q \in \mathcal{A}(1)).$$

Therefore,

$$\Delta(q) = \frac{\mathrm{MRS}^{\mathrm{soc}}_{q,2}}{H(c_2)} - \frac{\mathrm{MRS}^{\mathrm{soc}}_{q,1}}{H(c_1)} \ \ge\ \frac{1-\varepsilon}{H(c_2)} - \frac{1+\varepsilon}{H(c_1)} = \Big(\frac{1}{H(c_2)} - \frac{1}{H(c_1)}\Big) - \varepsilon\Big(\frac{1}{H(c_2)} + \frac{1}{H(c_1)}\Big).$$

Since $c_2 < c_1$ and $1/H$ is decreasing, the first bracket is strictly positive; choose ε small enough (via small R) so that the right-hand side is $\ge \eta > 0$ uniformly on $\mathcal{A}(1)$. Hence $\Delta_{\min} \ge \eta > 0$. $\square$

Finally, we close by noting sufficient model primitives which ensure the $c_2 < c_1$ restriction for optimally increasing taxes holds.

Setup at $\delta = 1$. Period constraints decouple except through the *revenue* constraint. Under flow taxes, period-t revenue is $R_t = \tau_t c_t$, so for small changes the first-order revenue raised by $d\tau_t$ is

$$dR_t = c_t^0\, d\tau_t + o(\|d\tau\|).$$

Let $V_t(\tau_t)$ be the period-t indirect utility; the social objective is

$$W(\tau_1, \tau_2) = V_1(\tau_1) + \beta V_2(\tau_2).$$

By the envelope theorem and the household budget $y_t \ge z_t + q_t + \tau_t c_t$ with $c_t = \gamma q_t$,

$$\frac{\partial V_t}{\partial \tau_t}\Big|_{\tau=\tau^0} \ = \ -\lambda_t^0\, c_t^0,$$

where $\lambda_t^0 = \partial u_t / \partial z_t\big|_{(z_t^0, c_t^0)}$ is the marginal utility of the outside good evaluated at the undistorted allocation (finite and strictly positive by Inada).

Lemma 5 (Local Ramsey revenue allocation). *At $\delta = 1$ and for $R > 0$ small, the planner chooses $(d\tau_1, d\tau_2)$ to* minimize *the first-order welfare loss*

$$dW \ = \ -\lambda_1^0 c_1^0\, d\tau_1 \ - \ \beta\, \lambda_2^0 c_2^0\, d\tau_2$$

subject to $c_1^0 d\tau_1 + c_2^0 d\tau_2 = dR$ and $d\tau_t \ge 0$. Consequently, revenue is raised entirely in the period with the lower cost-per-dollar

$$\phi_1 := \lambda_1^0, \qquad \phi_2 := \beta\, \lambda_2^0,$$

i.e., the optimal choice is

$$\textit{if } \phi_2 < \phi_1 :\ d\tau_2 = \frac{dR}{c_2^0},\ d\tau_1 = 0 \quad \textit{and} \quad \textit{if } \phi_1 < \phi_2 :\ d\tau_1 = \frac{dR}{c_1^0},\ d\tau_2 = 0.$$

Proof. Linear program: minimize a linear cost with a single linear equality and nonnegativity. The KKT conditions imply an interior split only if $\phi_1 = \phi_2$; otherwise, allocate all dR to the cheaper arm. $\square$

Proposition 6 (Sufficient primitives for $c_2 < c_1$)**.** *At $\delta = 1$ and for $R > 0$ small:*

1. *If $y_1 = y_2$ and $\beta < 1$, then $c_2 < c_1$.*
2. *If $\beta = 1$ and $y_2 > y_1$, then $c_2 < c_1$.*

Proof. Let (z_t^0, c_t^0) be the undistorted allocation ($R = 0$). With identical primitives and $y_1 = y_2$, symmetry implies $c_1^0 = c_2^0$ and $\lambda_1^0 = \lambda_2^0$. Then

$$\phi_1 = \lambda_1^0, \qquad \phi_2 = \beta\,\lambda_2^0 = \beta\,\lambda_1^0 < \phi_1.$$

By Lemma 5, the planner sets $d\tau_2 = dR/c_2^0$ and $d\tau_1 = 0$, so c_2 falls while c_1 is unchanged to first order; hence $c_2 < c_1$ for small R.

For $\beta = 1$ with $y_2 > y_1$: by monotonicity of marginal utility, $\lambda_2^0 < \lambda_1^0$ (the richer period has lower marginal utility of the outside good). Thus

$$\phi_2 = \beta\,\lambda_2^0 = \lambda_2^0 < \lambda_1^0 = \phi_1,$$

so the planner again raises all revenue in period 2, reducing c_2 while leaving c_1 unchanged to first order; hence $c_2 < c_1$ for small R. $\square$

Corollary 1 (From $c_2 < c_1$ to $\Delta(q) > 0$ when $p > 1$)**.** *If $p > 1$, then $H(c) := 1 + pc^{p-1}$ is strictly increasing, so*

$$\mathrm{MRS}_{q,t} = \frac{\mathrm{MRS}^{\mathrm{soc}}_{q,t}}{H(c_t)}.$$

At $\delta = 1$ and R small, $\mathrm{MRS}^{\mathrm{soc}}_{q,t} \approx 1$, hence

$$\Delta(q) := \mathrm{MRS}_{q,2} - \mathrm{MRS}_{q,1} \;\approx\; \frac{1}{H(c_2)} - \frac{1}{H(c_1)} \;>\; 0$$

whenever $c_2 < c_1$. Therefore, under the primitives of Proposition 6 and $p > 1$, we have $\Delta(q) > 0$.

Remarks. (i) The argument is local in R; by continuity of optimal allocations in R, the inequalities persist on a neighborhood of $R = 0$. (ii) If $\beta = 1$ and $y_2 < y_1$, the inequality reverses: the planner raises revenue in period 1 (the cheaper arm), and generically $c_1 < c_2$. (iii) No concavity of f is required; the local result relies only on envelope derivatives and the linear (to first order) revenue mapping under flow taxes at $\delta = 1$.

B Appendix: Proof of Theorem 4

Proof. *Step 1 (Separability at $\delta = 1$, interiority, and law of demand).* At $\delta = 1$ the household problems decouple across periods. In period t, the Marshallian demand $(z_t(\tau_t), c_t(\tau_t))$ solves

$$\max_{z,c\geq 0} \left[\theta z^\rho + (1-\theta) f(c,X)^\rho\right]^{1/\rho} \quad \text{s.t.} \quad z + (1+\tau_t)c \leq y_t, \;\; X = c,$$

which is strictly concave and admits a unique interior solution by the Inada properties of the CES aggregator for $\rho \in (0,1)$. Standard comparative statics imply the *law of demand*: $dc_t/d\tau_t < 0$ at $\tau_t = 0$.

Step 2 (Private envelope and Roy). Let $u_t^{\text{priv}}(z,c)$ denote the private period utility. Let $V_t^{\text{priv}}(\tau_t)$ be the induced private indirect utility. By the envelope theorem and Roy's identity,

$$\frac{d}{d\tau_t} V_t^{\text{priv}}(\tau_t)\Big|_{\tau_t=0} = -\lambda_t^0\, c_t^0, \qquad \frac{\partial u_t^{\text{priv}}}{\partial c}\Big|_{(z_t^0,c_t^0)} = \lambda_t^0 \cdot (1+\tau_t) \;\;\Rightarrow\;\; \frac{\partial u_t^{\text{priv}}}{\partial c}\Big|_{(z_t^0,c_t^0)} = \lambda_t^0 \;\text{ at } \tau_t = 0. \tag{21}$$

Along the Marshallian response,

$$\frac{d}{d\tau_t} u_t^{\text{priv}}(z_t(\tau_t), c_t(\tau_t)) = \frac{\partial u_t^{\text{priv}}}{\partial z}\frac{dz_t}{d\tau_t} + \frac{\partial u_t^{\text{priv}}}{\partial c}\frac{dc_t}{d\tau_t} = -\lambda_t^0\, c_t^0 \quad \text{at } \tau_t = 0, \tag{22}$$

by the envelope identity in (21).

Step 3 (Relating social and private marginals). Let $u_t^{\text{soc}}(z,c) := \left[\theta z^\rho + (1-\theta) g(c)^\rho\right]^{1/\rho}$ be the *social* period utility with $g(c) = c + c^p$. Because $g(c) = f(c,X)\big|_{X=c}$, we have

$$\frac{\partial u_t^{\text{soc}}}{\partial z}\Big|_{(z_t^0,c_t^0)} = \frac{\partial u_t^{\text{priv}}}{\partial z}\Big|_{(z_t^0,c_t^0)} = \lambda_t^0, \qquad \frac{\partial u_t^{\text{soc}}}{\partial c}\Big|_{(z_t^0,c_t^0)} = \frac{\partial u_t^{\text{priv}}}{\partial c}\Big|_{(z_t^0,c_t^0)} \cdot H(c_t^0),$$

where

$$H(c) := 1 + \frac{f_2(c,c)}{f_1(c,c)} = 1 + p\, c^{p-1} \;\; \geq 1,$$

and strictly > 1 for $p > 1$ and $c > 0$. Therefore, the social welfare derivative with respect to τ_t at $\tau = 0$ is

$$\begin{aligned}
\frac{d}{d\tau_t}\Big(\beta^{t-1} u_t^{\text{soc}}(z_t(\tau_t), c_t(\tau_t))\Big)\Big|_{\tau_t=0} &= \beta^{t-1}\Big[\underbrace{\frac{\partial u_t^{\text{soc}}}{\partial z}}_{=\lambda_t^0}\frac{dz_t}{d\tau_t} + \underbrace{\frac{\partial u_t^{\text{soc}}}{\partial c}}_{=H(c_t^0)\,\lambda_t^0}\frac{dc_t}{d\tau_t}\Big] \\
&= \beta^{t-1}\Big[\underbrace{\frac{\partial u_t^{\text{priv}}}{\partial z}}_{=\lambda_t^0}\frac{dz_t}{d\tau_t} + \underbrace{\frac{\partial u_t^{\text{priv}}}{\partial c}}_{=\lambda_t^0}\frac{dc_t}{d\tau_t}\Big] \;+\; \beta^{t-1}\Big(H(c_t^0) - 1\Big)\lambda_t^0 \frac{dc_t}{d\tau_t} \\
&= -\,\beta^{t-1}\lambda_t^0\, c_t^0 \;+\; \beta^{t-1}\Big(H(c_t^0) - 1\Big)\lambda_t^0 \frac{dc_t}{d\tau_t},
\end{aligned} \tag{23}$$

where we used (22) in the last step.

Step 4 (A conservative bound). By Step 1, $dc_t/d\tau_t < 0$ at $\tau_t = 0$. Hence, since $H(c_t^0) \geq 1$,

$$\frac{d}{d\tau_t}\Big(\beta^{t-1}u_t^{\text{soc}}(\cdot)\Big)\Big|_{\tau_t=0} \;\leq\; -\,\beta^{t-1}\lambda_t^0\, c_t^0 \quad \text{with strict inequality if } p > 1 \text{ (so } H > 1\text{)}. \tag{24}$$

Therefore, writing total social welfare as $W(\tau) = u_1^{\text{soc}} + \beta u_2^{\text{soc}}$, its first-order change at $\tau \equiv 0$ satisfies

$$dW \;\leq\; -\phi_1\, c_1^0\, d\tau_1 \;-\; \phi_2\, c_2^0\, d\tau_2 \;+\; o(\|d\tau\|), \qquad \phi_1 := \lambda_1^0,\;\; \phi_2 := \beta\,\lambda_2^0. \tag{25}$$

Step 5 (Revenue neutrality to first order). Government revenue is $R(\tau) = \tau_1 c_1(\tau_1) + \tau_2 c_2(\tau_2)$. Hence, at $\tau \equiv 0$,

$$dR \;=\; c_1^0\, d\tau_1 + c_2^0\, d\tau_2 \;+\; o(\|d\tau\|). \tag{26}$$

Choose a *revenue-neutral* direction by setting $d\tau_1 = -\eta$ and $d\tau_2 = \eta\,\frac{c_1^0}{c_2^0}$, so that $c_1^0 d\tau_1 + c_2^0 d\tau_2 = 0$.

Step 6 (Strict improvement when $\phi_2 < \phi_1$). Substitute the revenue-neutral direction into (25):

$$dW \;\leq\; -\phi_1 c_1^0(-\eta) \;-\; \phi_2 c_2^0\Big(\eta\,\frac{c_1^0}{c_2^0}\Big) \;+\; o(|\eta|) \;=\; c_1^0(\phi_1 - \phi_2)\,\eta \;+\; o(|\eta|).$$

If $\phi_2 < \phi_1$ and $\eta > 0$ is small, then $dW \geq c_1^0(\phi_1 - \phi_2)\eta/2 > 0$ for sufficiently small η (absorbing the $o(|\eta|)$ term). Because the inequality in (25) is *conservative* (we dropped the strictly negative Pigouvian term when $p > 1$), the actual dW is even larger when $p > 1$. Thus the revenue-neutral perturbation with $d\tau_1 < 0 < d\tau_2$ strictly raises welfare, proving that $(\tau_1, \tau_2) = (0, 0)$ is not locally optimal under $R(\tau) = 0$.

Sufficient primitives for $\phi_2 < \phi_1$. At $\tau \equiv 0$ with identical primitives across time, symmetry yields $(z_1^0, c_1^0) = (z_2^0, c_2^0)$ when $(y_1 = y_2)$; hence $\lambda_1^0 = \lambda_2^0$ and $\phi_2 = \beta\lambda_1^0 < \lambda_1^0 = \phi_1$ whenever $\beta < 1$. If $\beta = 1$ and $y_2 > y_1$, then by monotonicity of marginal utility, $\lambda_2^0 < \lambda_1^0$, so $\phi_2 = \lambda_2^0 < \lambda_1^0 = \phi_1$. This verifies the two sufficient conditions stated in the proposition. □

Remark 2 (Role of $p > 1$). The bound (25) already implies the result under $\phi_2 < \phi_1$ without using $p > 1$. When $p > 1$ (so $H(c) > 1$), the term $\big(H(c_t^0) - 1\big)\lambda_t^0\, dc_t/d\tau_t$ in (23) is strictly *negative* (since $dc_t/d\tau_t < 0$), making the improvement strictly stronger: the Pigouvian motive pushes both periods toward subsidies, and financing part of the period-1 subsidy with a period-2 tax that is cheaper at the margin further raises welfare.

C Comparative Statics in the Two-Period Model

Theorem 7 (Continuity and interior minimum of the depreciation threshold $\delta^*(\theta)$)**.** *Assume flow taxation with $\gamma = 1$ and the representative-agent equilibrium $X_t = c_t$. Fix $p > 1$, endowments (y_1, y_2) with either (i) $y_1 = y_2 > 0$ and $\beta \in (0, 1)$, or (ii) $\beta = 1$ and $y_2 > y_1 > 0$. Let $R > 0$ be small enough so that the feasible sets in the Ramsey problem are nonempty near $\delta = 1$ (Assumption 1). For each $\theta \in (0, 1)$,* *define*

$$D(\theta,\delta;q) := \chi_2(\theta,\delta;q) - \chi_1(\theta,\delta;q) = \Big[\mathrm{MRS}_{q,2} - \mathrm{MRS}_{q,1}\Big](\theta,\delta;q) \ - \ (1-\delta)\,\Lambda_{\mathrm{priv}}(\theta,\delta;q),$$

where (flow taxes) $\chi_2 = \mathrm{MRS}_{q,2}$ *and* $\chi_1 = \mathrm{MRS}_{q,1} + (1-\delta)\Lambda_{\mathrm{priv}}$*, and let* $\mathcal{A}(\theta,\delta)$ *denote the Ramsey argmax correspondence (Definition in Assumption 1). Define the value function*

$$F(\theta,\delta) \ := \ \min_{q\in\mathcal{A}(\theta,\delta)} D(\theta,\delta;q), \qquad \delta^*(\theta) \ := \ \inf\big\{\delta \in [0,1) : \ F(\theta,\delta) > 0\big\}.$$

Then:

(a) $\delta^*(\theta)$ *is continuous on* $(0,1)$*;*

(b) $\lim_{\theta\downarrow 0}\delta^*(\theta) = \lim_{\theta\uparrow 1}\delta^*(\theta) = 1$*;*

(c) For every $\theta \in (0,1)$*,* $\delta^*(\theta) < 1$*.*

Consequently, $\delta^*(\theta)$ *attains a strict interior minimum on* $(0,1)$ *and is not monotone in* θ*.*

Proof. We proceed in three steps. Throughout, we use Lemma 1 (existence and upper hemicontinuity of the Ramsey argmax correspondence $\mathcal{A}$), Lemma 2 (uniform interiority near $\delta = 1$), Lemma 3 (limit along optimal selections), and Proposition 5 (neighborhood positivity from $\Delta_{\min} > 0$), together with Lemma 4 (positivity of the MRS gap at $\delta = 1$ under the stated primitives).

Preliminaries and notation. Fix the parameter set

$$\mathcal{P} \ := \ \{(\theta,\delta) : \ \theta \in (0,1), \ \delta \in [0,1)\}.$$

For $(\theta,\delta) \in \mathcal{P}$, $\mathcal{A}(\theta,\delta)$ is nonempty, compact, and upper hemicontinuous in (θ,δ) by Lemma 1. All the primitives and the implementability objects $\mathrm{MRS}_{q,t}$ and Λ_{priv} are continuous in (θ,δ,q) and (by Lemma 2) uniformly bounded on a neighborhood of $\delta = 1$.

We will repeatedly use the continuity of $D(\theta,\delta;q)$ in (θ,δ,q) and the compactness of $\mathcal{A}(\theta,\delta)$.

Step 1: Lower semicontinuity of F and definition of δ^*. Define $F(\theta,\delta) = \min_{q\in\mathcal{A}(\theta,\delta)} D(\theta,\delta;q)$. Because D is continuous and $\mathcal{A}$ is compact-valued and upper hemicontinuous, the standard minimizing version of Berge's Theorem implies that the *minimum value function F* is *lower semicontinuous* (l.s.c.) on $\mathcal{P}$.

We henceforth take

$$\delta^*(\theta) = \inf\{\delta \in [0,1) : \ F(\theta,\delta) > 0\} = \sup\{\delta \in [0,1) : \ F(\theta,\delta) \le 0\}.$$

(The equality follows from l.s.c. of F and compactness of $[0,1)$.)

Step 2: Proof of (c) and (b). Fix any $\theta \in (0,1)$. By Lemma 4 (applied with the stated primitives: either $y_1 = y_2$ and $\beta < 1$, or $\beta = 1$ and $y_2 > y_1$), we have $\Delta_{\min}(\theta) := \inf_{q\in\mathcal{A}(\theta,1)} \big[\mathrm{MRS}_{q,2} -$

$\mathrm{MRS}_{q,1}\big](\theta, 1; q) > 0$. By Proposition 5 (neighborhood positivity from a positive $\Delta_{\min}$), there exists $\bar{\delta}(\theta) < 1$ such that

$$F(\theta, \delta) > 0 \quad \text{for all } \delta \in (\bar{\delta}(\theta), 1).$$

By the definition of δ^*, this implies $\delta^*(\theta) \leq \bar{\delta}(\theta) < 1$, proving (c).

For (b), we show $\lim_{\theta \downarrow 0} \delta^*(\theta) = \lim_{\theta \uparrow 1} \delta^*(\theta) = 1$. We give the argument at $\theta \uparrow 1$; the case $\theta \downarrow 0$ is analogous.

By Lemma 2, there exists a neighborhood $U_\delta = (\delta_0, 1)$ with $\delta_0 < 1$ and a constant $\underline{\Lambda} > 0$ such that

$$\Lambda_{\text{priv}}(\theta, \delta; q) \;\geq\; \underline{\Lambda} \qquad \text{for all} \quad \theta \in (0,1),\; \delta \in U_\delta,\; q \in \mathcal{A}(\theta, \delta). \tag{27}$$

(This follows because Lemma 2 yields uniform interior lower bounds on z_t and $f(c_t)$ on $\{(\theta, \delta) : \theta \in (0,1),\; \delta \in U_\delta\}$, hence, by continuity, a uniform lower bound on Λ_{priv} there.)

Next, by the analysis at $\delta = 1$ (see Lemma 4 and its proof), for each *fixed* $\delta \in U_\delta$ we have

$$\lim_{\theta \uparrow 1} \sup_{q \in \mathcal{A}(\theta, \delta)} \Big| \big[\mathrm{MRS}_{q,2} - \mathrm{MRS}_{q,1}\big](\theta, \delta; q) \Big| \;=\; 0,$$

since at $\delta = 1$ the MRS gap $\Delta(\theta, 1)$ tends to 0 as $\theta \uparrow 1$ (Lemma 4) and the map $(\theta, \delta, q) \mapsto \mathrm{MRS}_{q,2} - \mathrm{MRS}_{q,1}$ is continuous, so the same holds uniformly for δ in any compact subinterval of U_δ.

Fix any $\eta \in (0, 1 - \delta_0)$ and set $\hat{\delta} := 1 - \eta \in U_\delta$. Then, combining the two displays,

$$\limsup_{\theta \uparrow 1} F(\theta, \hat{\delta}) = \limsup_{\theta \uparrow 1} \min_{q \in \mathcal{A}(\theta, \hat{\delta})} \big\{ \big[\mathrm{MRS}_{q,2} - \mathrm{MRS}_{q,1}\big](\theta, \hat{\delta}; q) \;-\; \eta\, \Lambda_{\text{priv}}(\theta, \hat{\delta}; q) \big\} \;\leq\; 0 \;-\; \eta\, \underline{\Lambda} \;<\; 0.$$

Therefore there exists $\theta_\eta \in (0,1)$ such that for all $\theta \in (\theta_\eta, 1)$ we have $F(\theta, \hat{\delta}) < 0$. By the definition of δ^*, this implies $\delta^*(\theta) \geq \hat{\delta} = 1 - \eta$ for all $\theta \in (\theta_\eta, 1)$. Because $\eta \in (0, 1 - \delta_0)$ is arbitrary,

$$\lim_{\theta \uparrow 1} \delta^*(\theta) = 1.$$

The same argument with $\theta \downarrow 0$ proves the other endpoint limit.

Step 3: Continuity of δ^* on $(0,1)$. Fix $\bar{\theta} \in (0,1)$. We show both upper and lower semicontinuity of δ^* at $\bar{\theta}$.

Upper semicontinuity at $\bar{\theta}$. Let $\varepsilon > 0$ be arbitrary. By the definition of $\delta^*(\bar{\theta})$ and l.s.c. of F, there exists $\delta_+ > \delta^*(\bar{\theta})$ and $\kappa > 0$ such that $F(\bar{\theta}, \delta_+) \geq \kappa$. By l.s.c. of F, there is a neighborhood V of $\bar{\theta}$ with $F(\theta, \delta_+) \geq \kappa/2 > 0$ for all $\theta \in V$. Hence $\delta^*(\theta) \leq \delta_+ < \delta^*(\bar{\theta}) + \varepsilon$ for all $\theta \in V$. Therefore $\limsup_{\theta \to \bar{\theta}} \delta^*(\theta) \leq \delta^*(\bar{\theta})$.

Lower semicontinuity at $\bar{\theta}$. Let $\varepsilon > 0$. Pick $\delta_- < \delta^*(\bar{\theta})$. By the definition of $\delta^*(\bar{\theta})$, $F(\bar{\theta}, \delta_-) \leq 0$. By l.s.c. of F, for any sequence $\theta_n \to \bar{\theta}$ we have

$$F(\bar{\theta}, \delta_-) \;\leq\; \liminf_{n \to \infty} F(\theta_n, \delta_-) \;\;\Rightarrow\;\; \liminf_{n \to \infty} F(\theta_n, \delta_-) \;\geq\; F(\bar{\theta}, \delta_-) \;\leq\; 0.$$

Thus for all large n, $F(\theta_n, \delta_-) \leq \varepsilon'$, and in particular $F(\theta_n, \delta_-) \leq 0$ for infinitely many n. Hence $\delta^*(\theta_n) \geq \delta_-$ along a subsequence, and taking $\delta_- \uparrow \delta^*(\bar{\theta})$ yields $\liminf_{\theta \to \bar{\theta}} \delta^*(\theta) \geq \delta^*(\bar{\theta})$.

Combining upper and lower semicontinuity shows that $\delta^*(\theta)$ is continuous at $\bar{\theta}$. Since $\bar{\theta} \in (0, 1)$ was arbitrary, (a) holds.

Conclusion. Part (c) provides $\delta^*(\theta) < 1$ for each $\theta \in (0, 1)$. Part (b) gives $\delta^*(\theta) \to 1$ as $\theta \downarrow 0$ or $\theta \uparrow 1$. By (a), δ^* is continuous on $(0, 1)$. Therefore, by the extreme value theorem, δ^* attains a minimum on any compact subinterval $[\epsilon, 1 - \epsilon]$, and the limits at the endpoints force a *strict* interior minimum on $(0, 1)$. $\square$

D Additional Simulations

Panel A in Table 4 presents results from a simulation with the same parameters as 1 Panel B, but setting $r = 1$. This implies the externality generating good and the numeraire are perfect substitutes. Although we specify the disposal costs as prohibitive (the purchase is permanent) the gains from dynamic taxation are relatively small (1.15 percent), even when the network externalities are very strong ($\alpha = 2.0$)

Panel B depicts a scenario where the numerarire good is a far less perfect substitute for the externality generating good ($r = 0.5$) but also imposes that the externality generating good yields utility only through the network (private flow utility $\gamma = 0$). In this case, when $\alpha = 0.0$, consumer surplus is increased by sequences that *discourage* consumers from buying the good, as the good is essentially worthless. High tax rates in this scenario helps prevent agents from succumbing to preference shocks and purchasing a good that generates no utility. However, when $\alpha = 1$, dynamic taxation strongly improves welfare over the baseline cases of $\tilde{\tau} = 0.3$ and $\tilde{\tau} = 0.5$. These gains from dynamic taxation can be attributed purely to the presence of those network effects.

The simulation results in Panel C in Table 4 depict a case where dynamic taxation increases consumer surplus while remaining at least revenue neutral compared to the benchmark case. In this case, the network externality is *not* responsible for added value. Rather 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 taxes are pivotal and when there are positive network externalities. 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 network 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 α parameter is large enough. Whether those large values are reasonable depends on the particular good or market being considered.

Table 4: Simulation results under different baseline prices, flow utility, and substitution elasticity

Panel A: $(p = 1.5; \gamma = 0.6; r = 1.0; y = 4)$

		$\tilde{\tau}_t = 0.3 \forall t$			$\tilde{\tau}_t = 0.5 \forall t$		
α	Discardable	$\tau^\star$	%ΔCS	Normalized	$\tau^\star$	%ΔCS	Normalized
$\alpha = 0$	Permanent	0.9, 0.9, 0.6, 0.4, 0.3, 0.1	0.37	0.0550	0.9, 0.9, 0.8, 0.8, 0.7, 0.2	0.33	0.0494
$\alpha = 1.0$	Permanent	0.0, 0.0, 0.0, 0.0, 0.2, 0.9	0.20	0.0310	0.0, 0.0, 0.0, 0.1, 0.7, 0.9	0.18	0.0248
$\alpha = 2.0$	Permanent	0.0, 0.0, 0.1, 0.0, 0.1, 0.9	1.15	0.1334	0.0, 0.0, 0.0, 0.1, 0.7, 0.9	1.55	0.1659

Panel B: $(p = 0.9; \gamma = 0.0; r = 0.5; y = 4)$

		$\tilde{\tau}_t = 0.3 \forall t$			$\tilde{\tau}_t = 0.5 \forall t$		
α	Discardable	$\tau^\star$	%ΔCS	Normalized	$\tau^\star$	%ΔCS	Normalized
$\alpha = 0$	Permanent	0.8; 0.9; 0.9; 0.6; 0.0; 0.0	0.42	0.0195	0.8, 0.9, 0.9, 0.8, 0.5, 0.2	0.494	0.0217
$\alpha = 1.0$	Permanent	0.0; 0.0; 0.0; 0.3; 0.6; 0.6	4.34	0.1013	0.0, 0.0, 0.0, 0.5, 0.8, 0.9	12.015	0.2012

Panel C: $(p = 0.9; \gamma = 0.6; r = 0.5; y = 4)$— expanded range of taxes/subsidies

		$\tilde{\tau}_t = 0.3 \forall t$			$\tilde{\tau}_t = 0.5 \forall t$		
α	Discardable	$\tau^\star$	%ΔCS	Normalized	$\tau^\star$	%ΔCS	Normalized
$\alpha = 0$	Permanent	0.0; 0.0; 0.3; 0.4; 0.5; 0.5	2.03	0.0469	0.0, 0.0, 0.4, 0.8, 0.7, 0.8	5.22	0.1119
$\alpha = 1.0$	Permanent	0.0; 0.0; 0.4; 0.4; 0.5; 0.4	1.91	0.0477	0.0, 0.0, 0.0, 0.5, 0.8, 0.9	4.16	0.1031

Notes: Each row within a Panel presents the welfare-maximizing tax sequence $\tau^\star$ for different parameter specifications of unit price p, flow utility γ, elasticity of substitution r, and income y. The α 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.