THE OPTIMAL TAXATION OF NETWORK GOODS

Enda Patrick Hargaden
Andrew Hanson
Matthew Harris

August 2023

WORKING PAPER #2023-01

WORKING PAPER SERIES
DEPARTMENT OF ECONOMICS
HASLAM COLLEGE OF BUSINESS
http://econ.bus.utk.edu

HASLAM
COLLEGE OF BUSINESS
THE UNIVERSITY OF TENNESSEE, KNOXVILLE

THE OPTIMAL TAXATION OF NETWORK GOODS

Enda Patrick Hargaden
Andrew Hanson
Matthew Harris

August 2023

Abstract

We derive optimal tax formulas for network goods. The solution trades-off contemporaneous revenue collection against the discounted future flows of reduced network growth. We provide conditions under which the optimal tax sequence is time-invariant, and show that the rates should in general change over time. A quantitative model with consumer heterogeneity highlights patterns in these optimal sequences, and underscores the equity trade-offs.

JEL Codes: H21, H23.
Keywords: Optimal taxation, network goods, consumption externalities, atmospheric externalities.

The authors are grateful for comments and suggestions from Don Bruce, Naomi Feldman, Jim Hines, Wojciech Kopczuk, Ben Lockwood, David Madden, Ben Meiselman, Matt Murray, Bill Neilson, Cormac O'Dea, Daniel Reck, Charles Sims, and conference and seminar participants at the IIPF Annual Congress 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 to 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 like Twitter, autonomous vehicles, digital currencies like Bitcoin, and buyer-seller marketplaces like eBay and Airbnb. As of 2023, seven of the ten largest U.S. companies by market capitalization sell products that yield substantial network externalities.[1]

This paper studies the optimal taxation of such network goods. We incorporate principles from dynamic pricing models in the industrial organization literature (e.g. Katz and Shapiro, 1985) into the public finance literature on taxation with externalities. The distinguishing feature of networks when considering taxation is that the network externality directly affects the marginal utility of the network good. Therefore, any changes to tax rates not only affect revenue collection today, but by affecting the number of users, also affects the optimal Pigouvian correction in subsequent periods. We model a government levying a consumption tax to meet a revenue requirement, aware that higher tax rates discourage new users and therefore decrease network externalities. This government must thus trade-off the efficient revenue raising-rate with a dynamic Pigouvian externality correction.

In this paper, we establish foundational results on externality-producing network goods. We do so through three lenses, starting with a static framework and subsequently introducing dynamic aspects to illustrate how network and atmospheric externalities intersect. For the static lens, we formulate a model in the tradition of Sandmo (1975), and demonstrate that the formula for the optimal tax rate on network goods takes an additive, separable form — even when the good yields atmospheric externalities. In a static setting, embedding the network size (i.e., the total quantity of the network good) into the demand function for that good adjusts the optimal tax formula in an intuitive way: 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 result generalizes Kopczuk (2003) and extends it to the case where the size of the network externality affects demand.

The second lens through which we model the problem is that of a two-period theoretical model. Here, we incorporate insights from seminal work from the industrial organization (IO) dynamic pricing and network literature (e.g. Shapiro, 1983; Gul, Sonnenschein and Wilson, 1986; Jackson, 2010; Galeotti and Vega-Redondo, 2011; Fainmesser and Galeotti, 2016) and treat the size of a network as a stock variable (e.g., the number of people who subscribe to a cell phone plan). The treatment of the externality as a stock variable distinguishes this second lens from the static model outlined above. The intuition from the IO literature suggests that cultivating network growth can reduce the sensitivity of consumers to subsequent tax increases. The implications of dynamic pricing strategies when the price-setter is a profit-maximizing firm are well understood. However,

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

this concept is much more novel from the perspective of a social surplus-maximizing, revenue-constrained government. By modeling the number of owners (and hence the network externality) as a stock variable, we examine the conditions under which optimal tax rates for network goods may vary over time. In the model 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 use the implicit function theorem to derive closed-form solutions for the optimal tax rates and derive necessary conditions for the optimal tax rates to vary over time. We find that generally, the government should set a schedule of different tax rates over time for network goods. We derive sufficient conditions for an *infant industry*-style strategy to be optimal, i.e. where the government subsidizes the network good in early periods to facilitate higher taxes in later periods.

These sufficient conditions, and the bridging of the public finance and IO literatures on dynamic pricing form our second contribution. Because deadweight losses increase exponentially in taxes, conventional wisdom dating as far back as Ramsey (1927) holds that mixtures of high and low tax rates are less efficient than stable moderate rates. However, we show that when network externalities are present, initial subsidization can increase both total consumer surplus and collected revenues relative to a time-invariant baseline. Our result that the optimal tax schedule is time-varying is therefore quite unusual in commodity taxation.

While the derived conditions above show that time-varying tax schedules can increase overall surplus, they do not indicate when time-varying tax schedules *should* be implemented. A common feature of many new network goods is an externality on non-users. These "atmospheric" externalities could be of the same or opposite sign as the network externalities, and unlike in a static setting, the welfare effects of each externality are not additively separable. In other words, these potentially contrasting welfare effects on users and non-users raise the issue of whether a dynamic taxation strategy that improves revenue-generation is Pareto-improving.

The third lens through which we view the problem is a full quantitative model of the environment that allows more detailed examination of the welfare effects and efficiency of time-varying taxation on externality-producing network goods. Our model permits analysis of the intersecting externalities when some fraction of society does not use the network good, multiple time periods, and different market structures. We simulate a purchasing decision for a durable network good for $n = 10,000$ individuals over six periods. There is a one-time purchase price and a per-period consumption tax which can be negative (i.e., a subsidy). The government chooses a sequence of taxes from a set of possible sequences that satisfy an exogenous revenue requirement. The consumers in the quantitative model forecast the benefits of holding the good in future periods based on information available at the time of purchase. At any point, consumers are free to discard the good, narrowing the government's time inconsistency possibilities.[2] Mechanically, the government conducts a

[2] 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 sales tax settings, if consumers are locked into contracts or are otherwise constrained in their choices, the government may have a similar incentive to abruptly increase sales tax rates. Incorporating free disposal is a realistic way to add discipline to the model.

grid search over a discrete sequence of tax rates to maximize total surplus subject to the budget requirement.[3]

The results from the quantitative model fill in specifics that the two-period theoretical model cannot provide at such a level of generality and yield several implications for the optimal taxation of network goods. First, the stronger the positive network externalities, the greater the gains from subsidizing the good early and raising taxes later. This result holds even though consumers in the model have a Constant Elasticity of Substitution (CES) utility function.[4] Second, when there are positive network externalities but 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 also be higher taxes and compensatory transfers to non-users in later periods. Of equal or greater importance, we also emphasize that there are parts of the parameter space where time-varying taxation may *not* be useful even when the good yields network effects (e.g. if taxes are not pivotal), and can even make the public worse off. Third, we show that 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 from taxing an established network good. 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, infant-industry setting.

Our results build on several strands of the optimal tax literature. Sandmo (1975) found the optimal tax on an externality-generating good was an additive and separable weighted average of Pigouvian taxation and a form of the Ramsey rule for commodity taxation. This feature is important as it permits a taxation strategy that directly "targets" the externality-generating good under relatively general conditions (e.g. Kopczuk, 2003; Micheletto, 2008). We show the optimal tax rate of a dual atmospheric-consumption externality good has two components in common with Sandmo, and one new component. The common components are the Ramsey- and Pigou-like factors; and the new component captures the effects of the network externality on demand. Indeed our closed form expression of the optimal tax rate as a linear and separable combination of external effects can be viewed a generalization of Sandmo (1975). From a dynamic perspective, this paper fits into recent work in public finance that brings critical elements from other literatures to optimal taxation. These elements include: behavioral factors (e.g. Goldin, 2015; Farhi and Gabaix, 2020; Lockwood, 2020), dynamic settings (e.g. Barrage, 2020; Akcigit, Hanley and Stantcheva, 2022), and "modern" realities like robots and superstars (Guerreiro, Rebelo and Teles, 2022; Thuemmel, 2022; Scheuer and Werning, 2017). This paper also brings intuition from the IO literature on dynamic pricing: building network growth can reduce the sensitivity of consumers to subsequent price tax increases, but is ultimately is closer to the public finance literature like Aronsson and Johansson-

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

[4] A key feature of network good demand is that as the number of other consumers increases, demand becomes less elastic; we show that there are welfare and revenue gains from setting tax rates relatively low in early periods even with a constant elasticity of substitution.

Stenman (2018) and Eckerstorfer and Wendner (2013). The stock-like characteristics of a network also relates this paper to the large literature on optimal capital taxation (e.g. Golosov, Kocherlakota and Tsyvinski, 2003; Saez and Stantcheva, 2018). The emergence of network goods as a prominent phenomenon provides overlap to each of these areas.

From a public finance perspective, perhaps the most important conclusion of this paper is that the presence of network externalities implies a change in the optimal tax formula. A conventional wisdom among public finance economists holds that mixtures of high and low commodity tax rates are less efficient than stable moderate rates. However, we derive a sufficient condition for initial subsidization of networks to both increase total consumer surplus and collected revenues relative to a time-invariant baseline.

We believe this paper has obvious and immediate policy relevance. In the United States, sales taxes raise over \$500bn of revenue every year,[5] and more than €1,000bn is raised through the VAT in Europe.[6] Network goods constitute an increasing share of consumer expenditure. In a policy context, this paper asks if goods like stays on Airbnb should be taxed identically to stays in hotels and, if not, what are the differences. As an intermediary in a two-sided market, Airbnb generates network externalities. A public finance lens suggests goods that generate externalities should not be taxed identically to goods that do not. This paper formalises that notion.

2 Background: Examples and Context within Public Finance

This paper sits at the intersection of multiple subsets of the economics literature. It is useful therefore to clarify our contribution in context of prior work, particularly in public finance.

In this paper we examine the implications of atmospheric and network externalities for efficient taxation of commodities. Rather than restrict our analysis to finding Pigouvian solutions that encourage 'optimal' adoption rates, we focus on how a government may efficiently raise revenue from these goods. Because commodity taxation is a crucial source of revenue for many levels of government and because network-related goods in our economy are increasingly prevalent, efficiently raising revenue from network goods will become increasingly policy relevant in the coming decades.[7,8]

Indeed, goods that yield both consumption and atmospheric externalities are already common. For example, car exhausts have negative health effects (Knittel, Miller and Sanders, 2016; Currie and Walker, 2011), and higher levels of pollution may affect the willingness to be a pedestrian (Neidell, 2009). Cell phones are network goods by nature, but also have atmospheric effects: cell phones' ubiquity may affect you even if you do not own a phone.[9] The total number of cell phone users is thus a dual atmospheric-consumption externality, and the question of this paper is how

[5]U.S. Census Bureau, "2022 Quarterly Summary of State & Local Tax Revenue Tables", Table 1 for Quarter 4, `https://www.census.gov/data/tables/2022/econ/qtax/historical.Q4.html#list-tab-1312412369`.

[6]European Commission (2022) "VAT Gap in the EU", `https://data.europa.eu/doi/10.2778/01447`.

[7]It is reasonable to expect further growth in sectors such as autonomous (networked) vehicles.

[8]A simplifying assumption of a zero revenue requirement will tend to reduce the problem to a Pigouvian setup.

[9]Through increased use of video recording, or in an emergency where your phone is stolen, for example.

this duality affects the optimal tax rate for such goods. Even goods that have a nominal price of zero such as Twitter may have negative atmospheric effects (e.g. polarized political discourse, see Allcott, Braghieri, Eichmeyer and Gentzkow, 2020), but positive consumption externalities if utility is increasing in the number of users on the platform.

The tax treatment of these externalities resists a simple answer. Consider goods such as Airbnb. The cross-side externalities are obvious because a greater number and variety of available properties increases the appeal of potential Airbnb renters. A larger number of Airbnb consumers likely also yields benefits to the marginal consumer by generating property reviews and thus alleviating an asymmetric information problem. However, Airbnb also yields atmospheric externalities as visiting vacationers may be less than considerate of permanent residents' needs for peace and quiet. Because taxation of Airbnb affects its appeal to both renters and owners, taxation of Airbnb in a given period can affect the size of the market (and therefore demand for Airbnb) in subsequent periods. The presence of these network and atmospheric externalities imply that the tax treatment of Airbnb *should* differ from that of traditional hotels.[10] Similar arguments can be made network platforms such as Uber and Alibaba.

As an alternative example, consider Bitcoin. The production (mining) of Bitcoin generates pollution, and the currency is used as a medium of exchange in black market activity. Bitcoin, however, has positive network externalities as demand is increasing in the number of other users. Should we tax transactions conducted with Bitcoin differently than with traditional currencies?

An extensive IO literature has studied pricing and network effects, e.g. Shapiro (1983) and Candogan, Bimpikis and Ozdaglar (2012). Threads from this literature have been applied to other fields, e.g. health economics and the dynamic pricing of pharmaceuticals (Bhattacharya and Vogt, 2004). We investigate the effects of network externalities in an efficient commodity taxation/public finance context. Essentially, we ask "if dynamic network pricing is a well-known strategy for profit maximization, does that translate to a viable strategy for a budget constrained government seeking to raise revenues from commodity taxation?"

Per the introduction, we address this question through three lenses, the first of which is a general static optimal tax model in the spirit of Sandmo (1975). It is perhaps unsurprising that we find these goods should be taxed differently to goods with a simple external effect. The setup of the model is quite general, and we derive a closed form expression for the optimal tax rate as a linear and separable combination of atmospheric and network effects, a generalization of Kopczuk (2003). For example, for goods that yield negative atmospheric externalities, the optimal tax rate increases if the good also yields a negative network externality. When atmospheric and network externalities are of opposing directions, the sign of the optimal tax rate may be counterintuitive. Positive network effects reduce the optimal tax rate on a good, mitigating Pigouvian taxation of negative atmospheric externalities. When those consumption externalities are strong enough, we show that optimal tax policy may be a subsidy, even when the good generates negative atmospheric

[10] Acknowledging that zoning laws and hotel taxes are partially attributable to the atmospheric externalities from hotels.

externalities such as pollution or black market activity.

Second, recognizing the dynamic nature of network effects and the dynamic pricing literature, we extend our analysis to a two-period model. Focusing on the intuition of the inter-temporal problem, we impose a CES utility structure and make demand in the second period an increasing function of quantity consumed in the first period. In addition to the key takeaways from the introduction, we show time-varying tax sequences can be welfare-improving whether the government does or does not face an exogenous revenue constraint. That is, dynamic setting of rates can improve welfare over a zero tax default. This stands in contrast to a strand of the public finance literature de-emphasizing commodity taxation as a means for efficiently raising revenue. These results do not contradict this literature,[11] but highlight the potential gains from commodity taxation in this specific context.

Recognizing these potential limitations of the two-period model, our third framework is a more complete quantitative model of consumer choice, externalities, and tax rates. We again model representative consumers using a CES utility function, but this time allow utility to be stochastic. To better explore the dynamic implications of network externalities for efficient taxation, we extend the time horizon to six periods rather than two. We treat the network good as durable (matching many real world examples like cell phones) and as having an up-front cost. The consumers in our model forecast the benefits of holding the good in future periods, and exercise a free disposal option if taxes go too high. We model the surplus-maximizing sequence of tax rates that satisfy exogenous revenue constraints. We further model the supply-side as either perfectly competitive or monopolistic. The dynamic model yields several implications for efficient commodity taxation when network externalities are present. First, the stronger the positive network externalities, the greater the gains — while satisfying the revenue constraint — from subsidizing the good early and raising taxes later. Even with a constant elasticity of substitution between the network good and the numeraire good, there are welfare and revenue gains from setting tax rates relatively low in early periods. Of course, stronger network externalities give the government greater latitude to raise taxes in later periods without erstwhile consumers leaving the market. Second, the quantitative model with heterogeneous consumers shows necessary conditions for Pareto-improving policy changes. When a non-negligible fraction of the society does not use the good, they must be compensated for any negative atmospheric externalities induced by the efficient tax system. Third, we show that conditions can exist when the optimal strategy is to heavily tax the network good at all periods despite the long-run revenue potential.

This paper thus contributes to the literature in several ways. We add to the literature on optimal taxation by deriving the optimal tax rate for a commodity with dual atmospheric-consumption

[11] One of the most important results in public finance comes from Atkinson and Stiglitz (1976), which shows that under certain conditions (e.g., optimal non-linear income taxation) commodity taxation is redundant. In the presence of consumption externalities, income tax policy is unlikely to induce socially optimal consumption choices. We find that if the government has a revenue threshold it must meet over several periods, that a commodity tax policy of initial subsidization (relative to a time-invariant baseline) can be a revenue-neutral, welfare-improving strategy. As a practical matter, it seems unlikely that many governments impose an optimal non-linear income tax. Nine U.S. states, for example, have no income tax at all.

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

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

The rest of the paper proceeds as follows: Section 3.1 lays out the baseline static model, and derives the optimal tax rate when dual externalities are present. Section 3.2 derives the optimal tax rate froms a two-period model, characterizes when these rates follow an infant-industry structure, and gives an example of a numerical solution. Sections 4 and 5 clarify the restrictions imposed, the methodology of, and results from the dynamic Monte Carlo simulations. Section 6 concludes.

3 Closed-form Theory

3.1 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 4 . 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 in unaffected by pollution in a lake, the consumption response component is zero and the optimal tax rate collapses to that found by Sandmo (1975).

This can be shown more clearly by grouping the final terms in Equation (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.

3.2 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 dynamic taxation is welfare improving. However, 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. In each period the household receives and exogenous endowment of real goods labeled y_1 and y_2 respectively. The household may then choose to purchase and consume the network good (c_1 and c_2) 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 and taxed 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.

The household problem is thus,

$$\begin{aligned} \max_{z_1,z_2,c_1,c_2} & \left[\theta z_1^{\rho} + (1-\theta) f(c_1, X_1)^{\rho}\right]^{1/\rho} + \beta \left[\theta z_2^{\rho} + (1-\theta) f(c_2, X_2)^{\rho}\right]^{1/\rho} \\ & \text{subject to} \\ y_1 &\geq z_1 + (1+\tau_1) c_1 \\ y_2 + c_1 &\geq z_2 + (1+\tau_2) c_2 \end{aligned}$$

The solution to the household problem are the following necessary first order conditions:

$$\begin{aligned} \theta u_1^{1-\rho} z_1^{\rho-1} &= \lambda_1 \\ \theta u_2^{1-\rho} z_2^{\rho-1} &= \lambda_2 \\ (1-\theta) u_1^{1-\rho} f(c_1, X_1)^{\rho-1} \frac{\partial f(c_1, X_1)}{\partial c_1} &= (1+\tau_1) \lambda_1 - \beta \lambda_2 \\ (1-\theta) u_2^{1-\rho} f(c_2, X_2)^{\rho-1} \frac{\partial f(c_2, X_2)}{\partial c_2} &= (1+\tau_2) \lambda_2 \end{aligned}$$

Government Consider now the problem of a government seeking to raise revenue R by levying per-unit taxes on the network good. 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 using function $g(c)$ instead of $f(c, X)$. We thus follow a Ramsey primal approach with the government choosing τ_1 and τ_2 subject to the revenue constraint and household implementability constraints. Let μ be the Lagrange multiplier on the revenue constraint.

The government's problem is

$$\max_{\tau_1,\tau_2} \left[\theta z_1^{\rho}(\tau_1,\tau_2) + (1-\theta)\, g\left(c_1(\tau_1,\tau_2)\right)^{\rho}\right]^{1/\rho} + \beta\left[\theta z_2^{\rho}(\tau_1,\tau_2) + (1-\theta)\, g\left(c_2(\tau_1,\tau_2)\right)^{\rho}\right]^{1/\rho}$$

subject to

$$R \geq \tau_1 c_1(\tau_1,\tau_2) + \tau_2 c_2(\tau_1,\tau_2)$$

The solution consists of the following two necessary first order conditions:

$$\begin{aligned}
0 = {} & \theta u_1^{1-\rho} z_1^{\rho-1} \frac{\partial z_1(\tau_1,\tau_2)}{\partial \tau_1} \\
& + (1-\theta)\, u_1^{1-\rho} g\left(c_1(\tau_1,\tau_2)\right)^{\rho-1} \frac{\partial g\left(c_1(\tau_1,\tau_2)\right)}{\partial c_1} \frac{\partial c_1(\tau_1,\tau_2)}{\partial \tau_1} \\
& + \beta\theta u_2^{1-\rho} z_2^{\rho-1} \frac{\partial z_2(\tau_1,\tau_2)}{\partial \tau_1} \\
& + \beta(1-\theta)\, u_2^{1-\rho} g\left(c_2(\tau_1,\tau_2)\right)^{\rho-1} \frac{\partial g\left(c_2(\tau_1,\tau_2)\right)}{\partial c_2} \frac{\partial c_2(\tau_1,\tau_2)}{\partial \tau_1} \\
& - \mu\left[c_1(\tau_1,\tau_2) + \tau_1 \frac{\partial c_1(\tau_1,\tau_2)}{\partial \tau_1} + \tau_2 \frac{\partial c_2(\tau_1,\tau_2)}{\partial \tau_1}\right]
\end{aligned} \tag{16}$$

and

$$\begin{aligned}
0 = {} & \theta u_1^{1-\rho} z_1^{\rho-1} \frac{\partial z_1(\tau_1,\tau_2)}{\partial \tau_2} \\
& + (1-\theta)\, u_1^{1-\rho} g\left(c_1(\tau_1,\tau_2)\right)^{\rho-1} \frac{\partial g\left(c_1(\tau_1,\tau_2)\right)}{\partial c_1} \frac{\partial c_1(\tau_1,\tau_2)}{\partial \tau_2} \\
& + \beta\theta u_2^{1-\rho} z_2^{\rho-1} \frac{\partial z_2(\tau_1,\tau_2)}{\partial \tau_2} \\
& + \beta(1-\theta)\, u_2^{1-\rho} g\left(c_2(\tau_1,\tau_2)\right)^{\rho-1} \frac{\partial g\left(c_2(\tau_1,\tau_2)\right)}{\partial c_2} \frac{\partial c_2(\tau_1,\tau_2)}{\partial \tau_2} \\
& - \mu\left[\tau_1 \frac{\partial c_1(\tau_1,\tau_2)}{\partial \tau_2} + c_2(\tau_1,\tau_2) + \tau_2 \frac{\partial c_2(\tau_1,\tau_2)}{\partial \tau_2}\right]
\end{aligned} \tag{17}$$

Equilibrium We now solve for the equilibrium optimal tax rates. Begin with equation 16 and combine with the household first order conditions (suppressing functional dependence on taxes for notational simplicity):

$$
\begin{aligned}
0 = &\frac{\partial z_1}{\partial \tau_1} \\
&+ H(c_1, X_1)\frac{\partial c_1}{\partial \tau_1}\left[(1+\tau_1) - \Lambda\right] \\
&+ \Lambda \frac{\partial z_2}{\partial \tau_1} \\
&+ \Lambda H(c_2, X_2)\frac{\partial c_2}{\partial \tau_1}(1+\tau_2) \\
&- \mu\left[c_1 + \tau_1\frac{\partial c_1}{\partial \tau_1} + \tau_2\frac{\partial c_2}{\partial \tau_1}\right]
\end{aligned} \tag{18}
$$

Where $\Lambda \equiv \beta\frac{\lambda_2}{\lambda_1}$ is the stochastic discount factor. $H(c_1, X_1) \equiv \left[\frac{g(c_1)}{f(c_1,X_1)}\right]^{\rho-1}\left(\frac{\partial g(c_1)/\partial c_1}{\partial f(c_1,X_1)/\partial c_1}\right)$ captures a notion of the degree to which the externalities provide a non-pecuniary benefit. Assume that $X_i = c_i,\ i \in \{1,2\}$ in equilibrium so that $H(c_1, X_1) = H(c_1) = \frac{\partial g(c_1)/\partial c_1}{\partial f(c_1,X_1)/\partial c_1}$. For a network good which provides positive externalities, $H(c) > 1$. It is useful to note that the first partial $\partial H/\partial c$ depends on the relative semielasticities of the first partials.

Note that the market clearing conditions for goods are identical to the household budget constraints. Using the Implicit Function Theorem, we have the following four conditions:

$$0 = \frac{\partial z_1}{\partial \tau_1} + c_1 + (1+\tau_1)\frac{\partial c_1}{\partial \tau_1} \tag{19}$$

$$0 = \frac{\partial z_1}{\partial \tau_2} + (1+\tau_1)\frac{\partial c_1}{\partial \tau_2} \tag{20}$$

$$\frac{\partial c_1}{\partial \tau_1} = \frac{\partial z_2}{\partial \tau_1} + (1+\tau_2)\frac{\partial c_2}{\partial \tau_1} \tag{21}$$

$$\frac{\partial c_1}{\partial \tau_2} = \frac{\partial z_2}{\partial \tau_2} + c_2 + (1+\tau_2)\frac{\partial c_2}{\partial \tau_2} \tag{22}$$

Furthermore, we can assume that the government revenue constraint must bind with equality given the properties of marginal utility and free disposal. The Implicit Function Theorem thus gives

$$0 = \tau_1\frac{\partial c_1}{\partial \tau_1} + c_1 + \tau_2\frac{\partial c_2}{\partial \tau_1} \tag{23}$$

$$0 = \tau_1\frac{\partial c_1}{\partial \tau_2} + c_2 + \tau_2\frac{\partial c_2}{\partial \tau_2} \tag{24}$$

Combining equations 18, 19, 21, and 23 we get

$$
\begin{aligned}
0 = & -\frac{\partial c_2}{\partial \tau_1} - \{\Lambda [H(c_1, X_1) - 1] + 1\} \frac{\partial c_1}{\partial \tau_1} \\
& + H(c_1, X_1) \frac{\partial c_1}{\partial \tau_1} (1 + \tau_1) \\
& + \{\Lambda [H(c_2, X_2) - 1] + 1\} \frac{\partial c_2}{\partial \tau_1} (1 + \tau_2)
\end{aligned} \tag{25}
$$

Using 17, 20, 22, and 24 we get a similar equation:

$$
\begin{aligned}
0 = & -\Lambda \frac{\partial c_2}{\partial \tau_2} - \Lambda H(c_1, X_1) \frac{\partial c_1}{\partial \tau_2} \\
& + [H(c_1, X_1) - 1 + \Lambda] \frac{\partial c_1}{\partial \tau_2} (1 + \tau_1) \\
& + \Lambda H(c_2, X_2) \frac{\partial c_2}{\partial \tau_2} (1 + \tau_2)
\end{aligned} \tag{26}
$$

We can write these two equilibrium conditions in matrix form[12]:

$$
\begin{bmatrix} H_1 \frac{\partial c_1}{\partial \tau_1} & \{\Lambda [H_2 - 1] + 1\} \frac{\partial c_2}{\partial \tau_1} \\ [H_1 - 1 + \Lambda] \frac{\partial c_1}{\partial \tau_2} & \Lambda H_2 \frac{\partial c_2}{\partial \tau_2} \end{bmatrix} \begin{bmatrix} 1 + \tau_1 \\ 1 + \tau_2 \end{bmatrix} = \begin{bmatrix} \frac{\partial c_2}{\partial \tau_1} + \{\Lambda [H_1 - 1] + 1\} \frac{\partial c_1}{\partial \tau_1} \\ \Lambda \frac{\partial c_2}{\partial \tau_2} + \Lambda H_1 \frac{\partial c_1}{\partial \tau_2} \end{bmatrix}
$$

We solve for the optimal tax rates using Cramer's Rule.

$$
\tau_1^* = (\Lambda - 1) \frac{\Lambda [H_1 - 1] H_2 \frac{\partial c_1}{\partial \tau_1} \frac{\partial c_2}{\partial \tau_2} - \Lambda [H_2 - 1] \frac{\partial c_2}{\partial \tau_1} \frac{\partial c_2}{\partial \tau_2} - [H_1 - 1] [\Lambda [H_2 - 1] + 1] \frac{\partial c_2}{\partial \tau_1} \frac{\partial c_1}{\partial \tau_2}}{\Lambda H_1 H_2 \frac{\partial c_1}{\partial \tau_1} \frac{\partial c_2}{\partial \tau_2} - [H_1 - 1 + \Lambda] [\Lambda [H_2 - 1] + 1] \frac{\partial c_2}{\partial \tau_1} \frac{\partial c_1}{\partial \tau_2}} \tag{27}
$$

$$
\tau_2^* = \frac{-\Lambda H_1 [H_2 - 1] \frac{\partial c_1}{\partial \tau_1} \frac{\partial c_2}{\partial \tau_2} + \Lambda [H_1 - 1 + \Lambda] [H_2 - 1] \frac{\partial c_2}{\partial \tau_1} \frac{\partial c_1}{\partial \tau_2} - (\Lambda - 1)^2 (H_1 - 1) \frac{\partial c_1}{\partial \tau_1} \frac{\partial c_1}{\partial \tau_2}}{\Lambda H_1 H_2 \frac{\partial c_1}{\partial \tau_1} \frac{\partial c_2}{\partial \tau_2} - [H_1 - 1 + \Lambda] [\Lambda [H_2 - 1] + 1] \frac{\partial c_2}{\partial \tau_1} \frac{\partial c_1}{\partial \tau_2}} \tag{28}
$$

Along with the government revenue constraint $R = \tau_1^* c_1(\tau_1^*, \tau_2^*) + \tau_2^* c_2(\tau_1^*, \tau_2^*)$, this solves for the optimal tax rates.

Is there an equilibrium in which $(1 + \tau_2^*) > (1 + \tau_1^*)$? This occurs when

$$
\begin{aligned}
& \left[\Lambda^2 H_1 (H_2 - 1) - (H_1 - 1)(1 - \Lambda)\right] \frac{\partial c_1}{\partial \tau_2} \frac{\partial c_2}{\partial \tau_1} - (1 - \Lambda)^2 [H_1 - 1] \frac{\partial c_1}{\partial \tau_1} \frac{\partial c_1}{\partial \tau_2} > \\
& [\Lambda (1 - \Lambda) H_2 + \Lambda H_1 (\Lambda H_2 - 1)] \frac{\partial c_1}{\partial \tau_1} \frac{\partial c_2}{\partial \tau_2} + \Lambda (1 - \Lambda) [H_2 - 1] \frac{\partial c_2}{\partial \tau_1} \frac{\partial c_2}{\partial \tau_2}
\end{aligned}
$$

The following conjecture will limit the scope of our search.

Conjecture. *It must be the case that* $\frac{\partial c_1}{\partial \tau_2} \leq 0$.

[12]To further conserve notation, let $H_1 \equiv H(c_1)$ and likewise for H_2.

Proof. *From equation 20 we have that*

$$\frac{\partial z_1}{\partial \tau_2} = -\left(1+\tau_1\right)\frac{\partial c_1}{\partial \tau_2}$$

$\frac{\partial z_1}{\partial \tau_2} \geq 0$ *given both income and substitution effects operate in the same direction. Since* $1+\tau_1 > 0$, $\frac{\partial c_1}{\partial \tau_2} \leq 0$. □

The following propositions describe a sufficient, yet not necessary, set of conditions under which dynamic taxation is welfare improving. A more general set of conditions may be found in the appendix.

Proposition 1. *Dynamic taxation cannot be welfare improving if own-price elasticities are negative and* $\frac{\partial c_1}{\partial \tau_2} = 0$.

Proposition 2. *Assume* $\Lambda < 1$. *Further, assume that own-price and cross-price elasticities of the network good are negative. Then dynamic taxation is welfare improving if*

$$\begin{aligned} \frac{\partial c_1}{\partial \tau_2} &< -\frac{\Lambda\left(1-\Lambda\right)\left(H_2-1\right)\frac{\partial c_2}{\partial \tau_2}}{\left(1-\Lambda\right)\left(H_1-1\right)+\Lambda^2 H_1\left(H_2-1\right)} \\ \frac{\partial c_2}{\partial \tau_1}\times\frac{\partial c_1}{\partial \tau_2} &> \frac{1-\Lambda}{\Lambda\left(H_1H_2-1\right)-\left(H_1-1\right)} \\ \frac{\partial c_1}{\partial \tau_1} &> \frac{\left[\left(1-\Lambda\right)\left(H_1-1\right)+\Lambda H_1\left(H_2-1\right)\right]\frac{\partial c_1}{\partial \tau_2}-\Lambda\left(1-\Lambda\right)\left(H_2-1\right)\frac{\partial c_2}{\partial \tau_2}}{\left(1-\Lambda\right)^2\left(H_1-1\right)\frac{\partial c_1}{\partial \tau_2}+\Lambda\left(H_2-H_1+\Lambda\left(H_1-1\right)H_2\right)\frac{\partial c_2}{\partial \tau_2}} \end{aligned}$$

and

$$\begin{cases} H_2 \geq 1+\frac{\left(\Lambda-1\right)^2}{\Lambda} & \textit{or} \\ \begin{cases} H_1 < \frac{1-\Lambda}{1-\Lambda-\Lambda^2\left(H_2-1\right)} & \textit{and} \\ H_2 < 1+\frac{1}{\Lambda^2}-\frac{1}{\Lambda} \end{cases} \end{cases}$$

The condition cannot be satisfied if $\Lambda > 1$.

Proposition 2 states that dynamic taxation can be welfare improving if the public valuation of the good in the second period is large compared to the private valuation. In terms of our model, this is true if H_2 is large. Dynamic taxation could also be welfare improving for moderate public valuation excess in both periods. However, this condition is heavily limited by the stochastic discount factor Λ, which is in turn dependent on the own- and cross-period price elasticities of demand for the network good. Furthermore, these elasticities depend on the relative size of the public valuations H_1 and H_2. In either case it is clear that the relative magnitude of network effects plays a critical role in determining whether constant or dynamic taxation is optimal.

Although equations can be derived to solve for these elasticities (see Appendix A), the resulting system is highly nonlinear. Therefore we turn to a numerical solution for the model equilibrium.

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

$$f(c, X) = \gamma c + (1/2)\alpha X^2$$
$$g(c) = \gamma c + (1/2)\alpha c^2$$
$$\alpha \in (0, 1)$$

It is clear then that $H(c) = 1 + \frac{\alpha}{\gamma}c$, which is increasing in the amount of network good consumption. We choose this functional form to model the increasing returns to network density for consumers. Consider a situation with modest externality effects ($\alpha = 0.5$), exogenous income per period is 4 dollars, and the government needs to raise one percent of this (0.04 dollars) to meet the revenue requirement. Setting reasonable values for the other parameters of the environment (i.e. $\beta = 0.99, \theta = 0.8, \rho = 0.9, \gamma = 5$), then the welfare-maximizing tax sequence that satisfies the revenue requirement is a subsidy of 17.7 percent in the first period, followed by a tax of 204 percent in the second period. While the exact tax rates here are merely examples, and obviously dependent on parameterization, the pattern is illustrative. We see the optimal strategy is a modest initial subsidy followed by a significant subsequent tax. Further details regarding the numerical solution method may be found in the appendix.

4 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 in the tradition of Sandmo (1975), Kopczuk (2003), Micheletto (2008), and Aronsson and Johansson-Stenman (2018).[13] 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.

[13]There have been some attempts to derive closed-form solutions to related problems. For example, Greaker and Midttømme (2016) models optimal tax policy with two competing types of good, of which consumers must possess one. Their model is a (repeated) three-period game, and governments do not have an exogenous revenue constraint, so it is not clear the extent to which their findings would transfer to our setting.

4.1 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).[14] 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 3.1 and 3.2. 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. $\mathbb{E}[x_{t+1}] = x_t$. 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.

We simulate purchasing decisions for a durable good for $n = 10,000$ individuals over six periods.[15] 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 3.2, we select a CES functional form for the individuals' utility function. The individual's utility from purchasing the good can be expressed as:

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

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,

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

[15] 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.

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.[16] While we explore adoption as a permanent state in some specifications, consumers in our model generally have the option of free disposal and in this sense our results offer a lower-bound as the potential gains from time-varying taxes.

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

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

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

[16] 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.

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.[17] The probability of disposal increases if taxes are increased.[18] Under either set of assumptions about permanence or disposal cost, because consumers enjoy flow utilities from the good without having to 're-purchase' the good each period, dynamic taxation may be welfare improving without consumption/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.

4.2 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

[17] 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.

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

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 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.

5 Results from 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 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.

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.

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.

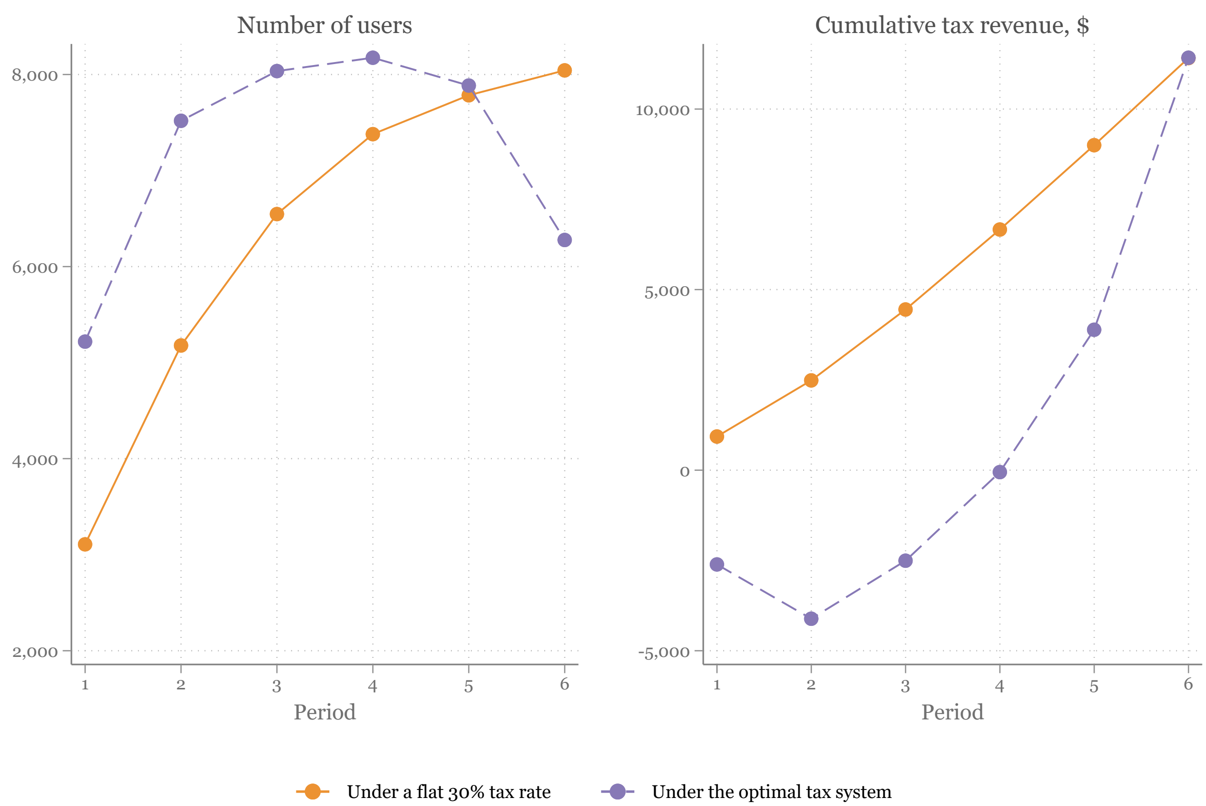
Because the optimal sequences in Panels A and B are bound by corner solutions, we relax the range of taxes in Panel C to span $[-0.5, 0.4]$ in the first three periods and $[0.3, 1.2]$ in the last three periods.[19] The bottom-left panel of Table 1 shows the optimal policy does induce government debt, and that debt persists until the fourth period. Initial adoption is very high, with about twice as many users consuming the good in the first period compared to the time-invariant baseline. Indeed, the optimal tax schedule induces more users than the time-invariant schedule through the fifth period. Only in the final period, when taxes are very high, does the number of users drop below the baseline. By this late stage, the increased number of users (and the utility boost the large network externality generates) has considerably increased welfare. The high tax rates in the final periods ensure cumulative revenue in the optimal schedule surpasses the baseline in the final period.

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 1 provides a graphical summary to supplement the findings in Table 1's Panel C. Figure 1 includes two charts, depicting behaviour in the default time-invariant case (solid line, a policy of

[19]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.

Figure 1: 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.

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 1 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 1 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 C.

5.1 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 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%

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.

(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.

5.2 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.

5.3 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:

$$\begin{aligned} \tau_t &= 0 \quad \forall t \in \{1, \ldots, 6\} \\ \tau_t &= \{-0.5, -0.5, 0, 0, 0.5, 0.5\} \end{aligned} \tag{29}$$

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 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 $\tau =$

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%

$\{-0.2, -0.2, 0.3, 0.3, 0.8, 0.8\}$. 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.

6 Conclusion

This paper studies how a planner would choose that tax rate for goods with dual atmospheric-network externalities. Beyond theoretical interest, it is likely that such products will be a large source of government revenue in coming decades.[20]

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

[20]For example, France began collecting a 3% digital services tax in December 2020.

be investigated with this model.

The solution to the model generalizes previous results from the literature, including those of Pigou (1920), Ramsey (1927), and Sandmo (1975). The tax rate comprises three additively separable factors related to substitution elasticities, the magnitude of the direct externality, and the effect of the externality on consumption behavior. Negative atmospheric externalities should be taxed, and we show that the optimal tax rate is higher if the externality lowers utility from private consumption. Equivalently, the optimal tax rate is lower if the externality increases the utility from private consumption. 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 dual atmospheric-consumption goods in the economy, we investigate if the government should tax early-stage goods differently to well-established ones. Alternatively stated, we ask if the optimal taxation of these goods is static through time. We find that it is not. 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. Incentivizing early adoption makes consumer less sensitive to subsequent tax increases, lowering excess burden in the long-run. Relative to a static baseline, initial subsidization can be revenue-neutral and welfare-enhancing.

We end on three points. Firstly, while our simulations have shown cases where time-varying taxes in general and initial subsidization in particular can delivery substantial gains, we also demonstrate several parameter values where this is not true. Secondly, the real-world pattern of internet goods only becoming part of the tax base after their widespread use can be interpreted as an approximation to the infant industry argument. 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.

References

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

Allcott, Hunt, Benjamin B Lockwood, and Dmitry Taubinsky, "Regressive sin taxes, with an application to the optimal soda tax," *The Quarterly Journal of Economics*, 2019, *134* (3), 1557–1626.

_ , 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, "Positional concerns in an OLG model: optimal labor and capital income taxation," *International Economic Review*, 2010, *51* (4), 1071–1095.

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

Atkinson, Anthony Barnes and Joseph E Stiglitz, "The design of tax structure: direct versus indirect taxation," *Journal of Public Economics*, 1976, *6* (1-2), 55–75.

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.

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

Candogan, Ozan, Kostas Bimpikis, and Asuman Ozdaglar, "Optimal pricing in networks with externalities," *Operations Research*, 2012, *60* (4), 883–905.

Currie, Janet and Reed Walker, "Traffic congestion and infant health: Evidence from E-ZPass," *American Economic Journal: Applied Economics*, 2011, *3* (1), 65–90.

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.

Farhi, Emmanuel and Xavier Gabaix, "Optimal taxation with behavioral agents," *American Economic Review*, 2020, *110* (1), 298–336.

Galeotti, Andrea and Fernando Vega-Redondo, "Complex Networks and Local Externalities: A Strategic Approach," *International Journal of Economic Theory*, 2011, *7* (1), 77–92.

Goldin, Jacob, "Optimal tax salience," *Journal of Public Economics*, 2015, *131*, 115–123.

Golosov, Mikhail, Narayana Kocherlakota, and Aleh Tsyvinski, "Optimal indirect and capital taxation," *The Review of Economic Studies*, 2003, *70* (3), 569–587.

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.

Gul, Faruk, Hugo Sonnenschein, and Robert Wilson, "Foundations of dynamic monopoly and the coase conjecture," *Journal of Economic Theory*, 1986, *39* (1), 155–190.

Jackson, Matthew O, *Social and economic networks*, Princeton University Press, 2010.

— **and Alison Watts**, "The evolution of social and economic networks," *Journal of Economic Theory*, 2002, *106* (2), 265–295.

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

Knittel, Christopher R, Douglas L Miller, and Nicholas J Sanders, "Caution, drivers! Children present: Traffic, pollution, and infant health," *Review of Economics and Statistics*, 2016, *98* (2), 350–366.

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

Lockwood, Benjamin B, "Optimal income taxation with present bias," *American Economic Journal: Economic Policy*, November 2020, *12* (4), 298–327.

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.

Neidell, Matthew, "Information, avoidance behavior, and health the effect of ozone on asthma hospitalizations," *Journal of Human Resources*, 2009, *44* (2), 450–478.

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

Ramsey, Frank P, "A contribution to the theory of taxation," *Economic Journal*, 1927, pp. 47–61.

Rees-Jones, Alex, "Quantifying loss-averse tax manipulation," *The Review of Economic Studies*, 2018, *85* (2), 1251–1278.

Saez, Emmanuel and Stefanie Stantcheva, "A simpler theory of optimal capital taxation," *Journal of Public Economics*, 2018, *162*, 120–142.

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.

Shapiro, Carl, "Optimal Pricing of Experience Goods," *The Bell Journal of Economics*, 1983, *14* (2), 497–507.

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

A Appendix: Numerical Solution to Two-Period Model

Finishing Equilibrium Characterization To fully solve the model we will need to use the Implicit Function Theorem on the implentability constraints. For ease of notation, define $E_y x$ as the elasticity of x with respect to y. For example, $E_{\tau_1} z_1 \equiv \frac{\partial z_1/\partial \tau_1}{z_1}$. Begin with

$$(1+\tau_1) = \frac{1-\theta}{\theta}\left(\frac{z_1}{f(c_1,X_1)}\right)^{1-\rho}\frac{\partial f(c_1,X_1)}{\partial c_1} + \beta\left(\frac{u_2}{u_1}\right)^{1-\rho}\left(\frac{z_2}{z_1}\right)^{\rho-1}$$

With respect to τ_1 we have

$$\begin{aligned}1 = &(1-\rho)\frac{1-\theta}{\theta}\left(\frac{z_1}{f(c_1,X_1)}\right)^{1-\rho}\frac{\partial f(c_1,X_1)}{\partial c_1}\left[E_{\tau_1}z_1 - E_{\tau_1}f(c_1,X_1) + \frac{1}{1-\rho}E_{\tau_1}\frac{\partial f(c_1,X_1)}{\partial c_1}\right] \\ &+\beta(1-\rho)\left(\frac{u_2}{u_1}\right)^{1-\rho}\left(\frac{z_2}{z_1}\right)^{\rho-1}\left[\underbrace{\frac{\partial u_2/\partial\tau_1}{u_2} - \frac{\partial u_1/\partial\tau_1}{u_1} + \frac{\partial z_2/\partial\tau_1}{z_2} - \frac{\partial z_1/\partial\tau_1}{z_1}}_{A}\right]\end{aligned} \tag{30}$$

In order to simplify A, we derive the following:

$$\begin{aligned}\frac{\partial u_2}{\partial \tau_1} &= \frac{\partial}{\partial \tau_1}\left[\theta z_2^\rho + (1-\theta) f(c_2,X_2)^\rho\right]^{1/\rho} \\ &= u_2^{1-\rho}\left[\theta z_2^\rho E_{\tau_1} z_2 + (1-\theta) f(c_2,X_2)^\rho E_{\tau_1} f(c_2,X_2)\right]\end{aligned}$$

so that

$$\frac{\partial u_2/\partial\tau_1}{u_2} = \frac{\theta z_2^\rho}{\theta z_2^\rho + (1-\theta) f(c_2,X_2)^\rho}E_{\tau_1}z_2 + \frac{(1-\theta) f(c_2,X_2)^\rho}{\theta z_2^\rho + (1-\theta) f(c_2,X_2)^\rho}E_{\tau_1}f(c_2,X_2)$$

Similarly

$$\frac{\partial u_1/\partial\tau_1}{u_1} = \frac{\theta z_1^\rho}{\theta z_1^\rho + (1-\theta) f(c_1,X_1)^\rho}E_{\tau_1}z_1 + \frac{(1-\theta) f(c_1,X_1)^\rho}{\theta z_1^\rho + (1-\theta) f(c_1,X_1)^\rho}E_{\tau_1}f(c_1,X_1)$$

Therefore A can be simplified into

$$A = \frac{(1-\theta) f(c_2,X_2)^\rho}{\theta z_2^\rho + (1-\theta) f(c_2,X_2)^\rho}\left[E_{\tau_1}f(c_2,X_2) - E_{\tau_1}z_2\right] - \frac{(1-\theta) f(c_1,X_1)^\rho}{\theta z_1^\rho + (1-\theta) f(c_1,X_1)^\rho}\left[E_{\tau_1}f(c_1,X_1) - E_{\tau_1}z_1\right]$$

Plugging this back into 30 and using $\beta\left(\frac{u_2}{u_1}\right)^{1-\rho}\left(\frac{z_2}{z_1}\right)^{\rho-1} = 1+\tau_1 - \frac{1-\theta}{\theta}\left(\frac{z_1}{f(c_1,X_1)}\right)^{1-\rho}\frac{\partial f(c_1,X_1)}{\partial c_1}$

we get

$$1-(1-\rho)(1+\tau_1)\Phi_{2,\tau_1} = (1-\rho)\frac{1-\theta}{\theta}\left(\frac{z_1}{f(c_1,X_1)}\right)^{1-\rho}\frac{\partial f(c_1,X_1)}{\partial c_1}\left[\Phi_{1,\tau_1}-\Phi_{2,\tau_1}\right] \tag{31}$$

$$\Phi_{1,\tau_1} \equiv E_{\tau_1}z_1 - E_{\tau_1}f(c_1,X_1) + \frac{1}{1-\rho}E_{\tau_1}\frac{\partial f(c_1,X_1)}{\partial c_1}$$

$$\Phi_{2,\tau_1} \equiv \frac{(1-\theta)f(c_2,X_2)^{\rho}}{\theta z_2^{\rho}+(1-\theta)f(c_2,X_2)^{\rho}}\left[E_{\tau_1}f(c_2,X_2)-E_{\tau_1}z_2\right] - \frac{(1-\theta)f(c_1,X_1)^{\rho}}{\theta z_1^{\rho}+(1-\theta)f(c_1,X_1)^{\rho}}\left[E_{\tau_1}f(c_1,X_1)-E_{\tau_1}z_1\right]$$

We have a similar result with respect to τ_2:

$$-(1-\rho)(1+\tau_1)\Phi_{2,\tau_2} = \frac{1-\theta}{\theta}\left(\frac{z_1}{f(c_1,X_1)}\right)^{1-\rho}\frac{\partial f(c_1,X_1)}{\partial c_1}\left[\Phi_{1,\tau_2}-\Phi_{2,\tau_2}\right] \tag{32}$$

$$\Phi_{1,\tau_2} \equiv E_{\tau_2}z_1 - E_{\tau_2}f(c_1,X_1) + \frac{1}{1-\rho}E_{\tau_2}\frac{\partial f(c_1,X_1)}{\partial c_1}$$

$$\Phi_{2,\tau_2} \equiv \frac{(1-\theta)f(c_2,X_2)^{\rho}}{\theta z_2^{\rho}+(1-\theta)f(c_2,X_2)^{\rho}}\left[E_{\tau_2}f(c_2,X_2)-E_{\tau_2}z_2\right] - \frac{(1-\theta)f(c_1,X_1)^{\rho}}{\theta z_1^{\rho}+(1-\theta)f(c_1,X_1)^{\rho}}\left[E_{\tau_2}f(c_1,X_1)-E_{\tau_2}z_1\right]$$

We now turn to the second implentability constraint:

$$(1+\tau_2) = \frac{1-\theta}{\theta}\left(\frac{z_2}{f(c_2,X_2)}\right)^{1-\rho}\frac{\partial f(c_2,X_2)}{\partial c_2}$$

with respect to τ_1 we get

$$0 = (1+\tau_2)\left[E_{\tau_1}z_2 - E_{\tau_1}f(c_2,X_2) + \frac{1}{1-\rho}E_{\tau_1}\frac{\partial f(c_2,X_2)}{\partial c_2}\right]$$

There is a similar condition with respect to τ_2:

$$1 = (1+\tau_2)\left[E_{\tau_2}z_2 - E_{\tau_2}f(c_2,X_2) + \frac{1}{1-\rho}E_{\tau_2}\frac{\partial f(c_2,X_2)}{\partial c_2}\right]$$

Assume that $z_2 > 0$, $\theta < 1$, and $\frac{\partial f(c_2,X_2)}{\partial c_2} \neq 0$. Then we have

$$E_{\tau_1}\frac{\partial f(c_2,X_2)}{\partial c_2} = (1-\rho)\left[E_{\tau_1}f(c_2,X_2)-E_{\tau_1}z_2\right]$$

$$1 = (1-\rho)(1+\tau_2)\left[E_{\tau_2}z_2 - E_{\tau_2}f(c_2,X_2) + \frac{1}{1-\rho}E_{\tau_2}\frac{\partial f(c_2,X_2)}{\partial c_2}\right]$$

Note that

$$E_{\tau_2} \frac{\partial f(c_2, X_2)}{\partial c_2} = \frac{\partial c_2/\partial \tau_2}{\partial c_2/\partial \tau_1} E_{\tau_1} \frac{\partial f(c_2, X_2)}{\partial c_2}$$
$$= (1-\rho)\left[E_{\tau_1} f(c_2, X_2) - E_{\tau_1} z_2\right] \frac{\partial c_2/\partial \tau_2}{\partial c_2/\partial \tau_1}$$

Thus

$$E_{\tau_1} \frac{\partial f(c_2, X_2)}{\partial c_2} = (1-\rho)\left[E_{\tau_1} f(c_2, X_2) - E_{\tau_1} z_2\right] \tag{33}$$
$$\frac{\partial c_2}{\partial \tau_1} = (1-\rho)(1+\tau_2)\left\{\frac{\partial c_2}{\partial \tau_1}\left[E_{\tau_2} z_2 - E_{\tau_2} f(c_2, X_2)\right] - \frac{\partial c_2}{\partial \tau_2}\left[E_{\tau_1} z_2 - E_{\tau_1} f(c_2, X_2)\right]\right\} \tag{34}$$

Equilibrium Characterization Given functional forms and parameters, equilibrium consists of endogenous variables $\left\{c_1, z_1, c_2, z_2, \tau_1, \tau_2, \frac{\partial c_1}{\partial \tau_1}, \frac{\partial c_2}{\partial \tau_1}, \frac{\partial z_1}{\partial \tau_1}, \frac{\partial z_2}{\partial \tau_1}, \frac{\partial c_1}{\partial \tau_2}, \frac{\partial c_2}{\partial \tau_2}, \frac{\partial z_1}{\partial \tau_2}, \frac{\partial z_2}{\partial \tau_2}\right\}$ which can be solved using

1. IFT on implementability constraints (x4)
2. Household budget constraints + IFT (x6)
3. Revenue constraint + IFT (x3)
4. Government Eulers (x2)

However in practice the system may be numerically solved searching over $\left\{\tau_1, \tau_2, \frac{\partial c_1}{\partial \tau_1}, \frac{\partial c_2}{\partial \tau_1}, \frac{\partial c_1}{\partial \tau_2}\right\}$. All others can be analytically written as functions of these five inputs. Note that equation 26 can be used to analytically solve for $\frac{\partial c_2}{\partial \tau_2}$. We assume that $\tau_1 \geq -1$, $\tau_2 > -1$, and that each of the three partials are less than zero. Equations used to solve this system are 25, 31, 32, 33, and 34. We use a derivative-free numerical optimization routine to minimize the sum of squared errors in the five equations subject to the nonlinear constraints given in Proposition 2.

B Appendix: Two Period Model Additional Propositions

Proposition 3. *Assume* $\Lambda < 1$, $\frac{\partial c_1}{\partial \tau_2} < 0$, *and* $\frac{\partial c_2}{\partial \tau_1} > 0$. *Further, assume that own-price elasticities of the network good are negative. Then dynamic taxation is welfare improving if*

$$\frac{\partial c_2}{\partial \tau_1} \times \frac{\partial c_1}{\partial \tau_2} < \frac{1-\Lambda}{\Lambda\left(H_1 H_2 - 1\right) - \left(H_1 - 1\right)}$$

$$\frac{\partial c_1}{\partial \tau_1} > \frac{\left[\left(1-\Lambda\right)\left(H_1-1\right) + \Lambda H_1\left(H_2-1\right)\right]\frac{\partial c_1}{\partial \tau_2} - \Lambda\left(1-\Lambda\right)\left(H_2-1\right)\frac{\partial c_2}{\partial \tau_2}}{\left(1-\Lambda\right)^2\left(H_1-1\right)\frac{\partial c_1}{\partial \tau_2} + \Lambda\left(H_2 - H_1 + \Lambda\left(H_1-1\right)H_2\right)\frac{\partial c_2}{\partial \tau_2}}$$

$$\frac{\partial c_1}{\partial \tau_1} < -\frac{\left(1-\Lambda\right)\left[\left(1-\Lambda\right)\left(H_1-1\right)\frac{\partial c_1}{\partial \tau_2} + \Lambda\left(H_2-1\right)\frac{\partial c_2}{\partial \tau_2}\right]\frac{\partial c_2}{\partial \tau_1}}{\left(1-\Lambda\right)^2\left(H_1-1\right)\frac{\partial c_1}{\partial \tau_2} + \Lambda\left(H_2 - H_1 + \Lambda\left(H_1-1\right)H_2\right)\frac{\partial c_2}{\partial \tau_2}}$$

$$H_2 < \frac{1}{\Lambda}$$

and

$$\begin{cases} \begin{cases} H_1 > \frac{(1-\Lambda)H_2}{1-\Lambda H_2} & \textit{and} \\ \frac{\partial c_1}{\partial \tau_2} < -\frac{\Lambda\left[\Lambda(H_1-1)H_2 + H_2 - H_1\right]\frac{\partial c_2}{\partial \tau_2}}{(1-\Lambda)^2(H_1-1)} & \textit{and} \end{cases} & \textit{or} \\ \frac{(1-\Lambda)H_2}{1-\Lambda H_2} \geq H_1 > \frac{1-\Lambda}{1-\Lambda H_2} \end{cases}$$

If $\Lambda > 1$, *the condition is satisfied if*

$$\frac{\partial c_2}{\partial \tau_1} \times \frac{\partial c_1}{\partial \tau_2} < \frac{1-\Lambda}{\Lambda\left(H_1 H_2 - 1\right) - \left(H_1 - 1\right)}$$

$$\frac{\partial c_1}{\partial \tau_1} > \frac{\left[\left(1-\Lambda\right)\left(H_1-1\right) + \Lambda H_1\left(H_2-1\right)\right]\frac{\partial c_1}{\partial \tau_2} - \Lambda\left(1-\Lambda\right)\left(H_2-1\right)\frac{\partial c_2}{\partial \tau_2}}{\left(1-\Lambda\right)^2\left(H_1-1\right)\frac{\partial c_1}{\partial \tau_2} + \Lambda\left(H_2 - H_1 + \Lambda\left(H_1-1\right)H_2\right)\frac{\partial c_2}{\partial \tau_2}}$$

$$\frac{\partial c_1}{\partial \tau_2} < \frac{\Lambda\left(H_2-1\right)\frac{\partial c_2}{\partial \tau_2}}{\left(\Lambda-1\right)\left(H_1-1\right)}$$

C 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

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.

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.