cesifo WORKING PAPERS

7852
2019

Original Version: September 2019
This Version: December 2019

Multi-Part Tariffs and Differentiated Commodity Taxation

Anna D'Annunzio, Mohammed Mardan, Antonio Russo

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7852*
Category 11: Industrial Organisation

Multi-Part Tariffs and Differentiated Commodity Taxation

Abstract

We study commodity taxation in markets where firms, such as Internet Service Providers, energy suppliers, and payment card platforms, adopt multi-part tariffs. We show that ad valorem taxes can correct underprovision and hence increase welfare, provided the government applies differentiated tax rates to the usage and access parts of the tariff. We obtain this result in different settings, including vertically interlinked markets, markets where firms adopt menus of tariffs to screen consumers and where they compete with multi-part tariffs. Our results suggest that exempting these markets from taxation may be inefficient.

JEL-Codes: D420, D610, H210.

Keywords: commodity taxation, multi-part tariffs, price discrimination.

Anna D'Annunzio
TBS Business School
1 Place Alphonse Jourdain – CS66810
France – 31068 Toulouse
dannunzio.anna@gmail.com

Mohammed Mardan
Norwegian School of Economics, Department of Business and Management Science
Helleveien 30
Norway – 5045 Bergen
mohammed.mardan@nhh.no

Antonio Russo
Loughborough University
School of Business and Economics
United Kingdom - Loughborough, LE11 3TU
a.russo@lboro.ac.uk

December 17, 2019
We thank David Agrawal, Jan Brueckner, Youssef Benzarti, Jacques Crémer, Sebastian Kessing, Marko Runkel, Stephen Smith for comments on an early version of this article. We also thank audiences at ETH Zurich 2016, University of Exeter 2016, Toulouse School of Economics 2017, LAGV conference 2016, IIPF conference 2016, CESifo PSE conference 2017, OFS Workshop on Indirect Taxes 2017, CESifo Summer Institue 2019 (Taxation in the Digital Economy), and CSEF-IGIER Symposium on Economics and Institutions 2019. We are grateful to the Editor, David Myatt, and to three anonymous referees for insightful comments that substantially improved the article. Part of this research was carried out while Anna D'Annunzio was at the CSEF (University Federico II, Naples). All errors are our own. The online appendix is available on the authors' website.

1 Introduction

Multi-part tariffs are common among telephone and Internet connection providers, energy distributors (electricity and gas), payment card platforms, and parking operators. These firms often charge consumers a fee for access in addition to a payment that depends on the amount or duration of usage. These markets are generally characterized by market power on the supply side, which is likely to result in underprovision.[1] Governments often apply indirect taxation (e.g., ad valorem and excise taxes) to the above industries and, given their importance in modern economies, the question arises of whether and how to design taxes without seriously distorting provision and reducing growth. As we argue shortly below, this question is part of an ongoing policy debate regarding the reform of indirect taxes applying to essential services. However, quite surprisingly, existing research has devoted little attention to the design of taxation in presence of multi-part tariffs.

Motivated by the above considerations, we study taxation of goods and services when providers charge multi-part tariffs. We explore a relevant dimension along which taxes can be designed to reduce their distortionary impact: the parts of the tariff to which they apply. This dimension has so far been ignored by the literature, although there are several examples of such differentiation in reality.[2] We show that the government can correct underprovision, and hence increase welfare, with a *positive* ad valorem tax on usage or access.

Our findings hinge on a simple observation. When a seller adopts linear pricing, it faces a trade-off because to gain revenue by selling marginal units, it must reduce its revenue from inframarginal ones. As a result, suppliers operate on the elastic part of the demand curve in equilibrium. Thus, if the government introduces a commodity tax (either unit or ad valorem), the suppliers' optimal response is to reduce provision. Consequently, if the good or service is underprovided (as it is usually the case with imperfect competition), taxation aggravates the distortion (Auerbach and Hines, 2002). However, when suppliers adopt multi-part tariffs, the logic governing their choice of prices is different. Typically, suppliers design the usage fees to induce the level of usage that maximizes the net surplus from consumption, which they can

[1]For example, the main U.S. cable operators hold de facto monopolies for high-speed services in several local markets. The Federal Communications Commission (FCC) reported that about 20% of households have access to a single broadband provider for a service of up to 4Mbits/s. This share rises to 30% and 55% for speeds up to 10Mbit/s and 25Mbits/s, respectively (see https://www.fcc.gov/document/chairman-remarks-facts-and-future-broadband-competition).

[2]For instance, excise taxes on either access or usage exist in the telecom sector. In several U.S. states, subscribers to wireless telecommunication services pay a separate per-line tax on top of ad valorem and other state-level taxes. Other examples include taxes on SMS, calls, handsets, and SIM cards, applied by countries such as Argentina, Brazil, Mexico, Greece, Turkey, Ukraine, and Pakistan (Katz, 2015; Matheson and Petit, 2017).

capture via the access fees (Oi, 1971). As we show, this logic implies that taxation can have counterintuitive effects, primarily because the supplier may not operate on the elastic part of demand. Hence, ad valorem taxes can increase output, provided that different tax rates are set on the "usage" and "access" part of the tariff. Thus, differentiated ad valorem taxes can produce a double dividend, by correcting the distortions due to market power while raising revenue for the government. In contrast, we do not find efficiency-enhancing effects for unit or uniform ad valorem taxes.

In Section 3, we introduce an ad-hoc model to convey the basic mechanism. We then provide foundations to this model, exploring several settings where underprovision takes place in presence of multi-part tariffs. First, we consider a model with identical consumers and a monopolist providing access to a piece of infrastructure which is essential to consume some final goods (Section 4.1). Examples include Internet service providers and payment-card systems. Although the infrastructure supplier can recover consumer surplus via the access fee, consumers are charged a usage fee higher than marginal cost in equilibrium. As shown by Economides and Hermalin (2015), restricting consumption at the margin allows the infrastructure supplier to capture part of the surplus that would otherwise accrue to the sellers of final goods. The combination of this fee with the mark-up set by final good suppliers implies that there is underprovision. We show that the government can correct this distortion with a positive ad valorem tax on usage, as long as marginal costs are not exceedingly large. If this condition holds, the equilibrium quantity lies on the inelastic part of consumer demand, implying that the supplier's optimal response to the tax is to decrease the usage fee and increase provision. We extend this model in Section 5 where we consider a duopoly of infrastructure providers, showing that our main results continue to hold.

Second, we consider taxation when the infrastructure provider offers a menu of tariffs to screen different consumer types (Section 4.2). We assume perfect competition in the final goods market and thus abstract from the distortion analyzed in the preceding setup. This setup also applies to situations where the infrastructure is not used to reach final good providers, as in the case of parking at a garage or consuming energy for heating and operating domestic appliances. In this context, the key distortion introduced by the provider is due to the provider's seeking to reduce the information rent of heavy users, which calls for underproviding the light users (Maskin and Riley, 1984). We find that an ad valorem tax on usage increases welfare as long as the supplier's gains from restricting the consumption of light users and its marginal cost are sufficiently low. When these conditions hold, the usage fees are such that consumers' demand is inelastic in equilibrium. Furthermore, taxing access can also increase efficiency in this context. Indeed, the access tax weakens the supplier's incentive to distort the consumption of light users, because the revenue collected with access fees is partly taxed away. When restricting consumption of light users has a strong effect on the information rent, the supplier responds to the access tax by reducing the usage fee intended for such users, alleviating the underprovision.

Our findings provide useful insights for the design of fiscal instruments applied to network services. In most countries, these industries are subject to ad valorem, sales, or excise taxes, but such levies typically apply uniformly to all parts of the tariffs. Our results suggest that such an implementation may worsen market distortions, so governments may consider adopting differentiated tax rates on separate tariff parts. More generally, our findings contribute to a lively policy debate on the restructuring of commodity taxation applying to essential services. A prime example is Internet access, which has attracted attention in light of the rapid digitization of economic activity (European Commission, 2014). The U.S. Congress recently passed the Permanent Internet Tax Freedom Act, which restricts taxation of Internet access services. One of the main arguments in favor of the ban was that taxes would discourage consumers from using Internet services, sapping their growth. In contrast, France has reportedly considered taxing Internet connections and downloaded data, to replace declining sources of revenue.[3] Our results show that, when appropriately designed, taxation need not result in reduced consumption.

2 Related literature

Our study relates to the longstanding literature that compares ad valorem to specific tax rates. These taxes are equivalent under perfect competition, but not under imperfect competition (Suits and Musgrave, 1955). Generally, the welfare dominance of either instrument depends on market conditions, such as the degree of concentration and whether one considers Cournot or Bertrand equilibrium (Delipalla and Keen, 1992; Skeath and Trandel, 1994; Anderson et al., 2001; Wang et al., 2018).[4] Recently, Peitz and Reisinger (2014) show that it is more

[3] The Permanent Internet Tax Freedom Act prohibits federal, state, and local governments from taxing Internet access, although some states have retained the right to impose preexisting taxes. See https://www.congress.gov/bill/114th-congress/house-bill/235. The French government proposed in 2008 "an infinitesimal sales tax on Internet access and mobile telephony" (see http://content.time.com/time/world/article/0,8599,1702223,00.html) and, more recently, a "new tax on the use of bandwidth by large operators."

[4] See Auerbach and Hines (2002) for an overview of commodity taxation in imperfectly competitive markets.

efficient to levy an ad valorem tax in the downstream than in the upstream market.[5] Wang and Wright (2017) show that ad valorem taxes allow efficient price discrimination across goods with different costs and values, unlike unit taxes. Differently from our article, this literature focuses on firms that are restricted to linear pricing. Furthermore, the comparison between tax instruments is generally about which instrument produces the smaller distortion. Instead, in our article, (differentiated) ad valorem taxes dominate because they reduce the distortions linked to the structure of the multi-part tariff.

Few articles on commodity taxation have considered multi-part tariffs. To our knowledge, none allows for different tax rates on different parts of the tariff, which is our key contribution.[6] Laffont (1987) studies taxation of a monopolist that discriminates consumers using non-linear tariffs. He shows that, when the government uses specific taxes, the welfare-optimal policy is a subsidy to increase production. Cheung (1998) shows that the dominance of ad valorem taxes established by Skeath and Trandel (1994) remains intact under a nonlinear pricing monopolist. Jensen and Schjelderup (2011) show that these results remain valid even if some consumers are excluded. They also find that ad valorem or specific taxation increases the usage fee for all consumers, but most likely reduces the access fee.[7]

An important finding of our article is that taxation can increase total surplus. Efficiency-enhancing effects of taxation have been shown in other articles, although the underlying mechanisms are different from ours. Hamilton (2009) considers multi-product transactions in retail markets, finding that excise ad valorem taxes decrease equilibrium output per product in the short-run, but increase it in the long-run (with endogenous entry). Carbonnier (2014) finds that nonlinear and price-dependent tax schedules can result in lower prices and thus higher welfare. Cremer and Thisse (1994) show, in a framework with endogenous vertical product differentiation, that a small uniform ad valorem tax lowers consumer prices and increases welfare.[8]

[5] A recent literature has pointed out an analogy between the effects of prices set by upstream firms and taxation on other firms in the supply chain (see Economides and Hermalin, 2015; Johnson, 2017; Gaudin and White, 2014). We contribute to this literature by studying the effect of fiscal policy on upstream firms.

[6] Some studies have explicitly considered differentiated taxation, though not in presence of nonlinear pricing. Cremer et al. (2001) prove that differentiated commodity tax rates are a relevant policy instrument besides income taxes for optimal tax policies, contrary to the well-known finding by Atkinson and Stiglitz (1976). Cremer and Thisse (1994) show that differentiating commodity tax rates according to the quality of the product possibly increases welfare.

[7] Some authors (e.g., De Borger, 2000) have also analyzed the design of two-part tariffs by public bodies, but not the effects of taxation on firms that adopt nonlinear pricing.

[8] The result that taxation can reduce the consumer price is referred to in the tax incidence literature as "undershifting." This phenomenon is what Edgeworth (1925) called the "Taxation Paradox." Hotelling (1932) illustrates that this result can hold with imperfect competition. Recently, Agrawal and Hoyt (2019) show in a multi-product setting that undershifting can occur when products are complements and with perfectly

The article also relates to the literature on taxation in two-sided markets. As pointed out by Rochet and Tirole (2006), in such markets "the volume of transactions [...] depends on the structure and not only on the overall level of the fees charged by the platform." Taxing one or both sides of the market (at different rates) can thus have counterintuitive effects, by affecting the structure of fees. Kind et al. (2008), show that the supply of goods provided by a two-sided platform may increase under ad valorem taxation. More recently, Bourreau et al. (2018) find that an ad valorem tax on subscriptions or on advertising may raise welfare. Belleflamme and Toulemonde (2018) show that taxation may either increase or decrease the profits of competing two-sided platforms. Tremblay (2018) distinguishes taxation at the access and the transaction level, and finds that either tax may increase welfare because of network effects. Bloch and Demange (2018) find that taxing access and data revenues at different rates is the most effective way to reduce a platform's incentive to collect data. We connect to this literature because, in our model, differentiated taxes correct distortions by altering the structure of fees as well. However, our findings do not hinge on two-sided effects.

3 A bare-bones model

In this section, we provide a stylized model of taxation in a market with multi-part tariffs, to convey the main message of our analysis. We provide foundations for this ad-hoc model in Section 4.

Consider a monopolist firm, I, providing access to a piece of infrastructure. There are Θ different types of consumers, indexed by $i = \{1, \ldots, \Theta\}$, and each group is of unit size. We assume I can observe the quantity consumed by each individual and is therefore able to charge a two-part tariff $A_i + p_i q$, where A_i is the access fee, p_i the usage fee charged to a type-i consumer, and q the quantity consumed. Firm I could be an Internet Service Provider (ISP), a distributor of natural gas or electricity, a payment card platform, or a parking operator. Accordingly, the quantity q can represent gigabytes of data, cubic meters of gas, kilowatts of electricity, card transactions, or the duration of parking, respectively.[9] The monopolist's marginal cost is $c \geq 0$.

Let $u_i(q)$ be the utility a type-i consumer derives from consuming q units, which is

competitive suppliers. The mechanism driving the results is very different than in our model.

[9]Two-part tariffs are common among energy distributors (Ito, 2014), parking operators (Inci, 2015), and payment card platforms (Bedre-Defolie and Calvano, 2013). Telecom suppliers often adopt more complex tariffs (Economides and Hermalin, 2015). However, we show in Appendices C.8 and D.1 that there is no loss of generality in restricting attention to two-part tariffs.

increasing and concave. We denote by q_i the quantities chosen (i.e., demand) by type-i consumers, given the usage fee p_i. These quantities are such that:

$$\frac{du_i}{dq} = p_i, \qquad i = 1, \ldots, \Theta. \tag{1}$$

The government can levy unit or ad valorem commodity taxes. We assume the latter can be differentiated according to the part of the tariff applied by I. Let τ be the unit tax rate, t_A the ad valorem tax rate on the access fee and t_p the ad valorem tax rate on the usage charge. The standard ad valorem tax $t_A = t_p = t$ is a special case of this tax system.[10] Firm I's net of tax profit (which we assume to be concave in p_i for $i = 1, \ldots, \Theta$) is

$$\pi_I \quad = \textstyle\sum_{i=1}^{\Theta} A_i \left(1 - t_A\right) + \left[p_i \left(1 - t_p\right) - \tau - c\right] q_i. \tag{2}$$

In a standard monopoly model (Oi, 1971), firm I can capture each consumer's net surplus, $u_i\left(q_i\right) - p_i q_i$, with the access fee. However, in more realistic settings there may be constraints on the monopolist's ability to extract consumer surplus. To convey the essence of our results, at this stage we model these constraints in an ad-hoc way, assuming there is an exogenous part of surplus, F_i, that I cannot extract from consumers of type i. A possible foundation for F_i is that part of the surplus may be captured by providers of complementary goods. For example, connection to an ISP enables consumers to reach online content (e.g., music, video, games, apps) and part of the surplus accrues to content providers (Economides and Hermalin, 2015). A second possible foundation relates to the presence of multiple consumer types: the supplier must leave an information rent to some types in order to obtain the intended self-selection (Maskin and Riley, 1984). We model the foundations for F_i in Sections 4.1 and 4.2, respectively. Given p_i and q_i, A_i satisfies the following equality

$$A_i = u_i\left(q_i\right) - p_i q_i - F_i, \qquad i = 1, \ldots, \Theta. \tag{3}$$

We assume F_i to be a function of the usage fees, p_j, for $i, j = 1, \ldots, \Theta$. Indeed, under reasonable conditions (see Section 4.1), the price of goods complementary to the network infrastructure, and thus the surplus extracted by their suppliers, decrease in I's usage fee. Furthermore, as we show in Section 4.2, when I proposes a menu of tariffs to screen consumers,

[10] Our choice of the tax base for t_A and t_p is based on several considerations. First, as will become clear, distortions in the market are due to the *structure* of the two-part tariff. Second, these taxes should be easy to calculate and implement for the tax administration. Third, excise taxes on either access or usage already exist in some countries for markets such as telecom, as we have argued in the Introduction (footnote 2).

raising the fee paid at the margin by certain types makes screening more effective, because information rents decrease. Accordingly, we assume that $\frac{\partial F_i}{\partial p_j} \leq 0,\ i,j = 1,\ldots,\Theta$.

We define welfare as the sum of profits, consumer surplus, and tax revenue, which simplifies to $\sum_{i=1}^{\Theta} u_i(q_i) - cq_i$. It is straightforward that the allocation maximizing welfare is such that $\frac{du_i}{dq} = c$. We denote the equilibrium usage fees (conditional on the tax rates) as p_i. Using equations (2) and (3), these fees satisfy the following first-order conditions:

$$\begin{aligned} \frac{d\pi_I}{dp_i} &= (1-t_A)\left[\left(\frac{du_i}{dq_i} - p_i\right)\frac{dq_i}{dp_i} - q_i - \sum_{j=1}^{\Theta}\frac{\partial F_j}{\partial p_i}\right] + \\ &+ (1-t_p)\left(p_i\frac{dq_i}{dp_i} + q_i\right) - (c+\tau)\frac{dq_i}{dp_i} = 0, \qquad i = 1,\ldots,\Theta. \end{aligned} \tag{4}$$

In the absence of taxation ($t_A = t_p = \tau = 0$), and given (1) and (4) we obtain

$$\frac{du_i}{dq} = p_i = c + \frac{\sum_{j=1}^{\Theta}\frac{\partial F_j}{\partial p_i}}{\frac{dq_i}{dp_i}}, \qquad i = 1,\ldots,\Theta, \tag{5}$$

This expression implies that, in equilibrium, consumers' marginal utility from consumption is weakly higher than the marginal cost c (given $\frac{\partial F_j}{\partial p_i} \leq 0$ and $\frac{dq_i}{dp_i} < 0$ for $i,j = 1,\ldots,\Theta$). As long as $\frac{\partial F_j}{\partial p_i} < 0$ for some j, raising p_i above c produces a net gain. Hence, I sets the usage fees in a way that restricts consumption with respect to the welfare-optimal level, to relax the constraint on the access fees, (3).

The market inefficiency identified above calls for government intervention. The focus of our subsequent analysis is whether taxation can alleviate this distortion. As a first step, we consider standard instruments that the literature has concentrated on so far (see Appendix B for the proofs of the statements that follow). Consider first the effect of unit taxes. Differentiating (4) delivers $\frac{dp_i}{d\tau} > 0$, and, hence $\frac{dq_i}{d\tau} < 0$. The intuition is that when τ rises, I faces higher costs for each unit provided. Thus, firm I's best response is to reduce total provision, which is achieved by increasing p_i.

Consider now uniform ad valorem taxation, i.e. $t_A = t_p = t$. Differentiating (4), one shows that $\frac{dp_i}{dt} \geq 0$ and therefore $\frac{dq_i}{dt} \leq 0$. The intuition is that the tax reduces the total revenue collected by the monopolist and, hence, makes provision implicitly more costly. Thus, neither unit nor uniform ad valorem taxation can alleviate the restrictions imposed by firm I, and both actually worsen them if $c > 0$. In other words, conditional on relying on these standard instruments, the welfare-maximizing policy is to subsidize the monopolist. This is a well-known result from the commodity taxation literature (Auerbach and Hines, 2002).

We now analyze the effects of differentiated ad valorem taxes. Differentiating (4) with respect to the usage tax, t_p, we get

$$\frac{dp_i}{dt_p} = \frac{\left(\frac{p_i}{q_i}\frac{dq_i}{dp_i} + 1\right) q_i}{\frac{d^2\pi_I}{dp_i^2}}, \qquad i = 1, \ldots, \Theta. \tag{6}$$

Because the denominator on the right hand side of (6) is negative by concavity of π_I, the tax t_p reduces p_i if and only if the equilibrium consumption level, q_i, is on the inelastic part of consumer demand, i.e. $-\frac{p_i}{q_i}\frac{dq_i}{dp_i} < 1$. To understand this condition, consider that t_p targets the revenue collected through the usage fee, $p_i q_i$. Therefore, firm I has an incentive to change p_i in a way that reduces such revenue and so cut its tax expenditures. If q_i is on the inelastic part of demand curve, the ensuing best response to t_p is to decrease p_i. Consequently, consumption by type-i individuals increases. We characterize the conditions on parameters such that ${dp_i}/{dt_p} < 0$ holds in Sections 4.1 and 4.2. However, to get a sense of when this inequality can hold, consider an initial equilibrium with zero taxes. Given q_i decreases with p_i, consumption by type-i individuals is more likely to be on the inelastic part of demand when neither term on the right hand side of (5) is large, that is, when the monopolist neither faces exceedingly large marginal costs nor large gains from distorting the consumption of type-i consumers.

Observe that ${dp_i}/{dt_p} < 0$ could not hold if firm I were restricted to charging linear tariffs (i.e., $A_i = 0$). Indeed, firms with market power that charge linear prices operate on the *elastic* part of demand and, thus, typically respond to ad valorem taxes by increasing prices (Auerbach and Hines, 2002).

Finally, consider the effect of t_A on p_i. Differentiating (4) and using (1), we obtain

$$\frac{dp_i}{dt_A} = \frac{\left(\frac{du_i}{dq} - p_i\right)\frac{dq_i}{dp_i} - q_i - \sum_{j=1}^{\Theta}\frac{\partial F_j}{\partial p_i}}{\frac{d^2\pi_I}{dp_i^2}} = \frac{-q_i - \sum_{j=1}^{\Theta}\frac{\partial F_j}{\partial p_i}}{\frac{d^2\pi_I}{dp_i^2}}, \qquad i = 1, \ldots, \Theta. \tag{7}$$

To understand this expression, note that an increase in t_A gives firm I an incentive to cut its tax liability by reducing its access fees. Given (3) holds in equilibrium, reducing such fees involves changing the usage fees. The sign of the required change in p_i depends on its effect on type-i consumers' net surplus $u_i - p_i q_i$ (this effect equals $-q_i$ by the envelope theorem), and on the term $\sum_{j=1}^{\Theta} F_j$, that captures the part of surplus that firm I cannot appropriate (e.g., consumers' information rent). Again, we postpone characterizing the conditions such that ${dp_i}/{dt_A} < 0$ holds to Sections 4.1 and 4.2. However, we expect that ${dp_i}/{dt_A} < 0$ when $\sum_{j=1}^{\Theta}\frac{\partial F_j}{\partial p_i}$

is large in magnitude. Summing up, to reduce the access fee in response to taxation, the firm finds it optimal to reduce the usage fee when the part of surplus that I cannot extract (and constrains the access fee) decreases with the usage fee.

To summarize, we have illustrated that differentiated ad valorem taxation can *increase* provision and welfare in presence of multi-part tariffs. To our knowledge, this is a previously unnoticed effect of commodity taxes.

4 Foundations of the bare-bones model

In this section, we provide foundations of the ad-hoc model presented above. In particular, we characterize the term F_i introduced in expression (3) and its relation to the tariffs set by the infrastructure provider. In Section 4.1, we consider a setting with identical consumers, where the infrastructure is essential to consume some final goods. In that setting, F_i is linked to the rent captured by the providers of final goods when the latter have market power. In Section 4.2, we consider taxation when the supplier offers a menu of tariffs, screening different consumer types (light and heavy users). In that setting, F_i originates from the information rent left to heavy users.

4.1 A representative consumer framework

This section builds upon the model of Economides and Hermalin (2015). We consider a unit mass of identical individuals who want to consume N goods, indexed by $j = 1, \ldots, N$. Each good is supplied by a different monopolist, also indexed by j. The utility function is

$$U = \sum_{j=1}^{N} u(q_j) + y, \tag{8}$$

where q_j are units of good j and y is a numeraire good.[11] We assume $\frac{du}{dq_j} > 0 > \frac{d^2u}{dq_j^2}$. To consume these goods, individuals need to access an essential infrastructure, provided by firm I.[12] For convenience, we refer to the good provided by I as the "infrastructure" good and to the N goods as "final" goods.

As in Economides and Hermalin (2015), firm I can be thought of as an ISP that connects

[11] We treat the N goods as independent. In Appendix C.2, we provide an alternative version of the model where these goods are imperfect substitutes.

[12] In Appendix C.3, we allow consumers to acquire goods without using firm I's infrastructure at some extra cost (for instance, consumers may prefer using cards to carrying a large amount of cash).

consumers to providers of digital content (e.g. music and video). I could also be a payment card platform that enables consumers to purchase products supplied by several retailers.[13]

Given all consumers acquire access to the infrastructure good in equilibrium (see below), the profit of final good provider j reads

$$\pi_j = (x_j - \phi)\, q_j, \quad j = 1, .., N, \tag{9}$$

where x_j denotes the price of a unit of good j and ϕ is the marginal cost, assumed symmetric for all final good providers. We assume for convenience that final good providers charge linear prices.[14]

For ease of exposition, we assume that for each unit of final good a consumer needs a fixed quantity of the infrastructure good, which we normalize to one. For example, watching a movie online entails downloading a given quantity of data (say, 1 GB). Furthermore, individuals need to make one transaction with the payment card provider per purchase. Given this assumption, consumption of $\sum_{j=1}^{N} q_j$ units of the final goods entails using the same quantity of the infrastructure good.

Firm I charges a two-part tariff of the form $T_I = A + p \sum_{j=1}^{N} q_j$, where A is the access fee and p is a per unit (usage) fee.[15] Thus, firm I's profits before tax are given by

$$\pi_I \;\; = \;\; A + (p - c) \sum_{j=1}^{N} q_j, \tag{10}$$

[13] We ignore payments from final goods providers to the infrastructure provider. In some circumstances, these payments may be limited by regulation (e.g., net neutrality rules for ISPs). However, they are common in some markets, e.g. payment card platforms charge merchants. In Appendix C.4.1, we let the usage fee p fall on final good providers. However, the physical incidence of p is irrelevant for the analysis, because this fee is similar to a tax by the infrastructure provider (Weyl and Fabinger, 2013). The provider could also charge a lump-sum access fee to the final goods suppliers. We consider this possibility in Appendix C.4.2.

[14] Linear prices are natural for final good providers such as brick-and-mortar retailers. Also, several digital content providers charge linear prices (e.g., the movie stores on iTunes and Google Play), though some adopt nonlinear tariffs (e.g., Netflix). We show in Appendix C.5 that our results are robust to this modification, as long as at least one final good provider charges linear prices. The linear formulation can also capture content providers that are ad-financed. In this interpretation, x_j can be seen as the quantity of ads per unit of content. Normalizing the advertising rate to one, $x_j q_j$ is the per-consumer ad revenue. Given that ads decrease utility, the effect on consumers is similar to that of a monetary price.

[15] In Appendix C.8, we show that firm I makes at least as much profit with a two-part tariff as with a three-part tariff of the form $T_I = A + p \cdot max \left\{0, \sum_{j=1}^{N} q_j - L\right\}$, where L is a consumption limit (i.e., a certain quantity of service bundled with access) and p is a per unit fee applying to all units over the limit (an "overage charge"). Such a three-part tariff encompasses most of the tariff structures encountered in reality. First, if $L > 0$ and p is finite, this tariff has a "loose" limit: the consumer can exceed the limit, but has to pay an overage charge per unit. Second, if $L > 0$ and $p = \infty$, the tariff has a "strict" limit, which cannot be exceeded. Obviously, when $L = 0$, the tariff has only two parts. See, e.g., arstechnica.com for examples of tariffs set by US residential ISPs that fit this description (http://tiny.cc/xp9h9y).

where c is the marginal cost of the infrastructure good.

Social welfare is defined as the sum of consumer surplus, firm profits and tax revenues. Taking into account the consumer's budget constraint, $M \geq A + p\sum_{j=1}^{N} q_j + \sum_{j=1}^{N} x_j q_j + y$, where M is the consumer's income, welfare is

$$W = \sum_{j=1}^{N} u(q_j) - (c + \phi) q_j + M. \tag{11}$$

From this expression, we obtain the socially optimal consumption levels q^* (we drop the index j because these quantities are symmetric for all final goods), characterized by $\frac{du}{dq_j} = c + \phi, \quad j = 1, .., N$. That is, the socially optimal quantities are such that marginal utility equals the combined marginal costs of provision of final and infrastructure goods.

We have shown in Section 3 that standard tax instruments (i.e. a uniform ad valorem tax or a unit tax on the infrastructure good) cannot increase welfare.[16] Because this result naturally extends to the current setting as well, we focus on differentiated ad valorem tax rates. That is, we allow for a "usage" tax rate t_p that applies to $p\sum_{j=1}^{N} q_j$, and an "access" tax rate t_A, which applies to the access payment A, where $t_k \in [-1, 1]$, $k = A, p$.

We assume the following timing of moves. First, the government sets t_A and t_p. Then, the infrastructure provider sets its tariff. Next, consumers decide whether to acquire access. Thereafter, the final good providers simultaneously decide on their prices, x_j. Finally, consumers choose q_j.[17]

We solve the model by backward induction, starting from the consumer's problem. Given the budget constraint binds, utility is

$$U = \begin{cases} \sum_{j=1}^{N} \left[u(q_j) - (p + x_j) q_j\right] - A + M & \text{if acquiring access} \\ M & \text{otherwise.} \end{cases} \tag{12}$$

To satisfy the consumer's participation constraint, the infrastructure provider can at most extract the consumer's net surplus from consumption, i.e. $A \leq \sum_{j=1}^{N} \left[u(q_j) - (p + x_j) q_j\right]$ in equilibrium. Under this condition, all consumers acquire access to the infrastructure.

[16] In Appendix C.6, we show that taxes on final goods cannot increase welfare.

[17] We consider an alternative timing, with sellers of final goods moving before I, in Appendix C.1. Changing the timing affects the pricing decisions by firm I, but not the main results regarding the effects of taxation.

Maximizing (12) with respect to q_j yields

$$\frac{du}{dq_j} = p + x_j,\ j = 1, .., N, \tag{13}$$

which defines a consumer's demand for good j, denoted $q_j(p, x_j)$. Note that, to save notation, in the following we omit this function's arguments. Clearly, the usage fee, p, and the final goods price, x_j, have the same effect on consumer demand, that is $\frac{\partial q_j}{\partial p} = \frac{\partial q_j}{\partial x_j} < 0$.

Each final good provider, $j = 1, .., N$, maximizes (9) with respect to x_j. Assuming concavity, the equilibrium price of good j (given p and the tax rates), is determined by

$$\frac{\partial \pi_j}{\partial x_j} = q_j + (x_j - \phi)\frac{\partial q_j}{\partial x_j} = 0 \Rightarrow x_j = \phi - \frac{q_j}{\frac{\partial q_j}{\partial x_j}}, \qquad j = 1, .., N. \tag{14}$$

Thus, x_j follows from the standard monopoly markup rule. Because consumers' utility is symmetric and separable, and final good suppliers are symmetric as well, the price defined by (14) is identical for all j. Henceforth, we denote the equilibrium price by x and the demand for each final good as q, dropping the index j.

Intuitively, because consumer demand for final goods depends on p as well as on x, the infrastructure provider can influence the price set by final good providers through its usage fee. The effect of p on x is ambiguous a priori. However, it is reasonable to expect that in most circumstances x decreases with p, because p reduces the surplus consumers get from each additional unit of final goods. A sufficient condition to ensure this intuitive outcome is that consumers' demand is not exceedingly convex.[18] In line with Economides and Hermalin (2015), we assume in what follows that $\partial x/\partial p < 0$. However, as we show in Appendix A.1, the combined price $p + x$ increases in p, i.e. $\frac{d(p+x)}{dp} = 1 + \frac{\partial x}{\partial p} > 0$. Thus, the overall effect of an increase in the usage fee on consumption, $\frac{dq}{dp} = \frac{\partial q}{\partial p}\left(1 + \frac{\partial x}{\partial p}\right)$, is negative.

Consider now the infrastructure provider's problem. The consumer's participation constraint, $A \leq N\left[u(q) - (p + x)q\right]$, must be binding in equilibrium, otherwise the provider could increase A without changing consumer behavior and make strictly higher profits. This constraint also implies that the infrastructure provider cannot extract the whole consumer surplus, because a part of it, Nxq, accrues to the suppliers of final goods. This latter term provides a foundation for the term F_i in expression (3) of Section 3 (given a single consumer type).

[18] Starting from (14) and using $\partial q/\partial p = \partial q/\partial x$, we get $\partial x/\partial p < 0$ if and only if $\partial q/\partial x + (x - \phi)\left(\partial^2 q/\partial x^2\right) < 0$. This condition holds as long as $\partial^2 q/\partial x^2$ is either non-positive or relatively small in magnitude.

Using the access fee we just characterized, the maximization problem of firm I can be written as

$$\max_{p} \pi_I = N\left[(1-t_A)\left(u\left(q\right)-pq-xq\right)+\left(\left(1-t_p\right)p-c\right)q\right]. \tag{15}$$

The equilibrium usage fee (conditional on t_A and t_p), satisfies the following first-order condition

$$N\left(1-t_A\right)\left[\left(\frac{du}{dq}-p\right)\frac{dq}{dp}-q-\left(x\frac{dq}{dp}+q\frac{\partial x}{\partial p}\right)\right]+N\left(1-t_p\right)\left(p\frac{dq}{dp}+q\right)-Nc\frac{dq}{dp}=0. \tag{16}$$

Note that the last term in square parentheses captures the effect of the usage fee on the part of consumer surplus, Nxq, that accrues to final goods providers. Using the equilibrium consumption behavior given by (13) in the above expression, we get

$$p=\frac{1}{1-t_p}\left(\left(t_p-t_A\right)\frac{q}{\frac{dq}{dp}}+\left(1-t_A\right)\frac{\frac{\partial x}{\partial p}q}{\frac{dq}{dp}}+c\right). \tag{17}$$

To analyze the effects of taxation, it is useful to first consider as a benchmark the case of zero taxes. Setting $t_A=t_p=0$ in (16) yields

$$\frac{du}{dq}=\frac{x\frac{dq}{dp}+q\frac{\partial x}{\partial p}}{\frac{dq}{dp}}+c. \tag{18}$$

The numerator on the right hand side of this expression corresponds to $\sum_{j=1}^{\Theta}\frac{dF_j}{dp_i}$ in expression (5). It captures the distortion that the infrastructure provider induces to extract part of the surplus otherwise accruing to the suppliers of final goods (Economides and Hermalin, 2015). Indeed, equation (18) simplifies to $p=\frac{\frac{\partial x}{\partial p}q}{\frac{dq}{dp}}+c$, which shows that, despite its ability to extract surplus via the access fee, the infrastructure provider charges a usage fee above the marginal cost. Doing so induces the suppliers of final goods to reduce their own prices (because $\partial x/\partial p<0$), so that the additional revenue from usage that firm I receives from increasing p exceeds the reduction in the access fee needed to maintain consumer participation.

Summing up, because $p>c$ and $x>\phi$ with no taxes, consumption is below the socially optimal level, i.e., $q<q^*$. We now analyze how the government can alleviate this distortion. First, we focus on the effects of the tax on access. Differentiating (16) with respect to p and t_A, using (13), yields

$$\frac{\partial p}{\partial t_A}=-\frac{dA/dp}{\partial^2\pi_I/\partial p^2}=-\frac{\left(1+\frac{\partial x}{\partial p}\right)qN}{\partial^2\pi_I/\partial p^2}>0. \tag{19}$$

Firm I responds to t_A by reducing A to cut its tax expenditures. In the current setting, the implication is that p increases. To see why, consider that, in equilibrium, A captures the consumer's net surplus, satisfying the consumer's participation constraint. This surplus decreases with p, given that the combined price of final goods, $p+x$, increases with the usage fee. In other words, I reacts to the access tax by relying less on the access fee and more on usage fee to extract surplus from consumers.

Consider now the effects of the usage tax. Totally differentiating (16) with respect to p and t_p delivers

$$\frac{\partial p}{\partial t_p} = \frac{\left(1+\frac{p}{q}\frac{dq}{dp}\right) qN}{{}^{\partial^2 \pi_I}/{}_{\partial p^2}} = \frac{\left(1+\frac{\partial x}{\partial p}\right)\left(\left(1-t_A\right)q + c\frac{\partial q}{\partial p}\right) N}{\left(1-t_p\right) {}^{\partial^2 \pi_I}/{}_{\partial p^2}}. \tag{20}$$

The first equality in (20) indicates that p decreases with t_p if and only if q lies on the *inelastic* part of consumer demand, i.e. $\frac{p}{q}\frac{dq}{dp} > -1$ holds. Under this condition, the tax base, Npq, decreases with p. Thus, *reducing* p is the provider's optimal response as t_p increases. The last equality in (20) follows from the equilibrium usage fee in (17) and from using $\frac{dq}{dp} = \frac{\partial q}{\partial p}\left(1+\frac{\partial x}{\partial p}\right)$. We therefore obtain

$$\frac{\partial p}{\partial t_p} < 0 \iff c < -\frac{\left(1-t_A\right)q}{\frac{\partial q}{\partial p}}. \tag{21}$$

Hence, if the marginal cost c is sufficiently low, an increase in the usage tax reduces the usage fee.[19] The reason is that higher marginal costs imply a higher usage fee, p, and a lower consumption level, q. Consequently, a higher marginal cost makes it less likely that q lies on the inelastic part of demand. Similarly, the usage fee is less likely to decrease with t_p the larger is t_A, because the access tax increases p (as pointed out above). Note also that the condition (21) tends to be more stringent when the cost of final goods, ϕ, increases, because this cost results in a higher price x and reduces q, all else equal.

These results imply that, in the current setting, consumption can be stimulated either by reducing the tax on access or, provided (21) holds, *increasing* the tax on usage.[20] Therefore, taxation can reduce the distortions stemming from suppliers' market power.

[19] The right hand side of (21) also depends on c, but remains strictly positive when c approaches zero. By continuity, the inequality holds when c is small enough.

[20] Although taxation affects the access fee as well, there is no effect on market participation because all consumers connect in equilibrium. In Appendix E, we consider an extension where only consumers with a sufficient valuation for access connect. Although the effects of taxing usage on the consumption of final goods do not change, there may be a reduction in the number of consumers connecting. However, a tax on access expands the number of connections. As a result, taxing usage (resp. access) is optimal when the demand for access tends to be inelastic (resp. elastic).

The effects of a change in either tax on welfare are given by

$$\frac{\partial W}{\partial t_k} = N\left(\frac{du}{dq} - c - \phi\right)\frac{dq}{dt_k} = N\left(p + x - c - \phi\right)\frac{dq}{dt_k}, \quad k = A, p. \tag{22}$$

Thus, because at equilibrium $p + x > c + \phi$ holds, the government can increase welfare by *raising* t_p, provided that the marginal cost c is small enough (as specified in (21)). This finding readily brings us to the optimal tax rates. Ideally, the government should implement the socially optimal allocation, characterized by $p + x = c + \phi$. However, given $x > \phi$, the optimum can only be achieved if p is below the marginal cost c. For the sake of exposition, we restrict our attention to equilibria where $p \geq c$.[21] Given $p+x$ increases with p, the constrained optimum is such that $p = c$. Using (17), the latter holds whenever $t_p = \frac{t_A\left(1+\frac{\partial x}{\partial p}\right)-\frac{\partial x}{\partial p}}{1+\frac{c}{q}\frac{dq}{dp}}$. There is an infinite set of pairs (t_p, t_A) that implements the constrained optimum, many of which involve positive tax rates on access and usage. However, if (21) holds, it is sufficient to tax usage only, setting $t_p^* = -\frac{\frac{\partial x}{\partial p}}{1+\frac{c}{q}\frac{dq}{dp}} > 0$ and $t_A^* = 0$.[22]

Proposition 1. *Consider a setting with homogeneous consumers, imperfect competition in the markets for final goods and a monopolist infrastructure provider adopting a multi-part tariff. If the marginal cost of infrastructure usage is sufficiently low (see* (21)*), an ad valorem usage tax increases consumption and social welfare.*

Note that a small marginal cost is fairly realistic for the main applications of this model. For instance, the cost for ISPs of delivering an additional gigabyte of data is close to zero. Similarly, there is virtually no cost of handling an additional transaction for payment card platforms.[23]

Our analysis focuses on efficiency, but we can also shed some light on the distributional consequences of taxation in this setting. It is straightforward to show that the tax on usage reduces the profit of the infrastructure supplier. The effect on final good providers' profits is instead positive as long as $\frac{\partial p}{\partial t_p} < 0$, because final good providers can charge higher prices and sell higher quantities if the usage fee decreases with t_p. To continue, the infrastructure provider captures the whole net consumer surplus. Therefore, the effect of taxation on consumers is

[21] Usage fees below marginal cost are possibly not feasible in applications such as telecom and payment cards, where marginal costs are most likely very small. Nevertheless, our conclusions about the optimal tax rates do not change if we allow for $p < c$ (see Appendix C.7).

[22] Note that when $t_A = 0$, (21) is sufficient to ensure that the denominator in the optimal tax rate, $1 + \frac{c}{q}\frac{dq}{dp}$, is positive, given $0 > \frac{dq}{dp} = \frac{\partial q}{\partial p}\left(1 + \frac{\partial x}{\partial p}\right) > \frac{\partial q}{\partial p}$.

[23] Telecom firms may face issues of network congestion, which could affect the optimal usage fees. We discuss these issues in Section 6.

zero. Note also that we have assumed that tax revenue has the same weight as the other components of social welfare. That is, the cost of public funds equals one. Assuming a larger weight would of course increase the welfare benefit of the taxes we consider. For example, the government could use their revenue to reduce other distortionary taxes in the economy. Thus, if provision increases with t_p, this tax would produce a double dividend.

4.2 Menus of tariffs and screening

In the industries where multi-part tariffs are common, such as telecom and energy distribution, firms often propose menus of tariffs with the goal of screening consumers. We now explore the effects of differentiated taxation in this context, characterizing the conditions such that imposing differentiated taxes on access and usage increases welfare. In so doing, we provide an additional foundation for the ad-hoc model in Section 3.

We again consider a monopolist infrastructure provider I that consumers use to access final goods. However, we relax the assumption of homogeneous consumers and allow them to differ in the utility they get from final goods. To keep the setup as simple as possible, we assume there is only one such good ($N = 1$). We consider two types of individuals, indexed by $i = h, \ell$, where h stands for "heavy user" and ℓ for "light user." We normalize the total number of consumers to one, denoting the share of type h by $\sigma \in (0, 1)$. Consumers' utility is $u(q, \alpha_i) + y, \quad i = h, \ell, \quad \text{with} \quad \frac{\partial^2 u}{\partial q \partial \alpha} > 0$, where the preference parameter α_i ($\alpha_h > \alpha_\ell > 0$) is private information and determines a consumer's utility from the final good. This parameter determines also the intensity of consumers' infrastructure network usage. The infrastructure provider engages in second-degree price discrimination, by proposing a menu of tariffs. We retain the same timing as in the previous section and focus again on equilibria with two-part tariffs.[24]

To concentrate on the effects of taxation when the infrastructure provider screens consumers, we assume there is perfect competition in the final good market. Therefore, the price of a unit of final good equals marginal cost, ϕ. Hence, unlike in Section 4.1, influencing the final good's price is not a motive driving the choice of tariffs by the infrastructure provider. In fact, in this setting we could ignore final goods altogether (dropping ϕ from the expressions that follow). Consequently, this model can also apply to situations where connecting to providers of final goods is not the main purpose of using the infrastructure, as in, e.g., the case of energy distributors or parking garages. Hence, one may also interpret q as the length of stay

[24] As in Section 4.1, the focus on two-part tariffs is without loss of generality (see Appendix D.1).

in a parking garage or as the amount of energy consumed for heating and operating domestic appliances.

The net utility of a type-i consumer, conditional on choosing the tariff intended for type $\tilde{i} = h, \ell$ and given the quantity q, is

$$u(q, \alpha_i) + M - A_{\tilde{i}} - (p_{\tilde{i}} + \phi) q, \quad i, \tilde{i} = h, \ell. \tag{23}$$

We denote the quantity chosen by such consumer as $q_{i\tilde{i}}$ and the ensuing gross utility, $u(q_{i\tilde{i}}, \alpha_i)$, as $u_{i\tilde{i}}$. We drop the double index when $i = \tilde{i}$ (i.e., when consumers choose the intended tariff). In equilibrium, participation and incentive compatibility constraints are satisfied and all consumers self-select into the tariff intended for their type. Hence, equilibrium consumption q_i is determined by $\frac{\partial u(q,\alpha_i)}{\partial q} = p_i + \phi$, $\forall i$. That is, marginal utility equals the sum of unit prices, $p_i + \phi$.

Tax revenue amounts to $R = t_p (p_h q_h \sigma + p_\ell q_\ell (1 - \sigma)) + t_A (A_h \sigma + A_\ell (1 - \sigma))$. Welfare is defined as the sum of tax revenue plus consumer surplus and firms' profits:

$$W = \sigma (u_h - (c + \phi) q_h) + (1 - \sigma)(u_\ell - (c + \phi) q_\ell) + M. \tag{24}$$

Hence, the socially optimal consumption levels, q_i^*, are such that the marginal utility equals the combined marginal costs, i.e. $\frac{\partial u(q,\alpha_i)}{\partial q} = c + \phi$.

The infrastructure provider's problem can be written as

$$\begin{aligned} \max_{A_h, p_h, A_\ell, p_\ell} \quad & \sigma [(1 - t_A) A_h + (1 - t_p) p_h q_h - c q_h] + (1 - \sigma) [(1 - t_A) A_\ell + (1 - t_p) p_\ell q_\ell - c q_\ell] \\ s.t. \quad & V_i \equiv u(q_i, \alpha_i) + M - A_i - (p_i + \phi) q_i \geq M, \quad i = h, \ell, \quad \text{and} \\ & V_i \geq u(q_{i\tilde{i}}, \alpha_i) + M - A_{\tilde{i}} - (p_{\tilde{i}} + \phi) q_{i\tilde{i}}, \quad i, \tilde{i} = h, \ell, \ i \neq \tilde{i}. \end{aligned}$$

The first set of constraints is the participation constraints and the second set represents the incentive compatibility constraints. We relegate the standard steps to solve this problem to Appendix A.2. In equilibrium, the participation constraint is binding for $i = \ell$, whereas the incentive compatibility constraint is binding for $i = h$. The other constraints are slack. The equilibrium access fees are thus given by

$$A_\ell = u_\ell - (p_\ell + \phi) q_\ell, \quad A_h = u_h - (p_h + \phi) q_h - (u_{h\ell} - u_\ell), \tag{25}$$

where $u_i \equiv u(q_i, \alpha_i)$ and $u_{h\ell} \equiv u(q_\ell, \alpha_h)$. The fee charged to the light users captures all their net surplus from consumption. By contrast, heavy users pay an access fee which does not capture their whole surplus: they receive some information rent to ensure they do not mimic the other type. This rent, captured by the last term in parentheses in (25), provides another foundation to the term F_i introduced in Section 3 (see (3)).

The equilibrium usage fees are

$$\begin{aligned} p_h &= \frac{1}{1-t_p}\left((t_p - t_A)\frac{q_h}{\frac{\partial q_h}{\partial p_h}} + c\right), \\ p_\ell &= \frac{1}{1-t_p}\left((t_p - t_A)\frac{q_\ell}{\frac{\partial q_\ell}{\partial p_\ell}} + (1-t_A)\frac{\sigma}{1-\sigma}\left(\frac{\partial u_{h\ell}}{\partial q_\ell} - \frac{\partial u_\ell}{\partial q_\ell}\right) + c\right). \end{aligned} \tag{26}$$

To analyze the effects of taxation, it is useful to first consider as a benchmark the equilibrium with zero taxes. Setting $t_A = t_p = 0$, we get

$$p_h = c, \quad \text{and} \quad p_\ell = \frac{\sigma}{1-\sigma}\left(\frac{\partial u_{h\ell}}{\partial q_\ell} - \frac{\partial u_\ell}{\partial q_\ell}\right) + c. \tag{27}$$

In the absence of taxation, firm I does not distort heavy users' consumption as compared to the socially optimal level. By contrast, the firm does impose a restriction to the light users, who pay a usage fee above marginal cost, as captured by the right hand side of the expression for p_ℓ in (27). This result stems from the trade-off between rent-extraction and efficiency (Maskin and Riley, 1984). By restricting light users' consumption, the infrastructure provider makes mimicking less appealing to heavy users, that value the marginal unit of consumption more than light users, given $\frac{\partial^2 u}{\partial q \partial \alpha} > 0$. In doing so, firm I reduces heavy users' information rent and can thus extract more revenue from them through the access fee, A_h (see (25)). As a result, there is underprovision (only) to the light users because $p_\ell > c$ and, in turn, $q_\ell < q_\ell^*$.

We now analyze whether taxation can correct this underprovision. The effects of the access tax, t_A, on the usage fees are given by

$$\frac{\partial p_h}{\partial t_A} = \sigma \frac{\partial A_h / \partial p_h}{\partial^2 \pi_I / \partial p_h^2} = -\frac{\sigma q_h}{\partial^2 \pi_I / \partial p_h^2} > 0, \tag{28}$$

$$\frac{\partial p_\ell}{\partial t_A} = \frac{(1-\sigma)\,\partial A_\ell / \partial p_\ell + \sigma \partial A_h / \partial p_\ell}{\partial^2 \pi_I / \partial p_\ell^2} = -(1-\sigma)\frac{q_\ell + \frac{\sigma}{1-\sigma}\left(\frac{\partial u_{h\ell}}{\partial q_\ell} - \frac{\partial u_\ell}{\partial q_\ell}\right)\frac{\partial q_\ell}{\partial p_\ell}}{\partial^2 \pi_I / \partial p_\ell^2}. \tag{29}$$

To understand these effects, consider that firm I responds to the access tax by reducing the access fees and thus cutting its tax liability. The access fees are pinned down, respectively, by the binding participation (for $i = \ell$) and incentive compatibility (for $i = h$) constraints, resulting in (25). Hence, to reduce A_h and A_ℓ, firm I has to modify the usage fees. Specifically, a higher p_h reduces A_h, because p_h reduces the net surplus of heavy users, $u_h - (p_h + \phi)\, q_h$. However, p_h has no effect on A_ℓ. Therefore, to reduce the access fee intended for heavy users, firm I finds it optimal to increase p_h and thus we obtain ${}^{\partial p_h}/_{\partial t_A} > 0$. Taxing access gives the infrastructure supplier an incentive to collect more revenue from heavy users through the usage fee.

By contrast, collecting less revenue through access fees does not necessarily entail an increase in the usage fee for light users. On the one hand, a higher p_ℓ reduces these users' net surplus, which in turn reduces A_ℓ. On the other hand, p_ℓ reduces consumption of heavy users who mimic as well. Therefore, the information rent decreases, which raises A_h. These opposing effects are captured, respectively, by the first and second terms in the numerator of the rightmost fraction in (29). The sign of ${}^{\partial p_\ell}/_{\partial t_A}$ depends on which of them dominates. When distorting the consumption of light users strongly affects the heavy users' information rent, firm I responds to the tax on access by reducing p_ℓ. In other words, the incentive to restrict the consumption of light users is weakened, because the gain from reducing heavy users' information rent is partly taxed away.

Consider now the effects of the tax on usage. Given (26), these can be written as

$$\frac{\partial p_h}{\partial t_p} = \sigma \frac{q_h + p_h \frac{\partial q_h}{\partial p_h}}{{}^{\partial^2 \pi_I}/_{\partial p_h^2}} = \sigma \frac{(1 - t_A)\, q_h + c \frac{\partial q_h}{\partial p_h}}{{}^{\partial^2 \pi_I}/_{\partial p_h^2}\, (1 - t_p)}, \tag{30}$$

$$\frac{\partial p_\ell}{\partial t_p} = (1-\sigma)\, \frac{q_\ell + p_\ell \frac{\partial q_\ell}{\partial p_\ell}}{{}^{\partial^2 \pi_I}/_{\partial p_\ell^2}} = -\,(1-\sigma)\, \frac{(1 - t_A)\left[q_\ell + \frac{\sigma}{1-\sigma}\left(\frac{\partial u_{h\ell}}{\partial q_\ell} - \frac{\partial u_\ell}{\partial q_\ell}\right)\frac{\partial q_\ell}{\partial p_\ell}\right] + c \frac{\partial q_\ell}{\partial p_\ell}}{{}^{\partial^2 \pi_I}/_{\partial p_\ell^2}\, (1 - t_p)}. \tag{31}$$

As in our previous settings, the effect of t_p on each usage fee, p_i, is negative if and only if q_i lies on the inelastic part of type-i consumers' demand. The condition such that demand of heavy users is locally inelastic is similar to (21): the marginal costs and the tax on access (which tends to increase p_h, as pointed out above) must be sufficiently small. The condition determining the sign of ${}^{\partial p_\ell}/_{\partial t_p}$ is more interesting. Expression (31) suggests that even if marginal costs and t_A are zero, p_ℓ can increase with t_p when the effect of this fee on the information rent left to heavy users, captured by $\frac{\sigma}{1-\sigma}\left(\frac{\partial u_{h\ell}}{\partial q_\ell} - \frac{\partial u_\ell}{\partial q_\ell}\right)\frac{\partial q_\ell}{\partial p_\ell}$, is large. Furthermore, a higher t_A makes it more likely that ${}^{\partial p_\ell}/_{\partial t_p} > 0$ if and only if the term in square parentheses in (31) is positive, i.e. $\frac{\sigma}{1-\sigma}\left(\frac{\partial u_{h\ell}}{\partial q_\ell} - \frac{\partial u_\ell}{\partial q_\ell}\right)\frac{\partial q_\ell}{\partial p_\ell}$ is small.

We now analyze whether taxation can increase welfare. Differentiating (24) with respect to t_k and given the equilibrium choice of q_i, we obtain

$$\frac{\partial W}{\partial t_k} = \sigma\left(p_h - c\right)\frac{dq_h}{dt_k} + (1-\sigma)\left(p_\ell - c\right)\frac{dq_\ell}{dt_k}, \quad k = A, p, \tag{32}$$

where $\frac{dq_i}{dt_k} = \frac{\partial q_i}{\partial p_i}\frac{\partial p_i}{\partial t_k}$, with $i = \ell, h$. Furthermore, setting $t_p = t_A = 0$ and using (27), we get

$$\left.\frac{\partial W}{\partial t_k}\right|_{t_A=0, t_p=0} = \sigma\left(\frac{\partial u_{h\ell}}{\partial q_\ell} - \frac{\partial u_\ell}{\partial q_\ell}\right)\frac{dq_\ell}{dt_k} > 0 \Longleftrightarrow \frac{\partial p_\ell}{\partial t_k} < 0, \quad k = A, p. \tag{33}$$

Because there is underprovision to the light users, the government can increase welfare by introducing either a tax on access or one on usage, provided p_ℓ decreases with such taxes. That is, the government can alleviate the restriction firm I imposes on light users by taxing access if the impact of p_ℓ on the information rent of heavy users is large. Instead, if the impact of p_ℓ on the information rent and the marginal cost are small, a similar effect can be obtained by taxing usage. Based on these results, we can state:

Proposition 2. *Consider a setting where a monopolist infrastructure provider implements a menu of multi-part tariffs to screen consumers. If the marginal cost of infrastructure usage and the effect of restricting the consumption of light users on heavy users' information rent are small, introducing an ad valorem usage tax increases welfare. Instead, if restricting the consumption of light users has a large effect on the information rent, welfare increases with an ad valorem access tax.*

A small marginal cost is a reasonable assumption for ISPs, credit card platforms, and parking providers. In line with Proposition 1, Proposition 2 suggests that this condition makes it more likely that a tax on usage increases output and welfare. However, the condition is not sufficient: the impact of restricting usage by light users on the information rent must also be relatively small. We can expect the latter condition to apply when either the share of heavy users in the population, σ, is small and/or when their marginal utility from consumption is similar to that of light users (i.e., the difference between α_h and α_ℓ is small). By contrast, a tax on access is more likely to increase welfare when restricting consumption by light users has a large effect on the information rent.

We can again shed some light on the distributional implications of taxation. It is straightforward to show that taxes reduce the profit of the infrastructure provider and leave the profit of final good providers unchanged. Moreover, because the surplus of light users is entirely captured by the provider, the net effect of taxation on light users is also zero. However, when taxes induce an increase in consumption by the light users, the information rent left to heavy users increases and they are better off.[25] Note also that, as in the previous section, we ignore the potential welfare benefits from the tax revenue that government could eventually use to reduce other distortionary taxes in the economy.

We conclude this part by discussing the implications of relaxing some of our assumptions. An infrastructure provider may also screen consumers on the quality of service (e.g. download speed in the case of ISPs). Adding this dimension to the model should not change the main conclusions. We expect the heavy users to have a higher willingness to pay for high-quality service. Hence, light users would be offered plans with lower quality as well as restricted volumes. The effect of the usage fee on information rents could be either magnified or reduced, depending on whether quality and volume are complements or substitutes.

Finally, we assumed I serves both consumer types. However, in some circumstances, the firm may prefer to exclude the light users, because serving them entails an exceedingly high rent to the heavy ones. This outcome can occur, for example, if the light users' willingness to pay is substantially smaller than that of the heavy users. Exclusion may also become more profitable with taxation. Nevertheless, provided both types are served in laissez faire, there should always exist positive (possibly small) tax rates such that welfare increases and exclusion does not take place.

5 Competition among infrastructure providers

We now extend the model of Section 4.1 to allow for competition among infrastructure providers. For brevity, we relegate much of the analysis to Appendix F, and focus here on describing the model setup and the results.

We consider a horizontally differentiated duopoly, with infrastructure providers located at the extremes of a Hotelling line, i.e. firm m is located at point 0 and firm n at point 1. The unit mass of consumers is uniformly distributed on the $[0,1]$ interval.[26] The utility a consumer located at point $z \in [0,1]$ gets from connecting via supplier $s = m, n$ and consuming q units of final good is

[25]To understand, replace A_h from (25) in (23), obtaining $U_h = u_{h\ell} - u_\ell$, and note that $\frac{\partial u_{h\ell}}{\partial q_\ell} - \frac{\partial u_\ell}{\partial q_\ell} > 0$.

[26]Horizontal differentiation may arise from differences in technologies (e.g. Cable vs DSL for ISPs, fossil vs. renewable sources for energy distributors) and other services bundled with access (e.g. TV channels bundled with Internet access). Horizontal differentiation may also capture heterogeneous coverage, which is quite common for network industries such as telecom. Depending on their physical location, some consumers may prefer using one provider rather than the other, because better coverage implies higher quality and reliability of service. See, e.g., Chen and Savage (2011), Dessein (2003), and Granier and Podesta (2010) for previous models of competition in the telecom and energy sectors that assume horizontally differentiated providers.

$$U_s = \int_0^q (\alpha - r)\, dr + V - \beta\,|z - l_s| + M - T_s - T_{f,s}, \quad s = m, n. \tag{34}$$

The first term of (34) is the utility from consuming the final good. To ease notation, let there be only one such good ($N = 1$). To simplify the analysis, we adopt a specific form for such utility (as in Economides and Hermalin, 2015). We assume $\alpha > \phi + c$. The second term in (34), $V - \beta\,|z - l_s|$, is the utility of a consumer at location z acquiring access from s, where V is the gross intrinsic surplus from acquiring access, β is the transportation cost, and l_s represents the firm's position on the Hotelling line, with $l_m = 0$ and $l_n = 1$. We assume consumers acquire access from exclusively one provider and focus on the case where their surplus is large enough that all acquire access in equilibrium. The last two terms in (34) are the tariffs paid to infrastructure provider s and to the supplier of the final good, respectively. We again focus on two-part tariffs set by infrastructure providers, $T_s = A_s + p_s q$. Instead, the provider of the final good sets a linear price, so $T_{f,s} = x_s q$. Note that we allow the latter firm to charge a different price to consumers according to which of the providers they subscribe to. This assumption is not essential, but simplifies the analysis (see below).[27]

Social welfare is the sum of consumer surplus, profits, and government revenue. This sum simplifies to

$$\begin{aligned} W \;=\; & \int_0^{D_m} \left(\int_0^{q_m} (\alpha - r)\, dr - z\beta - (c + \phi)\, q_m \right) dz + \\ & + \int_{D_m}^1 \left(\int_0^{q_n} (\alpha - r)\, dr - (1 - z)\,\beta - (c + \phi)\, q_n \right) dz + M. \end{aligned} \tag{35}$$

where D_s is the market share of infrastructure provider s and q_s the quantity consumed by an individual acquiring access from provider s.

To derive the socially optimal allocation, we maximize (35) with respect to q_m, q_n, and D_m, obtaining $q_s^* = \alpha - \phi - c$ and $D_m^* = D_n^* = \frac{1}{2}$. Consumers choose the infrastructure provider that is closest to their location, and their consumption of the final good is such that marginal utility equals the combined marginal cost of provision by the infrastructure and final good providers.

Setting $t_A = t_p = 0$, the equilibrium prices and quantities are such that: $p_s = \frac{\alpha - \phi + 2c}{3}$, $x_s = \frac{\alpha + 2\phi - c}{3}$, $q_s = \frac{\alpha - \phi - c}{3}$, and $D_s = \frac{1}{2}$, for $s = m, n$. In the absence of taxation, consumers

[27] There are examples of such price discrimination. Streaming music services Spotify and Deezer offer special rates to subscribers of specific ISPs (e.g. Orange in France), whereas Netflix allows Comcast's Xfinity subscribers free access to its content for limited time periods (such as the "Watchathon" week).

pay a usage fee above the marginal cost of infrastructure usage, c. The intuition is again that raising the usage fee allows to extract more surplus from consumers, because it induces the final good provider to reduce its own markup. Given this fee and the mark-up imposed by the final good supplier, there is underprovision in equilibrium ($q_s < q_s^*$).

We find that taxation produces similar effects as in the model of Section 4.1. In particular, introducing a tax on usage t_p increases consumption as long as the marginal costs are not exceedingly large. Although the infrastructure providers raise the access fee, this increase has only distributional but not welfare consequences, because the market is fully covered.

Proposition 3. *With competing infrastructure providers that adopt multi-part tariffs, an ad valorem usage tax increases consumption and social welfare as long as the marginal cost of infrastructure usage is sufficiently small.*

Before concluding, we briefly comment on relaxing the assumption of price discrimination by the final good supplier. When an infrastructure provider, say m, raises its usage fee above marginal cost, it induces a reduction in the price of the final good. Suppose the final good supplier does not discriminate consumers according to which infrastructure provider they connect to. Then, m does not gain from such price reduction, because consumers get to benefit even if they connect to the other provider, n. Thus, both providers would set $p_s = c$ in equilibrium. Nevertheless, because the final good supplier charges a monopoly markup, there is still underprovision without taxes. Consequently, provided it induces a reduction in p_s, the tax t_p still increases welfare. The analysis of this case is available upon request.

6 Concluding remarks

We studied commodity taxation with multi-part tariffs, allowing for differentiated ad valorem taxes on the various parts of the tariff. We have modeled different market situations where multi-part tariffs result in underprovision, showing that consumption and welfare can increase with differentiated taxes. Our results imply that tax exemptions for goods or services that are priced according to multi-part tariffs may be inefficient.

Although we have shown our results hold in several specifications, we briefly discuss some issues that were left out of the analysis. We have ignored consumption externalities such as pollution and congestion. The former is relevant for some applications of the model, such as energy. If the supplier does not internalize the externality, the welfare gains of increased provision must be weighed against the increased external costs. Congestion often takes place on digital and transport networks. Unlike pollution, this external cost should be internalized by the supplier (particularly in the case of a monopoly) and therefore reflected in the usage fee. This effect implies a larger usage fee, which in turn means that the conditions ensuring that the usage fee decreases with taxation become stricter (see (6)). Whether these conditions hold depends of course on how large congestion is, which is an empirical question. We note that, as shown by Economides and Hermalin (2015), underprovision may occur even if the supplier internalizes congestion. Hence, the taxes we study may still increase welfare.

Our analysis also abstracted from long-run issues. By reducing the net revenue collected from each consumer, taxation may reduce an infrastructure supplier's incentives to invest in capacity and service improvements. On the other hand, if taxation does increase usage, incentives to invest may be strengthened. Moreover, final good suppliers' investments may increase. In addition, tax revenue can be used to fund public initiatives to enhance infrastructure investment (e.g., universal service funds; see OECD, 2015 for an overview of these initiatives). We have also ignored how the market structure may be affected by taxes in the long run. Because taxation reduces profits, it may discourage entry in the infrastructure market. However, if taxes stimulate usage, they indirectly benefit the providers of the final goods, possibly leading to additional entry. A complete analysis of the effects of taxation in a dynamic context is left for future work.

References

[1] Agrawal D.R. and Hoyt W.H.. "Tax Incidence in a multi-product world: Theoretical foundations and empirical implications." Available at SSRN: https://ssrn.com/abstract=3173180, (2019).

[2] Anderson S., De Palma A. and Kreider B.. "The efficiency of indirect taxes under imperfect competition." Journal of Public Economics, Vol. 81 (2001), pp. 231-251.

[3] Atkinson A.B. and Stiglitz J.E.. "The design of tax structure: Direct versus indirect taxation." Journal of Public Economics, Vol. 6 (1976), pp. 55-75.

[4] Auerbach A.J. and Hines J.. "Taxation and economic efficiency, in A.J. Auerbach and M. Feldstein (eds.), Handbook of Public Economics, Amsterdam, Elsevier, 2002.

[5] Bedre-Defolie O. and Calvano E.. "Pricing payment cards." American Economic Journal: Microeconomics, Vol. 5 (2013), pp. 206-231.

[6] Belleflamme P. and Toulemonde E.. “Tax incidence on competing two-sided platforms: lucky break or double jeopardy.” Journal of Public Economic Theory, Vol. 20 (2018), pp. 9-21.

[7] Bloch F. and Demange G.. “Taxation and privacy protection on Internet platforms.” Journal of Public Economic Theory, Vol. 20 (2018), pp. 52-66.

[8] Bourreau M., Caillaud B. and de Nijs R.. “Taxation of a digital monopoly platform.” Journal of Public Economic Theory, Vol. 20 (2018), pp. 40-51.

[9] Carbonnier C.. “The incidence of non-linear price-dependent consumption taxes.” Journal of Public Economics, Vol. 118 (2014), pp. 111-119.

[10] Chen Y. and Savage S.. “The effects of competition on the price for cable modem Internet access.” Review of Economics and Statistics, Vol. 93 (2011), pp. 201-217.

[11] Cheung F.K.. “Excise taxes on a non-uniform pricing monopoly: Ad valorem and unit taxes compared.” Canadian Journal of Economics, Vol. 31 (1998), pp. 1192-1203.

[12] Cremer H., Pestieau P., and Rochet J.-C.. “Direct versus indirect taxation: The design of the tax structure revisited.” International Economic Review, Vol. 42 (2001), pp. 781-800.

[13] Cremer H. and Thisse J.-F.. “Commodity taxation in a differentiated oligopoly.” International Economic Review, Vol. 35 (1994), pp. 613-633.

[14] De Borger B.. “Optimal two-part tariffs in a model of discrete choice.” Journal of Public Economics, Vol. 76 (2000), pp. 127-150.

[15] Delipalla S. and Keen M.J.. “The comparison between ad valorem and specific taxation under imperfect competition.” Journal of Public Economics, Vol. 49 (1992), pp. 351-367.

[16] Dessein W.. “Network competition in nonlinear pricing.” RAND Journal of Economics, Vol. 34 (2003), pp. 593-611.

[17] Economides N. and Hermalin B.. “The strategic use of download limits by a monopoly platform.” RAND Journal of Economics, Vol. 46 (2015), pp.297-327.

[18] Edgeworth F.Y.. Papers Relating to Political Economy. London, UK: MacMillan, 1925.

[19] European Commission. “Report of the Commission expert group on taxation of the digital economy.” 2014.

[20] Gaudin G. and White A.. "The Antitrust Economics of the Electronic Books Industry." Mimeo, 2014.

[21] Granier L. and Podesta M.. "Bundling and Mergers in Energy Markets." Energy Economics, Vol. 32 (2010), pp. 1316-1324.

[22] Hamilton S.F.. "Excise taxes with multiproduct transactions." American Economic Review, Vol. 99 (2009), pp. 458-471.

[23] Hotelling H.. "Edgeworth's Taxation Paradox and the Nature of Demand and Supply Functions." Journal of Political Economy, Vol. 40 (1932), pp. 577-616.

[24] Inci E.. "A review of the economics of parking. Economics of Transportation." Vol. 4 (2015), pp. 50-63.

[25] Ito, K.. "Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing." American Economic Review, Vol. 104 (2014), pp. 537-63.

[26] Jensen S. and Schjelderup G.. "Indirect taxation and tax incidence under nonlinear pricing." International Tax and Public Finance, Vol. 18 (2011), pp. 519-532.

[27] Johnson J.. "The Agency Model and MFN Clauses." Review of Economic Studies, Vol. 84 (2017), pp. 1151–1185.

[28] Katz R.. "The impact of taxation on the digital economy." GSR discussion paper, ITU, 2015.

[29] Kind H.J., Köthenbürger M. and Schjelderup G.. "Efficiency-enhancing taxation in two-sided markets." Journal of Public Economics, Vol. 92 (2008), pp. 1531-1539.

[30] Laffont J.-J.. "Optimal taxation of a non-linear pricing monopolist." Journal of Public Economics, Vol. 33 (1987), pp. 137-155.

[31] Laffont J.-J. and Martimort D.. "The theory of incentives - The principal-agent model." Princeton University Press, 2001.

[32] Maskin E. and Riley E.. "Monopoly with incomplete information." RAND Journal of Economics, Vol. 2 (1984), pp. 171-196.

[33] Matheson, T. and Petit P.. "Taxing Telecommunications in Developing Countries." IMF Working Papers 17/247, 2017.

[34] OECD. "Development of high speed networks and the role of municipal networks." OECD Publishing, Paris, 2015.

[35] Oi W.. "A Disney dilemma: two-part tariffs for a Mickey Mouse monopoly." Quarterly Journal of Economics, Vol. 85 (1971), pp. 77-90.

[36] Peitz M. and Reisinger M.. "Indirect taxation in vertical oligopoly." Journal of Industrial Economics, Vol. 62 (2014), pp. 709-755.

[37] Rochet, J. C. and Tirole J.. "Two-Sided Markets: A Progress Report." RAND Journal of Economics, Vol. 35 (2006), pp. 645-67.

[38] Skeath S.E. and Trandel G.A.. "A Pareto comparison of ad valorem and unit taxes in noncompetitive environments." Journal of Public Economics, Vol. 53 (1994), pp. 53-71.

[39] Suits D.B. and Musgrave R.A.. "Ad valorem and unit taxes compared." Quarterly Journal of Economics, Vol. 67 (1955), pp. 598-604.

[40] Tremblay M.. "Taxing platform markets: Transaction vs. access taxes." Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2640248, 2018.

[41] Wang K.-C. A., Chou P.-Y., and Liang W.-J.. "Specific versus ad valorem taxes in the presence of cost and quality differences." International Tax Public Finance, Vol. 25 (2018), pp. 1197-1214.

[42] Wang Z. and Wright J.. "Ad valorem platform fees, indirect taxes, and efficient price discrimination." RAND Journal of Economics, Vol. 48 (2017), pp. 467-484.

[43] Weyl E.G. and Fabinger M.. "Pass-through as an economic tool: Principles of incidence under imperfect competition." Journal of Political Economy, Vol. 121 (2013), pp. 529-583.

Appendix

A Proofs

A.1 Proof that the combined price of final goods increases with the usage fee

We differentiate $p + x$ with respect to p, and get $d\,(p + x(p))\,/dp = 1 + \partial x/\partial p = (\partial q/\partial p)\,/\left(2\frac{\partial q}{\partial p} + (x - \phi)\,\frac{\partial^2 q}{\partial p^2}\right) > 0$. The last equality follows from totally differentiating (14) and the fact that $\partial q/\partial p = \partial q/\partial x < 0$. The denominator in the last expression is negative by second order conditions of the final good providers' problem.

A.2 Proof of (26)

The infrastructure provider's problem is

$$\max_{A_h, p_h, A_\ell, p_\ell} \quad \sigma\left((1 - t_A)\,A_h + (1 - t_p)\,p_h q_h - c q_h\right) + (1 - \sigma)\left((1 - t_A)\,A_\ell + (1 - t_p)\,p_\ell q_\ell - c q_\ell\right)$$
$$s.t. \qquad V_i \geq M \quad i = h, \ell \quad \text{and} \quad V_i \geq V_{i\tilde{i}} \quad i, \tilde{i} = h, \ell$$

where the first set of constraints are the participation constraints (PCs) and the second set of constraints are the incentive compatibility constraints (ICCs). Moreover,

$$V_i = u(q_i, \alpha_i) + M - A_i - (p_i + \phi)\,q_i \quad i = h, \ell,$$

$$V_{i\tilde{i}} = u(q_{i\tilde{i}}, \alpha_i) + M - A_{\tilde{i}} - (p_{\tilde{i}} + \phi)\,q_{i\tilde{i}} \quad i, \tilde{i} = h, \ell,\ i \neq \tilde{i},$$

are, respectively, the indirect utility levels of a type-i consumer adopting the intended tariff, and that of a mimicker. Following standard steps (Laffont and Martimort, 2001), it can be shown that the solution has to be such that the PCs are slack for $i = h$, and the ICCs are slack for $i = \ell$ and $\tilde{i} = h$, whereas the other constraints are binding. Furthermore, in order to relax the ICC for $i = h$, it is optimal to have $q_{h\ell}$ arbitrarily close to q_ℓ (this can be implemented by imposing an extra fee for usage immediately beyond q_ℓ, which would however not be paid in equilibrium). Hence, we rewrite the problem as

$$\max_{A_h,p_h,A_\ell,p_\ell} \quad \sigma\left(\left(1-t_A\right)A_h+\left(1-t_p\right)p_h q_h-cq_h\right)+\left(1-\sigma\right)\left(\left(1-t_A\right)A_\ell+\left(1-t_p\right)p_\ell q_\ell-cq_\ell\right)$$
$$s.t. \qquad V_\ell = M \quad \text{and} \quad V_h = V_{h\ell}.$$

From the equality constraints, we get

$$A_\ell = u(q_\ell,\alpha_\ell) - \left(p_\ell+\phi\right)q_\ell, \tag{36}$$

$$A_h = u(q_h,\alpha_h) - \left(p_h+\phi\right)q_h - \left(u(q_\ell,\alpha_h) - \left(p_\ell+\phi\right)q_\ell - A_\ell\right). \tag{37}$$

We can therefore rewrite the profit maximization problem as

$$\max_{p_h,p_\ell} \quad \sigma\left[\left(1-t_A\right)\left(u(q_\ell,\alpha_\ell)-u(q_\ell,\alpha_h)+u(q_h,\alpha_h)-\left(p_h+\phi\right)q_h\right)+\left(1-t_p\right)p_h q_h-cq_h\right]+$$
$$+\left(1-\sigma\right)\left[\left(1-t_A\right)\left(u(q_\ell,\alpha_\ell)-\left(p_\ell+\phi\right)q_\ell\right)+\left(1-t_p\right)p_\ell q_\ell-cq_\ell\right]. \tag{38}$$

Using the equilibrium conditions $\frac{\partial u(q,\alpha_i)}{\partial q} = p_i+\phi$, for $i=h,\ell$, we can write the first-order conditions (FOCs):

$$\frac{d\pi}{dp_h} = \left(t_A-t_p\right)q_h + \frac{\partial q_h}{\partial p_h}\left(\left(1-t_p\right)p_h-c\right) = 0, \tag{39}$$

$$\frac{d\pi}{dp_\ell} = -\sigma\left(1-t_A\right)\left(\frac{\partial u_h}{\partial q_\ell}-\frac{\partial u_\ell}{\partial q_\ell}\right)\frac{\partial q_\ell}{\partial p_\ell}+ \tag{40}$$
$$+\left(1-\sigma\right)\left(q_\ell\left(t_A-t_p\right)+\frac{\partial q_\ell}{\partial p_\ell}\left(\left(1-t_p\right)p_\ell-c\right)\right)=0.$$

We obtain (26) by rearranging the above FOCs. Furthermore, differentiating these FOCs, we obtain the comparative statics in (28) - (31).