THE NEW ECONOMY: ESSAYS IN NETWORK ECONOMICS AND TWO-SIDED MARKETS

A Dissertation Submitted In Partial

Fulfilment Of The Requirements

For The Degree Of

DOCTOR OF PHILOSOPHY IN ECONOMICS

by

David Henriques

NOVA SCHOOL OF BUSINESS AND ECONOMICS

December 23rd, 2013

ABSTRACT

David Henriques

The New Economy: Essays in Network Economics and Two-Sided Markets

Following the Introduction, which surveys existing literature on the technology advances and regulation in telecommunications and on two-sided markets, we address specific issues on the industries of the New Economy, featured by the existence of network effects. We seek to explore how each one of these industries work, identify potential market failures and find new solutions at the economic regulation level promoting social welfare.

In Chapter 1 we analyze a regulatory issue on access prices and investments in the telecommunications market. The existing literature on access prices and investment has pointed out that networks underinvest under a regime of mandatory access provision with a fixed access price per end-user. We propose a new access pricing rule, the indexation approach, i.e., the access price, per end-user, that network i pays to network j is function of the investment levels set by both networks. We show that the indexation can enhance economic efficiency beyond what is achieved with a fixed access price. In particular, access price indexation can simultaneously induce lower retail prices and higher investment and social welfare as compared to a fixed access pricing or a regulatory holidays regime. Furthermore, we provide sufficient conditions under which the indexation can implement the socially optimal investment or the Ramsey solution, which would be impossible to obtain under fixed access pricing. Our results contradict the notion that investment efficiency must be sacrificed for gains in pricing efficiency.

In Chapter 2 we investigate the effect of regulations that limit advertising airtime on advertising quality and on social welfare. We show, first, that advertising time regulation may reduce the average quality of advertising broadcast on TV networks. Second, an advertising cap may reduce media platforms' and firms' profits, while the net effect on viewers' (subscribers) welfare is ambiguous because the ad quality reduction resulting from a regulatory cap offsets the subscribers' direct gain from watching fewer ads. We find that if subscribers are sufficiently sensitive to ad quality, i.e., the ad quality reduction outweighs the direct effect of the cap, a cap may reduce social welfare. The welfare results suggest that a regulatory authority that is trying to increase welfare via regulation of the volume of advertising on TV might necessitate to also regulate advertising quality or, if regulating quality proves impractical, take the effect of advertising quality into consideration.

In Chapter 3 we investigate the rules that govern Electronic Payment Networks (EPNs). In EPNs the No-Surcharge Rule (NSR) requires that merchants charge at most the same amount for a payment card transaction as for cash. In this chapter, we analyze a three-party model (consumers, merchants, and a proprietary EPN) with endogenous transaction volumes and heterogenous merchants' transactional benefits of accepting cards to assess the welfare impacts of the NSR. We show that, if merchants are local monopolists and the network externalities from merchants to cardholders are sufficiently strong, with the exception of the EPN, all agents will be worse off with the NSR, and therefore the NSR is socially undesirable. The positive role of the NSR in terms of improvement of retail price efficiency for cardholders is also highlighted.

Acknowledgements

to my father (1933 – 2010), mother and Inês

Contents

Introduction

Network industries, i.e., industries that exhibit network effects, represent an important fraction of the world economy in the modern life. The new economy would be dreadfully diminished without, for example, the telecommunications and Internet, the media and the electronic payment networks. Network effects are a specific kind of externality in which agents' value functions are affected by the number of other agents using the same or a compatible good or service. The effects may be positive, when agents benefit from an increase in the number of other agents sharing the same or a compatible brand. In other words, there is complementarity on consumption of different types of agents that participate in the network. Network effects may also be negative,[1] in which case the number of agents makes the good or service less valuable for some consumers. This is evident in sectors, such as public transports or media, where congestion or advertisements can impose a negative effect on commuters or viewers, respectively.

In this thesis we explore two relevant topics in the Network Economics literature: (i) the technology advances and economic regulation in telecommunications and (ii) the two-sided markets.[2] More specifically, this dissertation deals with three industries of the so-called New Economy: telecommunications, broadcasting and electronic payments. A specific question is addressed in each industry and an analytical model will support the insights in each chapter. All the three chapters include welfare analysis and discuss economic regulation policies within the context of each industry.

Telecommunications

Telecommunications services are based on an increasingly sophisticated network that is able to offer a diverse variety of services that differ in nature of data, voice or video, requirement of "real time" delivery, among other aspects. This industry has been going through a very significant change namely in the US and in Europe. The provision of voice services, broadband Internet access and video services exclusively over cable lines in the "local loop" requires major technological advances and considerable investment just for the conversion of the cable network to Next Generation Networks (NGNs). The term NGN refers to the installation of high-speed physical infrastructures, mostly based on optical fiber for the transmission of voice, data and video. Under many respects the NGNs represent a dramatic technological shift in the provision of telecom services: new

---

[1] For example, negative network effects may result from snobbism or vanity, in that an agent loses the feeling to belong to an elite group when a product or service is extensively and commonly used in the population.

[2] Important topics in the Network Economics literature, among others, are: consumer demand under network externalities, information networks and intellectual property, consumer demand under network effects and social influence. See Shy (2010) for a survey on the network effects literature.

networks enable a bandwith up to 100 megabits per second, as compared to the maximum of 20 megabits on DSL platforms.

Regulation and investment. Considerable and detailed regulation is desirable in some telecommunications contexts, namely in NGNs, when the market forces are not sufficient to reach the (price and investment) efficient outcomes and social and private benefits differ. The public interest goal of telecommunications regulation is twofold. On the one hand it is to increase total surplus, i.e., the unweighted sum of consumers' surplus and firms' profits. On the other hand, it is generally accepted that the public interest should promote innovation and growth.

It is noteworthy that there are also drawbacks related to the economic regulation of telecommunication networks. First, regulators face several informational constraints regarding the costs related to new technologies, but also regarding the willingness to pay that consumers attribute to the new technologies. Second, in an industry that is technically changing at fast pace, it may be difficult to define the appropriate range of regulated services. Third, the regulatory setup may be slow, bureaucratic and politically biased. Individuals or organizations with high-stake interests in the regulatory decisions may center their attentions in attempting to gain their preferred policy outcome, while dispersed individual consumers, each with only a small individual stake in the outcome, will ignore it altogether. The regulatory capture refers to this disparity of attentions dedicated to a policy outcome that succeeds at "capturing" influence at the regulatory body and put into practice the preferred policy of the special interest. For more about the capture theory refer to Laffont and Tirole (1991), and Levine and Forrence (1990).

Investments in telecommunication networks have been characterized by a strong growth in the period up to 2000 and by a subsequent strong decrease. Such a decrease was mainly due to two reasons. First, the end of the substantial initial investments in access and backbone infrastructures, both fixed and mobile, by new entrants in the telecommunication market, led by overoptimistic expectations on the pick up of Internet services. Second, the end of the financial bubble in the telecommunication industry, that pressed operators and capital markets to be more focused on obtaining an adequate return on investment.

The excess capacity that has characterized the first decade of the 21st century as a consequence of the massive build-up of transmission infrastructures from 1995 to 2000 is being eroded by the increase in demand for broadband Internet. Not just the number of connections for broadband access has increased in Europe and in the US, but also end-users now require networks with larger bandwidth, i.e., faster Internet connections. NGNs will increase bandwidth, thus being the leading driver of the future investment in telecommunication networks.

Placing fiber up to the end-user's location represents a significant financial effort due to the cost of obtaining building permits and of engineering works in urban and rural areas,

which together represent from 50 to 80% of the overall capital expenditure (European Commission (2007)). The main potential benefits from NGNs are cost reduction in the services supply (data, HDTV and voice in a single network, instead of multiple networks) and increase revenues by delivering brand new services. The pace with which these benefits are made will depend on how NGNs are regulated.

The relevant trade-off in NGNs is between the incentive to investment and the degree of price competition in the future. On the one hand, incumbent operators that will likely make most of the investment are waiting to see whether the regulatory bodies decide to require permanent regulation wholesale obligations such as open access to other operators at a price equal to the cost of the service. If this were the case, one could infer that incumbents would have fewer incentives to build NGNs, as regulation would wipe out the quasi-rents arising from the fiber deployment. The existence of wholesale obligations will also condition the behavior of the new entrants networks, which may either, decide to invest or to make use of the incumbents' network as the latter are installed (free riding behavior), thus avoiding significant fixed costs. On the other hand, regulatory bodies are concerned about removing any initial conditions of major advantage to the incumbents that could prevent the rise of a competitive market. The potential advantages include (i) the exclusivity of some network elements (regulatory holidays) such as the engineering works to install the fiber, and (ii) the control of a large base of customers which could enable the incumbent to reach significant network economies before its competitors. Regulators such as the European Commission are usually against regulatory holidays, i.e., the absence of all obligations on NGNs for a pre-defined period of time.

The economic literature on the impact of regulation on investments is separated into two strands of research.

A research strand is the investment analysis where the impact of regulation, either rate-of-return or incentive regulation, is usually assessed in a static context. The rate-of-return regulation has been criticized because it encourages overinvestment if the regulated rate is set too high. This is the so-called Averch–Johnson (1962) effect. Also, some dynamic models on the timing of infrastructure investment are applied, such as in Gans and Williams (1999), Gans (2001) and Gans and King (2004). A survey of the static and dynamic models of investment under different forms of regulation and optimal (Ramsey) pricing may be found in Biglaiser and Riordan (2000).

Another research strand is the real option approach that captures the fact that demand, technology and other factors impacting on investment decisions are subject to uncertainty. As a result networks may have interest to delay the investment in order to obtain more information and to decrease risk. Authors such as Dobbs (2004) have integrated uncertainty and irreversibility in their models and have considered a more general

problem of setting regulated prices when faced with non-constant demand and technology.

In Chapter 1 we will focus our attention on the impact that a change in the access price rule (in a static context) has on the NGNs investments, on retail prices and on welfare.

Two-sided markets

In Chapter 2 and Chapter 3 we deal with specific questions in two-sided markets. Below we present a brief introduction on two-sided markets, including: (i) a general view of the economics of two-sided markets and its basic principles, (ii) a discussion of the definition of two-sided platforms, (iii) regulatory concerns and (iv) examples of two-sided markets.

A general view. Since Veblen (1899) it is established in the economic literature that consumers' choices may be affected not just by their own preferences and income but also by the consumption choices of other agents. These effects are significant in several industries where the choice to purchase from a particular brand is affected by the number of agents buying the same product or patronizing the same product or service.

A two-sided platform facilitates the members of two distinct groups of agents to be together in a way that adds value for the agents that could not get as efficiently as without the platform. The platform internalizes the network externalities among the groups of agents. Some platforms may face more than two sides (i.e., groups of agents) but the insights achieved in two-sided platforms to a large extent apply to multi-sided markets.

Two-sided markets are not a new sort of business. In fact, people have been using the two-sided logic for hundreds of years. One example that goes back centuries ago is the advertising-supported newspapers that help firms to advertise potential consumers.[3] Though, the explicit identification that businesses across different markets have two-sided characteristics, which have important economic implications, has started only after 2000. The theory of two-sided markets was initiated by Rochet and Tirole in the seminal article *Platform Competition in Two-Sided Markets* in 2003 and the subsequent article *Two-sided markets: a progress report* in 2006. Vital contributions to the burgeoning literature of two-sided markets are also Caillaud and Jullien (2003) and Armstrong (2006), among others.

Some factors have been bringing into play the two-sided market analysis to regulatory authorities, particularly the following two. First, the two-sided market literature has been developing notably since 2000 and has been a hot topic in economics since then. Second, there is the impression that a number of key industries recurrently under discussion, for

---

[3]For example, in France in 1836 the newspaper La Presse was the first to insert paid advertising on their pages lowering the price to readers, which extended its readership and increased its profitability. The new business model adopted by La Press was a success and copied by other newspapers.

example, software, communications, media or electronic payments, are either two-sided markets or take part in a business environment such that two-sided platforms play a substantial role.

Two-sided platforms must manage the demands of two distinct groups of agents that exert cross-group externalities. In order to internalize the network externalities, platforms choose pricing and non-pricing strategies that can be very different from firms that serve a single group of agents. In order to put both sides of the market on board a platform has to choose the price level and the price structure. If the cross-group externalities are strongly unbalanced between groups, the structure of prices that balances the demands in each side of the market can be extremely biased. For example, one side may be subsidized with price below the respective marginal cost or be charged nothing to participate, whilst platform's revenues are extracted from the other side of the market.[4] The profit maximizing price charged to agents on each side of the platform, in general, does not follow the standard mark-up formula or track the marginal cost. Typically, the type of agents that generates the highest level of network effects will be charged relatively less. This is the reason why agents on one of the sides might pay a price below marginal cost, or even below zero (e.g., card payment rewards), whereas agents on the other side will be charged prices considerably above marginal cost, generating most of the platform's revenues.

Provided that two-sided platforms must manage demands of agent groups that exert network effects among themselves, a price variation on one side of the market produces side effects on the other sides of the market. Hence, the analysis in two-sided markets must consider the network effects to measure the overall effect of a price variation on platform's profit.

Note that in the presence of network externalities across groups, the marginal revenue related with each group of agents has a direct and an indirect element. First, by joining the platform an agent generates directly revenues to the platform associated to the fees he pays. Second, by joining the platform an agent increases the value of the platform to consumers on the other side. This enables the platform to charge more to agents on the other side. Thus, the profit maximizing condition, marginal revenue equals marginal cost, for a two-sided platform, has to be adjusted such that the marginal revenue is corrected for the existence of indirect network externalities across the groups.

Also, note that due to the existence of network effects, if there are joint costs for providing services to both types of consumers, it is neither profit maximizing nor socially efficient to follow the standard rule "price equal to marginal cost" in each side of the market.

Definition. The term "two-sided markets" was established in the seminal article of Ro-

---

[4]For example, this may happen in night clubs where men pay a high price just to enter while women have some free drinks, i.e., women have their entry in the club subsidized.

chet and Tirole (2003). Rochet and Tirole created the expression to define circumstances in which platforms supply at the same time two mutually dependent groups of clients.

For the sake of clarity, it is convenient to distinguish between two-sided platforms and the markets in which they operate. Two-sided platforms may compete with single-sided firms or compete in given markets with other two-sided platforms that serve a different second side. The two-sided markets literature employs the term "multi-homing" when customers use two or more platforms for the same service, and "single-homing" when just one platform is in use. Multi-homing can occur on one or more sides of the platform. For example, platforms such as computer operating systems have multi-homing only on one side. Most end-users have a single operating system for their computers, while coders tend to produce applications to run on several operating systems.

Evans (2003) and Rochet and Tirole (2006), among others, have pointed out the following features of a two-sided platform. First, there must be considered two distinct groups of agents to whom the platform is an intermediary. A two-sided platform provides goods or services simultaneously to these two groups.

Second, indirect network effects exist across the two groups of agents. This means that the value that an agent on one side captures from the platform is dependent on the number of agents on the other side. For example, Sony PlayStation provides programming code that reduces the effort for game developers to write all code themselves and provides a standard environment for end-users to run games. These actions from Sony increase the pool of programmers and, thus, of games available, making more consumers willing to buy a PlayStation. Also, programmers are more willing to write their games in Playstation format when the number of players using that platform is higher. Another example is of a search platform that will be more valuable to advertisers if more potential buyers use the platform. Simultaneously, it will be more valuable to consumers looking to buy something if more firms advertise on that platform since will be more likely that the consumer sees a relevant advertisement.

Third, a two-sided platform is featured by the non-neutrality of the price structure. Let the price level charged by the two-sided platform be the sum of the per-interaction prices charged to each agent involved in the interaction. The price structure can be defined as the part of the price level that is paid by each type of agent. Let total welfare be the unweighted sum of the welfare of both groups of agents and the platform. The platform will be one-sided if total welfare changes with the price level but does not change with reallocations of the price structure between groups. For example, a tax on wheat charged to buyers has the same welfare effect as compared to the same tax charged to sellers. The platform will be two-sided if total welfare changes with both the price level and the price structure. Nonetheless, while useful to understand the two-sided logic, this is not a general definition.

Sometimes the two-sidedness of a market is a vital point for the analysis. Other times it is just an interesting, but not fundamental, aspect of the market. And still other times it is irrelevant. It is often the strength of the indirect network effects that decides whether the two-sidedness matters enough to have a substantive effect on the results of economic analysis, or whether it is only an interesting curiosity. Hence, two-sidedness is a question of degree.

Regulatory issues. Like in telecommunications networks, two-sided markets regulation can also be very challenging. In two-sided markets price variations may not imply welfare variations. In markets without cross-group externalities prices and social welfare loss move in the same direction, i.e., as price goes up, welfare loss goes up. Therefore, welfare changes can be inferred from price changes. In markets with externalities, such as two-sided markets, this may not be the case. In fact, prices and consumer welfare may move in the same direction, this implies that prices and social welfare may be positively related. Under network externalities prices do not serve for inferring about welfare and regulators have to measure welfare directly, which is a much more subjective and demanding task than measuring price differences.

Social welfare is harder to compute in two-sided markets given that platforms and consumers on both sides of the market should be taken into account. Moreover, the welfare effects on end-users on both sides of the market might change in different ways in response to a policy measure. The net effect from opposite welfare variations generated by a regulatory policy will be more difficult to predict as the number of groups involved with the platform increases.

Given the existence of indirect network effects across end-users the characterization of the social optimum can be very demanding in terms of information. Moreover, since in these markets the price structure plays a very important role, it might be also complex to find which directions policy measures should be taken to increase welfare.

Multi-sided platforms can and do follow anti-competitive behavior that can be equally damaging as the anti-competitive behavior of firms in markets without network effects. Competition analysis in two-sided markets must take into account the network effects to evaluate platforms' competitive actions. For example, competition authorities have commonly followed the principle of price regulation according to marginal costs. Nonetheless, due to the presence of network externalities, this principle should not be extended to regulate prices of two-sided platforms. In particular, the social optimal prices have to consider not only the cost-side but also the demand-side together with network effects.

Two examples. Advertising-supported media platforms, e.g., TV, magazines, newspapers and internet search engines, supply at the same time two distinct groups of agents: viewers (consumers) and advertising firms. The media platforms supply contents to attract viewers while the viewers are the bait to attract advertising firms.

Network effects are present between viewers and advertising firms. On the one hand, advertisers value platforms with larger audiences for the reason that they get more coverage. On the other hand, viewers value platforms with fewer advertisements due to the nuisance caused by the advertising time. Hence, the profit maximizing prices result from balancing the viewers and advertisers demands.

Advertising-supported media receives a large part of overall revenues from advertisers. Advertising firms are frequently charged based on the number of subscribers of the platform. The prices that platforms charge to advertising firms fulfill the purpose of subsidizing the contents that the platforms show to viewers. In fact, platforms such as free-to-air TV charge viewers only an implicit price which is the cost of watching advertisements and waiting for the show to resume.

Some relevant cases in the media market where the two-sided market analysis is important have been the European Commission against Newspaper Publishing, the Carlton Communications/Granada merger, and the acquisition of DoubleClick by Google which was studied by the FTC in US and by the European Commission.

Another example of a two-sided market is the card payment system. The payment card analysis has played an important role in the progress of two-sided markets. Payment systems, such as cash and payment cards, endow agents with the possibility of transacting goods and services without having to barter. A payment system has an increasing value to merchants as more consumers use it, and is more valuable to consumers as more merchants accept it, and is only possible if both types of agents use it. Thus, there are positive network effects between the two groups of agents.

Cash payments involve no explicit costs both to consumers and merchants, but might entail implicit costs such as the risk of theft or handling costs. With a card payment a fee might be charged to each side of the market by the electronic payment platform. Usually these platforms make most of their revenue from the merchant side of the market, i.e., merchants bear the entire cost per transaction while cardholders may pay a small fixed annual fee, in a number of cases even the annual fee is set down to zero and consumers have rewards per card transaction. For example, consider a monopolist credit card platform that charges a per transaction fee to both merchants and cardholders. The more it charges merchants, the greater is the incentive to persuade cardholders to use their cards, by reducing the usage fee or increasing their amenities (rewards). The merchant's fee works as a subsidy to the platform in serving consumers.

The two-sided market logic has been playing an important role on market regulation namely in terms of the pricing charged by the electronic payment networks. Some cases scrutinized by regulatory authorities comprise the analysis of the Reserve Bank of Australia on credit cards and the case of the European Commission versus MasterCard on

interchange fees.[5]

References

Armstrong, M. (2006), *Competition in two-sided markets*, RAND Journal of Economics, Vol. 37, No. 3, pp. 668–691.

Averch, H. and L. L. Johnson (1962), *Behaviour of the Firm under a Regulatory Constraint*, American Economic Review, Vol. 52, No. 5, pp. 1052–1069.

Biglaiser, G. and M. Riordan (2000), *Dynamics of Price Regulation*, RAND Journal of Economics, Vol. 31, Issue 4 (winter), pp. 744–767.

Caillaud, B. and Bruno Jullien (2003), *Chicken and Egg: Competition among Intermediation Service Providers*, RAND Journal of Economics, Vol. 34, No. 2, pp. 521–552.

Dobbs, Ian M. (2004), *Intertemporal price cap regulation under uncertainty*, Economic Journal, Vol. 114, No. 495, pp. 421–440.

Economides, Nicholas (2004), *Telecommunications Regulation: An Introduction*, available at http://www.stern.nyu.edu/networks/Economides_Telecommunications_Regulation.pdf

Economides, Nicholas and Fredrick Flyer, (1998), *Compatibility and Market Structure for Network Goods*, Discussion Paper EC-98-02, Stern School of Business, NYU, available at http://www.stern.nyu.edu/networks/98-02.pdf

European Commission (2007), I*mpact Assessment* (mimeo), Brussels.

Evans, D. (2003), *The Antitrust Economics of Multi-Sided Platform Markets*, Yale Journal of Regulation, Vol. 20, No. 2, pp. 325–381.

Gans, J. S. (2001), *Regulating Private Infrastructure Investment: Optimal Pricing for Access to Essential Facilities*, Journal of Regulatory Economics, Vol. 20, No. 2, pp. 167–89.

Gans, J. S. and S. P. King (2004), *Access Holidays and the Timing of Infrastructure Investment*, Economic Record, Vol. 80, No. 248, pp. 89–100.

Gans, J. S. and P. L. Williams (1999), *Access Regulation and the Timing of the Infrastructure Investment*, Economic Record, Vol. 79, No. 229, pp. 127–138.

---

[5]The issuing bank receives an interchange fee from the acquiring bank every time cardholders pay for purchases with a card.

Guthrie, G. A. (2006), *Regulating Infrastructure: The Impact on Risk and Investment*, Journal of Economic Literature, Vol. XLIV, pp. 925–72.

Katz, Michael and Carl Shapiro (1985), *Network Externalities, Competition and Compatibility*, American Economic Review, Vol. 75, No. 3, pp. 424–440.

Laffont, J. J., and Tirole, J. (1991), *The politics of government decision making. A theory of regulatory capture*, Quarterly Journal of Economics, Vol. 106, No. 4, pp. 1089–1127.

Levine, M. E., and Forrence, J. L. (1990), *Regulatory capture, public interest, and the public agenda. Toward a synthesis*, Journal of Law Economics & Organization Vol. 6, pp. 167–198.

Rochet, J-C and J. Tirole (2003), *Platform Competition in Two-Sided Markets*, Journal of the European Economic Association, Vol. 1, No. 4, pp. 990–1029.

Rochet, J-C and J. Tirole (2006), *Two-Sided Markets: A Progress Report*, RAND Journal of Economics, Vol. 37, No. 3, pp. 645–667.

Shy, O. (2011), *A Short Survey of Network Economics*, Review of Industrial Organization, Vol. 38, No. 2, pp. 119–149.

Veblen, T. (1899), *The Theory of the Leisure Class. An Economic Study of Institutions.* London: Macmillan.

1 Can access price indexation promote efficient investment in Next Generation Networks?

1.1 Introduction

Motivation. A key concern for the United States and Europe is the timely rollout of Next Generation Networks (NGNs).[6] Fiber optics technology is at the core of NGNs and is considered the future of telecommunications infrastructure, since it allows faster and wider transmission of all sorts of information than copper-based networks.[7] Significant investments are required to supply the necessary communications infrastructure that consumers and firms demand in order to effectively compete in nowadays' knowledge based society. While the technology exists today, it is uncertain when and to what extent it will be deployed by network operators. In 2009, fiber to the home (FTTH) had reached nearly 13% penetration of US households in terms of homes passed and 4% in terms of homes connected (RVA LLC Market Research & Consulting, April 2009), in Germany and Spain FTTH covered less than 5% of the households, and in Italy less than 10% (IDATE Consulting & Research, February 2010). At the end of 2010, the percentage of subscribers out of total homes passed by fiber was 17.5% in Europe, 34% in the United States and 39% in Japan (IDATE Consulting & Research, 9 February 2011).[8] These facts suggest that residential and business users, particularly in Europe, are unsure about the benefits of FTTH given the level of retail prices charged by the operators.

Telecommunications regulators usually have the task of encouraging investment and innovation and simultaneously ensuring that networks remain competitive, as competition is a vital matter for end-users and for businesses relying on the new networks. However, regulators and competition authorities seem to face a trade-off between static and dynamic efficiency. On the one hand, static regulation reduces the extent to which operators exert

---

[6]The idea behind the NGN is that a single network infrastructure transports all information and services (e.g. voice, data, high definition TV, interactive gaming) allowing increased transmission speeds by encapsulating information into packets.

[7]Aside from fiber, there are a number of alternative technologies capable of supporting NGN access such as: coaxial cable, mobile and fixed wireless networks. Since fiber is one of the fastest technologies for content transmission (both downloading and uploading), debates on wired NGN access have focused on fiber deployment.

[8]Asian carriers occupy eight out of the top 10 spots in terms of fiber subscriber numbers. Japan is ranked number 1 with 13,839,000 subscribers. None of the top 10 FTTH market players is from Europe. Asian operators were the first to strongly invest in fiber rollout, and have in 2011 achieved virtually complete coverage in their respective national markets.

market power on the downstream market, inducing retail prices to converge closer to marginal cost.[9] On the other hand, a fixed access price based on cost, while it may promote the statically efficient use of the network, discourages investment (dynamic inefficiency) since the returns that can be earned by investors are constrained by the access price set by the regulator.[10]

The question that telecommunications regulators face under the most used costing methodologies[11] is: how much price efficiency must be sacrificed to achieve a desired level of investment? In this chapter we question whether such a trade-off always exists. The main challenge is thus to create an access price rule to respond to the question: how to encourage investment in network (bottleneck) infrastructures without lessening downstream price competition relatively to a fixed access pricing methodology?

Description of the chapter. We consider a context of bilateral one-way access, i.e., there are two bottleneck facilities (networks) forced to provide access to each other under some regulatory conditions.[12] Underinvestment derives from the inability of networks to capture the full social benefit from investment. The problem of access obligations mandated by regulation is that they diffuse the investment benefits among operators and consumers while the investment cost is concentrated on the investor (network owner). Hence, underinvestment in infrastructure is aggravated by the non-exclusivity imposed by regulation together with the fact that investment is costly.

Let a denote the access price per subscriber under the fixed access price methodology, which is currently used by many regulators. We compare the socially optimal fixed access price a^* to a new access price rule, which we call of access price indexation, in terms of retail prices, fiber coverage and social welfare levels. Under the indexation approach the access price is defined by the regulator as a function of the operators' investments in fiber coverage.

The main purpose of the new access price proposal is to reward or punish operators

---

[9]Setting access conditions in network industries is an essential issue for regulators to avoid anti-competitive behavior on the part of the networks (bottleneck-facility owners). In particular, access regulation is important to avoid networks that deter entry by refusing access to competitors and to provide competitors with reasonable access prices, guaranteeing competitive parity among operators.

[10]Imposing open access with a fixed access price calls to mind the classical free-riding problem in static frameworks (see Olson (1965), Chamberlin (1974) and McGuire (1974)). The literature on free-riding points out that the investment level of equilibrium in public goods is lower than the Pareto efficient investment level. In a monopolistic market structure the free-riding problem vanishes; however, the retail price would become inflated, generating potential welfare losses.

[11]In 2009, the long-run-incremental cost was the costing approach most often applied to European markets for wholesale access at a fixed location (64%) and the second most used for wholesale broadband access (46%). The fully distributed cost approach had a share of 32% and 54%, respectively (ERG 2009). With respect to investment in NGN, the European Commission recommendation (European Commission, 20 September 2010) suggested a risk premium when setting access prices to the unbundled fiber loop in order to compensate the investor for bearing the risk of failure alone.

[12]This is different from two-way access since in this model end-users do not interact with each other, whilst two-way access environments are characterized by end-user interconnection.

depending on the investments made by each one. On the one hand, the indexation rule should reward an operator i for covering cities with fiber, lowering the access price a_j charged by operator j when i requests access to j's network. Thus, the indexation rule can grant a competitive advantage at the downstream level to operators investing relatively more. On the other hand, the new rule intends to punish the operator that invests relatively less by increasing its access price to the other network. The indexation can impose a competitive disadvantage at the downstream market to operators investing relatively less in fiber coverage. This is a solution that internalizes for operator i the positive spillovers exerted from its investment, I_i, in the sense that a_i should increase in I_i. With the access price indexation suitably chosen by the regulator, the "dilemma" faced by the networks is that, whatever the other does and as long profits are non-negative, each network is better off investing relatively more since investments are a source of a competitive advantage in retail prices. For example, by using a simple linear access pricing rule depending on investments (I_i, I_j) by operators i and j, $a_i(I_i, I_j) = xI_i - yI_j$, where $(x, y) \in \mathbb{R}_+^2$ are regulatory parameters, we can create a causal link from retail price competition to (investments in) fiber coverage. This rule incentivizes networks to compete more strongly in investments.

We show that the new rule increases economic efficiency as compared to the fixed access price methodology. The indexation methodology dominates both a fixed access pricing rule and a regulatory holidays policy in terms of retail price efficiency (or equivalently, the number of consumers served with a fiber connection), investment efficiency (i.e., the number of cities covered with fiber) and social welfare. Furthermore, we provide conditions under which the indexation rule can promote the socially optimal (first-best) investment or the Ramsey allocation, which would be unfeasible either with a fixed access price or with regulatory holidays.[13]

The intuition for these results is the following. Since part of the benefits generated by investments is retained by consumers, the monopolistic (regulatory holidays) outcome is not only inefficient in terms of retail prices but also inefficient in terms of investment. Under a fixed access price, the introduction of competition in the downstream market can only deteriorate investment efficiency, while under the indexation rule networks have incentive to compete in investments as a means to gain a competitive advantage in the downstream market. By choosing the proper access price indexation, the regulator can encourage operators to invest until a certain level, which under some conditions (set out in Proposition 3) can go up to the Ramsey investment or even to the first-best investment level, as long as networks have non-negative profits. In fact, under the indexation approach, a welfare-maximizing regulator sets (x, y) such that, in a symmetric equilibrium,

---

[13]According to our model results, under the assumption that the regulator knows with certainty all parameters, a fixed access price is condemned to be inefficient both in retail prices and investments.

networks' profits are zero (see Lemma 1). Therefore, if the regulator is able to incentivize networks to invest up to the Ramsey level, by the zero-profit condition, Ramsey pricing will be implemented as well. We conclude that in equilibrium the social welfare level under the new rule lies outside the previously perceived "second-best efficiency frontier" under a fixed access price approach.

The main contribution of this chapter is to show that an access price rule depending on investments, without being informationally more demanding, can improve economic efficiency both in terms of retail prices and investments as compared to a fixed access price rule. In a nutshell, the access price indexation is a feasible instrument that can enable a regulator to achieve higher social welfare.

Background. A crucial issue in the economics of regulation of NGN access is how to encourage operators to invest in infrastructure. Attempts to develop and invest in NGNs have been taken in many countries by National Regulatory Authorities (NRAs) and Governments. For example, in 2006, in Germany the incumbent operator Deutsche Telekom told the Government that would make these investments only if the Government granted regulatory holidays, i.e., the incumbent would be temporarily a monopolist without obligation to provide access to competitors at regulated prices.[14] In 2008, the Spanish NRA removed the requirement on Telefonica, the incumbent operator, to supply wholesale access service to its FTTH network. This verdict gave Telefonica a regulatory holiday on FTTH network access, similar to that held by Deutsche Telekom (ITU, 2009). The French model follows the cost-sharing perspective. It forces network operators, which may invest on their own, to make available access to ducts and supply information on planned civil works and fiber coverage, sharing the installation costs of additional fiber at other operators' request. Other options to stimulate the development on NGNs are the establishment of public-private partnerships (PPP), as has happened in Singapore and Australia, and the provision of credit lines and funds, for example in Portugal and, in a relatively small scale, in the United States.[15] In these cases Governments invest, provide funding or credit to kick-off projects on NGNs and accelerate fiber deployment.

The introduction of competition into historical monopolies in telecommunications has led to a number of research articles on access pricing issues, as regulators have been

---

[14]The European Commission opposed the adoption of regulatory holidays and sent Germany a formal caution in February 2007, after repeated warnings that had been ignored. The case was taken to European Court (European Comission vs. German Regulator and Deutsche Telekom) that ruled against German regulatory holidays in December 2009. See ITU (2009) and *EU court "sets precedent" in Germany telecoms ruling*, EurActiv, 4 December 2009, for further details.

[15]The Australian Government decided in 2009 to invest and to be the majority shareholder of a A$ 43 billion super-fast national broadband network. The US Government under President Barack Obama has allocated US$ 7.2 billion to support broadband build-up. In Portugal, in 2009, a protocol on NGN was signed between the Government and four operators (Portugal Telecom, Sonaecom, Zon and Oni Communications), in which there is a commitment of all parties to invest in NGN. The Portuguese Government is committed to make available a credit line of, at least, EUR 800 million.

confronted with the need to set the rules on which operators should have access to each other's network. The vast majority of articles on access pricing assume that access fees do not depend explicitly on investment levels.[16] Only recently some exceptions, as Hurkens and Jeon (2008), Nitsche and Wiethaus (2011), Klumpp and Su (2010) and Sauer (2012) have considered the idea of having access prices as a function of strategic variables, namely, retail prices, quantities or investments.

Gans (2004) presented a model to study the impact of access price regulation on investment timing. In particular, Gans investigated whether such regulation can improve investment timing on equilibrium outcomes, relatively to the social optimal, whilst encouraging price competition. First, it is shown that investment might be delayed vis-à-vis the socially efficient timing if one firm is "small". When two firms are "large", competition accelerates investment timing and the investment might be provided too rapidly at a higher cost than in the socially efficient solution. Second, the article shows that the regulator may use fixed access charges to induce the investment timing outcome to be socially efficient, by controlling the preemption incentives of other possible providers. Regulation may thus have an important role in preventing inefficient acceleration of facility investment.

De Bijl and Peitz (2004) explored situations of one-way access in which an integrated operator owns a network infrastructure and sells access directly to end-users and to a downstream operator. This article discusses the investment incentives of the integrated operator. In particular, De Bijl and Peitz show that it is possible to provide stronger incentives for the integrated operator to invest in infrastructure quality by increasing the sensitivity of the regulated access price to the network quality. Nonetheless, they do not consider any explicit form for how the access price should depend on quality.

Bourreau, Hombert et al. (2010) focused on industries in which an intermediate input (e.g. network access) is sold by vertically integrated firms that compete afterwards in prices with differentiated products in the downstream market with a non-integrated downstream firm. They show that upstream price competition with homogeneous inputs may not drive the input price down to marginal cost. The access price can be set at a level above marginal cost in order to lessen downstream competition between integrated and non-integrated firms. However, when final goods are strongly differentiated, downstream demands are practically independent among firms, and thus we are back to the classical Bertrand pricing result at the upstream level. The authors also derived conditions on the demand and cost functions under which an access price cap can repair the competitiveness in the upstream market.

Nitsche and Wiethaus (2011) analyzed investment incentives and consumer welfare

---

[16]See Valletti (2003), Guthrie (2006) and Cambini and Jiang (2009) for excellent reviews on how access pricing and network investments have been investigated by the theoretical literature. This literature points to the need to consider more deeply the impact of access regulation on investments and on welfare.

under different types of access regulation to NGNs. They show that for a given level of investment, risk-sharing (operators jointly deploy and share the costs of NGNs) induces the highest competitive intensity in the product market, followed by, respectively, long-run-incremental cost (LRIC), fully distributed costs (FDC) and regulatory holidays. They also show that, under uncertainty, FDC or regulatory holidays encourage the highest level of investments, followed by, respectively, risk-sharing and LRIC. Moreover, according to simulation results, risk-sharing induces the highest consumer surplus, since it puts together comparatively high ex-ante investment incentives with strong ex-post competitive intensity.

Hurkens and Jeon (2008), following a two-way access analysis with n network infrastructures, studied the retail benchmarking approach. They propose access pricing rules that determine the access price as function of the retail prices charged by both networks. They show that such a rule may induce the market outcome to achieve the socially efficient price at the retail level. Moreover, under two-part tariff competition, setting the access price paid by firm i to depend linearly on its average retail price and let networks invest in quality after the access pricing rule is determined and before they compete in two-part tariffs, it is possible to achieve both static and dynamic efficiency.

The closest independent research work to this chapter is Sauer (2012) which compares, from the social perspective, the performance of different regulatory access regimes. Sauer's research focuses on (i) the regime of endogenous access charges per user, contingent on networks' investment levels and (ii) the regime of investment cost-sharing with lump-sum charges, i.e., the access price is proportional to the investment costs of the competitor. Sauer shows that in the former it is possible to reach the socially efficient investment level without distorting downstream competition, whilst the latter is still below the socially efficient investment, despite the higher investment level than with fixed access charges. Our chapter is complementary since we focus on modelling techniques that differ at least in two major aspects. First, Sauer uses the Hotelling model with fully served consumers, while our model relies on the Hotelling model with hinterlands where consumers are fully served in the city center but may not be fully served in the hinterlands. Therefore, while in our model market power generates welfare effects, this does not happen in Sauer's model. Second, Sauer assumes that the access charge received by an operator is a non-negative function of its own investment. In this chapter access prices depend on investments of both networks and may be negative.

Our chapter is related to the theory of yardstick competition and tournaments, and incentives in teams. See Lazear and Rosen (1981), Green and Stockey (1983), Holmström (1982), Nalebuff and Stiglitz (1983), and Shleifer (1985) for relevant theory developments. Under a context of uncertainty, an agent's low performance may be due to an unfavorable state of nature rather than to low effort. Such effects can be detected, to some extent, by

comparing the agent's performance with that of other agents placed in similar conditions. The literature calls this scheme of "yardstick competition". Marino and Zábojník (2001) show that by organizing a tournament between two teams and transferring output from the team with inferior performance to the team with higher performance, this helps to solve (i) the free riding problem inside each team, and (ii) lessen the moral hazard problem. We use similar logic by creating a tournament between networks as a solution for an underinvestment problem in NGNs.

1.2   The model

We start by presenting the basic modelling structure and providing the social optimum as benchmark case. Then we solve the model for different regulatory regimes: (i) a fixed access price, (ii) access price indexation and (iii) regulatory holidays, and compare the equilibria in terms of fiber coverage, retail prices and social welfare levels.

Consider the market for fiber broadband service in which two networks labeled $i = 1, 2$ offer differentiated services. The timing of the model is summarized below in Table 1. First, the regulator sets the rule for pricing access to bottleneck facilities.[17]   Second, operators compete in investments (fiber coverage). In our framework this is the equivalent to each operator choosing the number of cities to cover by fiber. Third, operators compete in retail prices in the downstream market in all cities covered with fiber. Investments are made only once but operators compete in the downstream market over many periods. Therefore, the third stage of the game may be interpreted as a reduced form of a dynamic game of competition in the downstream market with a discounted stream of future profits. This structure of the game is natural as operators decide on prices in the short-run and on investments in the long-run, while regulators decide on access prices in the very long-run.

TABLE 1: Timing of the model

I.	The regulator defines the access price rule per end-user, a_i, which operator i must follow when j is the accessing network.
II.	Operators invest simultaneously and non-cooperatively in non-duplicable network infrastructure, which we interpret as NGN infrastructure (FTTH).[18]Immediately after, operators observe the investment outcome.
III.	Operators compete simultaneously and non-cooperatively in retail prices.

[17]Access pricing rules should be defined by the regulator as networks would otherwise have an incentive to set access prices too high.

[18] NGN access refers to the network segment connecting an end-user to the nearest location which houses the operator's equipment. In Europe, NGN access refers essentially to the introduction of fiber into the local loop.

Below we describe each of the participants in the model: the regulator, the networks, and the subscribers (consumers in each city) of fiber broadband services.

Regulator. The regulator can choose to fix the access price at some level $a_i = a^*$ or,[19] alternatively, to set an access price depending on operators' investment levels. For technical simplification, we assume a linear access price rule depending on investments defined by

$$a_i \left(I_i, I_j \right) \equiv x I_i - y I_j, \tag{1}$$

where $(x, y) \in \mathbb{R}_+^2$ are the regulatory parameters, and $I_i \geq 0$, $I_j \geq 0$ denote the number of cities covered by fiber by operator i and j, respectively.[20] The total number of cities covered by fiber is denoted by I, where $I \equiv I_1 + I_2$. Since the investment level corresponds to the number of cities covered by fiber we assume that investments are perfectly observable by the regulator. For example, by observing the duct construction and networks' physical infrastructures for fiber optic deployment in cities. We note that civil works cost of network deployment are the most significant in new build network construction.

We assume that the regulator is benevolent, i.e., maximizes social welfare, and can credibly commit ex-ante to impose an access price rule.[21] Otherwise, networks would infer that once the investments had been made the regulator would set a new access price rule stimulating competition in retail prices. Without a regulatory commitment networks would be less prone to invest.

Operators. Network operators are profit maximizers. We assume that operators invest in different regions, i.e., network infrastructures are non-overlapping.[22] Fixing infrastructure duplication at zero favors technical simplicity and allows focusing our attention on the static and dynamic efficiencies without considering potential inefficiencies regarding infrastructure duplication.[23]

---

[19]Regulatory holidays may be interpreted as the case when the regulator sets $a_i = \infty$ for a period of time.

[20]Under a linear indexation of access prices, the regulator will choose (x, y) such that in equilibrium networks' participation constraint binds, as shown in Lemma 1. Therefore, it is not possible to improve on investment efficiency without further distortions on retail prices. The linear indexation suits to show the main goal of this chapter: an access price indexation is better than a fixed access price in terms of retail price efficiency, investment efficiency and social welfare.

[21]For example, in the UK, in March 2009, Ofcom published a policy statement setting out a regulatory framework for Next Generation Access. Ofcom is the Independent Regulator and Competition Authority for the UK Communications Industries. This gave sufficient regulatory certainty for BT to proceed with the initial phase of super-fast broadband roll out. BT has invested £2.5 billion to make fiber broadband available to around two-thirds of UK premises by the end of 2014 (RFS 2012).

[22]For example, if $I_1 = 10$ and $I_2 = 2$, we interpret this as operator 1 covering ten cities in the north part of the country; while operator 2 covers two cities in the south part of the country. It is implicitly assumed that cities are identical with regard to their population, however they differ with respect to their cost of fiber coverage.

[23]We acknowledge, though, that different regulatory regimes may result in different levels of infrastructure duplication. By design, indexation brings further incentives for investment when compared to a fixed

The network installation cost (i.e., the cost of covering cities with fiber) is convex in the sense that it is more expensive to connect subscribers in peripheral cities.[24] This captures the fact that operators start investing from cities where fiber is relatively cheaper to install. For sake of simplicity, we assume that the investment cost follows the form

$$C(I_i) \equiv cI_i^2/2, \qquad (2)$$

where $c > 0$ is a constant.[25]

We assume that subscribers pay independently of the traffic volume exchanged in the communications, i.e., they only pay for accessing the network, e.g. a periodical subscription fee. This reflects the fact that currently in the United States and in Europe a number of broadband offers are essentially flat rates per month. Let then p_i denote i's retail price to provide broadband access to one subscriber. The respective mass of subscribers using i's service in one city is denoted by q_i.

Network i faces a marginal cost, per subscriber, for serving broadband access equal to

$$\begin{cases} 0, & \text{if subscriber is in } i\text{'s area} \\ a_j, & \text{if subscriber is in } j\text{'s area} \end{cases}.$$

Subscribers. For each city covered by fiber we assume a "Hotelling model with hinterlands" specification regarding subscribers' choices.[26]

---

access price. Therefore, indexation may incentivize more infrastructure duplication than a fixed access price. However, note that when the access price is higher, operators have further incentives to deploy their own network, which may result in inefficient duplication of infrastructures. In Proposition 2, we argue that, in a symmetric equilibrium, the access price resulting from the indexation approach can be lower than the one resulting from a fixed access approach. Thus, it may be also the case that indexation is more efficient than a fixed access price in terms of duplication of infrastructures.

[24]Network installation cost also varies significantly between areas due to differences in geology and ground cover. The convexity of costs also applies to postal services and third generation mobile telephone systems. See Foros and Kind (2003).

[25]The main results in this chapter are not dependent on the quadratic form of $C(I_i)$. Results hold as long $C(I_i)$ is sufficiently convex such that it guarantees that networks' profits as a reduced-form function of investments are concave with respecto to own investment. This is shown in Theorem 2.

[26]See the section *Mobile market expansion* in Amstrong and Wright (2009) for another application of the Hotelling model with hinterlands.
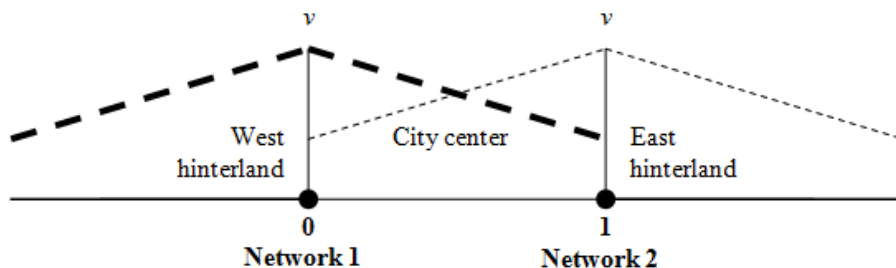
Figure 1: Hotelling model with hinterlands - the representative city. Consumers' willingness to pay for broadband service offered by network 1 and 2 represented by bold and thin dashes, respectively.

From the subscriber perspective there is some service differentiation among networks for reasons such as technical support, proximity to clients, marketing campaigns, advertising or switching costs. Each city comprises by the center plus two symmetric hinterlands (West and East sides of the city center) as in Figure 1. We assume that subscribers located in the city center, indexed by $\tilde{x} \in [0,1]$, are fully served, while consumers in the hinterlands, indexed by $\tilde{y}$, are partially served. In a representative city the surplus of a consumer indexed by $\tilde{x}$ and $\tilde{y}$ is defined by, respectively, $CS_{\tilde{x}}$ and $CS_{\tilde{y}}$

$$CS_{\tilde{x}} \equiv \begin{cases} v - t\tilde{x} - p_1 & \text{if operator 1} \\ v - t(1-\tilde{x}) - p_2 & \text{if operator 2} \end{cases}, \tag{3}$$

$$CS_{\tilde{y}} \equiv \begin{cases} v - t\tilde{y} - p_i & \text{if nearest operator } i = 1,2 \\ 0 & \text{if no service} \end{cases} \tag{4}$$

where v is the intrinsic value from subscribing to the service and t measures the subscriber's disutility of not being connected to their ideal taste network. We assume that $v > t > 0$, i.e., service differentiation is sufficiently small as compared to the intrinsic value v of the service. There is a total mass 2 of consumers in a representative city. In the city center represented by the unit interval $[0,1]$ there is a mass 1 of consumers uniformly distributed with density 1, while in each hinterland there is a mass $1/2$ of consumers uniformly distributed with density $t/2v$ on intervals $\left[0, \frac{v}{t}\right]$, i.e., $\tilde{y} \in \left[0, \frac{v}{t}\right]$.[27] Gross consumer surplus (utility), U, and consumer surplus, CS, in a city are thus

$$U(x_1, z_1, x_2, z_2) \equiv v(x_1 + x_2 + z_1 + z_2) - \frac{t(x_1^2 + x_2^2) + 2v(z_1^2 + z_2^2)}{2} \tag{5}$$

$$CS \equiv U - \sum_{i=1}^{2} p_i(x_i + z_i) \tag{6}$$

---

[27] By assuming density $t/2v$ in the hinterlands we guarantee a fixed mass $1/2$ of consumers in each hinterland. Otherwise the number of consumers in hinterlands would depend on t and v.

28

where $0 \leq x_i \leq 1$ and $0 \leq z_i \leq \frac{1}{2}$ denote the mass of subscribers located in the city center and hinterlands, respectively, using network i's service. Note that $z_i \equiv y_i \frac{t}{2v}$, where y_i is the distance to the nearest operator (city center) and z_i may be interpreted as the mass of subscribers, using network i's service, along that distance. Since consumers are fully served in the city center we have then $x_1 + x_2 = 1$.

The individual consumer surpluses from (3) and (4) imply that

$$x_i = \frac{1}{2} - \frac{p_i - p_j}{2t}, \; z_i = \frac{v - p_i}{2v} = \frac{1}{2} - \frac{p_i}{2v}. \tag{7}$$

Therefore,

$$q_i \equiv x_i + z_i = 1 - \frac{(v + t)\,p_i - vp_j}{2tv} \text{ and } Q \equiv q_i + q_j = 2 - \frac{p_i + p_j}{2v}. \tag{8}$$

A summary of the model's notation follows in Table 2.

<div align="center">TABLE 2: Notation</div>

a_i	Access price, per subscriber, charged by network i.
x, y	Regulatory parameters under the indexation rule.
v	Intrinsic value from subscribing to a fiber service.
t	Service differentiation parameter.
c	Investment cost parameter.
I_i	Number of cities covered by fiber installed by network i.
I	Total number of cities covered by fiber, defined as $I \equiv I_1 + I_2$.
p_i	Retail price charged by network i for fiber optic broadband service.
x_i	Mass of subscribers located in the city center using network i's service.
z_i	Mass of subscribers located in the hinterlands of a city using network i's service.
q_i	Total mass of subscribers using network i's service in a city, defined as $q_i \equiv x_i + z_i$.
Q	Total mass of fiber broadband subscribers in a city, defined as $Q \equiv q_1 + q_2$.
U	Gross consumer surplus in a city.
CS	Consumer surplus in a city.

1.3   The social optimum benchmark

In order to assess the fixed and the indexation access price rules from the social welfare standpoint we compute, as benchmark, the first-best solution that a benevolent planner could achieve. Let the social value of providing fiber broadband be defined as the sum of consumer surpluses, CS, in all cities covered by fiber, plus networks' subscription revenues

minus the costs with regard to fiber coverage. Access prices are mere transfers among networks, therefore access revenues minus the access costs across operators sum up to zero. For that reason access prices are not relevant in the first-best analysis. In other words, the measure of social welfare taken is the unweighted sum of consumer surplus in all cities covered by fiber and the industry profit.

Given that the cities covered by fiber are identical, x_1, z_1, x_2, z_2 must be the same across them. Hence, in the first-best a benevolent regulator would solve

$$\max_{x_1,z_1,x_2,z_2,I_1,I_2} W \equiv (I_1 + I_2) U - c \left(\frac{I_1^2}{2} + \frac{I_2^2}{2} \right) \tag{9}$$
$$\text{subject to } x_1 + x_2 = 1.$$

From the first-order conditions (FOCs) of the problem it follows that

$$\begin{cases} x_i^{opt} = \frac{1}{2}, \ z_i^{opt} = \frac{1}{2} \\ \qquad p_i^{opt} = 0 \\ I_i^{opt} = \frac{1}{4c} (6v - t) \\ U^{opt} = \frac{1}{4} (6v - t) \\ \qquad W^{opt} = \frac{(6v-t)^2}{16c} \end{cases} \tag{10}$$

The efficient retail prices correspond to the social marginal cost of serving a fiber subscriber, i.e., zero by assumption. Thus, it is socially optimal to supply FTTH to all consumers where fiber broadband is available. Due to symmetry of willingness to pay for fiber broadband service between networks, the welfare-maximizing market shares in the city center and hinterlands are given by $x_i^{opt} = 1/2$ and $z_i^{opt} = 1/2$, respectively. The efficient network size (investment) is driven by parameters v and t, which affect the gross consumer surplus in a city covered by fiber, and the cost of covering an additional city with fiber, which is affected by parameter c. It is noteworthy that in the absence of lump-sum transfers the social optimum is not feasible under any access price rule per subscriber. In the social optimum $p_i^{opt} = 0$ and as a consequence networks would not extract revenues from subscribers, while the access revenue $aq_j^{opt} I_i^{opt}$ is equal to access cost $aq_i^{opt} I_j^{opt}$ under symmetry. Profits would be then negative

$$\Pi_i^{opt} = I^{opt} \times p_i^{opt} q_i^{opt} + aq_j^{opt} I_i^{opt} - aq_i^{opt} I_j^{opt} - c \left(I_i^{opt} \right)^2 /2 = -c \left(I_i^{opt} \right)^2 /2 < 0$$

and networks would prefer to exit the market. Therefore, the first-best solution is not feasible without lump-sum transfers that cover the networks' investment cost. We can conclude from here that to maximize social welfare subject to non-negative profits we would derive strictly positive Ramsey prices.

Bearing in mind the first-best benchmark in (10), in the following section we establish

comparisons between the equilibria under: a fixed access price, an indexation access price rule and a regulatory holidays regime.

1.4 The subgame perfect Nash equilibria

In this section we compare equilibria where networks operate under a fixed access price, an indexation access price rule and a regulatory holidays regime. We use backward induction to solve the model for a symmetric subgame perfect Nash equilibrium under each regulatory regime. First, given a regulatory regime and investment levels, we solve the networks' problem for the profit maximizing retail prices. Second, given a regulatory regime we solve the networks' problem for optimal investments. Third, we solve the regulator's problem for welfare-maximizing access pricing rules (fixed and indexed access prices). Technical details and calculations follow in an appendix.

1.4.1 The fixed access price approach

Below we solve the model for the symmetric subgame perfect Nash equilibrium under a fixed access price rule. We identify economic inefficiencies related to this type of regulation. In particular, we stress that neither investment efficiency nor retail price efficiency is feasible under a fixed access price.

Stage III: retail price choices. In the retail pricing stage, operator i's problem is, given an access price a, a pair of investment levels (I_i, I_j) and p_j,

$$\max_{p_i} \Pi_i = I \times p_i q_i + a q_j I_i - a q_i I_j - c I_i^2/2, \tag{11}$$

where $I \times p_i q_i$ corresponds to i's subscription revenues and $a q_j I_i - a q_i I_j$ represents the access revenue received from network j subtracted from the access cost paid. Term $c I_i^2/2$ is the cost of covering I_i cities with fiber. From the FOC of the problem in (11), in equilibrium we get

$$p_i^* = \frac{\left(3v\left(2t + a\right) + 4t^2\right)v\left(I_i + I_j\right) + at\left(v\left(3I_i + 4I_j\right) + 2tI_j\right)}{\left(2t + v\right)\left(2t + 3v\right)\left(I_i + I_j\right)}, \; i, j = 1, 2, \; j \neq i, \tag{12}$$

as long $p_i^*(a) \leq v - t/2$. In plain words, in equilibrium the retail price must be below the willingness to pay of the consumer in the middle of the city center. Otherwise, the full coverage assumption of the city center would not hold. Note that in the case of symmetric investments, i.e., $I_i = I_j$, the price equilibrium in (12) will be valid under the

constraint $a \leq \bar{a} \equiv \left(2v^2 - tv - 2t^2\right) / \left(t + 2v\right)$. Plugging (12) into (8) we get that the mass of consumers subscribing to i's broadband fiber broadband in a city is

$$q_i^* = \frac{\left(6v^2 + 10tv + 4t^2\right) v \left(I_i + I_j\right) - 2a \left(t^2 I_j + v^2 I_i\right) - 4v^2 a I_j - atv \left(I_i + 6I_j\right)}{2v \left(2t + v\right) \left(2t + 3v\right) \left(I_i + I_j\right)} \quad (13)$$

and the total mass of fiber subscribers in a city is

$$Q^* = \frac{4v \left(t + v\right) - a \left(t + 2v\right)}{2v \left(2t + v\right)}.$$

Stage II: investment choices. In the investment stage under a fixed access price, a, network i's maximization problem is

$$\max_{I_i} \Pi_i^* = I \times p_i^* q_i^* + aq_j^* I_i - aq_i^* I_j - cI_i^2/2,$$

where p_i^* is defined by (12) and q_i^* by (13) in the retail price stage.

From the FOCs of networks' problems in the investment stage, $\partial \Pi_i^* / \partial I_i = 0$, in equilibrium we reach an investment level (per network) of

$$I_i^* = \frac{\left(t + 2v\right) \left(8v \left(6tv + 4t^2 + 3v^2\right) - a \left(25tv + 14t^2 + 12v^2\right)\right) a + 16 \left(t + v\right) \left(2t + 3v\right) tv^2}{8cv \left(2t + 3v\right) \left(2t + v\right)^2}. \quad (14)$$

Stage I: access price regulation. We compare now the first-best solution to the equilibrium outcome under a fixed access price rule. We claim that under a fixed access price rule a regulator cannot induce the socially optimal level of investment, regardless of how much static efficiency is sacrificed. Moreover, given that negative access prices are not implemented in practice, we note that retail price efficiency cannot be achieved.

Proposition 1 (underinvestment) *Under a fixed access price (i) it is not possible to implement the socially optimal investment, i.e., there is underinvestment $I_i^* < I_i^{opt}$, and (ii) retail price efficiency would require a negative access price.*
Proof All proofs are in an appendix. $\square$

The intuition for the underinvestment result with a fixed access price comes from the fact that networks are unable to capture the full social benefit of investment. This inability stems from retail price competition and uniform pricing. The fixed access price is

a regulatory tool that may incentivize investments but sacrifice retail price competition, i.e., by increasing the access price. Nonetheless, even if the access price were set to maximize the investment outcome, there would be benefits captured by the subscribers due to their heterogeneity in the willingness to pay for the fiber broadband service and the fact that networks are, generally, unable to price discriminate to extract the subscribers' full valuations. Moreover, if networks were able to practice first-degree price discrimination, retail price competition would imply positive surplus to subscribers. Hence, in general, networks cannot fully internalize the benefits from their investments which implies a choice that is necessarily inefficient. With regard to retail price (in)efficiency, due to the existence of market power in the downstream market, the access price would have to be negative to counterbalance the market power effect. In a nutshell, bearing in mind that negative access prices are not implemented in practice, the equilibrium outcome is condemned to underinvestment and retail price inefficiency. This result holds regardless of the access price being privately bargained between networks in an unregulated market or being set by a benevolent regulator.[28]

Theorem 1 extends Proposition 1 to a set of more general assumptions. The model discussed earlier in this chapter satisfies all the assumptions in Theorem 1. Note that the demand for fiber broadband in a city does not change with consumers' income. This follows the Marshallian notion that when a good represents a small fraction of the total expenditure of a consumer then income effects become negligible.[29] Let $\max U = U^{opt}$ where U^{opt} is a positive constant denoting the maximum gross consumer surplus (willingness to pay) derived from the fiber broadband usage in a city. Hence, the gross consumer surplus level U^{opt} can be interpreted as the maximum revenue that networks can collect from fiber broadband users in a city, thus, $U^{opt} \geq \sum_{i=1}^{2} p_i^* q_i^*$.

Theorem 1 (underinvestment) *Consider a sequential game such that the regulator chooses the access charge a^* before networks compete first in investments and second in retail prices, and the following conditions hold: (a) network i's profit is defined by $\Pi_i \equiv (I_i + I_j) \times p_i q_i + a q_j I_i - a q_i I_j - C(I_i)$, where $C(I_i)$ is an increasing, twice differentiable and sufficiently convex cost function ensuring that network i's profit as a reduced-form function of investments is strictly concave in I_i; (b) social welfare measure is $W \equiv \sum_{i=1}^{2} [\Pi_i + I_i \times CS]$ where $CS \equiv U - \sum_{i=1}^{2} p_i q_i$ and U is a twice differentiable, strictly concave function in (q_i, q_j) such that $U^{opt} \geq p_i^* q_i^* + a q_j^* + \frac{\partial \Pi_i}{\partial p_j} \frac{\partial p_j^*}{\partial I_i}$; and (c) $q_i(p_i, p_j)$*

is twice differentiable and non-increasing in $(p_i, -p_j)$.

Thus, under a fixed access price (i) is not possible to implement the socially optimal investment, i.e., there is underinvestment $I_i^* < I_i^{opt}$, and (ii) retail price efficiency requires a negative access price.

To find the socially optimal number of cities covered by fiber, the regulator equates the marginal social benefit from covering an additional city by fiber, U^{opt}, to the marginal cost $C'(I_i)$. However, network i equates the marginal private benefit $p_i^* q_i^* + a q_j^* + \frac{\partial \Pi_i}{\partial p_j} \frac{\partial p_j^*}{\partial I_i}$ to the marginal cost $C'(I_i)$. Note that $p_i^* q_i^*$ is the revenue from selling fiber broadband to final consumers in a city, $a q_j^*$ is the revenue from selling access to network j in a city and $\frac{\partial \Pi_i}{\partial p_j} \frac{\partial p_j^*}{\partial I_i}$ is the strategic effect that i's investment has on j's retail price and consequently on i's profit. This strategic effect is due to the fact that j's best-reply in terms of retail price depends on the investment levels (I_i, I_j). Note that the marginal cost for network j to serve fiber broadband to a customer is zero if the customer is in a city covered by j, otherwise j has to pay an access charge a. Nonetheless, the retail price is set nationwide, i.e., network j charges the same retail price across cities. This means that if network i increases its fiber coverage and j responds by increasing the retail price in all cities (because, on average, there is an increase in the marginal cost per customer) network i may increase profits via this strategic effect. We assume that the strategic effect is relatively small such that $U^{opt} \geq p_i^* q_i^* + a q_j^* + \frac{\partial \Pi_i}{\partial p_j} \frac{\partial p_j^*}{\partial I_i}$ holds. Since networks do not fully internalize the benefits from their investments this implies a choice that is necessarily inefficient. The intuition for retail price (in)efficiency is the same as discussed for Proposition 1.

1.4.2   The new rule: access price indexation

Given the inefficiencies in the use of a fixed access price, we consider a new rule by indexing access prices to networks' investments. This new indexation rule has the purpose of increasing investment incentives without sacrificing static efficiency and, ultimately, boosting social welfare. In particular, the new access price rule is defined in (1) where $(x, y) \in \mathbb{R}_+^2$ is the pair of regulatory parameters to be determined. Under access price indexation, besides the impact via retail and access revenues, investments affect networks' profits via changes in access prices.

We solve the three-stage game under the new rule and compare it to the equilibrium obtained under a fixed access price. We also solve the game for the case where the regulator's goal is to implement the socially efficient (first-best) level of fiber coverage with the lowest possible retail pricing. In Proposition 2 we show that access price indexation can increase social welfare relatively to a fixed access pricing. Theorem 2 extends Proposition 2 to a set of more general assumptions. In response to the title of this chapter, in

Proposition 3 we show that under certain conditions the access price indexation can promote the socially optimal investment. This may be particularly useful, for example, to meet a universal service obligation. Moreover, we show that under certain conditions the Ramsey outcome is feasible with the access price indexation.

Stage III: retail price choices under indexation. With access price indexation, operator i's optimization problem in the retail pricing stage is

$$\max_{p_i} \Pi_i = I \times p_i q_i + a_i q_j I_i - a_j q_i I_j - cI_i^2/2. \tag{15}$$

Note that the problem in (15) is different from the one in (11) since access prices may now differ among operators depending on investment levels.

Taking the FOC of the problem in (15) and solving for the equilibrium retail prices we get

$$p_i^{**} = \frac{v\left(3ta_i + 3va_i + 6tv + 4t^2\right)I_i + \left(6tv^2 + 4t^2v + 2t^2a_j + 3v^2a_j + 4tva_j\right)I_j}{(2t + v)(2t + 3v)(I_i + I_j)}. \tag{16}$$

Plugging (16) into q_i in (8) we reach

$$q_i^{**} = \frac{v\left(10tv + 4t^2 + 6v^2 - ta_i - 2va_i\right)I_i + 2\left(t + v\right)\left(2tv + 3v^2 - ta_j - 2va_j\right)I_j}{2v(2t + v)(2t + 3v)(I_i + I_j)}. \tag{17}$$

Stage II: investment choices under indexation. In the investment stage, network i's maximization problem is

$$\max_{I_i} \Pi_i^{**} = I \times p_i^{**} q_i^{**} + a_i q_j^{**} I_i - a_j q_i^{**} I_j - cI_i^2/2,$$

where p_i^{**} is defined in (16) and q_i^{**}, and analogously q_j^{**}, in (17). The network's optimal investment is now defined by

$$\frac{d\Pi_i^{**}}{dI_i} \equiv \underbrace{\frac{\partial \Pi_i^{**}}{\partial I_i}}_{\text{"direct effect"}} + \underbrace{\sum_{k=1}^{2} \frac{\partial \Pi_i^{**}}{\partial a_k} \frac{\partial a_k}{\partial I_i}}_{\text{"indexation effect"}} = 0, \tag{18}$$

while under the fixed access approach only the "direct effect" exists. The "direct effect" accounts for the marginal private benefit and marginal cost of covering cities with fiber assuming that access prices are held constant. The "indexation effect" accounts for the

network's profit variation due to changes in access prices (both paid and received by the network) that come via the network's investment.

Under a fixed access price $a_k = a^*$ does not depend on investments, thus $\partial a_k / \partial I_i = 0$ and the "indexation effect" is zero. Under the indexation approach the "indexation effect" is positive (see the proof of Proposition 2). In particular, for $(x, y) \in \mathbb{R}_+^2$

$$
\underbrace{\frac{\partial \Pi_i^{**}}{\partial a_i}}_{(+)} \underbrace{\frac{\partial a_i}{\partial I_i}}_{(+)} + \underbrace{\frac{\partial \Pi_i^{**}}{\partial a_j}}_{(-)} \underbrace{\frac{\partial a_j}{\partial I_i}}_{(-)} > 0
$$

meaning that network i's profit increases (decreases) if i charges (pays) a higher access price, and the access price charged (paid) increases (decreases) in i's investment. In plain English, the marginal benefit from investment is higher when networks are under an access charge indexed to investments than under a fixed access charge. As a consequence, networks have incentive to invest more under the indexation approach than under fixed access.

Assuming investment symmetry, operator i's FOC for investment can be written as

$$
\left\{
\begin{array}{c}
3\left(t+2v\right)\left(y-x\right)\left[\left(10\left(t^2+v^2\right)+19tv\right)x-\left(2t+3v\right)\left(t+2v\right)y\right]I_i^2+ \\
-8v\left[\left(2t+3v\right)\left(c\left(2t+v\right)^2+v\left(t+2v\right)y\right)-2\left(t+2v\right)\left(6tv+4t^2+3v^2\right)x\right]I_i+ \\
+16tv^2\left(t+v\right)\left(2t+3v\right)
\end{array}
\right\} = 0,
$$

$$(19)$$

while the following inequality has to be satisfied for the second-order condition (SOC) to hold,

$$
S \equiv - \left\{
\begin{array}{c}
8v\left(2t+3v\right)\left[c\left(2t+3v\right)\left(2t+v\right)^2-2\left(t+2v\right)\left(6tv+4t^2+3v^2\right)x\right]+ \\
+\left(t+2v\right)\left(
\begin{array}{c}
\left(132t^3+162v^3+433tv^2+404t^2v\right)x^2+ \\
+\left(t+2v\right)\left(2t+3v\right)^2y^2-6\left(2t+3v\right)\left(19tv+10t^2+10v^2\right)xy
\end{array}
\right)
\end{array}
\right\}_{I_i^{**}} < 0.
$$

$$(20)$$

Stage I: regulator's choice under indexation. The welfare-maximizing regulator solves the following problem

$$\max_{x,y} W \equiv (I_i + I_j) U - c \left(I_i^2 + I_j^2 \right) / 2 \text{ subject to} \tag{21}$$

$$x_i^{**} (I_i, I_j, x, y) = x_i^{**} \text{ (Stage III)}$$

$$z_i^{**} (I_i, I_j, x, y) = z_i^{**} \text{ (Stage III)}$$

$$q_i^{**} (I_i, I_j, x, y) = q_i^{**} \text{ (Stage III)}$$

$$p_i^{**} (I_i, I_j, x, y) = p_i^{**} \text{ (Stage III)}$$

$$d\Pi_i^{**}/dI_i = 0 \text{ (Stage II)}$$

$$d^2\Pi_i^{**}/dI_i^2 \leq 0 \text{ (Stage II)}$$

$$\Pi_i^{**} \geq 0 \text{ (PC)},$$

where PC denotes network i's participation constraint. Plugging the restrictions from stage III into the objective function in (21) and assuming investment symmetry, the regulator's problem under the indexation approach can be rewritten as

$$\max_{x,y} W = I_i \frac{\left(\begin{array}{c} 2v \left(23tv^2 + 12t^2v - 4t^3 + 6v^3 \right) + \\ -4v \left(c \left(2t + v \right)^2 + 2t \left(t + 2v \right) \left(x - y \right) \right) I_i + \\ - \left(x - y \right)^2 \left(t + 2v \right)^2 I_i^2 \end{array} \right)}{4v \left(2t + v \right)^2} \text{ subject to}$$

$$d\Pi_i^{**}/dI_i = 0 \text{ (Stage II)}$$

$$d^2\Pi_i^{**}/dI_i^2 \leq 0 \text{ (Stage II)}$$

$$I_i \frac{16tv^2 \left(v + t \right) - 2v \left(4ct \left(v + t \right) + cv^2 + v \left(2t + 4v \right) \left(y - x \right) \right) I_i - \left(x - y \right)^2 \left(t + 2v \right)^2 I_i^2}{4v \left(2t + v \right)^2} \geq 0 \text{ (PC)}.$$

Lemma 1 (participation constraint binds) *In a symmetric equilibrium, networks' participation constraint will be binding, i.e., $\Pi_i = 0$, with a welfare-maximizing regulator using an access price indexation yielding $(x, y) \neq (0, 0)$ and $a^* \geq 0$.*

Recall that in the social optimum (first-best) networks' would have negative profits. Given that the social optimum is not feasible in the absence of transfers, the best that a welfare-maximizing regulator can do under the indexation approach is to choose a regulatory regime (x, y) subject to networks' zero-profit condition. If networks presented positive profits in equilibrium, the regulator could enhance the social welfare by choosing (x, y) such that retail prices were lower (i.e., improving static efficiency) and/or investments in fiber coverage were higher (i.e., improving dynamic efficiency).

Proposition 2 (indexation vs fixed access) *In a symmetric equilibrium, a linear access pricing rule depending on investments with $(x, y) \in \mathbb{R}_+^2$ can simultaneously (i)*

expand total investment in fiber coverage, (ii) expand the mass of subscribers in each city,
and (iii) enhance social welfare, as compared to a fixed access price $a^* > 0$.

The introduction of an access price indexation can create a scheme of rewards to in-
vestors and punishment to those who do not invest or invest relatively less. In particular,
the network that invests relatively more will benefit from a lower access price when ac-
cessing other network, while benefiting from a higher price when providing access. As a
result of the additional incentives to investment derived from the indexation approach,
for a same access price, a^*, in equilibrium networks invest more than under a fixed access
price a^*.

The total mass of fiber subscribers depends on the retail price level which in turn
depends on the access price level. Therefore, if the equilibrium access price under the
indexation rule is below the one defined under a fixed access price rule, the mass of
subscribers will be higher under the former, rather than under the latter rule. Suppose
that with a fixed access price rule the access price is set at a^*. Under the access price
indexation the regulator can choose (x, y) such that $a_i = xI_i - yI_j = a^* - \varepsilon$, $\varepsilon > 0$, while
for ε sufficiently small the investments $\left(I_i^{**}, I_j^{**}\right)$ are above the equilibrium investment
levels under a fixed access. With investment symmetry in equilibrium, we can write
$x - y < a^*/I_i$. Note that x and y can be set at a level such that the difference $x - y$ is
sufficiently small to assure that the access price will be below a^*. However, x and y must
be set sufficiently high to incentivize more investment than under a fixed access charge.

In relation to social welfare, we note that, in equilibrium, the mass of subscribers
in each city is higher under the indexation approach than under a fixed access charge.
Consequently, both the gross consumer surplus in each city and the marginal social benefit
from investment increase. In fact, in the fixed access equilibrium the marginal social
benefit from investment is positive, implying that further investment would enhance the
social welfare level. Hence, if the fixed access rule is the *status quo*, the social welfare
variation due to the implementation of access price indexation must be positive. This is
explained by the increase of gross consumer surplus in each city together with the increase
in the number of cities covered by fiber (while the cost of covering an additional city by
fiber is lower than the gross consumer surplus generated).

In Theorem 2 below we claim that Proposition 2 is robust to a set of more general
assumptions.

Theorem 2 (indexation vs fixed access) *Consider a sequential game such that the
regulator chooses the access price a_i before networks compete first in investments and
second in retail prices, and the following conditions hold: (a) network i's profit is defined
by $\Pi_i = (I_i + I_j) \times p_i q_i + a_i q_j I_i - a_j q_i I_j - C(I_i)$, where $C(I_i)$ is an analytic,[30] increasing*

---

[30] An analytic function is a function that is locally given by a convergent power series. Typical examples

and sufficiently convex cost function ensuring that network i's profit as a reduced-form function of investments is strictly concave in I_i; (b) $\partial \Pi_i / \partial p_j \geq 0$ in equilibrium; (c) Π_i is strictly concave in p_i; (d) $q_i(p_i, p_j)$ is an analytic function and non-increasing in $(p_i, -p_j)$; (e) social welfare measure is $W \equiv \sum_{i=1}^{2} [I_i U - C(I_i)]$; and (f) gross consumer surplus in a city, U, is differentiable and decreasing in $(p_i, p_j) \in \mathbb{R}^2_+$.

Thus, in a symmetric equilibrium, an access pricing rule depending on investments can simultaneously (i) expand total investment in fiber coverage, (ii) expand the mass of subscribers in each city, and (iii) enhance social welfare, as compared to a fixed access price $a^* > 0$.

Under the assumptions set out in Theorem 2, for a same access price a^*, the marginal private benefit from investment for a network is higher under the indexation approach than under a fixed access charge. This comes as a consequence of the additional incentives to investment generated by the indexation approach. Therefore, in a symmetric equilibrium where the access charge is a^*, networks have incentives to cover more cities with fiber under the indexation approach, rather than under a fixed access charge.

Under the indexation approach it is feasible to expand of the number of cities covered with fiber even if in equilibrium the access price, $a^* - \varepsilon$, $\varepsilon > 0$, is lower than the fixed access price a^*. To check this we rely on the assumption that Π_i is an analytic function so that it can be rewritten as a Taylor series around a fixed access price a^*. By doing this we can isolate the effect of ε and show that for a sufficiently small (but strictly positive) ε, the marginal private benefit from investment for a network is still higher under the indexation approach than under a fixed access charge at a^*. Thus, networks will have further incentives to cover more cities with fiber under the indexation approach even if, in equilibrium, the access price is below (but sufficiently close to) a^*.

In relation to social welfare, a reduction in retail prices diminishes static inefficiency while enhancing the gross consumer surplus, U, in each and every city covered with fiber. Given that U increases when the regulator imposes an indexation rule, this means that the social marginal benefit from investment, $U - C'(I_i)$, increases as well (when compared to the use of a fixed access price). If the social marginal benefit from investment in equilibrium under a fixed access price is positive (i.e., $U^* - C'(I_i^*) > 0$), then an increase in investment will boost social welfare. Otherwise, the regulator would not need to incentivize more investment and the social welfare level would increase via access price reduction only.

In the following section we show that with an indexation rule, contrarily to a fixed access price, a regulator may aspire to achieve objectives such as the socially efficient (first-best) investment level and the Ramsey outcome.

---

of analytic functions are: polynomial, exponential, trigonometric, logarithm, and power functions. Any analytic function is smooth, that is, infinitely differentiable.

Stage I revisited: universal service and the Ramsey solution. A regulator may choose a regulatory policy (x, y) with the purpose of implementing the first-best investment level, $I_i^{opt} = (6v - t) / (4c)$. This action may result, for instance, from the existence of a universal service requirement. Universal service is an economic, legal and business term used mostly in regulated industries, referring to the practice of providing a baseline level of services to all residents of a country at an affordable price. Examples of this concept may be found in the Telecommunications Act of 1996[31] and the Directive (2002/22/EC) of the European Parliament and of the Council of the European Union of 7 March 2012. In this section we suggest that the access price indexation may be used to promote an objective of universal service. In particular, we argue that if the degree of service differentiation, t, is sufficiently small, then there will exist a regulatory policy (x, y) such that the first-best investment level can be implemented.

Within the context of this chapter, the Ramsey solution is a policy concerning what price and investment a monopolist would set, in order to maximize social welfare, subject to a constraint of non-negative profit. In Proposition 3 we set out the conditions under which the Ramsey solution is feasible.

Proposition 3 (first-best investment level and Ramsey solution) *In a symmetric equilibrium, if service differentiation, t, is sufficiently small, a linear access pricing rule depending on investments can implement (i) the first-best investment level or (ii) the Ramsey solution.*

A regulatory regime (x, y) will implement the first-best investment if, at $I_i = I_i^{opt}$, the pair (x, y) passes three tests: (i) the network FOC in (19), (ii) the SOC, which is characterized by inequality (20), and (iii) non-negative profits. In equilibrium, for $I_i = I_i^{opt}$, the regulator chooses (x, y) such that networks have zero profits and simultaneously satisfy the FOC. A small differentiation parameter is a sufficient (but not necessary) condition to ensure that the SOC is satisfied in networks' problems. Intuitively, if service differentiation is small it implies fiercer price competition between the two networks. When price competition is more intense, networks have further incentives to cut prices in the sense that a small price cut shifts a large mass of subscribers towards the network with the lowest price. Therefore, if price competition becomes fiercer, operator i will have more incentives to invest under the indexation approach as a means to inflate a_i and reduce a_j, achieving a competitive advantage at the retail pricing stage. If service differentiation decreases, it will be easier for the regulator to ensure the implementation of higher investment levels, namely the first-best investment level.

---

[31] The US Telecommunications Act of 1996, set out the following goals: (i) to promote the availability of quality services at just, reasonable, and affordable rates; (ii) to increase access to advanced telecommunications services throughout the Nation; and (iii) to advance the availability of such services to all consumers, including those in low income, rural, insular, and high cost areas at rates that are reasonably comparable to those charged in urban areas.

The Ramsey investment, I_i^{Ramsey}, is lower than I_i^{opt},[32] and at $(x,y) = (0,0)$ the investment of equilibrium under the indexation approach is lower than I_i^{Ramsey}. Therefore, by continuity of the FOC that defines I_i^{**} under the indexation approach, we can guarantee the existence of a pair (x,y) such that the investment in equilibrium is equal to the Ramsey investment. Recall that, from Lemma 1, the participation constraint binds, and at $I_i^{**} = I_i^{Ramsey}$ the retail price in equilibrium shall be set such that networks have zero-profit. Hence, in a symmetric equilibrium, p_i^{**} corresponds to Ramsey pricing.

Note that under symmetry, as $p_j = p_i$ and $I_j = I_i$, the Ramsey outcome is simply characterized by two instruments: p_i^{Ramsey} and I_i^{Ramsey}. Under the indexation approach, p_i^{**} and I_i^{**} are functions of two regulatory instruments: x and y. Thus, for a sufficiently small degree of differentiation between fiber broadband services, we can find a mapping between (x,y) and $\left(p_i^{Ramsey}, I_i^{Ramsey}\right)$. However, in the event of an asymmetric Ramsey outcome characterized by four different instruments $\left(p_i^{Ramsey}, p_j^{Ramsey}, I_i^{Ramsey}, I_j^{Ramsey}\right)$, the Ramsey outcome may offer a higher social welfare level than an access price indexation with only two regulatory instruments (x,y).

1.4.3 Regulatory holidays

In this section we show that the indexation rule can perform better than regulatory holidays with regard to fiber coverage, retail prices and social welfare.[33] In a city monopolized by operator i (located at point 0 and point 1) that is unable to price discriminate, the demand function faced by the monopolist serving both hinterlands is defined by[34]

$$q_i = \begin{cases} 2\left(\frac{v-p_i}{2v} + \frac{v-p_i}{t}\right) & \text{if } v \geq p_i > v - \frac{t}{2} \\ \frac{v-p_i}{v} + 1 & \text{if } 0 \leq p_i \leq v - \frac{t}{2} \end{cases}.$$

Operator i chooses p_i and I_i solving the following maximization problem

$$\max_{p_i, I_i} \Pi_i^{mon} = I_i p_i q_i - c I_i^2/2.$$

We demonstrate in an appendix that each monopoly network chooses to charge the retail price $p_i^{mon} = v - t/2$ serving $q_i^{mon} = 1 + t/(2v)$ subscribers in each city. Each network covers $I_i^{mon} = (4v^2 - t^2)/(4cv)$ cities with fiber and attains a profit level of $\Pi_i^{mon} =$

---

[32] The social marginal benefit from investment is maximal when U is maximized, which happens in the first-best solution when retail prices are set at zero.

[33] The regulatory holidays case can also be seen as a special case of the indexation rule, for example, by setting $x = \infty$, $y = 0$. In this case, for any investment $I_i > 0$ the access price, under the indexation rule, becomes infinite which is equivalent to granting local monopolies.

[34] In an appendix we solve for both the cases where the monopolist offers one or two brands, i.e., it is present in one or both hinterlands of a city. Assuming that $v > 2t$, the results in Proposition 4 hold regardless of the monopolist's presence in one or both hinterlands. Here we present the two-brand case which only requires $v > t$ to verify Proposition 4.

$(2v - t)^2 (2v + t)^2 / (32cv^2)$. The inefficiency of the monopoly with respect to retail prices and investments is clear since $p_i^{mon} = v - t/2 > 0 = p_i^{opt}$ and $I_i^{mon} = (4v^2 - t^2) / (4cv) < (6v - t) / (4c) = I_i^{opt}$, provided that $v > t$. In a nutshell, the retail price inefficiency derives from networks' market power, while investment inefficiency is due to a part of the surplus generated by the fiber service being captured by consumers (given uniform pricing and no lump-sum transfers). We conclude that the indexation approach can do better than regulatory holidays in terms of social welfare.

Proposition 4 (regulatory holidays) *A linear access pricing rule depending on investments can simultaneously decrease retail prices and increase both investment and social welfare levels as compared to the regulatory holidays regime (i.e., local monopolies).*

Granting a local monopoly expands total investment relative to a fixed access price but at the cost of a retail price distortion which reduces the mass of fiber subscribers. The regulatory holidays regime is dominated by the proposed access price indexation rule, both in terms of investment (broadband coverage by fiber) and retail price efficiency, resulting in higher social welfare with the indexation approach rather than with regulatory holidays. Intuitively, the regulatory holidays policy consists of alleviating price competition pressure in order to increase the investment rewards as a way to encourage further investment. The indexation approach works in the opposite direction proposing a "tournament" where networks have incentives to compete in investments.

1.5 Informational issues under access price indexation

The assumption that investments (i.e., fiber broadband coverages) are observable and verifiable to a third party is fundamental for the access price indexation to fulfill its intended outcomes (as set out in previous sections). In this section we discuss the reasonableness of this assumption, bearing in mind that networks may have an incentive to use the access price indexation to increase their own profits. In particular, in order to gain a competitive advantage at the downstream level and, ultimately, increase profits, networks may have an incentive to report wider fiber broadband coverage than they actually have.

We argue that regulators may, at least in part, observe and verify (at some cost) the fiber optic infrastructures. The economics of FTTH network deployment is usually characterized by high fixed costs of which the dominant component is the civil works: digging the roads (including construction permits) and laying ducts, whose existence is observable and verifiable.[35] Moreover, regulators engage with stakeholders in the industry, therefore,

---

[35] In some areas, pole distribution may be the norm; while in others direct buried cable can be used as well.

if a network reports a fiber coverage that does not effectively own, eventually another network will become aware of that fact[36] and expose such type of misconduct. Note that, by the definition of access price indexation, networks have an incentive to expose such type of misconduct from competitors.[37] Also, the quality of fiber networks may be inferred from fault rate information, consumer complaints and from a number of websites that allow to freely test broadband speeds (the process is as easily as clicking one button).[38] Some regulators also produce maps showing accurate information on broadband take-up, speeds and availability.[39]

An additional way to tackle potential unintended consequences, due to lack of perfect observability and verifiability of investments, is to attach a price-floor and a price-cap to the access price indexation rule. For example, a regulator may impose a price cap equal to a^* (optimal fixed access price) combined with a price floor at a_i^{**} (the intended access price under the indexation approach). This guarantees that in the event of a mistaken (x, y) choice, the access price in equilibrium will still be within a deemed reasonable interval.

To conclude, we also point out that the need for information it is just as much a problem for the indexation approach as it is for the fixed access approach. Under the fixed access approach the optimal level of a^* requires to the regulator information on consumers' willingness to pay and networks' costs. A regulator may obtain relevant information in a number of ways. For example, Ofcom has available the following instruments and sources of information.

- **Regulatory financial statements**. British Telecom Group plc (BT) has a regulatory obligation to prepare and publish audited Financial Statements and their associated documents.[40] Relevant, reliable and timely regulatory information informs many of Ofcom's decisions. Ofcom requires this information in order to monitor and enforce various obligations that are placed on BT and as a source of data for setting and monitoring charge controls. It is also a tool of assurance that BT is complying with its regulatory obligations.

- **Engagement with stakeholders**. This includes devising questions and respond-

---

[36]For example, when a network requests access in an area that is allegedly covered with fiber by another network while, in fact, it is not.

[37]Networks may try to collude on fiber coverage reports, agreeing on not to disclose rivals' misconducts. However, such an agreement shall hardly be sustainable, namely when the degree of service differentiation is small such that a small price cut drives a large proportion of consumers to switch operators.

[38]For example, see `http://www.speedtest.net`.

[39]See, for example, `http://maps.ofcom.org.uk/broadband`, where a map showing broadband information by administrative authority in the UK has been produced by Ofcom. This forms part of Ofcom's ongoing work to provide useful information about broadband in the UK. Ofcom also produces regular infrastructure report update, see for example `http://stakeholders.ofcom.org.uk/market-data-research/other/telecoms-research/broadband-speeds/infrastructure-report-2012/`.

[40]See `http://www.btplc.com/Thegroup/RegulatoryandPublicaffairs/Financialstatements/2013/index.htm`.

ing to stakeholder questions, asking and providing further explanation (e.g. in the form of Call for inputs and Consultation Documents) and disclosure where possible. Ofcom's responses to individual stakeholder queries are published on Ofcom's website,[41] in order to provide transparency and to ensure that all stakeholders are provided with the information and data. During these processes Ofcom may hold bilateral and multilateral meetings with stakeholders upon request. Ofcom seeks stakeholders' views and experiences both on specific products and the industry in general. Communications providers (CPs) are usually aware of the market practices of other CPs and are able to identify key issues and provide industry information contributing to an improvement of observability and verifiability.

- **Formal information request under section 135 of the Communications Act 2003**. A person required to provide information under this section must provide it in such manner and within such reasonable period as may be specified by Ofcom. Moreover, the person is required to ensure that the response is on time, complete and accurate. Failure to comply with a formal information request may result in enforcement action being taken by Ofcom (e.g. financial penalties, suspension of entitlement to provide network services, prosecution).

1.6   Conclusions

Investment incentives have been at the core of the access debate. Some authors argue that networks will not invest in facilities subject to strong access regulation (e.g. Sidak and Spulber (1996) on open access). Others have supported the idea of forced access because of the gains in static efficiency, but advise that the access price must take into account investment incentives (e.g. Laffont and Tirole (2001)). This chapter contributes to this debate with the formulation of a new rule for access pricing. We have shown that access pricing rules depending on the investment level of each network, without being informationally more demanding, can boost investment efficiency without sacrificing retail price efficiency and ultimately enhance social welfare vis-à-vis the rules of fixed access price.

Under the proposed indexation rule operators are aware that by investing less they will pay (receive) a higher (lower) access price when competing in the downstream market. Free riders on network investment will become less competitive in the downstream market, thus being punished with a lower profit level relatively to an operator that invested more and consequently is awarded with a competitive advantage. By setting the appropriate

---

[41]Ofcom's web-site at `http://www.ofcom.org.uk`.

indexation rule, regulators can open an important avenue for harsher competition in investment. We have shown that the proper calibration of the indexation rule can induce to achieve the socially efficient level of investment or the Ramsey solution, which are impossible to reach with a fixed access price. Moreover, access price indexation can perform better in terms of social welfare than granting regulatory holidays. While granting regulatory holidays consists of a temporary reduction of retail competition to stimulate investments, the indexation rule works in the opposite way, enticing competition among operators beginning from the investment stage.

While our model is placed within the NGNs context, namely in the fiber deployment problem, the logic of our results goes beyond particular cases. In general, the results herein presented should remain valid to any infrastructure facilities facing an underinvestment problem and whose operators have to choose non-cooperatively the investment levels and compete in retail prices or, equivalently, in quantities.

There are some issues which we do not address in this chapter but that may be of interest for future research. First, we have assumed full information over the analysis, in particular in the decision-making process of the regulator. A question for future research is whether results will hold when the regulator faces informational constraints, e.g. uncertainty on a set of parameters with regard to demand or costs. We note, as discussed in the previous section, that the need for information it is just as much a problem for the indexation approach as it is for the fixed access approach. The estimation of the relevant parameters is inevitably imperfect, and estimation errors may imply efficiency losses under both methodologies. Second, we do not model the entry decisions made by networks, as we assume, for sake of technical simplicity, that there are two symmetric networks. We believe, though, that results and intuitions on the indexation rule should extend on a similar logic to non-symmetric cases and to the N-operator case as well.[42] Third, we do not consider what happens if the networks' facilities are subject to congestion. While this is not currently a concern for NGNs since these are considered high-speed networks, one may want to relax the non-rivalry assumption in applications to other type of infrastructures. Despite the shortcomings, this chapter demonstrates the potential benefits of a new access pricing rule that welfare dominates both the regulatory holidays and the fixed access pricing solutions.

---

[42] De Bijl and Peitz (2004) studied a case of extreme asymmetry where an integrated operator owns a network infrastructure and sells access directly to end-users and to a downstream operator. De Bijl and Peitz reached a conclusion consistent with this chapter that it is possible to provide stronger incentives for the integrated operator to invest in infrastructure quality by increasing the sensitivity of the regulated access price to the network quality.

1.7 References

Armstrong, M. and David Sappington (2007), *Recent Developments in the Theory of Regulation* in Handbook of Industrial Organization, Elsevier, edition 1, Vol. 3, No. 1.

Armstrong, M. and David Sappington (June 2006), *Regulation, Competition, and Liberalization*, Journal of Economic Literature, Vol. XLIV, pp. 325-366.

Armstrong, M. and Julian Wright (June 2009), *Mobile Call Termination*, The Economic Journal, Vol. 119, Issue 538, pp. F270-F307.

Armstrong, M., Chris Doyle, and John Vickers (1996), *The Access Pricing Problem: A Synthesis*, Journal of Industrial Economics, 44 (2), pp. 131–150.

Bourreau, M., Cambini, C. and Hoernig, S. (2010), *National FTTH plans in France, Italy and Portugal*, Communications & Strategies, IDATE, Com&Strat dept., Vol. 1(78), pp. 107–126, 2nd quart.

Bourreau, M., Hombert, J., Pouyet, J., and Schutz, N. (2011), *Upstream competition between vertically integrated firms*, Journal of Industrial Economics, Vol. 59(4), pp. 677–713.

Cabral, L. and Ben Polak (January 2007), *Dominant Firms, Imitation, and Incentives to Innovate*, CEPR DP 4577, NYU Stern WP EC-07-06, available at `http://pages.stern.nyu.edu/~lcabral/workingpapers/CabralPolakDec06.pdf`.

Cambini, C. and Jiang, Y. (2009), *Broadband investment and regulation: A literature review*, Telecommunications Policy, 33(10-11), pp. 559-574.

Chamberlin, J. (1974), *Provision of Collective Goods As a Function of Group Size*, American Political Science Review, Vol. 68, No. 2.

De Bijl, P. and M. Peitz (2004), *Unbundling the local loop: One-way access and imperfect competition*, Technical Report 2004-025, TILEC.

Directive 2002/22/EC of the European Parliament and of the Council of 7 March 2002, on universal service and users' rights relating to electronic communications networks and services (Universal Service Directive), Official Journal of the European Communities, available at `http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2002:108:0051:0051:EN:PDF`.

Economides, N. and Lawrence J. White (1996), *Access and Interconnection Pricing: How Efficient is the "Efficient Component Pricing Rule"?*, Antitrust Bulletin Vol. XL, No. 3, pp. 557-579.

EU court "sets precedent" in Germany telecoms ruling, EurActiv, 4 December 2009, available at `http://www.euractiv.com`.

European Commission (2010), "Commission recommendation of 20 September 2010 on regulated access to Next Generation Access Networks (NGA)", Official Journal of the European Union, L 251, pp. 35-48.

European Regulatory Group (2009), "Regulatory Accounting in Practice 2009", ERG (09) 41.

Foros, O. and Kind, H. J. (2003), *The broadband access market: Competition, uniform pricing and geographical coverage,* Journal of Regulatory Economics, 23(3), pp. 215-35.

Gans, J. (2004), *Access Pricing and Infrastructure Investment*, working paper.

Green, J., and N. Stokey (1983), *A Comparison of Tournaments and Contests*, Journal of Political Economy 91, pp. 349-364.

Guthrie, G. (2006), *Regulating infrastructure: The impact on risk and investment*, Journal of Economic Literature, 44(4), pp. 925-972.

Hoernig, S., Valletti, T. (2002), *The Interplay between Regulation and Competition: The Case of Universal Service Obligations*, Ifo-Studien, 48(1), pp. 57-71.

Holmström, B. (1982), *Moral Hazard in Teams*, The Bell Journal of Economics 13, No. 2, pp. 324–340.

Hurkens, S., and Jeon, D-S (2008), *A Retail Benchmarking Approach to Efficient Two-Way Access Pricing: No Termination-Based Price Discrimination,* RAND Journal of Economics, Vol. 39, No. 3, pp. 822-849.

IDATE Consulting & Research (February 2010), *FTTH Panorama: EU36 & ME14.*

IDATE Consulting & Research (9 February 2011), IDATE News 544.

International Telecommunication Union (ITU) (2009), *Developments of Next Generation Networks: country case studies.*

Klumpp, T. and Su, X. (May 2010), *Open Access and Dynamic Efficiency*, American Economic Journal: Microeconomics, Vol. 2, No. 2, pp. 64-96.

Kotakorpi, K., (2006), *Access price regulation, investment and entry in telecommunications*, International Journal of Industrial Organization 24, pp. 1013-1020.

Laffont, J-J, and Jean Tirole (2001), *Competition in Telecommunications*, MIT Press.

Lazear, E., and S. Rosen (1981), *Rank-Order Tournaments as Optimal Labor Contracts*, Journal of Political Economy 89, pp. 841-864.

Leite, A. N., Santos, V. and Seabra, M. C. (April 1999), *A New, Simple and More-Efficient Access Pricing Rule*, available at `http://docentes.fe.unl.pt/~vascos/`.

Marino, A. and J. Zábojník (2001), *Profit Centers and Incentives in Teams*, USC CLEO Research Paper No. C01-22; and USC Finance & Business Econ. Working Paper No. 01-17.

McGuire, M. (1974), *Group Size, Group Homogeneity, and the Aggregate Provision of Pure Public Goods Under Cournot Behavior*, Public Choice, 18(1), pp. 107–26.

Nalebuff, B. and J. Stiglitz (1983), *Prices and Incentives: Towards a General Theory of Compensation and Competition*, Bell Journal of Economics, Vol. 14(1), pp. 21–43.

Nitsche, R. and Lars Wiethaus (March 2011), *Access regulation and investment in Next Generation Networks – a ranking of regulatory regimes*, International Journal of Industrial Organization, Vol. 29, Issue 2, pp. 263–272.

Olson, M. (1965), *The Logic of Collective Action*, Harvard University Press.

Ovum Consulting (June 2008), *Study about the impact of Next Generation Networks in the market*, report to ANACOM.

Regulatory financial statements (RFS) 2012, available at `http://www.btplc.com/Thegroup/RegulatoryandPublicaffairs/Financialstatements/2012/index.htm`.

RVA LLC Market Research & Consulting (April 2009), *Fiber-To-The-Home: North American market update*.

Sauer, D. (2012), *Access rules and investment incentives*, available at `https://sites.google.com/site/davidfsauer/research`.

Shleifer, A., (1985), *A Theory of Yardstick Competition*, RAND Journal of Economics, Vol. 16(3), pp. 319-327.

Sidak, Gregory, and Daniel Spulber (1996), *Deregulatory Takings and Breach of the Regulatory Contract*, New York University Law Review, 71 (4), pp. 851–999.

Telecommunications Act of 1996, available at `http://transition.fcc.gov/telecom.html`.

Valletti, T. (2003), *The theory of access pricing and its linkage with investment incentives*, Telecommunications Policy 27, pp. 659-675.

Valletti, T., Hoernig, S., and Barros, P. P. (2002), *Universal service and entry: The role of uniform pricing and coverage constraints*, Journal of Regulatory Economics, 21(2), pp. 169-90.

Vareda, J., and Hoernig, S. (2010), *Racing for Investment under Mandatory Access*, B.E. Journal of Economic Analysis & Policy (Contributions), 10(1), article 67.

Vives, X. (1987), *Small Income Effects: A Marshallian Theory of Consumer Surplus and Downward Sloping Demand*, Review of Economic Studies LIV, pp. 87-103.

1.8   Appendix

1.8.1   Hotelling model with hinterlands

A short list of main assumptions. (i) In a given city, the surplus of a consumer indexed by $\tilde{x}$ and $\tilde{y}$ is defined by $CS_{\tilde{x}}$ and $CS_{\tilde{y}}$, respectively, where

$$
CS_{\tilde{x}} \;\equiv\; \begin{cases} v - t\tilde{x} - p_1 & \text{if operator 1} \\ v - t(1-\tilde{x}) - p_2 & \text{if operator 2} \end{cases},
$$

$$
CS_{\tilde{y}} \;\equiv\; \begin{cases} v - t\tilde{y} - p_i & \text{if nearest operator } i = 1,2 \\ 0 & \text{if no service} \end{cases}.
$$

(ii) Each city comprises the center and two hinterlands (West and East side of the city center). In the city center there is a mass 1 of consumers (indexed by $\tilde{x}$) uniformly distributed with density 1 in the unit interval $[0,1]$. Each hinterland has a mass $1/2$ of consumers (indexed by $\tilde{y}$) uniformly distributed with density $t/2v$.

The gross consumer surplus U and the consumer surplus CS in a representative city are, respectively

$$
\begin{aligned}
U(x_1, y_1, x_2, y_2) \;&\equiv\; \int_0^{x_1} (v - t\tilde{x})\, d\tilde{x} + \int_0^{y_1} (v - t\tilde{y}) \frac{t}{2v} d\tilde{y} + \int_0^{x_2} (v - t\tilde{x})\, d\tilde{x} + \int_0^{y_2} (v - t\tilde{y}) \frac{t}{2v} d\tilde{y} \\
&=\; v(x_1 + x_2 + z_1 + z_2) - \frac{t(x_1^2 + x_2^2) + 2v(z_1^2 + z_2^2)}{2}, \text{ and} \\
CS \;&\equiv\; U - \sum_{i=1}^{2} p_i (x_i + z_i),
\end{aligned}
$$

where $z_i \equiv y_i \frac{t}{2v}$.

(iii) The city center is fully served, i.e., $x_1 + x_2 = 1$.

(iv) I_i corresponds to the number of cities covered by fiber by operator $i = 1, 2$. The total number of cities covered by fiber is I, where $I \equiv I_1 + I_2$.

(v) Investment cost for operator i is given by technology

$$C(I_i) \equiv cI_i^2/2.$$

(vi) The marginal cost of serving subscribers is zero (except access charges, when applicable).

The first-best solution. The first-best solution is obtained by solving

$$
\begin{aligned}
\max_{x_1,z_1,x_2,z_2,I_1,I_2} W &\equiv (I_1 + I_2) U - c \left(I_1^2 + I_2^2 \right)/2 \\
\text{subject to } x_1 + x_2 &= 1
\end{aligned}
$$

FOC :

$$
\begin{cases}
\frac{dW}{dx_1} = -t \left(I_1 + I_2 \right) \left(2x_1 - 1 \right) = 0 \\
\frac{dW}{dz_1} = -\left(I_1 + I_2 \right) v \left(2z_1 - 1 \right) = 0 \\
\frac{dW}{dz_2} = -\left(I_1 + I_2 \right) v \left(2z_2 - 1 \right) = 0 \\
\frac{dW}{dI_1} = -\frac{1}{2} \left(t - 2v + 2cI_1 - 2tx_1 - 2vz_1 - 2vz_2 + 2tx_1^2 + 2vz_1^2 + 2vz_2^2 \right) = 0 \\
\frac{dW}{dI_2} = -\frac{1}{2} \left(t - 2v + 2cI_2 - 2tx_1 - 2vz_1 - 2vz_2 + 2tx_1^2 + 2vz_1^2 + 2vz_2^2 \right) = 0
\end{cases}
,
$$

thus

$$
\begin{cases}
x_i^{opt} = \frac{1}{2} \\
z_i^{opt} = \frac{1}{2} \\
I_i^{opt} = \frac{1}{4c} \left(6v - t \right) \\
U^{opt} = \frac{1}{4} \left(6v - t \right) \\
W^{opt} = \frac{(6v-t)^2}{16c}
\end{cases}
.
$$

Consumer demand functions. The individual consumer surplus defined by (3) implies that

$$x_i = \frac{1}{2} - \frac{p_i - p_j}{2t},$$

and from (4) we get

$$z_i = \frac{v - p_i}{\tau} \times \frac{\tau}{2v} = \frac{v - p_i}{2v}.$$

Hence,

$$q_i \equiv x_i + z_i = \frac{1}{2} - \frac{p_i - p_j}{2t} + \frac{v - p_i}{2v} = 1 - \frac{(v+t)\,p_i - vp_j}{2tv} \text{ and } Q \equiv q_1 + q_2 = 2 - \frac{p_1 + p_2}{2v}.$$

The fixed access price approach. We compare below the first-best solution with a symmetric subgame Nash equilibrium under a fixed access price.

Stage III: Retail Price Competition

Operator 1's problem

$$\max_{p_1} \Pi_1 = (I_1 + I_2) \times p_1 q_1 + aq_2 I_1 - aq_1 I_2 - cI_1^2/2$$

$$FOC : \frac{atI_2 + v\,(a + 2t)\,(I_1 + I_2) - 2\,(t + v)\,(I_1 + I_2)\,p_1 + v\,(I_1 + I_2)\,p_2}{2tv} = 0.$$

Operator 2's problem

$$\max_{p_2} \Pi_2 = (I_1 + I_2) \times p_2 q_2 + aq_1 I_2 - aq_2 I_1 - cI_2^2/2$$

$$FOC : \frac{atI_1 + v\,(a + 2t)\,(I_1 + I_2) - 2\,(t + v)\,(I_1 + I_2)\,p_2 + v\,(I_1 + I_2)\,p_1}{2tv} = 0.$$

In equilibrium

$$p_i^* = \frac{2at^2 I_j + 3av^2\,(I_i + I_j) + 6tv^2\,(I_i + I_j) + 4t^2 v\,(I_i + I_j) + 3atvI_i + 4atvI_j}{(2t + v)\,(2t + 3v)\,(I_i + I_j)},$$

thus

$$q_i^* = \frac{(6v^2 + 10tv + 4t^2)\,v\,(I_i + I_j) - 2a\,(t^2 I_j + v^2 I_i) - 4v^2 aI_j - atv\,(I_i + 6I_j)}{2v\,(2t + v)\,(2t + 3v)\,(I_i + I_j)},$$

$$Q^* = \frac{4tv + 4v^2 - at - 2av}{2v\,(2t + v)}.$$

Stage II: Investment

$$\max_{I_i} \Pi_i^* = I \times p_i^* q_i^* + aq_j^* I_i - aq_i^* I_j - cI_i^2/2$$

$$FOC \quad : \quad \frac{\partial \Pi_i^*}{\partial I_i} = 0.$$

In a symmetric equilibrium

$$I_i^* = \frac{(t+2v)\left(48tv^2 + 32t^2 v + 24v^3 - a\left(25tv + 14t^2 + 12v^2\right)\right)a + 16\left(t+v\right)\left(2t+3v\right)tv^2}{8cv\left(2t+3v\right)\left(2t+v\right)^2}.$$

The Socially Optimal Solution vs Equilibrium Under a Fixed Access Price

The socially optimal (first-best) solution is characterized by

$$
\begin{aligned}
I_i^{opt} &= (6v - t)/(4c), \\
x_i^{opt} &= 1/2, \; z_i^{opt} = 1/2, \; q_i^{opt} = 1, \; Q^{opt} = 2, \\
p_i^{opt} &= 0, \; U^{opt} = (6v - t)/4, \; W^{opt} = (6v - t)^2/(16c).
\end{aligned}
$$

Hence, in the first-best solution operators would present negative profits

$$\Pi_i^{opt} = I^{opt} \times p_i^{opt} q_i^{opt} + aq_j^{opt} I_i^{opt} - aq_i^{opt} I_j^{opt} - c\left(I_i^{opt}\right)^2/2 = -c\left(I_i^{opt}\right)^2/2 < 0.$$

A symmetric equilibrium under the fixed access price approach is characterized by

$$I_i^* = \frac{(t+2v)\left(48tv^2 + 32t^2 v + 24v^3 - a\left(25tv + 14t^2 + 12v^2\right)\right)a + 16\left(t+v\right)\left(2t+3v\right)tv^2}{8cv\left(2t+3v\right)\left(2t+v\right)^2},$$

$$p_i^* = \frac{1}{4t+2v}\left(at + 2av + 4tv\right), \; x_i^* = \frac{1}{2}, \; z_i^* = \frac{2v^2 - (2v+t)a}{4v\left(2t+v\right)},$$

$$q_i^* = \frac{1}{2} + \frac{2v^2 - (2v+t)a}{4v\left(2t+v\right)}, \; Q^* = 1 + \frac{2v^2 - (2v+t)a}{2v\left(2t+v\right)},$$

$$\Pi_i^* = \frac{\left(\begin{array}{c} 48av^4 + 144tv^4 + 6a^2t^3 - 24a^2v^3 + 240t^2v^3 + 96t^3v^2 + \\ +9a^2t^2v - 8atv^3 - 32at^3v - 80at^2v^2 - 18a^2tv^2 \end{array}\right) \times}{128cv^2 \left(2t + 3v\right)^2 \left(2t + v\right)^4} ,$$

$$\times \left(\begin{array}{c} 48av^4 + 48tv^4 - 14a^2t^3 - 24a^2v^3 + 80t^2v^3 + 32t^3v^2 + \\ +120atv^3 + 32at^3v + 112at^2v^2 - 62a^2tv^2 - 53a^2t^2v \end{array}\right)$$

$$U^* = \frac{46tv^3 - 8t^3v - a^2t^2 - 4a^2v^2 + 24t^2v^2 + 12v^4 - 16atv^2 - 8at^2v - 4a^2tv}{8v \left(2t + v\right)^2} ,$$

$$CS^* = \frac{14tv^3 - 16av^3 - 8t^3v + a^2t^2 + 4a^2v^2 - 8t^2v^2 + 12v^4 - 24atv^2 - 8at^2v + 4a^2tv}{8v \left(2t + v\right)^2} ,$$

$$W^* = \frac{\left(\begin{array}{c} 48av^4 - 276tv^4 + 32t^4v - 10a^2t^3 - 248t^2v^3 - 16t^3v^2 - 72v^5 + \\ +216atv^3 + 64at^3v + 224at^2v^2 - 22a^2tv^2 - 31a^2t^2v \end{array}\right) \times}{64cv^2 \left(2t + 3v\right)^2 \left(2t + v\right)^4}$$

$$\times \left(\begin{array}{c} 14a^2t^3 + 24a^2v^3 - 48av^4 - 48tv^4 + 62a^2tv^2 + \\ +53a^2t^2v - 80t^2v^3 - 32t^3v^2 - 120atv^3 - 32at^3v - 112at^2v^2 \end{array}\right)$$

Retail price efficiency requires

$$p_i^* = \frac{1}{4t + 2v} \left(at + 2av + 4tv\right) = 0 = p_i^{opt}$$

$$\Leftrightarrow a^{efficient} = -\frac{4tv}{t + 2v} < 0.$$

Investment efficiency requires

$$I_i^* = I_i^{opt},$$

which is impossible to achieve under the fixed access price approach, since

$$\max_a I_i^* = \frac{\left(t + 2v\right)\left(48tv^2 + 32t^2v + 24v^3 - a\left(25tv + 14t^2 + 12v^2\right)\right)a + 16\left(t + v\right)\left(2t + 3v\right)tv^2}{8cv \left(2t + 3v\right)\left(2t + v\right)^2}$$

$$FOC : -\frac{1}{4}\left(t + 2v\right)\frac{14at^2 + 12av^2 - 24tv^2 - 16t^2v - 12v^3 + 25atv}{cv \left(2t + 3v\right)\left(2t + v\right)^2} = 0$$

$$\Leftrightarrow a^{invest} = 4v\frac{6tv + 4t^2 + 3v^2}{25tv + 14t^2 + 12v^2},$$

$$SOC : -\frac{1}{4}\left(t + 2v\right)\frac{25tv + 14t^2 + 12v^2}{cv \left(2t + v\right)^2\left(2t + 3v\right)} < 0,$$

and

$$I_i^*\left(a^{invest}\right) - I_i^{opt} = 2v\frac{45tv^2 + 39t^2v + 11t^3 + 18v^3}{c\left(2t + 3v\right)\left(25tv + 14t^2 + 12v^2\right)} - \frac{1}{4c}\left(6v - t\right)$$

$$= -\frac{\left(6v^2 - 2t^2 + 3tv\right)\left(27tv + 14t^2 + 12v^2\right)}{4c\left(2t + 3v\right)\left(25tv + 14t^2 + 12v^2\right)} < 0$$

provided that $v > t$ by assumption.

The access price indexation approach. Let the access price charged by operator i, per subscriber of operator j using i's infrastructure, be defined by $a_i \equiv xI_i - yI_j$, where (x, y) is the pair of regulatory parameters.

Stage III: Retail Price Competition Under Indexation

Operator 1's problem

$$
\max_{p_1} \Pi_1 \;=\; I \times p_1 q_1 + a_1 q_2 I_1 - a_2 q_1 I_2 - cI_1^2/2
$$

$$
FOC \quad : \quad \frac{\left(\begin{array}{c} 2tvI_1 + 2tvI_2 + tI_2 a_2 + vI_1 a_1 + vI_2 a_2 + \\ -2tI_1 p_1 - 2tI_2 p_1 - 2vI_1 p_1 - 2vI_2 p_1 + vI_2 p_2 + vI_1 p_2 \end{array} \right)}{2tv} = 0
$$

$$
SOC \quad : \quad -\left(I_1 + I_2\right)\frac{t + v}{tv} < 0.
$$

Operator 2's problem

$$
\max_{p_2} \Pi_2 \;=\; I \times p_2 q_2 + a_2 q_1 I_2 - a_1 q_2 I_1 - cI_2^2/2
$$

$$
FOC \quad : \quad \frac{\left(\begin{array}{c} 2tvI_1 + 2tvI_2 + tI_1 a_1 + vI_1 a_1 + vI_2 a_2 + \\ -2tI_1 p_2 - 2tI_2 p_2 - 2vI_2 p_2 - 2vI_1 p_2 + vI_1 p_1 + vI_2 p_1 \end{array} \right)}{2tv} = 0
$$

$$
SOC \quad : \quad -\left(I_1 + I_2\right)\frac{t + v}{tv} < 0.
$$

In equilibrium

$$
p_i^{**} = \frac{v\left(3ta_i + 3va_i + 6tv + 4t^2\right)I_i + \left(6tv^2 + 4t^2 v + 2t^2 a_j + 3v^2 a_j + 4tva_j\right)I_j}{\left(2t + v\right)\left(2t + 3v\right)\left(I_i + I_j\right)}
$$

and

$$
\begin{cases}
x_i^{**} = \frac{(2t+3v+a_i)I_i + (2t+3v-a_j)I_j}{2(I_i+I_j)(2t+3v)} \\[2mm]
z_i^{**} = \frac{v\left(2tv+3v^2-3ta_i-3va_i\right)I_i + \left(2tv^2-2t^2 a_j-3v^2 a_j-4tva_j+3v^3\right)I_j}{2v(2t+v)(2t+3v)(I_i+I_j)} \\[2mm]
q_i^{**} = \frac{v\left(10tv+4t^2+6v^2-ta_i-2va_i\right)I_i + 2(t+v)\left(2tv+3v^2-ta_j-2va_j\right)I_j}{2v(2t+v)(2t+3v)(I_i+I_j)}
\end{cases} .
$$

Stage II: Investments Under Indexation

$$\max_{I_i} \Pi_i^{**} = I \times p_i^{**} q_i^{**} + a_i q_j^{**} I_i - a_j q_i^{**} I_j - cI_i^2/2$$

$$FOC \quad : \quad \frac{d\Pi_i^{**}}{dI_i} = \frac{\partial \Pi_i^{**}}{\partial I_i} + \sum_{k=1}^{2} \frac{\partial \Pi_i^{**}}{\partial a_k} \frac{\partial a_k}{\partial I_i} = 0.$$

Assuming investment symmetry in equilibrium, $I_i = I_j$, we get that $a_i = a_j = (x - y) I_i$. Plugging $I_i = I_j$ and $a_i = a_j$ into the operator i's FOC we reach

$$\frac{\left(\begin{array}{c} 3\left(t+2v\right)\left(y-x\right)\left(10t^2x - 2t^2y + 10v^2x - 6v^2y + 19tvx - 7tvy\right)I_i^2 + \\ -8v \left(\begin{array}{c} 8ct^3 + 3cv^3 - 8t^3x - 12v^3x + 6v^3y + 14ctv^2 + \\ +20ct^2v - 30tv^2x - 28t^2vx + 7tv^2y + 2t^2vy \end{array} \right) I_i + \\ +16tv^2\left(t+v\right)\left(2t+3v\right) \end{array} \right)}{8v\left(2t+3v\right)\left(2t+v\right)^2} = 0,$$

while the SOC, in the equilibrium, must hold the following inequality

$$-\frac{\left(\begin{array}{c} 8v\left(2t+3v\right)\left(8ct^3 + 3cv^3 - 8t^3x - 12v^3x + 14ctv^2 + 20ct^2v - 30tv^2x - 28t^2vx\right) + \\ +\left(t+2v\right) \left(\begin{array}{c} 132t^3x^2 + 4t^3y^2 + 162v^3x^2 + 18v^3y^2 - 120t^3xy - 180v^3xy + \\ +433tv^2x^2 + 404t^2vx^2 + 33tv^2y^2 + 20t^2vy^2 - 462tv^2xy - 408t^2vxy \end{array} \right) I_i \end{array} \right)}{8v\left(2t+v\right)^2\left(2t+3v\right)^2} < 0.$$

Stage I: Regulatory Regime Under Indexation

Suppose that the regulator maximizes the social welfare under the indexation approach, i.e., solves the following problem

$$\max_{x,y} W \equiv (I_i + I_j) U - c \left(I_i^2 + I_j^2\right)/2$$

$$= (I_i + I_j) \left(v (1 + z_i + z_j) - \frac{t \left(x_i^2 + (1 - x_i)^2\right) + 2v \left(z_i^2 + z_j^2\right)}{2} \right) - c \left(\frac{I_i^2}{2} + \frac{I_j^2}{2} \right)$$

subject to

$$x_i^{**} (I_i, I_j, x, y) = x_i^{**} \text{ (Stage III)}$$

$$z_i^{**} (I_i, I_j, x, y) = z_i^{**} \text{ (Stage III)}$$

$$q_i^{**} (I_i, I_j, x, y) = q_i^{**} \text{ (Stage III)}$$

$$p_i^{**} (I_i, I_j, x, y) = p_i^{**} \text{ (Stage III)}$$

$$d\Pi_i^{**}/dI_i = 0 \text{ (Stage II)}$$

$$d^2\Pi_i^{**}/dI_i^2 \leq 0 \text{ (Stage II)}$$

$$\Pi_i^{**} \geq 0 \text{ (PC)}.$$

The regulator's problem under the indexation approach can be rewritten as

$$\max_{x,y} W = I_i \frac{\left(\begin{array}{c} 2v \left(23tv^2 + 12t^2v - 4t^3 + 6v^3\right) + \\ -4v \left(4ct^2 + cv^2 + 2t^2 (x - y) + 4tv (c + x - y)\right) I_1 + \\ - (x - y)^2 (t + 2v)^2 I_1^2 \end{array} \right)}{4v (2t + v)^2} \text{ subject to}$$

$$\begin{cases} d\Pi_i^{**}/dI_i = 0 & \text{(Stage II)} \\ d^2\Pi_i^{**}/dI_i^2 \leq 0 & \text{(Stage II)} \\ I_i \frac{16tv^2(v+t) - 2v\left(4ct(t+v) + cv^2 + v(2t+4v)(y-x)\right)I_i - (x-y)^2(t+2v)^2 I_i^2}{4v(2t+v)^2} \geq 0 & \text{(PC)}. \end{cases}$$

The regulatory holidays case. Below we derive two equilibria under regulatory holidays.

Monopolist offering two brands: two hinterlands served

In a city monopolized by operator i (located at point 0 and point 1) that is unable to price discriminate, the demand function faced by the monopolist is defined by

$$q_i = \begin{cases} 2 \left(\frac{v - p_i}{2v} + \frac{v - p_i}{t} \right) & \text{if } v \geq p_i > v - \frac{t}{2} \\ \frac{v - p_i}{v} + 1 & \text{if } 0 \leq p_i \leq v - \frac{t}{2} \end{cases}.$$

Operator i chooses p_i and I_i by solving the following maximization problem

$$\max_{p_i, I_i} \Pi_i^{mon} = I_i p_i q_i - c I_i^2/2.$$

Suppose that $v > t$ and $p_i^{mon} = v - \frac{t}{2}$. We check now whether the monopolist has an incentive to deviate the price by an $\varepsilon > 0$. If the monopolist increases the price by ε it will get

$$\Pi_i^{mon} = I_i \left(v - \frac{t}{2} + \varepsilon \right) 2 \left(\frac{v - \left(v - \frac{t}{2} + \varepsilon \right)}{2v} + \frac{v - \left(v - \frac{t}{2} + \varepsilon \right)}{t} \right) - c \frac{I_i^2}{2}$$

where

$$\frac{d\Pi_i^{mon}}{d\varepsilon} = I_i \left(t + 2v \right) \frac{t - v - 2\varepsilon}{tv} < 0, \text{ for } v > t,$$

therefore, the monopolist does not have incentive to increase the price above $p_i = v - \frac{t}{2}$ given that $v > t$. If the monopolist decreases the price by ε it will get

$$\Pi_i^{mon} = I_i \left(v - \frac{t}{2} - \varepsilon \right) \left(\frac{v - \left(v - \frac{t}{2} - \varepsilon \right)}{v} + 1 \right) - c \frac{I_i^2}{2}$$

where

$$\frac{d\Pi_i^{mon}}{d\varepsilon} = -I_i \frac{t + 2\varepsilon}{v} < 0.$$

Therefore, the monopolist does not have incentive to decrease the price below $p_i = v - \frac{t}{2}$. We conclude that $p_i^{mon} = v - \frac{t}{2}$ and $q_i^{mon} = 1 + \frac{t}{2v}$.

The monopolist chooses its level of investment by solving

$$\max_{I_i} \Pi_i^{mon} = I_i \left(v - \frac{t}{2} \right) \left(1 + \frac{t}{2v} \right) - cI_i^2/2$$

$$FOC \quad : \quad -\frac{1}{4} \frac{t^2 - 4v^2 + 4cvI_i}{v} = 0 \Leftrightarrow I_i^{mon} = \frac{4v^2 - t^2}{4cv},$$

and obtains a total profit of

$$\Pi_i^{mon} = \frac{(2v - t)^2 (2v + t)^2}{32cv^2}.$$

With regard to social welfare, in the monopoly equilibrium with both hinterlands being served we have

$$x_i^{mon} = \frac{1}{2}, \quad z_i^{mon} = \frac{v - p_i}{2v} = \frac{t}{4v},$$

$$U^{mon} = v \left(x_1^{mon} + x_2^{mon} + z_1^{mon} + z_2^{mon} \right) - \frac{t \left((x_1^{mon})^2 + (x_2^{mon})^2 \right) + 2v \left((z_1^{mon})^2 + (z_2^{mon})^2 \right)}{2}$$

$$= v \left(1 + \frac{t}{2v} \right) - \frac{\frac{t}{2} + 4v \left(\frac{t}{4v} \right)^2}{2} = \frac{(4v - t)(t + 2v)}{8v},$$

$$W^{mon} \equiv \left(I_1^{mon} + I_2^{mon} \right) U^{mon} - c \left(\frac{(I_1^{mon})^2}{2} + \frac{(I_2^{mon})^2}{2} \right)$$

$$= \frac{4v^2 - t^2}{2cv} \frac{(4v - t)(t + 2v)}{8v} - c \left(\frac{\left(\frac{4v^2 - t^2}{4cv} \right)^2}{2} + \frac{\left(\frac{4v^2 - t^2}{4cv} \right)^2}{2} \right) = \frac{(2v - t)(t + 2v)^2}{8cv}.$$

Monopolist offering one brand: one hinterland served

Assume that $v > 2t$ and the monopolist (located only at point 0 or only at point 1) is unable to price discriminate. In this case the demand function faced by the monopolist is defined by

$$q_i = \begin{cases} \frac{v - p_i}{t} \left(1 + \frac{t}{2v} \right) & \text{if } v \geq p_i > v - t \\ 1 + \frac{v - p_i}{2v} & \text{if } 0 \leq p_i \leq v - t \end{cases}.$$

Operator i chooses p_i and I_i by solving the following maximization problem

$$\max_{p_i, I_i} \Pi_i^{mon} = I_i p_i q_i - c I_i^2 / 2.$$

Suppose that $p_i^{mon} = v - t$ and check now if the monopolist has an incentive to deviate the price by an ε. If the monopolist increases the price by ε it will get

$$\Pi_i^{mon} = I_i (v - t + \varepsilon) \left(\frac{t - \varepsilon}{t} \left(1 + \frac{t}{2v} \right) \right) - c \frac{I_i^2}{2}$$

where

$$\frac{d\Pi_i^{mon}}{d\varepsilon} = \frac{1}{2} I_i (t + 2v) \frac{2t - v - 2\varepsilon}{tv} < 0, \text{ for } v > 2t,$$

therefore, the monopolist does not have incentive to increase the price above $p_i = v - t$ given that $v > 2t$. If the monopolist decreases the price by ε it will get

$$\Pi_i^{mon} = I_i (v - t - \varepsilon) \left(\frac{t + \varepsilon}{2v} + 1 \right) - c \frac{I_i^2}{2}$$

where

$$\frac{d\Pi_i^{mon}}{d\varepsilon} = -I_i \frac{2t + v + 2\varepsilon}{2v} < 0,$$

therefore, the monopolist does not have incentive to decrease the price below $p_i = v - t$. We conclude that $p_i^{mon} = v - t$ and $q_i^{mon} = 1 + \frac{t}{2v}$.

The monopolist chooses its level of investment by solving

$$
\begin{aligned}
\max_{I_i} \Pi_i^{mon} &= I_i (v - t) \left(1 + \frac{t}{2v}\right) - c\frac{I_i^2}{2} \\
FOC \quad : \quad &-\frac{tv + t^2 - 2v^2 + 2cvI_i}{2v} = 0 \Leftrightarrow I_i^{mon} = \frac{(v - t)(t + 2v)}{2cv},
\end{aligned}
$$

and obtains a total profit of

$$
\Pi_i^{mon} = \frac{(t + 2v)^2 (v - t)^2}{8cv^2}.
$$

With regard to social welfare, in the monopoly equilibrium with one hinterland served we have

$$
\begin{aligned}
x_i^{mon} &= 1, \; x_j^{mon} = 0, \; z_i^{mon} = \frac{t}{2v}, \; z_j = 0, \\
U^{mon} &= v \left(x_1^{mon} + x_2^{mon} + z_1^{mon} + z_2^{mon}\right) - \frac{t\left((x_1^{mon})^2 + (x_2^{mon})^2\right) + 2v\left((z_1^{mon})^2 + (z_2^{mon})^2\right)}{2} \\
&= v\left(1 + \frac{t}{2v}\right) - \frac{t + 2v\left(\frac{t}{2v}\right)^2}{2} = \frac{(2v - t)(t + 2v)}{4v}, \\
W^{mon} &\equiv \left(I_1^{mon} + I_2^{mon}\right) U^{mon} - c\left((I_1^{mon})^2 + (I_2^{mon})^2\right)/2 \\
&= \frac{2v^2 - tv - t^2}{cv} \frac{(2v - t)(t + 2v)}{4v} - c\left(\frac{2v^2 - tv - t^2}{2cv}\right)^2 = \frac{(v - t)(t + 2v)^2}{4cv}.
\end{aligned}
$$

1.8.2 Proofs

Proof of Proposition 1 (i) The socially efficient investment is defined by (10). Investment efficiency under the fixed access price rule requires that I_i^*, defined by (14), satisfies

$I_i^* = I_i^{opt}$. However, $\max_a I_i^* < I_i^{opt}$ as is shown below

$$\max_a I_i^* = \frac{(t+2v)\left(\begin{array}{c}48tv^2 + 32t^2v + 24v^3 + \\ -a\left(25tv + 14t^2 + 12v^2\right)\end{array}\right)a + 16\left(t+v\right)\left(2t+3v\right)tv^2}{8cv\left(2t+3v\right)\left(2t+v\right)^2}$$

$$FOC \quad : \quad -\frac{1}{4}\left(t+2v\right)\frac{14at^2 + 12av^2 - 24tv^2 - 16t^2v - 12v^3 + 25atv}{cv\left(2t+3v\right)\left(2t+v\right)^2} = 0$$

$$\Leftrightarrow \quad a^{invest} = 4v\frac{6tv + 4t^2 + 3v^2}{25tv + 14t^2 + 12v^2},$$

$$SOC \quad : \quad -\frac{1}{4}\left(t+2v\right)\frac{25tv + 14t^2 + 12v^2}{cv\left(2t+v\right)^2\left(2t+3v\right)} < 0,$$

and

$$\begin{aligned} I_i^*\left(a^{invest}\right) - I_i^{opt} &= 2v\frac{45tv^2 + 39t^2v + 11t^3 + 18v^3}{c\left(2t+3v\right)\left(25tv + 14t^2 + 12v^2\right)} - \frac{1}{4c}\left(6v - t\right) \\ &= -\frac{\left(3tv - 2t^2 + 6v^2\right)\left(27tv + 14t^2 + 12v^2\right)}{4c\left(2t+3v\right)\left(25tv + 14t^2 + 12v^2\right)} < 0 \end{aligned}$$

provided that $v > t$ by assumption.

(ii) Retail price efficiency requires

$$p_i^* = \frac{1}{4t+2v}\left(at + 2av + 4tv\right) = 0 = p_i^{opt}$$

$$\Leftrightarrow \quad a^{efficient} = -\frac{4tv}{t+2v} < 0. \ \square$$

Proof of Theorem 1 (i) Given a pair $\left(p_i^*, p_j^*\right)$, network i chooses the investment level by solving

$$\frac{\partial \Pi_i^*}{\partial I_i}\left(I_i^*\right) = \frac{\partial \Pi_i}{\partial I_i} + \frac{\partial \Pi_i}{\partial p_j}\frac{\partial p_j^*}{\partial I_i} = 0$$

$$\Leftrightarrow p_i^* q_i^* + aq_j^* + \frac{\partial \Pi_i}{\partial p_j}\frac{\partial p_j^*}{\partial I_i} = C'\left(I_i\right).$$

According to (a) and (b) the social welfare measure can be written as $W = \left(I_i + I_j\right)U - C\left(I_i\right) - C\left(I_j\right)$ and in the first-best the regulator equates

$$U^{opt} = C'\left(I_i\right). \tag{22}$$

By assumption (b) $U^{opt} > p_i^* q_i^* + aq_j^* + \frac{\partial \Pi_i}{\partial p_j}\frac{\partial p_j^*}{\partial I_i}$. Hence, $I_i^{opt} > I_i^*$, given that by assumption (a) $C'\left(I_i\right)$ is an increasing function.

(ii) The marginal cost of providing fiber to a subscriber in a covered city is zero, thus

$p_i^{opt} = 0$. Networks choose retail prices by solving

$$\frac{\partial \Pi_i}{\partial p_i} = 0 \Leftrightarrow (I_i + I_j)\left(q_i + p_i^* \frac{\partial q_i}{\partial p_i}\right) + a\left(\frac{\partial q_j}{\partial p_i} I_i - I_j \frac{\partial q_i}{\partial p_i}\right) = 0$$

$$\Leftrightarrow p_i^* = \frac{a\left(\frac{\partial q_j}{\partial p_i} I_i - I_j \frac{\partial q_i}{\partial p_i}\right) + q_i(I_i + I_j)}{-(\partial q_i/\partial p_i)(I_i + I_j)}.$$

Therefore

$$p_i^* = p_i^{opt} \Leftrightarrow \frac{a\left(\frac{\partial q_j}{\partial p_i} I_i - I_j \frac{\partial q_i}{\partial p_i}\right) + q_i(I_i + I_j)}{-(\partial q_i/\partial p_i)(I_i + I_j)} = 0$$

$$\Leftrightarrow a^{efficient} = -\frac{q_i(I_i + I_j)}{\frac{\partial q_j}{\partial p_i} I_i - I_j \frac{\partial q_i}{\partial p_i}} < 0,$$

since $\partial q_j/\partial p_i > 0$ and $\partial q_i/\partial p_i < 0$ by assumption (c). $\square$

Proof of Lemma 1 In a symmetric equilibrium, the Lagrangean function of the regulator's problem is

$$\mathcal{L} = W(x, y, I_i) + \lambda_1 [\Pi_i(x, y, I_i)] + \lambda_2 [S(x, y, I_i)] + \lambda_3 [F(x, y, I_i)],$$

where $\Pi_i(x, y, I_i)$, $S(x, y, I_i)$ and $F(x, y, I_i)$ denote the network i's profit, and the second and the first order conditions with respect to investment, respectively. The optimality conditions to the regulator's problem are

$$\begin{cases} \mathcal{L}'_x \leq 0,\ x\mathcal{L}'_x = 0 \\ \mathcal{L}'_y \leq 0,\ y\mathcal{L}'_y = 0 \\ \mathcal{L}'_{\lambda_1} \geq 0,\ \lambda_1 \mathcal{L}'_{\lambda_1} = 0,\ \lambda_1 \geq 0 \\ \mathcal{L}'_{\lambda_2} < 0,\ \lambda_2 \mathcal{L}'_{\lambda_2} = 0,\ \lambda_2 = 0 \\ \mathcal{L}'_{\lambda_3} = 0,\ \lambda_3 \mathcal{L}'_{\lambda_3} = 0 \end{cases}.$$

To show that networks' participation constraint is binding we need to check that the respective Lagrange multiplier, λ_1, is non-zero. Suppose that $x \neq 0$ and $y \neq 0$, thus, $\mathcal{L}'_x = 0$ and $\mathcal{L}'_y = 0$. Solving the system of simultaneous equations

$$\begin{cases} \mathcal{L}'_x = W'_x + \lambda_1 \Pi'_x + \lambda_3 F'_x = 0 \\ \mathcal{L}'_y = W'_y + \lambda_1 \Pi'_y + \lambda_3 F'_y = 0 \end{cases} \Leftrightarrow \begin{cases} \lambda_1^* = \frac{F'_x W'_y - F'_y W'_x}{\Pi'_x F'_y - \Pi'_y F'_x} \\ \lambda_3^* = \frac{\Pi'_y W'_x - \Pi'_x W'_y}{\Pi'_x F'_y - \Pi'_y F'_x} \end{cases} \tag{23}$$

for non-negative access prices, $a_i \geq 0$, i.e., $x \geq y$, we have

$$F'_x W'_y - F'_y W'_x = \frac{I_i^3(t+2v)^2(4tv+(t+2v)(x-y)I_i)(8tv+6v^2-3I_i(t+v)(x-y))}{4v^2(2t+3v)(2t+v)^3} \neq 0,$$

since $4tv + (t + 2v)(x - y)I_i > 0$ and $8tv + 6v^2 - 3I_i(t + v)(x - y) \neq 0 \Leftrightarrow I_i \neq \frac{8t+6v}{3(t+v)(x-y)}v$. To see that $I_i \neq \frac{8t+6v}{3(t+v)(x-y)}v$, suppose by contradiction that $I_i = \frac{8t+6v}{3(t+v)(x-y)}v$ and plug the expression into the first-order condition $F(x, y, I_i) = 0$. We get then

$$F\left(x, y, \frac{8t + 6v}{3(t + v)(x - y)}v\right) = -\frac{\left(\begin{array}{c} 4cv(4t + 3v)(2t + v)(t + v) + \\ +v(28tv^2 + 17t^2v + 2t^3 + 12v^3)(x - y) \end{array}\right)}{6(2t + v)(t + v)^2(x - y)} < 0,$$

which means that the FOC with respect to investment is not satisfied and $I_i \neq \frac{8t+6v}{3(t+v)(x-y)}v$ must hold. Provided that $F'_x W'_y - F'_y W'_x \neq 0$ and $\lambda_1 \geq 0$, we conclude that in equilibrium $\lambda_1^* > 0$ and, thus, the participation constraint binds. $\square$

Proof of Proposition 2 (i) Under a fixed access price the regulator sets $a_i = a^*$. We can show that for a given access price $a^* > 0$, in a symmetric equilibrium, networks invest more under indexation than under fixed access. Under a fixed access price, networks choose the investment level in accordance with the condition $\partial \Pi_i^* / \partial I_i = 0$, since $a_i = a_j = a^*$ is fixed and, thus, $\partial a_i / \partial I_i = \partial a_j / \partial I_i = 0$. Under access price indexation networks choose the investment level in accordance with

$$\frac{\partial \Pi_i^{**}}{\partial I_i} + \sum_{k=1}^{2} \frac{\partial \Pi_i^{**}}{\partial a_k} \frac{\partial a_k}{\partial I_i} = 0, \tag{24}$$

where $\partial a_k / \partial I_i = \begin{cases} x & \text{if } i = k \\ -y & \text{if } i \neq k \end{cases}$. We can show that $\sum_{k=1}^{2} \frac{\partial \Pi_i^{**}}{\partial a_k} \frac{\partial a_k}{\partial I_i} > 0$ provided that

$$\sum_{k=1}^{2} \frac{\partial \Pi_i^{**}}{\partial a_k} \frac{\partial a_k}{\partial I_i} = I_i(t + 2v) \frac{\left(\begin{array}{c} 12v^3x - 9av^2x + 3av^2y - 2atvy + 24tv^2x - 16atvx + \\ +16t^2vx - 8at^2x + 16t^2vy - 4at^2y + 16tv^2y \end{array}\right)}{4v(2t + 3v)(2t + v)^2}$$

and $12v^3x > 9av^2x$, $3av^2y > 2atvy$, $24tv^2x > 16tvax$, $16t^2vx > 8t^2ax$, $16t^2vy > 4t^2ay$, for $(x, y) \in \mathbb{R}_+^2$, $v > a > 0$ and $v > t$ by assumption. By (24) and the fact that $\sum_{k=1}^{2} \frac{\partial \Pi_i^{**}}{\partial a_k} \frac{\partial a_k}{\partial I_i} > 0$ then $\partial \Pi_i^{**} / \partial I_i < 0$. For $a_i = a^*$, (12) and (16) are identical, thus $\partial \Pi_i^* / \partial I_i = \partial \Pi_i^{**} / \partial I_i < 0$, and by the (SOC) concavity of the profit function with respect to I_i, i.e., $\partial^2 \Pi_i^* / \partial I_i^2 < 0$, we conclude that $I_i^{**} > I_i^*$.

(ii) The total mass of subscribers in a representative city is determined by (8), thus the mass of subscribers will expand if retail prices decrease. Retail prices will decrease if a_i decreases. Suppose that the regulator, in equilibrium, would like to attain, under the indexation approach, $a_i = a^* - \varepsilon$, where $\varepsilon > 0$ is arbitrarily small and a^* is the optimal access charge under a fixed access approach. We can show that this is compatible with

$I_i^{**} > I_i^*$ for ε sufficiently small. Replacing a_i by $a^* - \varepsilon$ in $\partial \Pi_i^{**}/\partial I_i$ we can show that

$$\frac{\partial \Pi_i^{**}}{\partial I_i}\left(a^* - \varepsilon\right) = \frac{\partial \Pi_i^{**}}{\partial I_i}\left(a^*\right) - \frac{\left(\begin{array}{c} \left(14t^3 + 24v^3 + 62tv^2 + 53t^2v\right)\left(\varepsilon - 2a\right) + \\ + \left(48v^4 + 120tv^3 + 32t^3v + 112t^2v^2\right) \end{array}\right)}{8v\left(2t + 3v\right)\left(2t + v\right)^2}\varepsilon$$

and in the limit

$$\lim_{\varepsilon \to 0} \frac{\left(14t^3 + 24v^3 + 62tv^2 + 53t^2v\right)\left(\varepsilon - 2a\right) + \left(48v^4 + 120tv^3 + 32t^3v + 112t^2v^2\right)}{8v\left(2t + 3v\right)\left(2t + v\right)^2}\varepsilon = 0. \tag{25}$$

Hence, by continuity of the expression in (25) there exists $\varepsilon > 0$ such that

$$-\frac{\left(\begin{array}{c} \left(14t^3 + 24v^3 + 62tv^2 + 53t^2v\right)\left(\varepsilon - 2a\right) + \\ + \left(48v^4 + 120tv^3 + 32t^3v + 112t^2v^2\right) \end{array}\right)}{8v\left(2t + 3v\right)\left(2t + v\right)^2}\varepsilon + \sum_{k=1}^{2}\frac{\partial \Pi_i^{**}}{\partial a_k}\frac{\partial a_k}{\partial I_i} > 0$$

provided that $\sum_{k=1}^{2}\frac{\partial \Pi_i^{**}}{\partial a_k}\frac{\partial a_k}{\partial I_i} > 0$ as shown in (i). Condition (24) implies that $\partial \Pi_i^{**}\left(a^*\right)/\partial I_i < 0$, which combined with the concavity of Π_i^{**} with respect to I_i results in $I_i^{**}\left(a^* - \varepsilon\right) > I_i^*\left(a^*\right)$ for $\varepsilon > 0$ sufficiently small.

(iii) The social welfare measure is

$$W = \left(I_i + I_j\right)U\left(q_i, q_j\right) - \frac{c}{2}\left(I_i^2 + I_j^2\right),$$

where

$$\frac{\partial W}{\partial I_i} = U - cI_i,$$
$$\frac{\partial W}{\partial q_i} = \frac{\partial W}{\partial U}\frac{\partial U}{\partial q_i} = \left(I_i + I_j\right)\frac{\partial U}{\partial q_i}.$$

We can show that $\partial W/\partial I_i > 0$ under the fixed access price approach, where the expressions $U^*\left(a\right)$ and $I_i^*\left(a\right)$ are defined under the heading "The fixed access price approach" in an appendix. We note that $\frac{d\left(U^* - cI_i^*\right)}{da} < 0$, and $a \leq \frac{2v^2 - tv - 2t^2}{t + 2v}$ in order to guarantee full participation in the city center. Given that $U^*\left(\frac{2v^2 - tv - 2t^2}{t + 2v}\right) - cI_i^*\left(\frac{2v^2 - tv - 2t^2}{t + 2v}\right) = \frac{1}{4}t\frac{24tv^2 + 19t^2v + 6t^3 + 12v^3}{v\left(2t + 3v\right)\left(t + 2v\right)} > 0$, thus, $U^* - cI_i^* > 0$ for any $0 \leq a \leq \frac{2v^2 - tv - 2t^2}{t + 2v}$.

It is trivial that $\partial W/\partial q_i > 0$ when there are consumers without a broadband fiber connection in a city. Recall that, by assumption, all consumers have a non-negative willingness to pay for a broadband fiber connection. Therefore, as the mass of subscribers increases, the social welfare level will rise.

Provided that both the total investment in fiber coverage and the mass of subscribers expand under a linear access pricing rule, as shown in (i) and (ii), thus, the social welfare

level will be higher when compared to the case where a fixed access price is used. $\square$

Proof of Theorem 2 (i) We can show that for any given access price $a_i\left(I_i^{**}, I_j^{**}\right) = a_j\left(I_j^{**}, I_i^{**}\right) = a^* > 0$ networks invest more under indexation than under a fixed access price. Under a fixed access price networks choose the investment level according to condition $\partial\Pi_i^* / \partial I_i = 0$ while under access price indexation networks choose the investment level in accordance with

$$\frac{\partial\Pi_i^{**}}{\partial I_i} + \sum_{k=1}^{2} \frac{\partial\Pi_i^{**}}{\partial a_k}\frac{\partial a_k}{\partial I_i} = 0 \tag{26}$$

where $\partial\Pi_i^{**} / \partial I_i = \partial\Pi_i^* / \partial I_i$ if $a_i\left(I_i^{**}, I_j^{**}\right) = a^*$.

By assumption (a)

$$\frac{\partial\Pi_i^{**}}{\partial a_i} = \underbrace{I_i q_j^{**}}_{\geq 0} + \underbrace{\left.\frac{\partial\Pi_i}{\partial p_i}\right|_{p_i=p_i^{**}}}_{\substack{=0,\text{ by FOC}\\\text{in price stage}}} \cdot \frac{\partial p_i^{**}}{\partial a_i} + \underbrace{\left.\frac{\partial\Pi_i}{\partial p_j}\right|_{p_j=p_j^{**}}}_{\geq 0,\text{ by (b)}} \cdot \underbrace{\frac{\partial p_j^{**}}{\partial a_i}}_{\geq 0} \geq 0, \text{ where}$$

$$\frac{\partial p_j^{**}}{\partial a_i} = -\frac{\partial^2\Pi_j / \partial p_j \partial a_i}{\partial^2\Pi_j / \partial p_j^2} \geq 0 \text{ since } \partial^2\Pi_j / \partial p_j^2 < 0 \text{ by (c) and}$$

$$\partial^2\Pi_j / \partial p_j \partial a_i = -I_i.\partial q_j / \partial p_j \geq 0 \text{ by (d).}$$

Given that Π_i is infinitely differentiable,[43] both $\partial\Pi_i^{**} / \partial a_i$ and $\partial\Pi_i^{**} / \partial a_j$ must be finite. Hence, the regulator can choose $\partial a_i / \partial I_i$ and $\partial a_j / \partial I_i$ such that $\sum_{k=1}^{2} \frac{\partial\Pi_i^{**}}{\partial a_k}\frac{\partial a_k}{\partial I_i} > 0$. Given $\sum_{k=1}^{2} \frac{\partial\Pi_i^{**}}{\partial a_k}\frac{\partial a_k}{\partial I_i} > 0$ and $\partial\Pi_i^{**} / \partial I_i = \partial\Pi_i^* / \partial I_i$ when $a_i = a^*$, thus $\partial\Pi_i^* / \partial I_i < 0$ by (26). Due to sufficient convexity of $C(I_i)$, Π_i^* is concave in I_i (SOC in the investment stage). Therefore, the investment solution I_i^{**} defined by (26) must be higher than I_i^* which is defined by $\partial\Pi_i^* / \partial I_i = 0$.

(ii) Suppose that the regulator intends to implement an access price $a_i\left(I_i^{**}, I_j^{**}\right) = a^* - \varepsilon$, for $\varepsilon > 0$. We show that this is compatible with having $I_i^{**}(a^* - \varepsilon) > I_i^*(a^*)$ for ε sufficiently small, while equilibrium prices decrease with ε.

In relation to retail prices

$$\partial p_j^{**} / \partial a_i \geq 0 \text{ as shown in (i),}$$

$$\frac{\partial p_i^{**}}{\partial a_i} = -\frac{\partial^2\Pi_i / \partial p_i \partial a_i}{\partial^2\Pi_i / \partial p_i^2} \geq 0 \text{ since } \partial^2\Pi_i / \partial p_i^2 < 0 \text{ by (c) and}$$

$$\partial^2\Pi_i / \partial p_i \partial a_i = I_i.\partial q_j / \partial p_i \geq 0 \text{ by (d).}$$

In relation to investments, replacing a^* by $a^* - \varepsilon$ in $\partial\Pi_i^* / \partial I_i$ and taking the Taylor

---

[43]The sums, products, and compositions of analytic functions are analytic. Any analytic function is infinitely differentiable.

series[44] we get

$$\frac{\partial \Pi_i^* (a^* - \varepsilon)}{\partial I_i} = \frac{\partial \Pi_i^* (a^*)}{\partial I_i} + \sum_{n=1}^{\infty} \frac{\frac{\partial^{1+n} \Pi_i^*(a^*)}{\partial I_i \partial a^n}}{n!} (-\varepsilon)^n \qquad (27)$$

where $\frac{\partial^{1+n} \Pi_i^*(a^*)}{\partial I_i \partial a^n}$ is finite, since $\partial \Pi_i^* / \partial I_i$ is differentiable infinitely many times, and independent of ε. Thus,

$$\lim_{\varepsilon \to 0} \sum_{n=1}^{\infty} \frac{\frac{\partial^{1+n} \Pi_i^*(a^*)}{\partial I_i \partial a^n}}{n!} (-\varepsilon)^n = \sum_{n=1}^{\infty} \frac{\frac{\partial^{1+n} \Pi_i^*(a^*)}{\partial I_i \partial a^n}}{n!} \lim_{\varepsilon \to 0} (-\varepsilon)^n = 0.$$

Given the continuity of the expression in (27) and the fact that $\sum_{k=1}^{2} \frac{\partial \Pi_i^{**}}{\partial a_k} \frac{\partial a_k}{\partial I_i} > 0$ as shown in (i), there exists an $\varepsilon > 0$ such that

$$\sum_{n=1}^{\infty} \frac{\frac{\partial^{1+n} \Pi_i^*(a^*)}{\partial I_i \partial a^n}}{n!} (-\varepsilon)^n + \sum_{k=1}^{2} \frac{\partial \Pi_i^{**}}{\partial a_k} \frac{\partial a_k}{\partial I_i} > 0.$$

By (26) we get that $\partial \Pi_i^{**} (a^*) / \partial I_i = \partial \Pi_i^* (a^*) / \partial I_i < 0$ and conclude that, by concavity of Π_i^* with respect to I_i, $I_i^{**} (a^* - \varepsilon) > I_i^* (a^*)$ for $\varepsilon > 0$ sufficiently small.

(iii) The social welfare measure is

$$W = \sum_{i=1}^{2} [I_i U - C(I_i)]$$

where

$$\frac{\partial W}{\partial I_i} = U - C'(I_i),$$
$$\frac{\partial W}{\partial p_i} = (I_i + I_j) \frac{\partial U}{\partial p_i}.$$

It is straightforward that $\partial W / \partial p_i < 0$ given $\partial U / \partial p_i < 0$ by assumption (e). Provided that the retail prices will decrease under a linear access pricing rule (when compared to the use of a fixed access price, as shown in (ii)), *ceteris paribus*, this reduction will enhance the social welfare level.

Given that U increases when the regulator employs a linear access pricing rule (compared to the use of a fixed access price), this means that the social marginal benefit from investment increases as well. Note that if $U^* - C'(I_i^*) > 0$, then an increase in investment at level I_i^* (derived from the fixed access case) will increase social welfare. Otherwise, if $U^* - C'(I_i^*) \leq 0$, the regulator would not need to incentivize more investment. In the latter case, the social welfare level would increase via a retail price reduction. $\square$

---

[44] An analytic function is infinitely differentiable and is equal to its Taylor series.

Proof of Proposition 3 (i) For $I_i = I_i^{opt} = (6v - t)/(4c)$ to be implemented with the indexation rule a regulatory regime (x, y) has to pass three tests: (a) the network FOC in (19), (b) the SOC whose signal is defined by (20), and (c) profits, defined by (15), have to be non-negative, otherwise networks exit the market. Therefore, the efficient investment can be implemented if there exists a regulatory policy (x, y) that satisfies

(a)
$$\left\{ \begin{array}{c} 3\left(t + 2v\right)\left(y - x\right)\left(10t^2x - 2t^2y + 10v^2x - 6v^2y + 19tvx - 7tvy\right)\left(\frac{6v-t}{4c}\right)^2 + \\ -8v\left(\begin{array}{c} 8ct^3 + 3cv^3 - 8t^3x - 12v^3x + 6v^3y + 14ctv^2 + \\ +20ct^2v - 30tv^2x - 28t^2vx + 7tv^2y + 2t^2vy \end{array} \right)\frac{6v-t}{4c} + \\ +16tv^2\left(t + v\right)\left(2t + 3v\right) \end{array} \right\} = 0,$$

(b)
$$-\left\{ \begin{array}{c} 8v\left(2t + 3v\right)\left[c\left(2t + 3v\right)\left(2t + v\right)^2 - 2\left(t + 2v\right)\left(6tv + 4t^2 + 3v^2\right)x\right] + \\ +\left(t + 2v\right)\left(\begin{array}{c} \left(132t^3 + 162v^3 + 433tv^2 + 404t^2v\right)x^2 + \\ +\left(t + 2v\right)\left(2t + 3v\right)^2 y^2 - 6\left(2t + 3v\right)\left(19tv + 10t^2 + 10v^2\right)xy \end{array} \right) \end{array} \right\}\frac{6v-t}{4c} < 0,$$

(c) $16tv^2\left(v + t\right) - 2v\left(\begin{array}{c} 4ct\left(t + v\right) + cv^2 + \\ +v\left(2t + 4v\right)\left(y - x\right) \end{array} \right)\dfrac{6v - t}{4c} - \left(x - y\right)^2\left(t + 2v\right)^2\left(\dfrac{6v - t}{4c}\right)^2 \geq 0.$

If the participation constraint is active (see Lemma 1) and parameters (v, t, c) satisfy the SOC whose signal is defined by (20), i.e.,

$$S \equiv c\dfrac{\left(\begin{array}{c} \left(\begin{array}{c} -444\,320t^{10}v - 4272\,480t^9v^2 - 13\,002\,620t^8v^3 + \\ -12\,144t^{11} - 139\,968v^{11} + 16\,472\,160tv^{10} + \\ +103\,878\,864t^2v^9 + 269\,586\,480t^3v^8 + 373\,756\,596t^4v^7 + \\ +289\,781\,298t^5v^6 + 111\,303\,761t^6v^5 + 3797\,783t^7v^4 \end{array} \right) + \\ -\dfrac{\left(\begin{array}{c} 16v\left(1056tv^4 + 1617t^2v^3 + 1028t^3v^2 + 216t^4v + 252v^5 - 8t^5\right) + \\ -\left(4\left(t + 2v\right)\left(\begin{array}{c} 1092tv^4 + 412t^4v + 891t^2v^3 + \\ +563t^3v^2 + 132t^5 + 468v^5 \end{array} \right)\right) \end{array} \right)\sqrt{\frac{2v}{t+2v}}}{51tv^2 + 54v^3 - 4t^2v - 9t^3} \times \\ \times \left(t + v\right)\left(\begin{array}{c} 442\,440tv^6 - 12\,212t^6v + 638\,472t^2v^5 + \\ +384\,866t^3v^4 + 40\,005t^4v^3 - 47\,334t^5v^2 + 600t^7 + 112\,752v^7 \end{array} \right) \end{array} \right)}{16\left(6v - t\right)\left(\begin{array}{c} 51tv^2 - 4t^2v + \\ -9t^3 + 54v^3 \end{array} \right)\left(\begin{array}{c} 1092tv^4 + 412t^4v + 891t^2v^3 + \\ +563t^3v^2 + 132t^5 + 468v^5 \end{array} \right)\left(2t + v\right)^2} < 0 \tag{28}$$

together with $v > t > 0$ and $c > 0$ by assumption, then the regulatory solution (x, y) will be defined by the zero-profit condition and (19). Taking the limit of S

$$\lim_{t \to 0} S = -\frac{11}{36}c < 0$$

clearly satisfies condition (28). Therefore, if service differentiation, t, is sufficiently small,

the regulatory regime (x, y) defined by the zero-profit condition and (19), by continuity of S, will implement the efficient level of investment.

(ii) The Ramsey problem is

$$
\max_{x_i, x_j, z_i, z_j, I_i, I_j} W = (I_i + I_j) U - \frac{c}{2} \left(I_i^2 + I_j^2 \right)
$$
$$
\text{subject to } \Pi_i = \Pi_j = 0
$$

where $U \equiv v(x_1 + x_2 + z_1 + z_2) - \frac{t\left(x_1^2 + x_2^2\right) + 2v\left(z_1^2 + z_2^2\right)}{2}$ and $\Pi_i = (I_i + I_j) \times p_i q_i + a q_j I_i - a q_i I_j - c I_i^2 / 2 = 0$.

Appealing to symmetry of the problem, the Ramsey solution yields $I_i = I_j$, $x_i = x_j = \frac{1}{2}$, $z_i = z_j$. Thus, $U = v(1 + 2z_i) - \frac{\frac{t}{2} + 4vz_i^2}{2}$ and $\Pi_i = 0 \Leftrightarrow I_i = \frac{4 \times p_i q_i}{c}$, where $q_i = \frac{1}{2} + z_i$ and $z_i = \frac{v - p_i}{2v}$. Note that the zero-profit condition can be rewritten as $I_i = \frac{2v - 8vz_i^2}{c}$ and the Ramsey problem can be rewritten as

$$
\max_{z_i} W = 2I_i \left(v(1 + 2z_i) - \frac{\frac{t}{2} + 4vz_i^2}{2} \right) - c I_i^2, \text{ s. to } I_i = \frac{2v - 8vz_i^2}{c}.
$$

Incorporating the restriction into the objective function we get $W = v(2z_i - 1)(2z_i + 1)\frac{t - 8vz_i - 8vz_i^2}{c}$. Maximizing in order to z_i

$$
FOC : 8v \frac{v + tz_i + 2vz_i - 12vz_i^2 - 16vz_i^3}{c} = 0,
$$
$$
SOC : 8v \frac{t + 2v - 24vz_i - 48vz_i^2}{c} < 0.
$$

Under the indexation approach, in a symmetric equilibrium, z_i and I_i are defined by

$$
\begin{cases}
z_i = \frac{1}{4} \frac{2v^2 - (t + 2v)(x - y)I_i}{v(2t + v)} \\
3(t + 2v)(y - x)\left[(10(t^2 + v^2) + 19tv) x - (2t + 3v)(t + 2v) y \right] I_i^2 + \\
-8v \left[(2t + 3v)\left(c(2t + v)^2 + v(t + 2v) y \right) - 2(t + 2v)(6tv + 4t^2 + 3v^2) x \right] I_i + \\
+16tv^2(t + v)(2t + 3v) = 0
\end{cases} .
$$

Consider the system of equations formed by the FOC and restriction from the Ramsey

problem and the equilibrium expressions for z_i and I_i from the indexation approach

$$
\begin{cases}
8v\frac{v+tz_i+2vz_i-12vz_i^2-16vz_i^3}{c} = 0 \\
z_i - \frac{1}{4}\frac{2v^2-(t+2v)(x-y)I_i}{v(2t+v)} = 0 \\
\begin{aligned}
&3(t+2v)(y-x)\left[(10(t^2+v^2)+19tv)x-(2t+3v)(t+2v)y\right]I_i^2+ \\
&-8v\left[(2t+3v)\left(c(2t+v)^2+v(t+2v)y\right)-2(t+2v)(6tv+4t^2+3v^2)x\right]I_i+ \\
&+16tv^2(t+v)(2t+3v) = 0
\end{aligned} \\
I_i - \frac{2v-8vz_i^2}{c} = 0
\end{cases} \quad . \quad (29)
$$

In order to prove that the Ramsey solution is feasible under the indexation approach we need to show that there is at least one point (x,y,z_i,I_i) such that all conditions in (29) are satisfied.

Note that: (a) according to (i), for t sufficiently small, the indexation approach can implement the first-best investment level $I_i^{opt} = (6v-t)/(4c)$; (b) at $(x,y) = (0,0)$ the investment of equilibrium under the indexation approach is $I_i(0,0) = \frac{2tv(t+v)}{c(2t+v)^2}$; and (c) the Ramsey solution with respect to investment, I_i^{Ramsey}, yields $I_i(0,0) < I_i^{Ramsey} < I_i^{opt}$.[45] Therefore, by continuity of the FOC for I_i under the indexation approach, we can guarantee the existence of a pair (x,y) such that

$$
\begin{cases}
z_i^{Ramsey} = z_i^{**}(x,y) \\
I_i^{Ramsey} = I_i^{**}(x,y)
\end{cases} \quad . \quad \square
$$

Proof of Proposition 4 (i) Suppose that both hinterlands are served by a monopolist offering two differentiated services (one product at point 0 and another product at point 1 in each city) and the regulator, using indexation, intends to implement the retail price $p_i^{**} = v - t/2 - \varepsilon_p < p_i^{mon}$ and the investment level $I_i^{**} = \frac{4v^2-t^2}{4cv} + \varepsilon_I > I_i^{mon}$, where $\varepsilon_p, \varepsilon_I > 0$. This proof consists in verifying if it is possible to find a regulatory regime (x,y) such that networks have non-negative profits, and (19) and (20) are satisfied for some $\varepsilon_p > 0$ and $\varepsilon_I > 0$.

By (16) and investment symmetry, retail prices in equilibrium follow $p_i^{**} = \frac{(t+2v)a_i+4tv}{2(2t+v)}$. In order to implement a retail price $p_i^{**} = v - \frac{t}{2} - \varepsilon_p$, the access price must satisfy

$$
a_i = (x-y)(I_i^{mon} + \varepsilon_I) = \frac{2v^2 - tv - 2t^2}{t+2v} - \varepsilon_a
$$

where $\varepsilon_a \equiv \frac{4t+2v}{t+2v}\varepsilon_p$. Moreover, the network choice regarding the investment level has to

---

[45] The inequality $I_i(0,0) < I_i^{Ramsey}$ can be shown plugging $I_i = I_i(0,0)$ into the zero-profit condition and solving in order to z_i. Then, plug z_i derived from the zero-profit condition into the FOC for z_i derived from the rewritten Ramsey problem. The left hand side of the FOC for z_i becomes negative when evaluated at that level. This together with the zero-profit condition imply that the Ramsey solution yields $I_i(0,0) < I_i^{Ramsey}$.

suffice (19). Solving the system of simultaneous equations in order to (x, y)

$$\begin{cases} (x-y)\left(\frac{4v^2-t^2}{4cv}+\varepsilon_I\right) = \frac{2v^2-tv-2t^2}{t+2v}-\varepsilon_a \\ \begin{pmatrix} 3(t+2v)(y-x)\left(10t^2x-2t^2y+10v^2x-6v^2y+19tvx-7tvy\right)\left(\frac{4v^2-t^2}{4cv}+\varepsilon_I\right)^2+ \\ -8v\begin{pmatrix} 8ct^3+3cv^3-8t^3x-12v^3x+6v^3y+ \\ +14ctv^2+20ct^2v-30tv^2x-28t^2vx+7tv^2y+2t^2vy \end{pmatrix}\left(\frac{4v^2-t^2}{4cv}+\varepsilon_I\right)+ \\ +16tv^2(t+v)(2t+3v) \end{pmatrix} = 0 \end{cases}$$

we get

$$\Leftrightarrow \begin{cases} x^* = \dfrac{cv(2t+3v)\begin{pmatrix} 3t^2\varepsilon_a^2+12v^2\varepsilon_a^2+12tv^3+4t^3v+12t^3\varepsilon_a-8v^3\varepsilon_a+9t^2v^2+ \\ +4t^4+4v^4+12tv\varepsilon_a^2+8tv^2\varepsilon_a+30t^2v\varepsilon_a+8cv^3\varepsilon_I+32ctv^2\varepsilon_I+32ct^2v\varepsilon_I \end{pmatrix}}{(2t+v)(4v^2+4cv\varepsilon_I-t^2)(19tv^2+17t^2v+3t^2\varepsilon_a+6v^2\varepsilon_a+6t^3+6v^3+9tv\varepsilon_a)} \\[2em] y^* = \dfrac{cv\begin{pmatrix} 30t^4\varepsilon_a^2+120v^4\varepsilon_a^2-124tv^5+404t^5v+120t^5\varepsilon_a-48v^5\varepsilon_a+ \\ -122t^2v^4+259t^3v^3+574t^4v^2+104t^6-24v^6+348tv^3\varepsilon_a^2+748t^2v^3\varepsilon_a+ \\ +177t^3v\varepsilon_a^2+1042t^3v^2\varepsilon_a+378t^2v^2\varepsilon_a^2+120tv^4\varepsilon_a+592t^4v\varepsilon_a+ \\ +48cv^5\varepsilon_I+432ct^2v^3\varepsilon_I+288ct^3v^2\varepsilon_I+248ctv^4\varepsilon_I+64ct^4v\varepsilon_I \end{pmatrix}}{(2t+v)(t+2v)(4v^2+4cv\varepsilon_I-t^2)(19tv^2+17t^2v+3t^2\varepsilon_a+6v^2\varepsilon_a+6t^3+6v^3+9tv\varepsilon_a)} \end{cases}.$$

(30)

Plugging the previous regulatory regime (x^*, y^*) and $I_i^{**} = \frac{4v^2-t^2}{4cv}+\varepsilon_I$ into the SOC whose signal is defined by (20) and taking the limit for $(\varepsilon_I, \varepsilon_a) \to (0,0)$ we get

$$\lim_{\varepsilon_a\to 0}\lim_{\varepsilon_I\to 0}\frac{S}{8v(2t+v)^2(2t+3v)^2} = \frac{c\begin{pmatrix} 2548t^7v-5984tv^7-9136t^2v^6-6512t^3v^5+ \\ +604t^4v^4+5938t^5v^3+5687t^6v^2+476t^8-1728v^8 \end{pmatrix}}{8(2v-t)(t+2v)(7tv+3t^2+6v^2)^2(2t+v)^2} < 0,$$

since $2548t^7v < 9136t^2v^6$, $604t^4v^4 < 1728v^8$, $5938t^5v^3 + 476t^8 < 6512t^3v^5$, $5687t^6v^2 < 5984tv^7$, provided that $v > t > 0$ by assumption. Hence, by continuity of the SOC we can assure that there exists $\varepsilon_p > 0$ and $\varepsilon_I > 0$ sufficiently small such that (x^*, y^*) defined by (30) can decrease retail prices and increase the investment relatively to the regulatory holidays regime. Furthermore, since $\Pi_i^{mon} > 0$ for $(p_i, I_i) = (p_i^{mon}, I_i^{mon})$, by continuity of the profit function, for $\varepsilon_p > 0$ and $\varepsilon_I > 0$ sufficiently small we can guarantee that profits are still non-negative with the implementation of (x^*, y^*).

With regard to social welfare,

$$W^{mon} \equiv (I_1^{mon} + I_2^{mon})U^{mon} - c\left(\frac{(I_1^{mon})^2}{2}+\frac{(I_2^{mon})^2}{2}\right)$$

and

$$U^{mon} = v\left(x_1^{mon} + x_2^{mon} + z_1^{mon} + z_2^{mon}\right) - \frac{t\left(\left(x_1^{mon}\right)^2 + \left(x_2^{mon}\right)^2\right) + 2v\left(\left(z_1^{mon}\right)^2 + \left(z_2^{mon}\right)^2\right)}{2},$$

$$x_i^{mon} = \frac{1}{2}, \quad z_i^{mon} = \frac{v - p_i}{2v} = \frac{t}{4v}.$$

Taking the derivatives of welfare in order to investments and retail prices

$$\frac{\partial W^{mon}}{\partial I_i} = U^{mon} - cI_i^{mon} = \frac{(4v - t)(t + 2v)}{8v} - c\frac{4v^2 - t^2}{4cv} = t\frac{t + 2v}{8v} > 0,$$

$$\frac{\partial W^{mon}}{\partial p_i} = \left(I_1^{mon} + I_2^{mon}\right)\frac{\partial U^{mon}}{\partial p_i} = \frac{4v^2 - t^2}{2cv} \times \sum_{i=1}^{2}\frac{\partial U^{mon}}{\partial z_i}\frac{\partial z_i}{\partial p_i}$$

$$= -\frac{4v^2 - t^2}{2cv} \times \frac{v - 2vz_i^{mon}}{v} = -\frac{(t + 2v)(2v - t)^2}{4cv^2} < 0,$$

since $v > t > 0$ and $c > 0$. Therefore, for a sufficiently small increase in investments and a sufficiently small decrease in retail prices, the welfare level increases relatively to the regulatory holidays case.

(ii) Suppose that only one hinterland is served by a monopolist that offers only one service and that the regulator, using indexation, intends to implement the retail price $p_i^{**} = v - t - \varepsilon_p < p_i^{mon}$ and the investment level $I_i^{**} = \frac{2v^2 - tv - t^2}{2cv} + \varepsilon_I > I_i^{mon}$, where $\varepsilon_p, \varepsilon_I > 0$. In order to implement a retail price $p_i^{**} = v - t - \varepsilon_p$, the access price must satisfy

$$a_i = (x - y)\left(I_i^{mon} + \varepsilon_I\right) = \frac{2(v - 2t)(t + v)}{t + 2v} - \varepsilon_a$$

where $\varepsilon_a \equiv \frac{4t + 2v}{t + 2v}\varepsilon_p$. Solving the system of simultaneous equations in order to (x, y)

$$
\begin{cases}
(x - y)\left(\frac{2v^2 - tv - t^2}{2cv} + \varepsilon_I\right) = \frac{2(v - 2t)(t + v)}{t + 2v} - \varepsilon_a \\
\left(\begin{array}{c} 3(t + 2v)(y - x)\left(10t^2 x - 2t^2 y + 10v^2 x - 6v^2 y + 19tvx - 7tvy\right)\left(\frac{2v^2 - tv - t^2}{2cv} + \varepsilon_I\right)^2 + \\ -8v\left(\begin{array}{c} 8ct^3 + 3cv^3 - 8t^3 x - 12v^3 x + 6v^3 y + \\ +14ctv^2 + 20ct^2 v - 30tv^2 x - 28t^2 vx + 7tv^2 y + 2t^2 vy \end{array}\right)\left(\frac{2v^2 - tv - t^2}{2cv} + \varepsilon_I\right) + \\ +16tv^2(t + v)(2t + 3v) \end{array}\right) = 0
\end{cases}
$$

$$
\Leftrightarrow
\begin{cases}
x^* = \dfrac{cv(2t+3v)\left(\begin{array}{l} 3t^2\varepsilon_a^2 + 12v^2\varepsilon_a^2 + 4tv^3 + 16t^3v + 24t^3\varepsilon_a - 8v^3\varepsilon_a - 8t^2v^2 + 32t^4 + \\ +4v^4 + 12tv\varepsilon_a^2 + 20tv^2\varepsilon_a + 60t^2v\varepsilon_a + 8cv^3\varepsilon_I + 32ctv^2\varepsilon_I + 32ct^2v\varepsilon_I \end{array}\right)}{2(2t+v)(2v^2+2cv\varepsilon_I-tv-t^2)(22tv^2+26t^2v+3t^2\varepsilon_a+6v^2\varepsilon_a+12t^3+6v^3+9tv\varepsilon_a)} \\[3em]
y^* = \dfrac{cv\left(\begin{array}{l} 30t^4\varepsilon_a^2 + 120v^4\varepsilon_a^2 - 172tv^5 + 1472t^5v + 240t^5\varepsilon_a - 48v^5\varepsilon_a - 204t^2v^4 + \\ +640t^3v^3 + 1728t^4v^2 + 448t^6 - 24v^6 + 348tv^3\varepsilon_a^2 + 1276t^2v^3\varepsilon_a + \\ +177t^3v\varepsilon_a^2 + 1852t^3v^2\varepsilon_a + 378t^2v^2\varepsilon_a^2 + 240tv^4\varepsilon_a + 1120t^4v\varepsilon_a + \\ +48cv^5\varepsilon_I + 432ct^2v^3\varepsilon_I + 288ct^3v^2\varepsilon_I + 248ctv^4\varepsilon_I + 64ct^4v\varepsilon_I \end{array}\right)}{2(2t+v)(t+2v)(2v^2+2cv\varepsilon_I-tv-t^2)(22tv^2+26t^2v+3t^2\varepsilon_a+6v^2\varepsilon_a+12t^3+6v^3+9tv\varepsilon_a)}
\end{cases}.
$$

Plugging the regulatory regime (x^*, y^*) and $I_i^{**} = \frac{2v^2-tv-t^2}{2cv} + \varepsilon_I$ into the SOC whose signal is defined by (20) and taking the limit for $(\varepsilon_I, \varepsilon_a) \to (0,0)$ we get

$$
\lim_{\varepsilon_a \to 0} \lim_{\varepsilon_I \to 0} \frac{S}{8v(2t+v)^2(2t+3v)^2} = c\frac{\left(\begin{array}{l} 4128t^6v^2 - 108v^8 + 2808t^5v^3 - 338tv^7 - 629t^2v^6 + \\ +247t^4v^4 - 812t^3v^5 + 656t^8 + 2624t^7v \end{array}\right)}{4(2t+v)^2(v-t)(t+2v)(5tv+3t^2+3v^2)^2} < 0
$$

since $4128t^6v^2 < 108v^8$, $2808t^5v^3 < 338tv^7$, $247t^4v^4 < 629t^2v^6$, $2624t^7v + 656t^8 < 812t^3v^5$, given that $v > 2t > 0$. Hence, a linear access pricing rule depending on investments can decrease retail prices and increase investments as compared to the regulatory holidays regime.

With regard to social welfare, taking the derivatives in order to investments and retail prices

$$
\begin{aligned}
\frac{\partial W^{mon}}{\partial I_i} &= U^{mon} - cI_i^{mon} = \frac{(2v-t)(t+2v)}{4v} - c\frac{2v^2-tv-t^2}{2cv} = t\frac{t+2v}{4v} > 0, \\
\frac{\partial W^{mon}}{\partial p_i} &= (I_1^{mon} + I_2^{mon})\frac{\partial U^{mon}}{\partial p_i} = \frac{(v-t)(t+2v)}{cv} \times \frac{\partial U^{mon}}{\partial z_i}\frac{\partial z_i}{\partial p_i} \\
&= -\frac{(v-t)(t+2v)}{cv} \times \frac{(v-2vz_i^{mon})}{2v} = -\frac{(t+2v)(v-t)^2}{2cv^2} < 0,
\end{aligned}
$$

since $v > t > 0$. Therefore, for a sufficiently small increase in investments and a sufficiently small decrease in retail prices, the welfare level increases relatively to the regulatory holidays case. $\square$

2 How does airtime regulation influence advertising quality? A two-sided market perspective

2.1 Introduction

Motivation. Individuals in developed countries spend a significant share of their time connected to mass media platforms.[46] In 2009, the average US American spent almost five hours per day watching TV,[47] while in Japan, the average time is three hours and thirty minutes per day. In the UK, the average viewer aged 4+ watched more than four hours of television per day in 2012. This has increased from three hours and forty-two minutes in 2004 (Communications Market Report 2013).

Advertising plays a significant role in the TV broadcasting business model in most western countries. Mass media platforms offer an opportunity for firms to advertise to a large pool of consumers. Some firms spend billions of dollars per year in advertising,[48] an industry that reached a revenue of over US$ 780 billion worldwide in 2010, with the largest share of it going to TV broadcasting.[49]

In many countries regulatory authorities limit advertising airtime on TV networks. Time restrictions (advertising caps) are generally intended to ensure that viewers are not exposed to excessive amounts of advertising, and that the quality of the overall viewing experience is maintained. With the conspicuous exception of the US, where the frequency and length of commercial breaks are generally unregulated, a number of examples of regulatory constraints on advertising time on TV arise in developed countries. For instance, advertising is limited to an average of six minutes per hour in France; the limit goes up to nine minutes in Germany, while English regulators impose a seven-minute

---

[46] "TV is the dominant medium for media consumption and advertising. Computer usage has supplanted radio as the second most common media activity and print ranks fourth," The New York Times, *8 Hours a Day Spent on Screens, Study Finds*, March 27, 2009.

[47] See *Television, Internet and Mobile Usage in the U.S., Three Screen Report Volume 7 - 4th Quarter 2009*, The Nielsen Company, 2010.

[48] "*For example, Advertising Age (2005) reports that, in 2003 in the U.S., General Motors spent $3.43 billion to advertise its cars and trucks; Procter and Gamble devoted $3.32 billion to the advertisement of its detergents and cosmetics; and Pfizer incurred a $2.84 billion dollar advertising expense for its drugs. Advertising is big business indeed.*" (Bagwell, 2005).

[49] See Karawang Business, Information, Tips and Solutions for Business and Finance, *World Cup 2010: World Advertising Expenditures, Translucent US $780 Billion*, June 13, 2010. In the UK, income from broadcast-based TV advertising declined in 2012 by 2% to just over £3.5 billion (Communications Market Report 2013).

ceiling (off-peak) for public service broadcasters[50] and nine minutes an hour for all other broadcasters.[51]

The amount of advertising watched by consumers has increased over time in a number of countries.[52] In fact, advertising time represents a remarkable proportion of the total airtime of some TV networks. For instance, some programs on major TV networks in the US have recorded advertising levels in excess of twenty minutes per hour.[53] This suggests that the quality of advertisements should matter not only for commercial purposes, e.g., to convince consumers to buy more products in the market, but also because it affects the quality of the viewing experience. Advertising caps may drive firms to change the quality of adverts. In this chapter we are concerned with the impact of advertising airtime restrictions on advertising quality and, ultimately, on social welfare.

Advertising quality is hard to verify and quantify. In this chapter, the term "advertising quality" refers to the nuisance of watching an advert and the viewer's probability of purchasing an advertised product. Within this context, "higher quality" could result from celebrity endorsement of the product to be advertised. This may not necessarily be synonymous with a higher art form. In practice, institutions such as the Australian Broadcasting Tribunal (ABT) appear to be increasingly using costs as a proxy for quality. However, the ABT acknowledges that cost should not necessarily be equated with quality (Wright (1994)). Since 2007, Google has been exploring ways to measure the quality of TV ads. Google aggregates data describing the precise second-by-second tuning behavior for millions of TV set-top boxes, covering millions of US households, doing so for several thousand TV ads every day. From this data, Interian et al. (2009) developed measures that can be used to gauge how appealing and relevant commercials appear to be to TV subscribers.[54] In 2013, VideoHub launched eQ in the US, a new quality score for video advertising. VideoHub's eQ score is a patent-pending formula to determine the potential of a video ad campaign to grab and keep viewers' attention.[55]

---

[50]In the UK no commercial advertising is allowed on BBC. Public television is funded by TV licences. TV viewers in the UK are required to have a TV licence.

[51]The maximum average number of minutes per hour in peak time (6pm – 11pm) is eight minutes for public service broadcasters in the UK. For more on the regulation of the quantity of advertising on television in the UK, see *Regulating the quantity of advertising on television* (Ofcom 2011). In 1992, the Australian Broadcasting Tribunal introduced regulations limiting the amount of non-programme (i.e., promotions and advertisements) to thirteen minutes per hour during prime time and fifteen minutes per hour at other times (Wright (1994)).

[52]In the UK, between 2006 and 2010, the number of different advertisements watched by a viewer rose by 20.9%. In the US, overall, advertising time on TV has been steadily increasing since 1982. This may be due to the fact that there has been a general increase in the overall level of TV viewing. See *Regulating the quantity of advertising on television* (Ofcom 2011).

[53]It is estimated that the average US American is exposed to 61 minutes of TV ads per day. The New York Times, *8 Hours a Day Spent on Screens, Study Finds*, 27 March 2009.

[54]One such measure is the percentage of initial audience retained: how much of the audience, tuned in to an ad when it began airing, remained tuned to the same channel until the ad finishes. TV retention scores are used to determine how highly an ad ranks.

[55]According to Videohub, eQ is the first scoring method of its kind and is the only measurement tool

Description of the chapter. A distinctive aspect of the mass media industry is that it simultaneously serves two groups of agents mutually linked by cross-group network externalities: the subscribers (consumers)[56] who may or may not be sensitive to the volume and quality of advertisements, and the advertising firms whose profits increase with the number of subscribers watching commercials. We utilize a model of subscriber-advertiser supported broadcasting in a two-sided market[57] framework that yields predictions on how advertising quality is determined by firms. The aim of this chapter is to formally investigate the link between regulations limiting the advertising airtime and advertising quality.

The main features of the model are as follows. There are two profit maximizing media platforms competing non-cooperatively in prices by setting them simultaneously and independently, selling ad-airtime to firms and content to subscribers. The advertising quantity is measured as the number of time units dedicated to advertising per time unit of overall broadcasting. We assume that platforms face the same level of costs regardless of whether the content is produced in-house or bought from a third party (e.g. a studio producer).

A mass of subscribers (normalized to one), who are also consumers in the goods market, extract a benefit from the content of media platforms, e.g., information or entertainment, and differ in their attitudes towards the number and quality of advertisements. We assume that a proportion λ of subscribers are ad-sensitive, while the remaining $(1 - \lambda)$ are ad-indifferent. This assumption is crucial in the chapter as the main results are related to the proportion of ad-sensitive subscribers.[58] Furthermore, regardless of type, every individual subscriber has an idiosyncratic preference for his favorite media platform, i.e., his favorite type of programming.

The existence of a proportion of ad-sensitive subscribers is supported by the advertising economics literature, namely, the persuasive and the informative views on ads. The persuasive view states that advertisements alter consumers' preferences and augment product differentiation and brand loyalty. As a result, advertising boosts firms' profits. The informative view holds that many markets suffer from imperfect consumer information because searching costs may prevent consumers from learning of a product's existence,

---

that weighs viewability throughout an ad's duration, player size, and completion rate. See `http://www.videohub.tv/news/dated/2013-02` for an interview with Greg Smith, General Manager of VideoHub Marketplace, on the launch of eQ score.

[56]We will use the terms "subscriber" and "consumer" interchangeably.

[57]In a two-sided market, two different groups of agents relate to each other through a platform. The latter sets access prices taking into account the cross-group externalities. For a general introduction to the theory of two-sided markets, see the seminal papers of Rochet and Tirole (2003, 2006) and Armstrong (2006).

[58]Ad-sensitive subscribers value increases in ad quality but simultaneously dislike ad-airtime. Without the presence of ad-sensitive subscribers, advertising quality would not play any role and would be optimally set to zero.

quality and price. Advertising comes out as one of the endogenous answers to imperfect information, supplying consumers with further information at low cost, e.g., regarding firm location, product description or prices. Both advertising views will be considered in our model.[59]

Wilbur (2008) estimated a two-sided model of the TV industry in the US and found that viewers tend to be averse to commercials. In our model the ad-sensitive subscribers are averse to advertising airtime, while also appreciating the quality of advertisements. For example, ad-sensitive subscribers enjoy the participation in ads of famous performers or athletes.[60] The nuisance perceived by ad-sensitive subscribers is related to the duration or number of commercials. In particular, this negative effect of commercials may be understood as the boredom and wasted time that the ad-sensitive subscribers bear each time there is a commercial break on TV. It is implicitly assumed in our framework that ad-sensitive subscribers have no way to receive the media platform contents while skipping advertisements.

There is empirical evidence that TV subscribers attempt to avoid the advertising time.[61] We also consider a group of ad-indifferent subscribers who are insensitive both to the number of ads and their quality. The proportion of ad-indifferent subscribers may be interpreted, for example, as the percentage of multi-taskers who browse the Internet during the advertising airtime on TV. These subscribers capture the informative part of the adverts (e.g. existence of a new product). However, ad-indifferent subscribers do not pay enough attention to ads and ignore their quality.

We consider a mass of advertising firms (normalized to one) that obtain a benefit from informing potential customers about their products (the informative view). Advertising products to consumers increases the probability of those products being purchased. Additionally, firms may upgrade their ad quality in order to increase the purchase probability of ad-sensitive consumers (the persuasive view).

We show that the average ad quality in a media platform may be increasing in the volume of ads broadcast. This will be the case when firms with a higher informative effect

---

[59]Another theory holds that advertising is a complement to the consumption of the advertised good. According to this perspective advertising does not transform consumers' preferences and need not supply any information. For example, this happens if the consumption of a good generates more prestige to consumers when the good is advertised. See Bagwell (2005) on the economic analysis of advertising.

[60]Nike's *Write the Future* commercial campaign during Fifa World Cup 2010 had the participation of some of the best soccer players in the world. The commercial hit almost 20 million views in only two months after its release (on May 17, 2010) on YouTube.

[61]Speck and Elliott (1997) explain that there are at least three possible ways to avoid advertising: a cognitive strategy (ignoring it), a behavioral strategy (e.g., leaving the room, multi-tasking), and a mechanical strategy (e.g., switching channels, DVRs). Multi-tasking, i.e., conducting more than one activity at the same time, is becoming commonplace. Just over half (53%) of all UK adults are regular media multi-taskers. Moreover, 49% of all UK adults are regularly media-stacking (conducting unrelated media tasks, such as surfing the net, social networking or shopping online while watching TV). See the Communications Market Report 2013 and *Regulating the quantity of advertising on television* (Ofcom 2011).

76

are simultaneously those with fewer incentives to invest in ad quality. In other words, the marginal gain of investing in ad quality is lower for firms with a higher informative effect, where the informative effect is measured in terms of consumer purchase probability of the advertised product. Therefore, the marginal advertiser on a platform exhibits higher ad quality compared to firms with a higher informative effect and, consequently, if platforms sell more advertising slots, doing so will increase their average ad quality. Within this context, an advertising cap confines advertisement slots to firms with a higher informative effect, which are also the firms with higher willingness to pay for an ad slot, to the detriment of firms that would invest in ads of higher quality but are now excluded from the advertising market due to the cap. We found that an advertising cap may cause the average advertising quality to decrease.

Also, we show that an advertising cap may result in the following welfare effects. Media platforms become worse off when their advertising airtime is constrained. Given that a lower advertising cap incentivizes platforms to set higher advertising fees, as a lower advertising fee cannot increase the volume of advertising sales in view of the cap, a tighter advertising cap will necessarily hurt advertisers' profits. The net effect on subscribers' welfare is ambiguous. Although there are fewer ads when an advertising cap is imposed, advertising quality is also reduced. We found that if ad-sensitive subscribers are sufficiently sensitive to advertising quality, a cap may lower social welfare. The welfare results suggest that a regulatory authority that is trying to increase welfare via regulation of the volume of advertising on TV might necessitate to also regulate advertising quality or, if regulating quality proves impractical, take the effect of advertising quality into consideration.

Related literature. Seminal normative work on advertising, such as Steiner (1952) and Spence and Owen (1977), tended to focus on the benefits that commercials generate to the audience but ignored the surplus obtained by the advertising firms. The assumptions of fixed levels of advertising airtime and prices prevent the analysis of whether the market under- or over-provisions advertisements.

More recently, Wright (1994) examined the effect that an advertising time ceiling has on programming quality and viewer welfare. Wright showed that regulations that limit the amount of advertising content per hour of television broadcasts can reduce programming quality and that this effect on viewer welfare is ambiguous. Under some conditions fostering competition can both reduce the number of advertisements and increase program quality, being preferable to an advertisement time ceiling.

Anderson and Coate (2005) explored the market failure in the broadcasting industry by modeling how media platforms fulfill their role of providing content to subscribers and simultaneously supplying eye-balls to advertising firms. Their work connects the goods market to the advertising market and analyzes the trade-off between the nuisance

stemming from commercial breaks during the broadcasts and the informational gains generated by the content of these commercials. Nonetheless, the authors ignored the possibility of firms investing in ad quality, which we consider in this chapter. They show that the market equilibrium may under- or over-provide advertising airtime, depending on the nuisance cost to viewers, the substitutability of programs, and the expected benefits to advertising firms from contacting viewers.

Gabszewicz, Laussel and Sonnac (2005) studied whether advertising subsidizes the newspaper prices charged to readers. They showed that in a two-sided market framework with advertisers on one side and readers on the other, the answer depends on the readership's attitude towards advertising, i.e., it depends on the proportion of readers that are ad-lovers or ad-avoiders. Dukes (2004) showed that less product differentiation or more media differentiation leads to higher market levels of advertising. In particular, if media is sufficiently differentiated, the advertising levels will surpass the socially optimal solution. Dukes (2006) investigated how competition in the media market shapes decisions about advertising and program quality. Dukes showed that product differentiation using advertising is more effective when media markets are less competitive, increasing the prices for advertised products. Gantman and Shy (2007) used an advertising-supported media model (free-to-air broadcasting) to study the firms' incentives to improve the quality of their advertisements. They showed that if improving ads' quality is profitable to firms, then it will be unprofitable to broadcasters.

This chapter is also related to the two-sided markets literature. The seminal articles by Rochet and Tirole (2003) and Armstrong (2006) investigate the determinants of the price balance between two groups of end-users when each group exerts an externality on the other, and both are intermediated by a platform. Some of the discussed determinants of the price balance are: (i) possibility of multi-homing (i.e., some end-users subscribe or use more than one platform), (ii) platform differentiation, (iii) presence of same-side externalities, (iv) platform compatibility, (v) per-transaction (or lump-sum) pricing and relative size of cross-group externalities.

To the author's knowledge there has been no previous work on the link between regulation limiting the advertising airtime and advertising quality. In part, the lack of published research no doubt reflects the scarcity of data with which to undertake formal analysis of this topic. This may be due to the fact that quality is a hard to measure concept. Despite the difficulties involved in measuring quality, Google has been trying to do so since 2007 and Videohub launched an "Effective Quality" (eQ) score for video ad placements in 2013. In the near future, as big data sets on advertising quality become available across countries with different ad ceilings, it may become possible to empirically test the theoretical results presented in this chapter, in particular those on the relationship between ad quality and ad ceilings.

2.2 The model

In this section we present a subscriber-advertiser supported media model, characterize each participating agent (platforms, subscribers and advertising firms) and describe how they interact in a three-stage game. For an illustration, think of media platforms as TV broadcasters. In Europe, an application of our model can be the direct broadcast satellite channels such as Canal Plus that are partially financed by subscription pricing. In the US, premium channels such as HBO and Showtime frequently have an individual price. Also, our qualitative results extend to the case where programmes are broadcast over the air and consumers can costlessly access such programming.

Media platforms.   There are two media platforms indexed by $i = 1, 2$ competing simultaneously and independently in two markets: (i) content subscription to subscribers and (ii) advertising airtime to firms whose profit level increases in the number of subscribers (potential customers). We assume that media platforms charge a fixed price to agents on each side of the market, e.g., a monthly flat rate to subscribers and a fee per 30-second advertising slot to firms. Platforms provide horizontally differentiated contents and each individual platform has the capacity to fully cover both sides of the market.

Each advertisement takes a fixed amount of time which will be deducted from the programming time. We assume that platforms face a fixed cost, K, regardless of the broadcast mix of advertising and regular programming.[62] Let f_i denote the advertising fee per slot charged to firms, and s_i the subscription fee charged to subscribers by platform i. Platform i chooses a pair of access prices $(f_i, s_i) \in \mathbb{R}^2_+$ that maximizes profit. Platform i's profit is defined as follows

$$\Pi_i (f_i, s_i) \equiv f_i \hat{x}_i + s_i D_i - K, \tag{31}$$

where $0 \leq D_i \leq 1$ denotes the mass of subscribers on platform i and $0 \leq \hat{x}_i \leq 1$ is the mass of firms advertising on the same platform.

Subscribers.   There is a mass one of subscribers, each of whom subscribes at most one platform, i.e., subscribers do not watch more than one TV channel simultaneously.

---

[62]As a matter of technical simplicity, platforms have constant marginal costs normalized to zero in providing their services. Alternatively, f_i and s_i may be interpreted as markups over constant marginal costs. Our qualitative results are unaffected if programming costs more than advertisements, i.e., there may be higher costs associated with acquiring or producing media content. In that instance, more advertisements (and, thus, less programming) reduce total cost for media platform i. Thus, f_i may be interpreted as the net markup over a negative marginal cost of ads (representing the cost savings from not having to purchase or produce further content) while $K > 0$ may be interpreted as the cost of content for an ad-free station. The amount K must be sufficiently small such that, in equilibrium, $\Pi_i^* \geq 0$. Otherwise, platform i would exit the market.

We assume that all subscribers already have the necessary hardware (e.g., televisions)[63] to allow them to receive the service. Subscribers are heterogeneous in two dimensions: (i) with respect to content in each platform, and (ii) regarding their attitude towards advertising. In particular, a proportion λ, $0 < \lambda < 1$, of subscribers is ad-sensitive, i.e., their utility depends on the ad-airtime and on the average ad quality.[64] The remaining $1 - \lambda$ are ad-indifferent, i.e., their utility does not change with either the duration, or the quality of advertisements. We will refer to ad-sensitive consumers as S-type consumers and to ad-indifferent ones as being I-type.

Formally, the utility derived by an S-type subscriber, indexed by $y \in [0, 1]$, from subscribing to network i at a subscription price s_i, is given by

$$
U_S(y) \equiv \begin{cases}
v - \gamma \hat{x}_1 (1 - \delta q_1) - s_1 - ty & \text{if platform 1} \\
v - \gamma \hat{x}_2 (1 - \delta q_2) - s_2 - t(1 - y) & \text{if platform 2} \\
0 & \text{if no service}
\end{cases} \tag{32}
$$

Subscriber's gross benefit of accessing an ad-free platform broadcasting his preferred program is denoted by $v > 0$. We will assume that subscribers' benefit of accessing a platform is sufficiently high to ensure full participation.[65] Parameter $\gamma > 0$ measures the nuisance cost of ads and is the same for all viewers. The term δq_i, where $\delta \geq 0$, is an ad quality evaluation factor or, alternatively, the discount on the nuisance cost of ads airtime due to the average quality of advertising. This means that for a given volume of ads on platform i, $\hat{x}_i$, an increase in the average ad quality, q_i, will attenuate the nuisance costs of advertising airtime.[66] The average ad quality is defined by

$$
q_i(\hat{x}_i) \equiv \begin{cases}
\frac{\int_{1-\hat{x}_i}^1 q_x^* dx}{\hat{x}_i} & \text{if } \hat{x}_i > 0 \\
0 & \text{if } \hat{x}_i = 0
\end{cases}, \tag{33}
$$

where $0 \leq q_x^* < 1$ is firm x's ad quality in equilibrium.[67] The formula in (33) is the unweighted average of ad quality across firms that advertise on platform i. As will be

---

[63] According to the Nielsen Company (July 2009), the average US American household in 2009 had 2.86 TV sets but only 2.5 individuals. Article *More than Half the Homes in U.S. Have Three or More TVs* available at `http://www.nielsen.com/us/en/newswire/2009/more-than-half-the-homes-in-us-have-three-or-more-tvs.html`.

[64] Previous research by Ofcom has shown that a majority of viewers sometimes or often see advertising on TV as "interfering" with their enjoyment of content, but they also see adverts as "informative" and "clever". See *Regulating the quantity of advertising on television* (Ofcom 2011).

[65] This assumption is realistic for some markets. For example, in 2013 the proportion of UK homes with digital TV is 97% (Communications Market Report 2013).

[66] In the limit case $\delta = 0$ subscribers would not care about the average ad quality. Note that for δ sufficiently high and depending on the average ad quality level, subscribers may enjoy advertising.

[67] We assume that in an ad-free platform the average quality of advertising is zero. This is an innocuous assumption because for $\hat{x}_i = 0$, we get that $\hat{x}_i (1 - \delta q_i) = 0$, for any $q_i \in \mathbb{R}$. See the section below on "advertising firms" for more on how a firm x chooses the ad quality q_x^*.

discussed further below, each firm is of some type x, where x is uniformly distributed on $[0, 1]$. Only the firms in the range $1 - \hat{x}_i$ to 1 will advertise their products on platform i. Thus, only these firms contribute for the ad quality average. The differentiation parameter $t > 0$ represents the degree to which the platforms are substitutes or subscribers' disutility of being prevented from watching their preferred programs. Subscribers' tastes are uniformly distributed on $[0, 1]$, so that the fraction of S-type subscribers with taste parameter less than y is simply y. For simplicity, we assume that not watching any program yields a zero net utility. Moreover, subscribers are aware of the advertising level and content type of each platform even before they subscribe (see "the three-stage game" description below for further details).

The utility of an I-type subscriber, indexed by z, where z is uniformly distributed on $[0, 1]$, is defined by

$$
U_I(z) \equiv \begin{cases} v - s_1 - tz & \text{if platform 1} \\ v - s_2 - t(1 - z) & \text{if platform 2} \\ 0 & \text{if no service} \end{cases} . \tag{34}
$$

The key differences between the utility functions of S-type and I-type subscribers are: (i) the effect that advertising volume exerts on S-type subscribers (but not on I-type subscribers), and (ii) the average advertising quality level, q_i, that positively affects $U_S(y)$ (but not $U_I(z)$) when ads are broadcasted.

Note that the subscribers' choice in the media market is independent of the goods market. Advertising provides product information (e.g., informs subscribers of the nature of new products in the market) that may influence consumers' shopping behavior. We assume, like Anderson and Coate (2005), that subscribers receive no other benefits from purchasing advertised products than those inherent to the product itself. In other words, firms extract the entire incremental surplus that advertising generates for their goods. This simplification allows us to focus on the media market without concerns about an endogenous distribution of informational gains between subscribers and advertising firms. Hence, subscriber's choice with respect to platforms does not depend on the information received, i.e., subscribers are solely interested in the programming contents, rather than the information conveyed by advertising.

Since the total measure of subscribers is equal to one, under the full coverage and single-homing assumptions the measure of subscribers on platform i corresponds to its market share $D_i \equiv \lambda \hat{y}_i + (1 - \lambda) \hat{z}_i$ on the subscription side of the market, where $\hat{y}_i$ and $\hat{z}_i$ are, respectively, the proportion of S-type and I-type subscribers on platform i. A summary of the notation for subscribers follows in Table 1.

λ	Proportion of S-type subscribers.
δ	Ad quality evaluation parameter.
q_i	Average ad quality on platform i.
$\hat{x}_i$	Advertising volume (time) on platform i.
y	Index for S-type subscribers.
$\hat{y}_i$	Proportion of S-type subscribers on platform i.
z	Index for I-type subscribers.
$\hat{z}_i$	Proportion of I-type subscribers on platform i.
v	Subscriber's gross benefit of accessing an ad-free platform broadcasting his preferred programme.
t	Subscriber's disutility from not being able to watch his preferred programme.
$U_S(y)$	Utility of an S-type subscriber, indexed by y.
$U_I(z)$	Utility of an I-type subscriber, indexed by z.
f_i	Advertising fee (per spot) charged by platform i.
s_i	Subscription price charged by platform i.
D_i	Platform i's market share on the subscription side of the market.

Advertising firms. There is a mass one of firms, each of which produces at most one new good. Each firm may advertise in more than one platform, i.e., multi-homing is feasible for firms. Firms may use media platforms as an advertising outlet to reach consumers and thus increase profits. We assume that ads are placed by monopoly firms of new products to inform potential customers about the existence, characteristics, and prices of the products that they offer as well as persuade them to buy. All viewers (subscribers) are homogeneous to advertisers so that there is no matching of advertisements to programming (e.g. tennis clubs advertising in a tennis program). Having received advertising for a particular new product, a consumer knows his willingness to pay for it and will purchase it with some probability if his willingness to pay is no less than its advertised price. New products are produced at a constant cost per unit $c > 0$. Each firm/new product is characterized by a purchase probability x, where x is uniformly distributed on $[0, 1]$. Products with a higher x are more likely to be attractive to consumers.

A firm x may wish to invest in advertising quality, q_x. The incentive for firm x to spend resources to improve ad quality is driven by the increase in the buying probability displayed by S-type consumers (perceived quality) for the advertised product. Moorthy and Zhao (2000) found evidence that advertising expenditure (taken as proxy for ad quality) and perceived quality of the underlying product are in general positively correlated

for both durable and nondurable goods, even after accounting for objective quality, price and market share.

One might expect that there exist diminishing returns to advertising quality, i.e., as the ad quality level for a given product increases, the increment in the probability of purchasing that product becomes smaller. We impose the following conditions on the ad quality effect $A(q_x)$: $A(0) = 0$, $0 \leq A(q_x) < 1$, $A'(q_x) > 0$, and $A''(q_x) < 0$, for $0 \leq q_x < 1$. In the analysis that follows, it will be useful to specify $A(q_x) \equiv \sqrt{q_x}$. Henceforth, if a firm with a new product of type x invests in q_x units of quality, an ad-indifferent consumer will purchase the product with probability x, while an ad-sensitive consumer will purchase the product with probability $x + (1-x)\sqrt{q_x}$.[68] Firm x pays a cost of implementing quality level q_x equal to βq_x, with $\beta > 0$. The higher β, the more expensive the ad quality technology.

A consumer will be willing to pay $\omega > c$ or 0 for advertised products. Thus, all firms will advertise price ω. Note that a lower price does not increase the probability of a sale (only better advertising will do so for ad-sensitive consumers). Let firm x's profit be defined as follows

$$
\Pi(x) \equiv \begin{cases}
(\omega - c) D_{x,1} - \beta q_x - f_1 & \text{if platform 1} \\
(\omega - c) D_{x,2} - \beta q_x - f_2 & \text{if platform 2} \\
(\omega - c)(D_{x,1} + D_{x,2}) - \beta q_x - f_1 - f_2 & \text{if platforms 1 and 2} \\
0 & \text{if no ads}
\end{cases}, \tag{35}
$$

where $D_{x,i} \equiv (1-\lambda)\hat{z}_i x + \lambda \hat{y}_i (x + (1-x)\sqrt{q_x})$ is the demand for firm x's product when it advertises in platform i.[69] The demand $D_{x,i}$ reflects the fact that $(1-\lambda)\hat{z}_i$ ad-indifferent consumers will buy product x with probability x, and $\lambda \hat{y}_i$ ad-sensitive consumers will buy product x with probability $x + (1-x)\sqrt{q_x}$, as previously discussed. We assume that the cost of ad quality, β, is sufficiently high. In particular, we assume that

$$
\beta > \max\left\{ \frac{(\omega - c)\lambda}{2}, (\omega - c)\lambda^2 \right\} \tag{36}
$$

to guarantee that in an equilibrium where firms multi-home (i) $q_x^* < 1$ and (ii) $\partial D_{x,i}^*/\partial x > 0$. Condition (i) guarantees that no ad-sensitive consumer purchases a product with

---

[68] Suppose that firm x's commercial featured a well-known public figure rather than an anonymous performer. Then, S-type subscribers would have purchased the advertised product with a higher probability, whereas I-type subscribers would not have purchased the advertised product with a higher probability as a result of the presence of the public figure. Recall that I-type subscribers are only concerned with the product's features (objective information), not with the way those features are presented.

[69] Note that (35) is compatible with both classical views of advertising: informative and persuasive. The term $((1-\lambda)\hat{z}_i + \lambda \hat{y}_i) x$ corresponds to the informative role of advertising since all subscribers of platform i learn about the existence and features of x's product upon watching the advertisement. The term $\lambda \hat{y}_i (1-x)\sqrt{q_x}$ captures the persuasive effect of advertising.

a probability exceeding 1. Condition (ii) guarantees that ad quality does not have an "explosive" effect on consumer demand. Despite the fact that firms with a low x have further incentives to invest in ad quality, the demand faced by a firm with a low x (low purchase probability) will not be higher than the demand faced by a firm with a higher x (high purchase probability). In other words, ad quality will help firms to sell more (specially those with a low x), though, not to a point where a low x firm sells more than a firm with a higher x.

If a firm does not advertise, we assume for simplicity that the product will not be known in the market and thus generates no profit. Note that the cost, βq_x, of producing an advertisement of quality q_x is only incurred once by the firm, regardless of the number of platforms broadcasting the advert.[70] A summary of the notation for advertising firms follows in Table 2.

TABLE 2: Notation for advertising firms

$\Pi(x)$	Firm x's profit.
q_x	Firm x's ad quality.
ω	Willingness to pay for a product by each consumer.
c	Marginal cost of producing a product.
x	Index for advertising firms.
$D_{x,i}$	Demand for firm x's product when it advertises in platform i.
f_i	Advertising fee charged by platform i.
β	Cost of one unit of ad quality.

The three-stage game. The participating agents interact according to the following three-stage game. Profit-maximizing platforms move first by choosing the subscription price and the fee per advertising spot. Second, the advertising slots are sold and firms with a slot choose on the level of advertising quality. Advertising firms produce their advertisements and decide the price of the products they sell. Platforms broadcast content and the advertising spots, and finally, consumers make their choices regarding the media subscription and the product market. A summary of the timing of the model follows in Table 3.

---

[70]Advertising firms generally use either a production company or an advertising agency to produce their commercials. The cost of the commercial will depend on the creative content, music, location, cast, etc., as well as the overall production quality of the commercial. Once the shooting is complete and the ad produced, the advertising firm needs to buy advertising slots on platforms in order to broadcast. More on how to get a commercial made at `http://www.thinkbox.tv`.

I.   Media platforms choose simultaneously and independently the pair of prices (f_i, s_i). Each platform i chooses (f_i, s_i) such that it maximizes (31).

II.   Firms decide if they want to buy advertising slots from one, two or none of the media platforms, depending on the advertising airtime prices and their (rational) expectation of how many subscribers there will be in each platform. Advertising firms choose the ad quality level that maximizes (35) and set the product price at ω.

III.   Subscribers, indexed by y and z, maximize (32) and (34), respectively, choosing between the two media platforms according to their idiosyncratic preferences regarding the programming type, the subscription prices, the advertising airtime and the average ad quality in each platform. Advertising airtime and average ad quality are only taken into consideration by S-type subscribers (and not by I-type subscribers).

2.3   The subgame perfect Nash equilibrium

The model is solved by backward induction in order to find a (symmetric) subgame perfect Nash equilibrium (SPNE). All computations are relegated to an appendix.

In stage III we solve the subscribers' problem, indexed by y and z, to maximize (32) and (34), respectively. Given that v is sufficiently high to ensure full participation of all consumers, an S-type subscriber at point y will choose to subscribe platform 1 if $v - \gamma \hat{x}_1 (1 - \delta q_1) - s_1 - ty > v - \gamma \hat{x}_2 (1 - \delta q_2) - s_2 - t(1 - y)$. Thus,

$$\hat{y}_i (\hat{x}_i, \hat{x}_j, q_i, q_j, s_i, s_j) = \frac{1}{2} + \frac{\gamma (\hat{x}_j (1 - \delta q_j) - \hat{x}_i (1 - \delta q_i)) + s_j - s_i}{2t}. \tag{37}$$

An I-type subscriber at point z will choose to subscribe platform 1 if $v - s_1 - tz > v - s_2 - t(1 - z)$. Thus

$$\hat{z}_i (s_i, s_j) = \frac{1}{2} + \frac{s_j - s_i}{2t}. \tag{38}$$

In stage II, the following result regarding ad quality choice emerges from firm x's problem.

Proposition 1 *If advertising firms buy advertising slots in both platforms, the ad quality chosen by an advertising firm x equals*

$$q_x^* = \left(\frac{(\omega - c) \lambda (1 - x)}{2\beta} \right)^2. \tag{39}$$

Proof All proofs are in an appendix. □

Proposition 1 underscores the driving forces that affect the decision of firm x to spend resources on advertising quality. First, as the proportion of subscribers liable to be persuaded by ads, λ, increases, the return to persuasive advertising also increases. As a consequence, firms have more incentive to invest in ad quality. Second, the quality increasing technology is crucial since it affects costs. Therefore, the incentive to improve quality increases with cheaper technologies (lower β). Third, higher profit margins in the goods market increase the return to persuasive ads and, thus, increase the incentive to invest in better ad quality. Fourth, intuitively firms with lower types (i.e., low x) have a higher incentive to invest in ad quality as a means to compensate for the weak informative effect of their ads. New products with a high probability of being bought may be those released by more established and better-known firms. Consumers will buy those products with higher probability. Firms with lower types may be interpreted as relatively unknown firms.[71]

Lemma 1 *In a symmetric equilibrium, where $f_i^* = f_j^*$ and firms expect $D_{x,i}^* = D_{x,j}^*$, firms advertise on both platforms (multi-homing) or do not advertise at all.*

The solution in (39) is conditional on firm x multi-homing, which assures that its advertisement will be watched by all consumers. In equilibrium, we need to verify that at $\left(s_i^*, s_j^*, f_i^*, f_j^*\right)$, the advertising firms will indeed multi-home, rather than single-home, while the remaining firms (with lower types) will not advertise at all. From (35) it is straightforward to conclude that in a symmetric equilibrium, where $D_{x,i}^* = D_{x,j}^*$ and $f_i^* = f_j^*$, if $(\omega - c) D_{x,i}^* - \beta q_x^* - f_i^* > 0$, then it must be the case that $(\omega - c) D_{x,j}^* - f_j^* > 0$ (recall that the cost βq_x^* is incurred only once at the time of the ad production) and the firm will choose to multi-home.

Also, if platforms exhibit a sufficiently small difference in advertising fees and advertising firms expect $D_{x,i}^*$ to be sufficiently close to $D_{x,j}^*$, each firm will multi-home or not advertise at all. We can show this in the following way. Without loss of generality, assume in what follows that $f_i^* < f_j^*$ and $D_i^* > D_j^*$. Take first the firm located at $x = 1$. This firm has the highest willingness to pay for an advert and will choose $q_1^* = 0$ regardless of the number of platforms broadcasting its advert. A necessary (but not sufficient) condition

---

[71] An example of this is Apple that, in 1984, had a relatively small market share in the personal computer market. In that year, Apple released the ad "1984" in the US television introducing the Apple Macintosh personal computer. The ad, directed by Ridley Scott, is considered a masterpiece in advertising and widely regarded as one of the most memorable and successful television commercials of all time in the US. A more recent example can be found in the UK where Sky exhibits a market share of 12% in the residential fixed line market, while BT (the incumbent) is the major operator with 45% market share. Contrarily to BT, Sky has hired the services of Hollywood stars such as Bruce Willis and Al Pacino to advertise its services in the UK.

for this firm to advertise only on platform i is $(\omega - c) D_{1,i}^* - f_i^* > 0$.[72] This implies that for $(f_j^*, D_{1,j}^*)$ sufficiently close to $(f_i^*, D_{1,i}^*)$, the firm will also attain a positive profit, $(\omega - c) D_{1,j}^* - f_j^* > 0$, from advertising on platform j.

Take now the advertising firms characterized by $0 \leq x < 1$. They will choose an ad quality $q_x^* > 0$. A necessary condition for these firms to advertise on platform i is $(\omega - c) D_{x,i}^* - \beta q_x^* - f_i^* > 0$. In addition, if a firm advertises on platform j, the profit from doing so will be $(\omega - c) D_{x,j}^* - f_j^*$, which will be positive for $(f_j^*, D_{x,j}^*)$ sufficiently close to $(f_i^*, D_{x,i}^*)$. Intuitively, given the absence of an advertising production cost when broadcasting the same advert via a second platform (e.g., platform j), a firm $x \in [0,1)$ is willing to pay a higher advertising fee (or pay the same fee for a smaller audience) to platform j, as compared to platform i. A costly advert to be broadcasted via only one platform (that will be watched by a fraction of consumers) may be unprofitable or will generate less profit compared to a multi-homing solution (which ensures that all consumers watch the advert).[73] Moreover, note that the optimal ad quality, q_x^*, increases when firms choose multi-homing, rather than single-homing. This is due to the fact that the benefit from an additional unit of ad quality increases with the size of the audience.

From (39) we can compute the average ad quality level on platform i and write it as a function of the advertising volume broadcasted by that platform.

Proposition 2 *If advertising firms purchase advertising slots in both platforms, the average ad quality on platform i equals*

$$q_i(\hat{x}_i) = \frac{1}{3} \left(\frac{(\omega - c) \lambda}{2\beta} \right)^2 \hat{x}_i^2. \tag{40}$$

Importantly, the average ad quality is increasing in the volume of ads broadcast. The advertising firms willing to pay more for an advertising slot are those with a higher informative effect (i.e., higher x). These are also the firms with fewer incentives to invest in ad quality. As a consequence, when platforms sell more advertising slots, the marginal advertiser will exhibit an above-average ad quality, thus increasing the average ad quality broadcast by the platform.

Plugging the ad quality solution q_x^* into $\Pi(x)$, we obtain

---

[72]Note that if $(\omega - c) D_{1,i}^* = f_i^*$, there is essentially no advertising, as firms with lower willingness to pay, i.e., those in the interval $0 \leq x < 1$, will choose not to advertise at all. We concentrate our analysis on the case where $(\omega - c) D_{1,i}^* > f_i^*$. Otherwise, advertising airtime regulation would not be a concern.

[73]The presence of an ad production cost that is paid only once regardless of how many stations air the advert induces scale economies, which lie at the root for this preference for multi-homing. That ad production costs may be significant is suggested by the following examples. In 2004, Chanel paid US\$ 33 million for a two-minute commercial. Nicole Kidman reportedly received \$3 million to act in this ad. Other big spenders include Guinness, which spent US\$ 16 million to create a domino effect through a small town in Argentina, and a British insurance company that paid US\$ 13 million for celebrities such as Ringo Starr and Bruce Willis. More at `http://www.businessinsider.com`.

$$\Pi\left(x\right)=\left(\omega-c\right)x+\frac{1}{\beta}\left(\frac{\left(\omega-c\right)\lambda\left(1-x\right)}{2}\right)^{2}-f_{1}-f_{2}.$$

The profit of the marginal advertiser is zero. From this zero-profit condition we can derive firms' demand function for advertising slots in platform i

$$\hat{x}_{i}\left(f_{i},f_{j}\right)=2\beta\frac{1-\sqrt{\frac{1}{\beta}\left(\beta-\lambda^{2}\left(\omega-c\right)+\lambda^{2}\left(f_{i}+f_{j}\right)\right)}}{\left(\omega-c\right)\lambda^{2}}, \tag{41}$$

as long $0\leq\hat{x}_{i}\leq 1$. As discussed further below, this will be the case in equilibrium.

In stage I we solve platform i's profit maximization problem

$$\max_{f_{i},s_{i}}\Pi_{i}\left(f_{i},s_{i}\right)\equiv s_{i}\left(\lambda\hat{y}_{i}+\left(1-\lambda\right)\hat{z}_{i}\right)+f_{i}\hat{x}_{i}$$

where $\hat{x}_{i}$, $\hat{y}_{i}$ and $\hat{z}_{i}$ are defined, respectively, by (41), (37) and (38).

Conventional wisdom and the standard IO literature suggest that prices are strategic complements. However, in this particular case, we can show (see the SPNE appendix) that the advertising fees (f_{i}, f_{j}) are strategic substitutes, i.e., platforms' reaction functions are downward sloping regarding the ad-fees. This is due to the one-off cost supported by advertisers when investing in ad quality. For the marginal firm, advertising is only profitable if the same commercial can be broadcasted via both platforms, rather than just one. If one of the platforms were to slightly increase its advertising fee, the marginal advertiser would be unable to fully recover the cost of producing the advert and, thus, would choose not to advertise at all. As a best-reply, the other platform chooses to slightly decrease its advertising fee in order to recover the marginal advertiser.

In a symmetric equilibrium, $s_{i}^{*}=s_{j}^{*}$ and $f_{i}^{*}=f_{j}^{*}$, in particular

$$\begin{cases} s_{i}^{*}=t \\ f_{i}^{*}=\frac{1}{25\lambda^{2}}\left(10\lambda^{2}\left(\omega-c\right)-6\beta+2\sqrt{\beta\left(9\beta-5\lambda^{2}\left(\omega-c\right)\right)}\right) \end{cases},$$

and

$$\hat{x}_{i}^{*}=2\beta\frac{1-\sqrt{\frac{1}{\beta}\left(\frac{13}{25}\beta-\frac{1}{5}\lambda^{2}\left(\omega-c\right)+\frac{4}{25}\sqrt{\beta\left(9\beta+5\lambda^{2}\left(\omega-c\right)\right)}\right)}}{\lambda^{2}\left(\omega-c\right)},$$

$$\hat{y}_{i}^{*}=\frac{1}{2}\ \text{and}\ \hat{z}_{i}^{*}=\frac{1}{2}.$$

The assumption in (36) yields $0<\hat{x}_{i}^{*}<1$.

The following comparative static results on the sensitiveness of consumers to the qual-

ity of advertising are of interest.

Proposition 3 *In a symmetric equilibrium, an increase in the proportion of ad-sensitive subscribers, λ, will: (i) expand firms' demand function for advertising slots; (ii) make platforms set lower advertising fees; and (iii) produce no impact on the subscription fees.*

An increase in the proportion of ad-sensitive subscribers means that there are more consumers willing to pay ω with probability $x + \sqrt{q_x}$, rather than just probability x. This implies that advertising firms have an additional incentive to upgrade their ad quality. The profits of advertising firms are increasing in the proportion of ad-sensitive subscribers, λ. Thus, firms are willing to pay more for an advertising slot as λ increases, which works out as an expansion of firms' demand for advertising slots.

As depicted in Figure 1 below, the firms' (inverse) demand function for advertising slots is convex[74] and has an intercept that is independent of λ. Note that an advertiser with the highest willingness to pay for a slot (i.e., an advertiser at $x = 1$) will sell its product with probability one to all consumers who watch its advert regardless of the ad's quality.
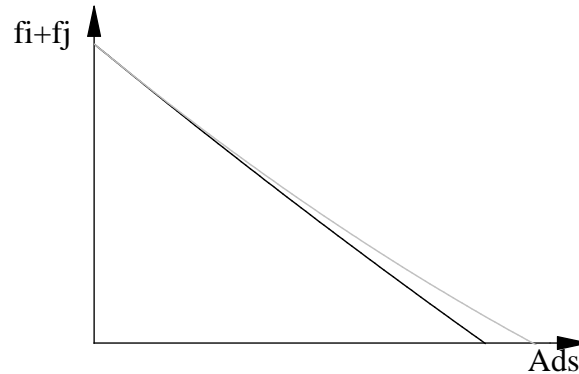
Figure 1: Firms' willingness to pay for advertising slots.

An increase in the quantity demanded of ads brought about by a change in λ is increasingly bigger the lower the willingness to pay of advertisers. As a result, an increase in λ induces to a flatter demand for ads, which in turn makes it more attractive to lower the price of slots since a given increase in sales can be obtained with a smaller price decrease and, hence, a smaller inframarginal reduction in revenue. In other words, the effect on platforms' profits of a slight price cut of advertising slots is more than compensated by the expansion of the quantity demanded of those slots. This explains why an increase in λ leads to a decrease in the price of ads.

In a symmetric equilibrium, both platforms exhibit the same volume of advertisements and average ad quality. As a consequence, from the subscribers' point of view (both

---

[74]This can be mathematically verified by checking that the signal of the second derivative with respect to advertising fees of (41) is positive.

89

ad-indifferent and ad-sensitive), platforms can only differentiate themselves in terms of content, with such differentiation measured by parameter t. In equilibrium, $s_i^* = t$ and, thus, the proportion of ad-sensitive subscribers does not have any impact on subscription fees.

Proposition 4 *In a symmetric equilibrium, an increase in the proportion of ad-sensitive subscribers, λ, will: (i) increase media platforms' profits; (ii) produce no impact on the utility of ad-indifferent subscribers; (iii) decrease (increase) the utility of ad-sensitive subscribers, if δ is sufficiently low (high); and (iv) increase advertisers' profits.*

Ad-indifferent subscribers are not affected either by advertising airtime or the average ad quality. In fact, as per the utility defined in (34), ad-indifferent subscribers are only concerned with their idiosyncratic preference with respect to content and the subscription fees. In equilibrium, the subscription fees equal t, thus being independent of λ. Hence, an increase in λ will not produce any effect on the utility of ad-indifferent subscribers.

In equilibrium, an increase in the proportion of ad-sensitive subscribers results in more advertisements but also in enhanced advertising quality on average. In general, the net effect on S-type subscribers welfare is ambiguous. On the one hand, if the subscribers' valuation of average ad quality, δ, is negligible, then, the nuisance effect of ads will dominate the quality effect, and S-type subscribers will be worse off. On the other hand, if δ is sufficiently high, the ad quality enhancing effect dominates the nuisance effect of ads and S-type subscribers will be better off.

Advertisers' profits increase due to two distinctive effects. First, as explained in Proposition 3, in a symmetric equilibrium platforms have an incentive to decrease advertising fees when the proportion of S-type subscribers expands. Thus, lower advertising fees increase the profits not just of firms that would advertise at higher advertising fees, but also allow more firms, that at higher prices would not advertise their products, to buy an advertising slot. Second, as λ increases, so does the marginal benefit of ad quality for firms. In fact, there is an increased fraction of S-type consumers willing to buy the advertised products at an enhanced probability. Thus, advertising firms increase sales and profits.

2.4   Advertising caps

As mentioned earlier, advertising airtime is regulated in a number of countries. In this section we address how the presence of advertising caps may affect the average quality of advertisements and, ultimately, platforms' profits, subscribers' surpluses and advertisers' profits. Let the regulated volume of advertisements be given by $\bar{x}$, where $\bar{x} < x_i^*$, i.e., the advertising cap is binding.

Proposition 5 *In a symmetric equilibrium, a lower advertising cap will reduce the average quality of advertisements.*

Equation (40) shows that a lower advertising cap will cause the average advertising quality to decrease. In Proposition 2 we highlighted that the average ad quality is increasing in the volume of ads broadcast. This is explained by the fact that firms with a higher informative effect (i.e., higher x and also higher willingness to pay for an advertising slot) are simultaneously the firms with fewer incentives to invest in ad quality. As the marginal advertiser exhibits higher ad quality as compared to firms with a higher informative effect, platforms selling more advertising slots will broadcast higher ad quality on average. A lower advertising cap will, thus, restrict advertisements to firms with a higher informative effect to the detriment of firms that would invest in ads of higher quality. In Proposition 6 we consider the welfare effects of an advertising cap.

Proposition 6 *In a symmetric equilibrium, a lower advertising cap will (i) decrease media platforms' profits; (ii) produce no impact on the utility of ad-indifferent subscribers; (iii) decrease (increase) the utility of ad-sensitive subscribers, if δ is sufficiently high (low); and (iv) decrease advertisers' profits.*

The platforms are clearly worse off through being constrained in advertising airtime given that the advertising cap $\bar{x} < x_i^*$ is binding by assumption.

As previously mentioned, ad-indifferent subscribers are concerned with their idiosyncratic preference with respect to content and the subscription fees, but are not concerned with the volume or quality of advertising broadcasted by each platform. Given that the subscription fees, set at t, will not be affected by the advertising cap $\bar{x}$, the utility of ad-indifferent subscribers will not change.

As ad-sensitive subscribers' utility is decreasing in the volume of advertisements but increasing in the average advertising quality it is impossible, without knowledge of the parameters of the model, to unambiguously rank the Nash equilibrium and advertising cap solution in terms of ad-sensitive subscribers' welfare. On the one hand, if ad-sensitive subscriber's valuation of average ad quality, δ, is negligible, then, the nuisance effect of ads airtime will dominate the ad quality effect, and ad-sensitive subscribers will be better off with an advertising cap $\bar{x} < x_i^*$. On the other hand, if δ is sufficiently high, the ad quality effect dominates the nuisance effect of ads and ad-sensitive subscribers will be worse off with an advertising cap (see the result in Proposition 5). Put differently, if δ is sufficiently high, a tighter cap will hurt S-type subscribers because the reduction of the average ad quality will more than offset the direct benefit of a reduced nuisance from advertisements themselves.

This suggests that a regulatory authority which is trying to increase viewer welfare via restrictions on the advertising airtime might also need to consider regulating adver-

tising quality or, at a minimum, take the effect of advertising quality into consideration. However, in practice, the implementation of quality regulation is likely to be difficult as quality is hard to clearly define and measure.

Under an advertising cap $\bar{x} < x_i^*$ a lower advertising fee does not expand the advertising airtime. As a result, a lower advertising cap will incentivize platforms to increase their advertising fees. From (35) it is clear that firms' profits are decreasing with respect to advertising fees. Since a lower advertising cap incentivizes platforms to set higher advertising fees, a tighter advertising cap necessarily hurts advertisers' profits.

The results in Proposition 5 and Proposition 6 suggest that the social effects of an advertising cap depend on both the advertising provision and the average level of ad quality. In fact, as discussed above, if δ is sufficiently high, the introduction of an advertising cap may make *all* market players (with the exception of ad-indifferent subscribers whose utility does not change with $\bar{x}$) worse off.

2.5   Robustness

This section addresses issues concerning the robustness of our conclusions. First, we address how our findings on average quality of advertisements (Proposition 5) and social welfare (Proposition 6) would differ under a case of asymmetric regulation. Second, we discuss how those findings depend upon our specific model.

2.5.1   Asymmetric regulation

In the basic model, broadcasters are constrained by the same advertising cap. Nonetheless, the regulation of the maximum volume of advertising may differ across platforms. For example, in the UK, the BBC is a well-known example of an ad free public TV financed only by licence fees and public transfers, while for other public service broadcasters the maximum average number of advertising minutes is eight per hour. We find other examples of asymmetric regulation in Germany and France where the Government decided to ban commercial advertising on public TV stations, while private platforms are allowed to broadcast some advertisements per hour.

Suppose now that there is an ad free public TV (network 1) financed only by licence fees and public transfers (i.e., $s_1 = \hat{x}_1 = 0$), and an unregulated free-to-air TV (network 2) where $s_2 = 0$ and $\hat{x}_2$ is chosen such that it maximizes profit. Given that advertising is only possible on network 2, as shown in the appendix (see "Stage II: firms' choice"), an advertiser x will choose an ad quality level

$$q_x^* = \left(\frac{(\omega - c) \lambda \hat{y}_2 (1 - x)}{2\beta} \right)^2,$$

and consequently the average ad quality on network 2 will be

$$q_2 \left(\hat{x}_2, \hat{y}_2 \right) = \frac{\int_{1-\hat{x}_2}^1 q_x^* dx}{\hat{x}_2} = \left(\frac{(\omega - c) \lambda \hat{y}_2}{2\beta} \right)^2 \frac{\hat{x}_2^2}{3}, \text{ for } \hat{x}_2 > 0. \tag{42}$$

Note that the average ad quality depends both on the advertising volume, $\hat{x}_2$, and the proportion of ad-sensitive consumers watching network 2, $\hat{y}_2$. In Proposition 5 we showed that, in a context of regulatory symmetry, a lower advertising cap reduces the average quality of advertisements broadcasted by a platform. Here, however, because of regulatory asymmetry we need to take into account the effect that an advertising cap may have on $\hat{y}_2$ as well. This effect will depend on how sensitive ad-sensitive consumers are to advertising volumes.

If a tighter advertising cap drives advertisers to expect a sufficiently large expansion in $\hat{y}_2$, then advertisers will choose a better quality for their ads. In this case, a tighter advertising cap may increase ad-sensitive consumers' surplus since fewer ads are broadcasted on network 2 and the average quality of those ads has increased. Ad-indifferent consumers will remain indifferent to advertising caps as they only take into account programming preferences and subscription prices (which are set at zero in this case). Platform 2 will increase the advertising fee up to the point where a higher advertising fee does not decrease the demand for ad slots below the cap. However, it will be clearly worse off as a result of being constrained regarding advertising airtime. The effect of a tighter advertising cap on advertisers is dubious. On the one hand, an advertising cap will prevent some firms from advertising their products. Thus, these firms will clearly be worse off. On the other hand, the firms that obtain a slot will pay a higher advertising fee while benefiting from a wider audience. Thus, the net effect for these firms is not clear.

A tighter advertising cap may, however, decrease the average ad quality. This will be the case if advertisers expect $\hat{y}_2$ to increase only by a sufficiently small amount, or to decrease. If δ is sufficiently high, the average ad quality effect dominates the nuisance effect of ad airtime, and ad-sensitive consumers watching network 2 will be worse off with a cap. Also, note that due to the decrease of the average ad quality, some consumers may switch from network 2 to network 1. Thus, this group of consumers must be worse off as well. As discussed above, ad-indifferent consumers will remain indifferent to caps, while platform 2 will increase advertising fees but will clearly be worse off through being constrained in advertising airtime. A tighter advertising cap may decrease firms' profits for two reasons. First, a set of firms will be unable to advertise their products as there are not enough advertising slots under a binding regulatory cap. Second, the firms that benefit from an advertising slot will pay higher advertising fees while obtaining a lower return on ads due to consumer switching (from network 2 to network 1) and the lower ad quality, which will reduce the purchase probability of ad-sensitive consumers.

In a nutshell, while the result that a lower cap may reduce the average ad quality and harm social welfare is still a possibility under asymmetric regulation, it is not clear that this must necessarily be the case.

2.5.2 Alternative models

We have adopted a model with specific assumptions regarding the demand for advertisements and the impact of ad quality on consumer behavior. Firms are monopolists of new goods who wish to inform consumers of their products' existence and characteristics. When trade occurs, firms extract all the gains from trade. The incentive for firm x to improve ad quality is driven by the increase in the buying probability displayed by S-type consumers, rather than an increase in their willingness to pay for the product. These are strong assumptions and it is important to discuss the sensitivity of our main results to our model specifications.

New goods may be substitutes for consumers. For example, consumers may purchase only one good from those they have been informed about. In this case there is a business stealing effect in placing an ad with better quality insofar as trade may come in detriment of the advertiser's competitors. A constraint in advertising airtime may restrict the degree of competition in the goods market, i.e., reduce the number of effective players in the goods market down to those who can advertise on TV. Once the cap comes into effect, the shift in market structure may be such that advertising firms have fewer incentives for investing in ad quality. In this case, a degree of substitution among goods reinforces our result that a lower advertising cap will reduce the average quality of advertisements and, in turn, may decrease the utility of ad-sensitive subscribers. However, depending on how market competition is modeled, advertising firms may have further incentives to invest in ad quality under a cap. In this latter case, while our result on the average ad quality is still a possibility, it is less likely than in the basic model when firms are monopolists of new goods.

Supposing that firms do not gain all the surplus $(\omega - c)$ from trade would reduce the marginal benefit of ad quality to firms and consequently would have a negative impact on the average ad quality q_i in (33). Note that this negative impact happens regardless of the imposition of an advertising cap. Our qualitative result regarding the impact of a lower advertising cap on the average quality of advertisements (Proposition 5) will hold even if we allocate a smaller proportion of the total surplus from trade to firms.[75] The qualitative welfare effects of an advertising cap on media platforms' profits and advertisers' profits will hold as previously explained and set out in Proposition 6. In relation to (both S-type and I-type) consumers, we need to take into account that the effect of a tighter advertising

---

[75]This can be shown, for example, if advertising firms retain a fraction $0 < \alpha < 1$ of the total surplus, i.e., retain $\alpha (\omega - c)$, rather than $\omega - c$.

cap will cause a reduction in the number of transactions. As a result, consumers will get a lower surplus in the goods market. In this case, if δ is sufficiently high, the introduction of an advertising cap may make *all* market players worse off, including the I-type consumers.

Assume now that ad quality increases S-type consumers' willingness to pay for the product, rather than increasing their buying probability. In particular, if a firm with a new product featured by type x invests in q_x units of ad quality, an I-type consumer will be willing to pay $(\omega - c)$ to purchase the product with probability x, while an ad-sensitive consumer will be willing to pay $(\omega - c + \sqrt{q_x})$ to purchase the product with probability x. The profit for a multi-homing firm x is

$$\Pi(x) \equiv (\omega - c + \sqrt{q_x}) \lambda x + (\omega - c)(1 - \lambda) x - \beta q_x - f_1 - f_2,$$

assuming that the firm is able to engage in a policy of price discrimination between ad-indifferent and ad-sensitive consumers. For example, this may consist in discriminating women against men, or single adult individuals against small children with their parents. Then,

$$q_x^* = \left(\frac{\lambda x}{2\beta}\right)^2 \text{ and}$$

$$q_i = \left(\frac{\lambda}{2\beta}\right)^2 \frac{\hat{x}_i^2 - 3\hat{x}_i + 3}{3}, \text{ for } \hat{x}_i > 0.$$

It is straightforward that $dq_i/d\hat{x}_i = 2\hat{x}_i - 3 < 0$, given that $0 < \hat{x}_i < 1$, meaning that an advertising cap may increase the average ad quality on platforms. On the one hand, the TV viewing experience for subscribers may be improved as fewer ads are broadcasted and the average ad quality is improved. On the other hand, platforms that rely only on advertising revenues will be worse off due to the regulatory constraint, and firms' profits will decrease as advertising fees are higher and some firms will be unable to advertise.

We can conclude that the impact of an ad airtime cap on the average ad quality depends on whether the ad quality increases the purchase probability or the willingness to pay of S-type consumers. In the latter case, regulating the advertising airtime may increase subscribers' welfare, even though it is not clear in which direction the aggregate social welfare will be affected.

2.6   Conclusions

Entertaining and informative contents are the bait to get prospective purchasers of consumer goods exposed to advertisements. What makes broadcasting different from other goods is that the broadcast delivers two goods, the contents to subscribers and

the audience to the advertising firms. Thus, it is useful to take a two-sided market's perspective when addressing economic issues on the TV broadcasting market.

In this chapter we showed that the average ad quality broadcast by a media platform may be increasing in the volume of ads. The marginal gain of investment in ad quality is lower for firms with a higher informative effect. Thus, the marginal advertiser on a platform exhibits higher ad quality compared to firms with a higher informative effect and, consequently, platforms selling more advertising slots will broadcast higher ad quality on average. We concluded that advertising caps may, thus, cause the average ad quality to decrease.

We showed that an advertising cap may generate the following welfare effects. Media platforms become worse off through being constrained in advertising airtime. Provided that a lower advertising cap incentivizes platforms to set higher advertising fees, as a lower advertising fee could not increase the volume of advertising sales, a tighter advertising cap will necessarily hurt advertisers' profits. The effect on subscribers' welfare is ambiguous because the ad quality reduction resulting from a cap offsets the direct subscribers' gain from watching fewer ads. We found that if ad-sensitive subscribers are sufficiently sensitive to ad quality (i.e., the quality reduction outweighs the direct effect of the cap), a cap may even reduce the social welfare level. The welfare results suggest that a regulatory authority that is trying to increase welfare via regulation of the volume of advertising on TV might necessitate to also regulate advertising quality or, if regulating quality proves impractical, take the effect of advertising quality into consideration.

Although this chapter does not provide the definitive answer to the question of whether regulating the advertising airtime increases subscriber surplus and the aggregate social welfare, it does provide a framework which can be used to answer this question. Such a framework, to the best of the author's knowledge, is currently absent in the economic analysis of advertising caps and ad quality.

2.7   References

Anderson, S., and J. Gabszewicz (2006), *The media and advertising: a tale of two-sided markets*, Handbook on the Economics of Art and Culture (Victor A. Ginsburg and David Throsby, eds.), North-Holland, Vol. 1, pp. 567–614.

Anderson, S., and S. Coate (2005), *Market Provision of Broadcasting: A Welfare Analysis*, Review of Economic Studies, Vol. 72, No. 4, pp. 947–972.

Armstrong, M. (2006), *Competition in two-sided markets*, RAND Journal of Economics, Vol. 37, No. 3, pp. 668–691.

Bagwell, K. (August 2005), *The Economic Analysis of Advertising*, Handbook of Industrial Organization, Vol. 3.

Cable TV Puts New Spin on Ads, Wall Street Journal, May 16, 2007.

Communications Market Report 2013, Ofcom, 1 August 2013, available at `http://stakeholders.ofcom.org.uk/binaries/research/cmr/cmr13/2013_UK_CMR.pdf`.

Dukes, A. (2004), *The Advertising Market in a Product Oligopoly*, Journal of Industrial Economics, Vol. 52, No. 3, pp. 327–348.

Dukes, A. (2006), *Media Concentration and Consumer Product Prices*, Economic Inquiry, Vol. 44, No. 1, pp. 128–141.

Dukes, A. and E. Gal-Or (2003), *Negotiations and Exclusivity Contracts for Advertising*, Marketing Science, Vol. 22, No. 2, pp. 222–245.

Epstein, G. (1998), *Network Competition and The Timing of Commercials*, Management Science, Vol. 44, No. 3, pp. 370–387.

Gabszewicz, J., D. Laussel, and N. Sonnac (1999), *TV-Broadcasting Competition and Advertising*, CORE Discussion Papers, Université Catholique de Louvain, Center for Operations Research and Econometrics (CORE).

Gabszewicz, J., D. Laussel, and N. Sonnac (2001), *Press advertising and the ascent of the Pensée Unique*, European Economic Review, Vol. 45(4-6), pp. 641–651.

Gabszewicz, J., D. Laussel, and N. Sonnac (2005), *Does advertising lower the price of newspapers to consumers? A theoretical appraisal* Economics Letters, Vol. 87, No. 1, pp. 127–134.

Gal-Or, E., and A. Dukes (2003), *Minimum Differentiation in Commercial Media Markets*, Journal of Economics and Management Strategy, Vol. 12, No. 3, pp. 291–325.

Gal-Or, E., M. Gal-Or, J. May, and W. Spangler (2006), *Targeted Advertising Strategies on Television*, Management Science, Vol. 52, No. 5, pp. 713–725.

Gantman, N., and O. Shy (May 2, 2007), *Broadcasting Competition and Advertising Quality: A Two-sided Market Approach.* Working paper available at http://ozshy.50webs.com.

Hotelling, H. (1929), *Stability in Competition*, Economic Journal, Vol. 39, No. 153, pp. 41–57.

Interian, Y., et al. (2009), *Ad Quality On TV: Predicting Television Audience Retention*, available at `http://static.googleusercontent.com/external_content/untrusted_dlcp/research.google.com/en//pubs/archive/35368.pdf`.

Kind, H., T. Nilssen, and L. Sørgard (2007), *Competition for Viewers and Advertisers in a TV Oligopoly*, Journal of Media Economics, Vol. 20, No. 3, pp. 211–233.

Liu, Q. and Konstantinos Serfes (2007), *Price Discrimination in Two-Sided Markets*, NET Institute working paper #07–25.

Masson, R., R. Mudambi, and R. Reynolds (1990), *Oligopoly in Advertiser-Supported Media*, Quarterly Review of Economics and Business, Vol. 30, No. 2, pp. 3–16.

Moorthy, S. and Zhao, H. (2000), *Advertising Spending and Perceived Quality*, Marketing Letters, Vol. 11, No. 3, pp. 221–233.

More than half the homes in U.S. have three or more TVs, The Nielsen Company, 20 July 2009.

Nelson, P. (1974), *Advertising as Information*, Journal of Political Economy, Vol. 82, No. 4, pp. 729–54.

Regulating the quantity of advertising on television, Ofcom, 15 December 2011, available at `http://stakeholders.ofcom.org.uk/binaries/broadcast/Advertising_minutage.pdf`.

Rochet, J-C and J. Tirole (2003), *Platform Competition in Two-Sided Markets*, Journal of the European Economic Association, Vol. 1, No. 4, pp. 990–1029.

Rochet, J-C and J. Tirole (2006), *Two-Sided Markets: A Progress Report*, RAND Journal of Economics, Vol. 37, No. 3, pp. 645–667.

Salop, S. (1979), *Monopolistic Competition with Outside Goods*, Bell Journal of Economics, Vol. 10, No. 1, pp. 141–156.

Speck, P. and Elliott, Michael T. (1997), *Predictors of Advertising Avoidance in Print and Broadcast Media*, Journal of Advertising, Vol. 26, No. 3, pp. 61–76.

Spence, M. and Owen, B. (1977), *Television Programming, Monopolistic Competition and Welfare*, Quarterly Journal of Economics, Vol. 91, No. 1, pp. 103–126.

Steiner, P. (1952), *Program patterns and Preferences, and the Workability of Competition in Radio Broadcasting*, Quarterly Journal of Economics, Vol. 66, No. 2, pp. 194–223.

Targeted advertising holds promise, Financial Times, August 31 2009.

Television, Internet and Mobile Usage in the U.S., Three Screen Report Vol. 7 – 4th Quarter 2009, The Nielsen Company, 2010.

Weyl, E. Glen (March 2009), *The Price Theory of Two-Sided Markets*, working paper at Harvard University.

Weyl, E. Glen (2010), *A Price Theory of Multi-sided Platforms*, American Economic Review, Vol. 100, No. 4, pp. 1642–72.

What would you do with an extra hour?, Wall Street Journal, June 23, 2010.

Wilbur, K. C. (2008), *A Two-Sided, Empirical Model of Television Advertising and Viewing Markets*, Marketing Science, Vol. 27, No. 3, pp. 356–378.

World Cup 2010: World Advertising Expenditures, Translucent US $780 Billion, Karawang Business, Information, Tips and Solutions for Business and Finance, June 13, 2010 in Business News.

Wright, D. (1994), *Television Advertising Regulation and Program Quality*, Economic Record, Vol. 70, Issue 211, pp. 361–367.

8 Hours a Day Spent on Screens, Study Finds, The New York Times, March 27, 2009.

2.8 Appendix

2.8.1 The SPNE

Stage III: subscribers' choice

An S-type subscriber located at y will choose to subscribe platform 1 if $v - \gamma \hat{x}_1 (1 - \delta q_1) - s_1 - ty > v - \gamma \hat{x}_2 (1 - \delta q_2) - s_2 - t(1 - y)$ i.e., if $y < \frac{1}{2} + \frac{\gamma(\hat{x}_2(1-\delta q_2) - \hat{x}_1(1-\delta q_1)) + s_2 - s_1}{2t}$ when v is sufficiently high to ensure full participation. Therefore,

$$
\begin{aligned}
\hat{y}_1 (\hat{x}_1, \hat{x}_2, q_1, q_2, s_1, s_2) &= \frac{1}{2} + \frac{\gamma (\hat{x}_2 (1 - \delta q_2) - \hat{x}_1 (1 - \delta q_1)) + s_2 - s_1}{2t}, \\
\hat{y}_2 (\hat{x}_2, \hat{x}_1, q_2, q_1, s_2, s_1) &= \frac{1}{2} - \frac{\gamma (\hat{x}_2 (1 - \delta q_2) - \hat{x}_1 (1 - \delta q_1)) + s_2 - s_1}{2t}.
\end{aligned}
$$

An I-type subscriber located at z will choose to subscribe platform 1 if $v - s_1 - tz > v - s_2 - t(1 - z)$, i.e., if $z < \frac{1}{2} + \frac{s_2 - s_1}{2t}$ when v is sufficiently high to ensure full participation. Therefore,

$$
\begin{aligned}
\hat{z}_1 (s_1, s_2) &= \frac{1}{2} + \frac{s_2 - s_1}{2t}, \\
\hat{z}_2 (s_2, s_1) &= \frac{1}{2} - \frac{s_2 - s_1}{2t}.
\end{aligned}
$$

Stage II: firms' choice

The profit of advertising firm x is defined as follows

$$\Pi(x) \equiv \begin{cases} (\omega - c)\left((1-\lambda)\hat{z}_1 x + \lambda\hat{y}_1\left(x + (1-x)\sqrt{q_x}\right)\right) - \beta q_x - f_1 & \text{if platform 1} \\ (\omega - c)\left((1-\lambda)\hat{z}_2 x + \lambda\hat{y}_2\left(x + (1-x)\sqrt{q_x}\right)\right) - \beta q_x - f_2 & \text{if platform 2} \\ (\omega - c)\left((1-\lambda)x + \lambda\left(x + (1-x)\sqrt{q_x}\right)\right) - \beta q_x - f_1 - f_2 & \text{if platforms 1 and 2} \\ 0 & \text{if no ads} \end{cases}.$$

If a firm chooses to advertise only via platform i, then

$$FOC \quad : \quad \frac{d\Pi(x)}{dq_x} = (\omega - c)\left(\frac{\lambda\hat{y}_i}{2}(1-x)q_x^{-0.5}\right) - \beta = 0 \Leftrightarrow q_x^* = \left(\frac{(\omega - c)\lambda\hat{y}_i(1-x)}{2\beta}\right)^2,$$

$$SOC \quad : \quad \frac{d^2\Pi(x)}{dq_x^2} = -(\omega - c)\frac{\lambda\hat{y}_i}{4q_x^{1.5}}(1-x) < 0 \text{ at } q_x = q_x^*, \text{ for } \hat{y}_i > 0.$$

If a firm chooses to multi-home, then

$$FOC \quad : \quad \frac{d\Pi(x)}{dq_x} = (\omega - c)\frac{\lambda}{2}(1-x)q_x^{-0.5} - \beta = 0 \Leftrightarrow q_x^* = \left(\frac{(\omega - c)\lambda(1-x)}{2\beta}\right)^2,$$

$$SOC \quad : \quad \frac{d^2\Pi(x)}{dq_x^2} = -(\omega - c)\frac{\lambda}{4q_x^{1.5}}(1-x) < 0 \text{ at } q_x = q_x^*.$$

From here onwards we focus our attention on the case where firms choose to multi-home (this assumption will be verified in equilibrium). Plugging the ad quality solution q_x^* in $\Pi(x)$ we get

$$\begin{aligned} \Pi(x) &= (\omega - c)\left(x + \lambda(1-x)\frac{(\omega - c)\lambda(1-x)}{2\beta}\right) - \frac{1}{\beta}\left(\frac{(\omega - c)\lambda(1-x)}{2}\right)^2 - f_1 - f_2 \\ &= (\omega - c)x + \frac{1}{\beta}\left(\frac{(\omega - c)\lambda(1-x)}{2}\right)^2 - f_1 - f_2. \end{aligned}$$

The profit of the marginal advertiser is zero, i.e.,

$$(\omega - c)(1-\hat{x}_i) + \frac{1}{\beta}\left(\frac{(\omega - c)\lambda\hat{x}_i}{2}\right)^2 - (f_i + f_j) = 0, \ i \neq j, \ i, j = 1, 2$$

where $\hat{x}_i \equiv 1 - x$ and x is the marginal advertiser on the interval $[0, 1]$ on which firms are uniformly distributed. Solving the zero-profit condition with respect to $\hat{x}_i$, we get firms'

demand function for advertising slots in platform i[76]

$$\hat{x}_i\left(f_i, f_j\right) = 2\beta \frac{1 - \sqrt{\frac{1}{\beta}\left(\beta - \lambda^2\left(\omega - c\right) + \lambda^2\left(f_i + f_j\right)\right)}}{\left(\omega - c\right)\lambda^2}$$

as long $0 \le \hat{x}_i \le 1$ is satisfied.

Stage I: platforms' choice

Platform i's problem is

$$\max_{f_i, s_i} \Pi_i\left(f_i, s_i\right) \equiv s_i\left(\lambda \hat{y}_i + \left(1 - \lambda\right)\hat{z}_i\right) + f_i \hat{x}_i \text{ subject to}$$

$$\hat{x}_i\left(f_i, f_j\right) = 2\beta \frac{1 - \sqrt{\frac{1}{\beta}\left(\beta - \lambda^2\left(\omega - c\right) + \lambda^2\left(f_i + f_j\right)\right)}}{\left(\omega - c\right)\lambda^2}$$

$$\hat{y}_i\left(\hat{x}_i, \hat{x}_j, q_i, q_j, s_i, s_j\right) = \frac{1}{2} + \frac{\gamma\left(\hat{x}_j\left(1 - \delta q_j\right) - \hat{x}_i\left(1 - \delta q_i\right)\right) + s_j - s_i}{2t}$$

$$\hat{z}_i\left(s_i, s_j\right) = \frac{1}{2} + \frac{s_j - s_i}{2t}$$

where

$$
\begin{aligned}
q_i\left(\hat{x}_i\right) &= \frac{\int_{1-\hat{x}_i}^{1}\left(\frac{\left(\omega - c\right)\lambda\left(1 - x\right)}{2\beta}\right)^2 dx}{\hat{x}_i} \\
&= \frac{-\left(\frac{\left(\omega - c\right)\lambda}{2\beta}\right)^2\left[\left(1 - x\right)^3\right]_{1-\hat{x}_i}^{1}}{3\hat{x}_i} \\
&= \frac{1}{3}\left(\frac{\left(\omega - c\right)\lambda}{2\beta}\right)^2 \hat{x}_i^2.
\end{aligned}
$$

Note that if firms choose to advertise in both platforms, we have $\hat{x}_i = \hat{x}_j$ and, therefore, $q_i = q_j$. This implies that $\hat{y}_i\left(\hat{x}_i, \hat{x}_j, q_i, q_j, s_i, s_j\right) = \frac{1}{2} + \frac{s_j - s_i}{2t}$. The FOCs for platform i's problem is

$$
\begin{cases}
\frac{\partial \Pi_i}{\partial s_i} = 0 \\
\frac{\partial \Pi_i}{\partial f_i} = 0
\end{cases}
\Leftrightarrow
\begin{cases}
\frac{1}{2} + \frac{s_j - 2s_i}{2t} = 0 \\
2\beta \frac{1 - \sqrt{\frac{1}{\beta}\left(\beta - \lambda^2\left(\omega - c\right) + \lambda^2\left(f_i + f_j\right)\right)}}{\left(\omega - c\right)\lambda^2} - \frac{\lambda^2 f_i\left(\frac{1}{\beta}\left(\beta - \lambda^2\left(\omega - c\right) + \lambda^2\left(f_i + f_j\right)\right)\right)^{-0.5}}{\left(\omega - c\right)\lambda^2} = 0
\end{cases}
$$

---

[76] We ruled out the solution $\hat{x}_i = 2\beta \frac{1 + \sqrt{\frac{1}{\beta}\left(\beta - \lambda^2\left(\omega - c\right) + \lambda^2\left(f_1 + f_2\right)\right)}}{\left(\omega - c\right)\lambda^2}$ since, by (36), we have that $\frac{\beta}{\left(\omega - c\right)\lambda^2} > 1$ implying $\hat{x}_i > 1$.

and platform i's best-reply functions are[77]

$$\begin{cases} s_i(s_j) = \frac{t+s_j}{2} \\ f_i(f_j) = \frac{1}{9\lambda^2}\left(6\lambda^2(\omega-c) - 4\beta - 6\lambda^2 f_j + 2\sqrt{\beta\left(4\beta - 3\lambda^2(\omega-c) + 3\lambda^2 f_j\right)}\right) \end{cases}.$$

Note that the reaction functions with respect to advertising fees are downward sloping,

$$\frac{df_i(f_j)}{df_j} = -\frac{1}{3}\frac{8\beta - 6\lambda^2(\omega-c) + 6\lambda^2 f_j - \sqrt{\beta\left(4\beta - 3\lambda^2(\omega-c) + 3\lambda^2 f_j\right)}}{4\beta - 3\lambda^2(\omega-c) + 3\lambda^2 f_j} < 0,$$

since inequalities $8\beta - 6\lambda^2(\omega-c) + 6\lambda^2 f_j - \sqrt{\beta\left(4\beta - 3\lambda^2(\omega-c) + 3\lambda^2 f_j\right)} > 0$ and $4\beta - 3\lambda^2(\omega-c) + 3\lambda^2 f_j > 0$ are ensured by condition (36). This implies that in this case advertising fees are strategic substitutes.

In a symmetric equilibrium, $s_i^* = s_j^*$ and $f_i^* = f_j^*$, thus[78]

$$\begin{cases} s_i^* = t \\ f_i^* = \frac{1}{25\lambda^2}\left(10\lambda^2(\omega-c) - 6\beta + 2\sqrt{\beta\left(9\beta - 5\lambda^2(\omega-c)\right)}\right) \end{cases}.$$

The SOCs are satisfied at (s_i^*, f_i^*),

$$H = \begin{bmatrix} -\frac{1}{t} & 0 \\ 0 & -\frac{\left(4\left(\beta-\lambda^2(\omega-c)\right)+\lambda^2\left(3f_i^*+4f_j^*\right)\right)\beta\sqrt{\frac{1}{\beta}\left(\beta-\lambda^2(\omega-c)+\lambda^2\left(f_i^*+f_j^*\right)\right)}}{2(\omega-c)\left(\beta-\lambda^2(\omega-c)+\lambda^2\left(f_i^*+f_j^*\right)\right)^2} \end{bmatrix}$$

$$|H_1| = -\frac{1}{t} < 0$$

$$|H_2| = \frac{\left(4\left(\beta-\lambda^2(\omega-c)\right)+\lambda^2\left(3f_i^*+4f_j^*\right)\right)\beta\sqrt{\frac{1}{\beta}\left(\beta-\lambda^2(\omega-c)+\lambda^2\left(f_i^*+f_j^*\right)\right)}}{2t(\omega-c)\left(\beta-\lambda^2(\omega-c)+\lambda^2\left(f_i^*+f_j^*\right)\right)^2} > 0$$

where $|H_2| > 0$ is guaranteed by condition (36).

---

[77] We ruled out the solution $f_i = \frac{6\lambda^2(\omega-c)-4\beta-6\lambda^2 f_j - 2\sqrt{\beta(4\beta-3\lambda^2(\omega-c)+3\lambda^2 f_j)}}{9\lambda^2}$ since $\frac{f_i}{\beta} <$ $\frac{6\frac{\lambda^2(\omega-c)}{\beta}-4-2\sqrt{4-3\frac{\lambda^2(\omega-c)}{\beta}}}{9\lambda^2} < 0$ for $0 < \frac{\lambda^2(\omega-c)}{\beta} < 1$, which must hold by assumption in (36).

[78] We ruled out the solution $f_i^* = \frac{10\lambda^2(\omega-c)-6\beta-2\sqrt{\beta(9\beta-5\lambda^2(\omega-c))}}{25\lambda^2}$ since it is negative for $0 < \frac{\lambda^2(\omega-c)}{\beta} < 1$, which must hold by condition (36).

In equilibrium

$$\hat{x}_i^* = 2\beta \frac{1 - \sqrt{\frac{1}{\beta}\left(\frac{13}{25}\beta - \frac{1}{5}\lambda^2\left(\omega - c\right) + \frac{4}{25}\sqrt{\beta\left(9\beta + 5\lambda^2\left(\omega - c\right)\right)}\right)}}{\lambda^2\left(\omega - c\right)},$$

$$\hat{y}_i^* = \frac{1}{2} \text{ and } \hat{z}_i^* = \frac{1}{2}.$$

Note that $0 < \hat{x}_i^* < 1$ must hold by condition (36). Also, according to Lemma 1, in a symmetric equilibrium, firms multi-home or do not advertise at all.

2.8.2 Proofs

Proof of Proposition 1 When broadcasting ads on both platform 1 and 2, the profit of firm x is defined as

$$\Pi\left(x\right) \equiv \left(\omega - c\right)\left(\left(1 - \lambda\right)x + \lambda\left(x + \left(1 - x\right)\sqrt{q_x}\right)\right) - \beta q_x - f_1 - f_2.$$

Therefore

$$FOC \quad : \quad \frac{d\Pi\left(x\right)}{dq_x} = \left(\omega - c\right)\frac{\lambda}{2}\left(1 - x\right)q_x^{-0.5} - \beta = 0 \Leftrightarrow q_x^* = \left(\frac{\left(\omega - c\right)\lambda\left(1 - x\right)}{2\beta}\right)^2,$$

$$SOC \quad : \quad \frac{d^2\Pi\left(x\right)}{dq_x^2} = -\left(\omega - c\right)\frac{\lambda}{4q_x^{1.5}}\left(1 - x\right) < 0 \text{ at } q_x = q_x^*. \ \Box$$

Proof of Lemma 1 This proof consists in showing that firms do not single-home if $f_i^* = f_j^*$ and $D_{x,i}^* = D_{x,j}^*$. A necessary (but not sufficient) condition for firm x to choose to single-home on platform i is

$$\left(\omega - c\right)D_{x,i} - \beta q_x^* - f_i \geq 0.$$

If firm x advertises on platform i, given the one-off cost of making an advertisement of quality q_x^*, the profit from advertising also on platform j will be $\left(\omega - c\right)D_{x,j} - f_j$. Firm x will multi-home if

$$\left(\omega - c\right)D_{x,j} - f_j \geq 0.$$

If

$$\left(\omega - c\right)D_{x,j}^* - f_j^* \geq \left(\omega - c\right)D_{x,i}^* - \beta q_x^* - f_i^* \Leftrightarrow$$
$$\beta q_x^* \geq \left(\omega - c\right)\left(D_{x,i}^* - D_{x,j}^*\right) - \left(f_i^* - f_j^*\right), \tag{43}$$

then, provided $\left(\omega - c\right)D_{x,i}^* - \beta q_x^* - f_i \geq 0$, firms will choose multi-homing. In a symmetric

equilibrium, where $f_i^* = f_j^*$ and firms expect $D_{x,i}^* = D_{x,j}^*$, (43) simplifies to

$$\beta q_x^* \geq (\omega - c) \times 0 - 0 = 0,$$

which is verified for any $q_x^* \geq 0$. Therefore, in a symmetric equilibrium, if it is profitable for firm x to advertise via platform i, then it is also profitable for firm x to advertise via platform j. $\square$

Proof of Proposition 2

$$
\begin{aligned}
q_i(\hat{x}_i) &= \frac{\int_{1-\hat{x}_i}^1 \left(\frac{(\omega-c)\lambda(1-x)}{2\beta}\right)^2 dx}{\hat{x}_i} \\
&= \frac{-\left(\frac{(\omega-c)\lambda}{2\beta}\right)^2 \left[(1-x)^3\right]_{1-\hat{x}_i}^1}{3\hat{x}_i} \\
&= \frac{1}{3}\left(\frac{(\omega-c)\lambda}{2\beta}\right)^2 \hat{x}_i^2. \quad \square
\end{aligned}
$$

Proof of Proposition 3 (i) We start by showing that $\partial \hat{x}_i(f_i, f_j)/\partial \lambda \geq 0$,

$$
\begin{aligned}
\frac{\partial \hat{x}_i(f_i, f_j)}{\partial \lambda} &= 2\frac{\left(\begin{array}{c}(2\beta - \lambda^2(\omega-c) + \lambda^2(f_i+f_j))\sqrt{\beta(\beta - \lambda^2(\omega-c) + \lambda^2(f_i+f_j))}+ \\ -2\beta(\beta - \lambda^2(\omega-c) + \lambda^2(f_i+f_j))\end{array}\right)}{\lambda^3(\omega-c)(\beta - \lambda^2(\omega-c) + \lambda^2(f_i+f_j))} \\
&= \frac{2\beta}{\lambda^3(\omega-c)}\frac{\left(\begin{array}{c}\left(1+1-\frac{\lambda^2(\omega-c)}{\beta} + \frac{\lambda^2(f_i+f_j)}{\beta}\right)\sqrt{1-\frac{\lambda^2(\omega-c)}{\beta} + \frac{\lambda^2(f_i+f_j)}{\beta}}+ \\ -2\left(1-\frac{\lambda^2(\omega-c)}{\beta} + \frac{\lambda^2(f_i+f_j)}{\beta}\right)\end{array}\right)}{1-\frac{\lambda^2(\omega-c)}{\beta} + \frac{\lambda^2(f_i+f_j)}{\beta}} \geq 0,
\end{aligned}
$$

since $\frac{2\beta}{\lambda^3(\omega-c)} > 0$ and $\frac{(1+X)\sqrt{X}-2X}{X} \geq 0$, where $X \equiv 1 - \frac{\lambda^2(\omega-c)}{\beta} + \frac{\lambda^2(f_i+f_j)}{\beta} > 0$ by condition (36).

(ii) Now we show that $\partial f_i^*/\partial \lambda \leq 0$,

$$\frac{\partial f_i^*}{\partial \lambda} = -\frac{2\beta}{25}\frac{9 + 9 - 5\frac{\lambda^2(\omega-c)}{\beta} - 6\sqrt{9 - 5\frac{\lambda^2(\omega-c)}{\beta}}}{\lambda^3\sqrt{9 - 5\frac{\lambda^2(\omega-c)}{\beta}}} \leq 0,$$

since $\frac{2\beta}{25} > 0$ and $\frac{9+X-6\sqrt{X}}{\sqrt{X}} \geq 0$, where $X \equiv 9 - 5\frac{\lambda^2(\omega-c)}{\beta} > 0$ by condition (36).

(iii) Given that $s_i^* = t$, it is straightforward that $\partial s_i^*/\partial \lambda = 0$. $\square$

Proof of Proposition 4 (i) In a symmetric Nash equilibrium, platform i's profit equals

$$\Pi_i^* = f_i^* \hat{x}_i^* + s_i^* D_i^* - K,$$

where $f_i^* = \frac{10\lambda^2(\omega-c)-6\beta+2\sqrt{\beta\left(9\beta-5\lambda^2(\omega-c)\right)}}{25\lambda^2}$, $s_i^* = t$, $\hat{x}_i^* = 2\beta\frac{1-\sqrt{\frac{1}{\beta}\left(\frac{13}{25}\beta-\frac{1}{5}\lambda^2(\omega-c)+\frac{4}{25}\sqrt{\beta\left(9\beta+5\lambda^2(\omega-c)\right)}\right)}}{\lambda^2(\omega-c)}$
and $D_i^* = \frac{1}{2}$. Note that

$$\frac{\partial \hat{x}_i^*}{\partial \lambda} = \frac{2}{5\lambda}\frac{72+30X+26\sqrt{9+5X}-5X\sqrt{9+5X}-10\sqrt{9+5X}\sqrt{13-5X+4\sqrt{9+5X}}}{X\sqrt{9+5X}\sqrt{13-5X+4\sqrt{9+5X}}} > 0,$$

where $0 < X \equiv \frac{\lambda^2(\omega-c)}{\beta} < 1$. We can write that

$$\frac{\partial \Pi_i^*}{\partial \lambda} = \frac{\partial f_i^*}{\partial \lambda}\hat{x}_i^* + f_i^*\frac{\partial \hat{x}_i^*}{\partial \lambda} =$$

$$= \frac{\beta}{\lambda^3}\left(\begin{array}{c} -\frac{4}{25}\frac{18-5X-6\sqrt{9-5X}}{\sqrt{9-5X}}\frac{1-\sqrt{\left(\frac{13}{25}-\frac{1}{5}X+\frac{4}{25}\sqrt{9+5X}\right)}}{X}+ \\ +\frac{2\left(10X+2\sqrt{9-5X}-6\right)\left(72+30X+26\sqrt{9+5X}-5X\sqrt{9+5X}-10\sqrt{9+5X}\sqrt{13-5X+4\sqrt{9+5X}}\right)}{125X\sqrt{9+5X}\sqrt{13-5X+4\sqrt{9+5X}}} \end{array}\right) > 0$$

since $\beta/\lambda^3 > 0$ and the expression in brackets is positive for $0 < X < 1$, where $X \equiv \frac{\lambda^2(\omega-c)}{\beta}$.

(ii) Note that s_i^* does not depend on λ. Therefore, the utility of ad-indifferent subscribers, as described in (34), will not change with λ as well.

(iii) In a symmetric Nash equilibrium, the utility of an ad-sensitive subscriber y is given by

$$U_S^*(y) = \begin{cases} v - \gamma\hat{x}_1^*(1-\delta q_1^*) - s_1^* - ty & \text{if platform 1} \\ v - \gamma\hat{x}_2^*(1-\delta q_2^*) - s_2^* - t(1-y) & \text{if platform 2} \end{cases} \Leftrightarrow$$

$$U_S^*(y) = \begin{cases} v - \gamma\hat{x}_1^*\left(1-\frac{\delta}{3}\left(\frac{(\omega-c)\lambda}{2\beta}\right)^2(\hat{x}_1^*)^2\right) - t - ty & \text{if platform 1} \\ v - \gamma\hat{x}_2^*\left(1-\frac{\delta}{3}\left(\frac{(\omega-c)\lambda}{2\beta}\right)^2(\hat{x}_2^*)^2\right) - t - t(1-y) & \text{if platform 2} \end{cases}$$

where $\hat{x}_i^* = 2\beta\frac{1-\sqrt{\frac{1}{\beta}\left(\frac{13}{25}\beta-\frac{1}{5}\lambda^2(\omega-c)+\frac{4}{25}\sqrt{\beta\left(9\beta+5\lambda^2(\omega-c)\right)}\right)}}{\lambda^2(\omega-c)}$. Therefore,

$$\frac{dU_S^*(y)}{d\lambda} = \frac{\partial U_S^*(y)}{\partial \lambda} + \frac{\partial U_S^*(y)}{\partial \hat{x}_i^*}\frac{d\hat{x}_i^*}{d\lambda},$$

where

$$\frac{\partial U_S^*(y)}{\partial \lambda} = \frac{2\delta}{3}\gamma (\hat{x}_i^*)^3 \left(\frac{(\omega - c)}{2\beta}\right)^2 \lambda > 0, \text{ for } \hat{x}_1^* > 0,$$

$$\frac{d\hat{x}_i^*}{d\lambda} > 0 \text{ as shown in (i)}$$

$$\frac{\partial U_S^*(y)}{\partial \hat{x}_i^*} = -\gamma + \delta\gamma \left(\frac{(\omega - c)\lambda}{2\beta}\right)^2 (\hat{x}_1^*)^2.$$

Note that $\hat{x}_i^*$ is independent of δ and that

$$\frac{dU_S^*(y)}{d\lambda} > 0 \Leftrightarrow \frac{2\delta\gamma (\hat{x}_i^*)^3}{3}\left(\frac{(\omega - c)}{2\beta}\right)^2 \lambda + \left(-\gamma + \delta\gamma \left(\frac{(\omega - c)\lambda}{2\beta}\right)^2 (\hat{x}_1^*)^2\right)\frac{d\hat{x}_i^*}{d\lambda} > 0 \Leftrightarrow$$

$$\delta > \frac{\frac{d\hat{x}_i^*}{d\lambda}}{\left(\frac{2}{3\lambda}(\hat{x}_i^*) + \frac{d\hat{x}_i^*}{d\lambda}\right)\left(\frac{(\omega - c)\lambda}{2\beta}\right)^2 (\hat{x}_1^*)^2} > 0.$$

Otherwise, the utility of ad-sensitive subscribers will decrease with λ.

(iv) In a symmetric Nash equilibrium, advertiser x's profit is

$$\Pi^*(x) \equiv \begin{cases} (\omega - c)(x + \lambda(1 - x)\sqrt{q_x^*}) - \beta q_x^* - f_1^* - f_2^* & \text{if platforms 1 and 2} \\ 0 & \text{if no ads} \end{cases},$$

where q_x^* as defined in (39) and $f_i^* = \frac{10\lambda^2(\omega - c) - 6\beta + 2\sqrt{\beta(9\beta - 5\lambda^2(\omega - c))}}{25\lambda^2}$. Hence, if firm x buys advertising slots both at platform 1 and 2,

$$\frac{d\Pi^*(x)}{d\lambda} = \frac{\partial \Pi^*(x)}{\partial \lambda} + \frac{\partial \Pi^*(x)}{\partial q_x^*}\frac{\partial q_x^*}{\partial \lambda} + \sum_{i=1}^{2}\frac{\partial \Pi^*(x)}{\partial f_i^*}\frac{\partial f_i^*}{\partial \lambda}$$

$$= (\omega - c)(1 - x)\sqrt{q_x^*} - 2\frac{\partial f_i^*}{\partial \lambda} > 0$$

given that $\partial \Pi^*(x)/\partial q_x^* = 0$ by the ad quality FOC and $\partial f_i^*/\partial \lambda < 0$ as shown in the proof of Proposition 3 (ii). $\square$

Proof of Proposition 5 From Proposition 2, the average advertising quality in platform i is

$$q_i(\hat{x}_i) = \frac{1}{3}\left(\frac{(\omega - c)\lambda}{2\beta}\right)^2 \hat{x}_i^2.$$

It is straightforward that

$$\frac{dq_i(\hat{x}_i)}{d\hat{x}_i} = \frac{2}{3}\left(\frac{(\omega - c)\lambda}{2\beta}\right)^2 \hat{x}_i > 0,$$

therefore, at $\hat{x}_i = \bar{x} < \hat{x}_i^*$, a lower advertising cap $\bar{x}$ will cause the average advertising quality in platform i to decrease. $\square$

Proof of Proposition 6 (i) Under an advertising cap, platform i's problem is

$$\max_{f_i, s_i} \Pi_i (f_i, s_i) \equiv s_i (\lambda \hat{y}_i + (1 - \lambda) \hat{z}_i) + f_i \hat{x}_i \text{ subject to}$$

$$\hat{x}_i (f_i, f_j) = 2\beta \frac{1 - \sqrt{\frac{1}{\beta} (\beta - \lambda^2 (\omega - c) + \lambda^2 (f_i + f_j))}}{(\omega - c) \lambda^2}$$

$$\hat{x}_i (f_i, f_j) \leq \bar{x}$$

$$\hat{y}_i (\hat{x}_i, \hat{x}_j, q_i, q_j, s_i, s_j) = \frac{1}{2} + \frac{\gamma (\hat{x}_j (1 - \delta q_j) - \hat{x}_i (1 - \delta q_i)) + s_j - s_i}{2t}$$

$$\hat{z}_i (s_i, s_j) = \frac{1}{2} + \frac{s_j - s_i}{2t}.$$

Hereafter, a variable with a bar on top refers to the model with a cap. In a symmetric equilibrium, where $\bar{f}_i = \bar{f}_j$, we get

$$\bar{x} = \hat{x}_i (\bar{f}_i, \bar{f}_j) \Leftrightarrow \bar{x} = 2\beta \frac{1 - \sqrt{\frac{1}{\beta} (\beta - \lambda^2 (\omega - c) + \lambda^2 (\bar{f}_i + \bar{f}_j))}}{(\omega - c) \lambda^2} \Leftrightarrow$$

$$\bar{f}_i = \frac{\beta}{2\lambda^2} \left(1 - \bar{x} \frac{(\omega - c) \lambda^2}{2\beta} \right)^2 - \frac{\beta}{2\lambda^2} + \frac{\omega - c}{2},$$

while $\bar{s}_i = \bar{s}_j = t$. Hence,

$$\overline{\Pi}_i \equiv \frac{t}{2} + \left(\frac{\beta}{2\lambda^2} \left(1 - \bar{x} \frac{(\omega - c) \lambda^2}{2\beta} \right)^2 - \frac{\beta}{2\lambda^2} + \frac{\omega - c}{2} \right) \bar{x}$$

where

$$\frac{d\overline{\Pi}_i}{d\bar{x}} = \frac{1}{8} (\omega - c) \left(4 - 8\bar{x} + 3\bar{x}^2 \frac{\lambda^2 (\omega - c)}{\beta} \right) > 0,$$

within the ranges $\bar{x} < \frac{1}{3 \frac{\lambda^2 (\omega - c)}{\beta}} \left(4 - 2\sqrt{4 - 3 \frac{\lambda^2 (\omega - c)}{\beta}} \right)$ and $\bar{x} > \frac{1}{3 \frac{\lambda^2 (\omega - c)}{\beta}} \left(4 + 2\sqrt{4 - 3 \frac{\lambda^2 (\omega - c)}{\beta}} \right)$. Since $\hat{x}_i^* < \frac{1}{3 \frac{\lambda^2 (\omega - c)}{\beta}} \left(4 - 2\sqrt{4 - 3 \frac{\lambda^2 (\omega - c)}{\beta}} \right)$ by condition (36), we can conclude that $d\overline{\Pi}_i / d\bar{x} > 0$ at $\bar{x} < \hat{x}_i^*$, i.e., a lower advertising cap $\bar{x}$ will cause platform i's profit to decrease.

(ii) Note that s_i^* does not depend on $\bar{x}$. Therefore, the utility of ad-indifferent subscribers, as described in (34), will not change with $\bar{x}$ as well.

(iii) The utility of an ad-sensitive subscriber y is given by

$$\bar{U}_S\left(y\right) = \begin{cases} v - \gamma\bar{x}\left(1 - \delta\bar{q}_1\right) - \bar{s}_1 - ty & \text{if platform 1} \\ v - \gamma\bar{x}\left(1 - \delta\bar{q}_2\right) - \bar{s}_2 - t\left(1 - y\right) & \text{if platform 2} \end{cases} \Leftrightarrow$$

$$\bar{U}_S\left(y\right) = \begin{cases} v - \gamma\bar{x}\left(1 - \frac{\delta}{3}\left(\frac{(\omega-c)\lambda}{2\beta}\right)^2\bar{x}^2\right) - t - ty & \text{if platform 1} \\ v - \gamma\bar{x}\left(1 - \frac{\delta}{3}\left(\frac{(\omega-c)\lambda}{2\beta}\right)^2\bar{x}^2\right) - t - t\left(1 - y\right) & \text{if platform 2} \end{cases}.$$

Therefore,

$$\frac{d\bar{U}_S\left(y\right)}{d\bar{x}} = -\gamma + \delta\gamma\left(\frac{(\omega-c)\lambda}{2\beta}\right)^2\bar{x}^2.$$

Note that

$$\frac{d\bar{U}_S\left(y\right)}{d\bar{x}} > 0 \Leftrightarrow -\gamma + \delta\gamma\left(\frac{(\omega-c)\lambda}{2\beta}\right)^2\bar{x}^2 > 0 \Leftrightarrow$$

$$\delta > \frac{1}{\left(\frac{(\omega-c)\lambda}{2\beta}\right)^2\bar{x}^2}.$$

Otherwise, the utility of ad-sensitive subscribers will increase with a lower advertising cap $\bar{x}$.

(iv) Note that

$$\frac{d\bar{f}_i}{d\bar{x}} = \frac{1}{4}\left(\omega - c\right)\left(\bar{x}\frac{\lambda^2\left(\omega-c\right)}{\beta} - 2\right) < 0$$

since $0 \leq \bar{x} < x_i^* \leq 1$ and $0 < \frac{\lambda^2(\omega-c)}{\beta} < 1$ by condition (36). This means that a lower advertising cap will incentivize platforms to increase their advertising fees. From (35) it is clear that firms' profits are decreasing in advertising fees. Thus, a lower advertising cap will decrease advertisers' profits due to the higher advertising fees set by platforms. $\square$

3 The No-Surcharge Rule and network effects: a welfare analysis

3.1 Introduction

Motivation. Payment cards[79] have been experiencing fast growth which has drawn attention to some of the contentious features of this industry, namely the No-Surcharge Rule (NSR).[80] The NSR means that a merchant charges at most the same amount for a payment card transaction as for cash.[81] In several countries, the NSR has been under examination by regulatory and competition authorities, central banks and courts. For example, in the US, on October 5th, 2010, Visa and MasterCard reached a settlement with the US DOJ that allows merchants to reward consumers for paying with credit or debit cards that charge the merchant lower fees, while American Express Co. (AmEx) vowed to fight a Government antitrust lawsuit.[82] In early 2010, the Portuguese Government decided to make the NSR mandatory by law claiming consumer protection and that the use of electronic payments is more efficient than cash and thus should be protected. Since April 2013, the UK Government has implemented a ban on payment card surcharges placed by some businesses such as flights, cinemas and hotels, as a means to protect consumers from paying excessive credit and debit card transactions fees.[83] In other countries, such as Australia, the Netherlands and Sweden, the NSR has been abolished (Prager et al.,

---

[79]The payment cards industry includes credit, debit, and prepaid cards. We will focus our attention on payment cards with rewards programs attached. With payment card rewards, consumers have an additional benefit of earning rewards virtually every time they use their payment card, rather than cash. For example, payment card rewards may comprise cashback or points which can be exchanged for travels and other goods and services. Rewards may also be construed as the features coupled with card usage such as theft-insurance for goods purchased with the card, or even dispute-resolution protection by electronic payment networks.

[80]In 2002, transactions done on electronic payment networks in the US exceeded $1.7 trillion (Schwartz and Vincent, 2004). In 2006, payment cards were used in 47 billion transactions for a total of $3.1 trillion (Shy and Wang, 2010). In 2008, debit and prepaid card purchases topped $3.285 trillion (almost a quarter of US GDP). In the UK it is expected that consumer card use will rise by 75% from nearly 10 billion payments in 2012 to nearly 17 billion in 2022 (see `http://www.paymentscouncil.org.uk`).

[81]Although infrequent, there have been cases where card payments were discounted relatively to cash, e.g., in Germany during the transition to the euro. Also in Argentina and Colombia since 2003 Governments have been providing VAT discounts to transactions processed with debit or credit cards.

[82]See *MasterCard, Visa Settle as AmEx Fights U.S. Lawsuit*, available at `http://www.bloomberg.com/news/2010-10-04/mastercard-visa-settle-antitrust-case-as-american-express-fights-lawsuit.html`.

[83]See *Ticket charges: No more "excessive" card surcharges*, available at `http://www.bbc.co.uk/news/business-22042309`.

2009). Critics of the NSR have claimed that it inefficiently encourages the use of more costly forms of payment (e.g., credit cards) over the less costly (e.g., cash), as well as more costly credit cards compared to less costly credit cards, leading to a *"Gresham's Law of Payments"*.

Description of the chapter. The main contribution of this chapter is to highlight potential (in)efficiencies of the NSR, in particular: (i) the improvement of retail price efficiency for cardholders, and (ii) the inefficiency of merchant acceptance. The inefficiency of merchant acceptance will be particularly overwhelming in the presence of strong network effects from merchants to cardholders.[84] We base our analysis on a three-party model with consumers (comprising cardholders and cash payers), merchants and a profit-maximizing EPN.[85]

There is a mass of consumers normalized to one. We assume that a proportion $0 < \alpha < 1$ of consumers are cardholders, i.e., have cash and card as feasible payment instruments, while the remaining $(1 - \alpha)$ are cash payers, i.e., only have cash as a feasible payment instrument. Cardholders care about the extent of merchant acceptance offered by the EPN (network effects from merchants to cardholders).

There is a mass one of profit-maximizing merchants that are local monopolists. A monopoly market structure provides a good first-order approximation to a number of markets where merchants have market power. Such market structure facilitates our welfare analysis of the NSR as we do not have to model potential strategic effects that typically arise in oligopoly markets. We note that the case of perfectly competitive merchants has been covered in previous articles, see for example, Gans and King (2003), and Wright (2003). Previous research has shown that under perfect competition the social surplus does not change regardless of the existence of the NSR. This is because perfectly competitive merchants will separate into those that accept cards and those that do not (see "related literature" below for further details). Merchants bear a fee as a supply cost for card transactions, while not facing explicit costs for cash transactions. The incentives for a merchant to accept card payments are: (i) transactional benefits from card usage (e.g. cash-handling costs' reduction), and (ii) a higher demand (depending on card rewards) if cardholders can pay with card, as opposed to cash.[86] Merchants are heterogeneous in their transactional benefits from card usage.

---

[84] These network effects may be justified by a preference that cardholders have for not carrying cash due to convenience and security reasons.

[85] Our analysis primarily addresses a closed network, but it may also characterize a four-party network if acquirers (issuers) are identical and perfectly competitive, while issuers (acquirers) are identical and collude when setting the fees to cardholders (merchants). One advantage of a three-party model is that we do not need to be concerned with the interchange fee, which, in a four-party setup compensates the issuing bank each time cardholders use the card in a purchase.

[86] Another reason why a merchant may accept card payments is due to the "business stealing" effect. However, since by assumption merchants are local monopolists, such an effect is disregarded from our model.

There is one profit-maximizing EPN setting, per card transaction, a fee to merchants and a reward to cardholders. As the number of cardholders is fixed we assume that there are no fixed fees (e.g., an annual membership fee).

We provide an efficiency justification for the implementation of the NSR. We show that the imposition of the NSR causes the volume of card transactions to increase and the volume of cash transactions to fall. Under the NSR, cardholders will pay a lower price (closer to marginal cost) which improves retail price efficiency for goods sold to cardholders. An important aspect in our model is that a transfer of a unit of a good from a cash payer to a cardholder (choosing to pay with card, rather than cash) implies a reduction in the marginal cost of providing that good. Thus, a transfer of cash transactions into card transactions, as a consequence of the NSR implementation, will be more cost-effective.

Also, we discuss the welfare variations introduced by the NSR in the presence of network effects. If network effects are sufficiently strong, then, with the exception of the EPN, all groups of agents (i.e., cash payers, cardholders and merchants) will be worse off with the NSR implementation. We show that the NSR will be socially undesirable if network effects from merchants to cardholders are sufficiently strong. In our model, the NSR implementation reduces card acceptance. Therefore, if network effects on cardholders are sufficiently strong, the NSR destroys value in the cardholder side of the market. This is the case provided that the network size of card acceptance matters to cardholders and under the NSR fewer merchants will accept payment cards.

Related literature. Formal economic analysis of electronic payment systems was initiated by Baxter (1983) with an analysis of the NaBanCo litigation.[87] The theoretical payment card literature has been growing, especially during the last decade, by addressing the issue of how costs of payment cards are and might be divided among EPNs, merchants and cardholders. The models considered in this literature point out that EPNs may charge fees significantly in excess of their costs to merchants and provide incentives to cardholders to increase card adoption and usage. To a great extent, this literature has not distinguished prepaid cards from debit or credit cards. Usually these models (e.g., Rochet and Tirole (R&T) (2002, 2003), Cabral (2006), Wright (2010)) focus on the adoption and usage of payment cards versus all other payment instruments and have showed that competition levels among merchants and among EPNs, along with consumer and merchant demand elasticities, are relevant factors in determining model outcomes.[88]

EPNs are a type of two-sided markets. The two-sided markets literature has been employed to investigate the structure of fees paid by cardholders and merchants. This

---

[87] See Frankel and Shampine (2006) for a summary on the NaBanCo case (National Bancard Corporation vs. Visa US Inc.).

[88] See Chakravorti (2010) for an excellent review of the growing payment card literature and discussion of the impact of regulatory interventions on card adoption, usage, and welfare.

strand of literature combines the network economics literature, which studies how agents' utility changes with participation of other agents in the network, and the multiproduct firm literature, which investigates how firms choose prices when offering more than one product.

The seminal articles in two-sided markets by R&T (2003, 2006) and Armstrong (2006) investigate the determinants of the price balance between two groups of end-users (e.g., consumers and merchants) when each group exerts a network effect on the other, and both are intermediated by a platform (e.g., an EPN). Some of the discussed determinants of the price balance are: the possibility of multi-homing (access to more than one platform), platform differentiation, presence of same-side externalities, platform compatibility, per-transaction or lump-sum pricing and relative size of cross-group externalities. However, as far as we know, the two-sided markets literature has been silent about the NSR implications on platform fees, profits and welfare, since it assumes that end-users are not allowed to negotiate prices of platform services.

Chakravorti and Roson (2006) compare the welfare level when two networks operate as competitors and as a cartel. One of their findings corroborates the conclusion of R&T (2003) that network competition does not imply, from a social standpoint, a better or worse balance of fees between consumers and merchants. Chakravorti and Roson show that, in general, the welfare gain of a drop in the total network fee more than compensates the deterioration in the efficiency of the fee balance. Moreover, network competition unambiguously increases consumer and merchant surpluses.

Gans and King (2003) show that, under a general four-party model of a payment system, abolishing the NSR is one sufficient condition to reach the neutrality of the interchange fee (IF), i.e., variations in the IF do not lead to changes in consumers' decisions on purchases, consumers' and merchants' adoption decisions nor issuers', acquirers' and merchants' profits. However, Gans and King did not do a welfare analysis.

Wright (2003) undertakes the welfare analysis of the NSR under two-merchant competition extremes: monopoly and Bertrand competition. The author shows that (i) the NSR is socially desirable when merchants operating in a monopoly EPN engage in price discrimination based on payment instruments, and (ii) under Bertrand competition among merchants, the social surplus does not change regardless the existence of the NSR. Wright explains that if merchants are monopolists, the imposition of the NSR prevents them from surcharging excessively, therefore, the NSR increases social surplus. If merchants compete *à la* Bertrand, they pass to consumers the full benefits and costs associated with the payment instruments used to complete the transaction. Hence, under the NSR, competitive merchants only accept cash or only accept card payments, and prices in the goods market are equal to the respective marginal cost net of benefits. Under surcharging, competitive merchants accept both types of payment and price discriminate. However, Wright con-

sidered the total quantity of transactions fixed, as all other literature (with the exception of Schwartz and Vincent (2006)) to our knowledge.

Schwartz and Vincent (2006) investigate the NSR welfare distribution effects among cash users and cardholders when a merchant is a local monopolist. Although the authors allow for elastic demand in the goods market, they assume that consumers are exogenously divided between a group that use only cards (i.e., cardholders cannot use cash), while the others use only cash. They conclude that the NSR harms cash users and merchants, benefits cardholders, and is profitable to the EPN. However, Schwartz and Vincent, considered the existence of only one merchant in their analysis, excluding the study of variations in the merchant acceptance network caused by the NSR implementation, as we do in this chapter.

This chapter builds on the existing literature by providing a model that provides new insights on the welfare effects of the NSR. In particular, we argue that the NSR: (i) promotes retail price efficiency for cardholders, and (ii) generates inefficiency in merchant acceptance. We note that the inefficiency in merchant acceptance is particularly harmful to social welfare in the presence of sufficiently strong network effects. Our model differs from the existing literature, at least, in two main aspects. First, to the best of our knowledge, articles studying the NSR have not considered network effects in the analysis, whilst in our results network effects play an important role. Second, we consider, in a same model, merchant heterogeneity with respect to transactional benefits from card usage, together with endogenous transaction volumes. This allows us to study in a same model the impacts of the NSR both in terms of the merchant acceptance and transaction volumes in the economy.

3.2   The model

In this section we present a model with two payment instruments (cash and cards),[89] characterize each participating agent (consumers, merchants and a proprietary EPN) and describe how they interact in a three-stage game. The elements of our model are as follows.

Consumers.   There is a mass one of consumers split in two types: E-type (i.e., cardholders) and C-type (i.e., cash payers) consumers.[90] E-type consumers hold cards from

---

[89]Cash is the default payment instrument accessible to all consumers and merchants at no cost. As compared to cash, the EPN offers a service for electronic transactions that may yield positive benefits for consumers and merchants.

[90]In the US, the fraction of households who have a credit card has been steady at about 70 to 75 percent during the past two decades, suggesting the maturity of the market (Schuh, Shy and Stavins (2010)). About 24% of US households do not hold cards of any kind (Schwartz and Vincent (2006)).

an EPN and have mass α, where $0 < \alpha < 1$.[91] E-type consumers can purchase goods using either cash or card (electronic transactions). C-type consumers, with mass $1 - \alpha$, do not hold a payment card and purchase goods using only cash.[92]

The demand for a good, per cash payer, is

$$q_c(p_c) \equiv v - p_c, \tag{44}$$

where $v > 0$ is the consumers' maximal willingness to pay for the good itself, q_c is the number of transactions per cash payer and p_c is the price per transaction paid by such a consumer. The consumer surplus per cash payer is then given by $CS_c \equiv \int_{p_c}^{v} q_c(x)\,dx$.

When using a payment card with reward, r, per unit transacted, the demand for a good, per cardholder, is

$$q_e(p_e) \equiv v + r - p_e, \tag{45}$$

where q_e is the number of transactions per cardholder and p_e is the price per transaction paid by such a consumer. The consumer surplus per cardholder paying with card (rather than cash) is then given by $CS_e \equiv \int_{p_e}^{v+r} q_e(x)\,dx$. If a cardholder pays cash to purchase a good, for example, because a merchant accepts only cash, the cardholder's demand in that instance (at that specific merchant) will be given by (44), rather than (45).

We assume that cardholders care about the extent of merchant acceptance offered by the EPN. The larger the merchant acceptance by the EPN, the larger will be the benefit of holding a payment card from it (network effect).[93] Let $D \geq 0$ denote the number of merchants accepting payments processed under the EPN. Let $b_B D$ measure the benefit, in addition to the total consumer surplus, per cardholder, from accessing to an acceptance network of D merchants, where $b_B \geq 0$. Note that b_B is unrelated with the consumers' willingness to pay (demand) for the goods themselves.

These families may be unable to get payment cards or have a preference for anonymity when making a transaction.

[91] Presently, cashless payments represent 92 percent of transactions in France, 89 percent in the UK, 62 percent in Japan and 31 percent in Russia. See MasterCard report (September 2013) at `http://www.mastercardadvisors.com/cashlessjourney`.

[92] Presently, cash payments represent 45 percent of transactions in China and 20 percent in the US. Only 15 percent of payments made in Australia are cash, while in Egypt cash represents 93 percent of transactions. Worldwide, around 85 percent of all retail payment transactions are done with cash. See MasterCard report (September 2013) at `http://www.mastercardadvisors.com/cashlessjourney`.

[93] A reason for this may be the security of paying by card, rather than cash. If cash is stolen, will be hardly recovered. Whereas if a payment card is stolen, the cardholder can cancel the card immediately before fraudulent purchases are made. Moreover, EPNs have sophisticated safeguards (e.g., real time active fraud detection system, chip & PIN authentication) in place to protect cardholders in the event of unauthorized use. Thus, a larger merchant acceptance may lead cardholders to carry a smaller amount of cash. Also, cash may imply costs to get, e.g., time to withdraw cash from banks.

Merchants. There is a mass one of merchants, each of whom supplying an independent good. Merchants are profit-maximizing local monopolists. The marginal cost of producing the goods demanded by consumers is c, where $0 < c < v$. Merchants bear a fee, m, as a supply cost for card transactions, while not facing explicit costs for cash transactions. Merchants are heterogeneous in their transactional benefits from card usage, b_S, where b_S is uniformly distributed on $[0, \bar{b}_S]$. Merchants know their own b_S, which reflects potential savings from cash-handling costs' reduction or from increased security. The EPN knows the distribution of b_S but not the transactional benefit from card usage for individual merchants. We assume that $\bar{b}_S \leq c$, otherwise the net marginal cost of a card transaction $(c - b_S)$ would be negative for merchants with b_S sufficiently high. Additionally, we assume that $\bar{b}_S \leq v - c$ in order to ensure that cash payers are also served when the NSR is in place.

The profit of a merchant with transactional benefit b_S is $(p_e - c + b_S - m)\, q_e\,(p_e)$ from a cardholder paying with card and $(p_c - c)\, q_c\,(p_c)$ from a cash payer, where $q_e\,(p_e)$ and $q_c\,(p_c)$ are defined, respectively, by (45) and (44). Assuming that cardholders will pay with card,[94] for given values of (m, r), merchants' profit will be

$$\Pi\,(b_S) = \begin{cases} (p_c - c)\,(v - p_c) & \text{if only cash} \\ \alpha\,(p_e - c + b_S - m)\,(v + r - p_e) + (1 - \alpha)\,(p_c - c)\,(v - p_c) & \text{if cash and card} \end{cases}.$$

$$(46)$$

We assume that all merchants must accept cash, due to its status as legal tender. A merchant will accept to run transactions under the EPN if and only if, in equilibrium, the option of accepting cash and card is at least as profitable as accepting only cash (default option).

Electronic Payment Network. There is one profit-maximizing EPN.[95] The EPN charges a merchant fee, m, per card transaction, to merchants accepting card payments. Simultaneously, the EPN finances a reward, r, per card transaction, to cardholders. Hence, the EPN will choose $r \leq m$, otherwise would have negative profit. Rewards can be thought of as a negative price of card use. For example, can take the form of direct monetary rebates (cashback), or goods, such as frequent flyer points. A card payment requires the merchant (payee) and the cardholder (payor) to have a common electronic payment network, i.e., the EPN.

Without loss of generality, EPN's marginal cost of servicing a card transaction is

---

[94]This assumption will be verified in equilibrium, both when merchants can surcharge and under the NSR.

[95]Although the monopoly case does not occur in practice, Cabral (2006) suggests that it may "provide a good first-order approximation to the reality of a number of countries". The assumption of a monopolist EPN has been used in a number of research articles, e.g., Schwartz and Vincent (2006).

normalized to zero.[96] We assume, for simplicity, like Schwartz and Vincent (2006), that only linear pricing is feasible for the EPN. Therefore, there are no membership fees (e.g., no cardholder fixed fees) in our model. The EPN solves the following maximization problem,

$$\max_{m,r} \Pi_{EPN}(m,r) = (m-r)T(m,r),$$

where $T(m,r) \equiv \frac{\alpha}{b_S} \int_{b_S^*}^{\bar{b}_S} q_e^*(b_S) \, db_S$ is the total volume of card transactions and let b_S^* denote the transactional benefit at which a merchant is indifferent between accepting cash and card payments or only cash.

A summary of the model's notation is shown in Table 1.

TABLE 1: Notation

α	Mass of cardholders in the population.
v	Consumers' maximal willingness to pay for a unit of a good.
c	Marginal cost of producing a good.
b_B	Cardholder's benefit (network effect) from having an additional merchant accepting card payments.
b_S	Merchant benefit of a card transaction relatively to cash.
$\bar{b}_S$	Maximum value of b_S.
D	Mass of merchants accepting card payments processed under the EPN.
q_e	Demand, per cardholder, for a good.
q_c	Demand, per cash payer, for a good.
p_e	Price of a unit of a good with card payment processed under the EPN.
p_c	Price of a unit of a good when cash is used for payment.
CS_c	Consumer surplus, per cash payer, for a good.
CS_e	Consumer surplus, per cardholder paying with card, for a good.
m	Merchant fee per card transaction processed under the EPN.
r	Cardholder reward per card transaction processed under the EPN.
T	Total number of card transactions processed under the EPN.

Timing of the game. The participating agents interact according to the following sequential game. First, the EPN sets the electronic payment system rule. In particular, a rule is set whereby merchants are either allowed to set a surcharge for a card payment, or not. Second, given the rule, the EPN sets the merchant fee, m, and the cardholder reward, r, per transaction processed. Third, merchants observe (m, r) and decide whether

---

[96]Nonetheless, the EPN may have to support a fixed cost, which must be sufficiently small such that, in equilibrium, profit is non-negative. Otherwise, the EPN would exit the market and there would be no alternative to cash payments.

to accept card transactions or not. Fourth, merchants accepting only cash define the price $p_c(b_S)$, while merchants accepting both cash and card payments define, respectively, the prices $p_c(b_S)$ and $p_e(b_S)$. Fifth, consumers without a payment card will make all purchases with cash, while cardholders will be able to choose between cash or card, if a merchant accepts both payment methods. If a merchant chooses to accept only cash, all consumers (regardless of holding a payment card) may only purchase goods with cash at that merchant. A summary of the timing of the model follows in Table 2.

TABLE 2: The timing of the model

I.	The payment system rule (surcharging or NSR) is set.
II.	The EPN sets the merchant fee, m, and the cardholder reward, r, per transaction processed.
III.	Merchants decide whether to join the EPN, or not.
IV.	Merchants set prices for goods ($p_c(b_S)$ and $p_e(b_S)$ if a merchant accepts both card and cash, or only $p_c(b_S)$ if a merchant accepts only cash).
V.	Consumers decide which payment instrument to use at each merchant, given the set of payment instruments accepted by each merchant.

3.3 The social optimum benchmark

We set out, as benchmark, the first-best solution. Note that merchants' fees and cardholders' rewards are mere transfers from merchants to the EPN and from the EPN to cardholders, respectively. For that reason card fees and rewards are not relevant in the first-best analysis. In the first-best solution, merchants should join the EPN whenever the social benefit arising from accepting payment cards (e.g., cash-handling costs' reduction, increased security in payments) exceeds the social cost of doing so (costs are zero by assumption), i.e., $b_S \geq 0$. Thus, in the first-best solution, all cardholders should use their cards and all merchants should accept card payments, i.e., $D^{opt} = 1$.

The level of production for each good that achieves the maximum total surplus is such that the marginal social benefit equals the marginal social cost. Thus, social optimality involves prices being set at $p_c^{opt} = c$ and $p_e^{opt} = c - b_S$.

Total surplus in the economy is then[97]

$$ TS^{opt} = \alpha \int_0^{\bar{b}_S} \left(\int_{c-b_S}^v (v - x)\, dx \right) db_S + (1-\alpha) \int_0^{\bar{b}_S} \left(\int_c^v (v - x)\, dx \right) db_S. $$

One distortion that will prevent the social optimum from being achieved in the market solution is the extent to which the EPN sets merchant fees above costs. Similarly, monopoly

---

[97]EPN's and merchants' profits are zero. Consumers' surplus (i.e., cash payers' and cardholders' joint consumer surplus) corresponds to total surplus in the economy.

merchants have an incentive to charge prices above the respective marginal cost. As in Wright (2003), the NSR will be powerless to eliminate such distortions.

In the following sections we present the market equilibria when merchants are allowed to surcharge and when the NSR is in place. We compare the levels of consumers' surpluses (cash payers and cardholders), merchants' profits and the EPN's profit obtained in equilibrium under each one of the payment system rules. Technical details and calculations follow in an appendix.

3.4   Equilibrium with merchant surcharging

In this section we set out the equilibrium when merchant surcharging is allowed, i.e., merchants may price discriminate consumers according to the payment instrument. Given the demands in (44) and (45) for a cash payer and a cardholder, respectively, a merchant with transactional benefit b_S will set

$$p_c^* = \frac{v+c}{2} \text{ and} \tag{47}$$

$$p_e^*(b_S) = \frac{v+c+m+r-b_S}{2}. \tag{48}$$

A merchant will accept to run transactions under the EPN if and only if the option of accepting cash and card is at least as profitable as accepting only cash. Only merchants with a transactional benefit b_S sufficiently high will accept card payments, in particular $b_S \geq b_S^* \equiv m - r$. Thus, the mass of merchants accepting card payments will be

$$D\left(m-r\right) = \int_{m-r}^{\bar{b}_S} \frac{1}{\bar{b}_S} db_S = 1 - \frac{m-r}{\bar{b}_S}.$$

Given the prices in (47) and (48), the consumers' surpluses, per merchant, are

$$CS_c^* = \frac{(v-c)^2}{8} \text{ per cash payer, and}$$

$$CS_e^* = \frac{(v-c+b_S-(m-r))^2}{8} \text{ per cardholder.}$$

A cardholder will choose to pay with card, rather than cash, if and only if

$$CS_c^* \leq CS_e^* \Leftrightarrow b_S \geq m - r.$$

Hence, if merchants accept card payments (i.e., merchants with a transactional benefit in the range $m - r \leq b_S \leq \bar{b}_S$), a cardholder will choose to pay with card, rather than cash,

at those merchants' shops. The volume of card transactions will be given by

$$T\left(m-r\right) = \frac{\alpha}{\bar{b}_S} \int_{m-r}^{\bar{b}_S} \frac{v - c - (m-r) + b_S}{2} db_S.$$

Note that the volume of card transactions is a function of $m - r$, but not m or r separately. Therefore, the EPN's problem can be written as

$$\max_{m - r \leq \bar{b}_S} \Pi_{EPN}\left(m - r\right) = \left(m - r\right) T\left(m - r\right).$$

This leads to the following neutrality result.[98]

Proposition 1 *In an equilibrium with merchant surcharging, the volume of card transactions, merchants' profits and EPN's profit depend only on the EPN's margin $m - r$, and not on m and r individually. That is, if (m^*, r^*) maximizes the EPN's profit, then so does any pair (m', r') where $m' - r' = m^* - r^*$.*
Proof All proofs are in an appendix. $\square$

Proposition 1 is in line with the result in standard Microeconomics literature that the effective incidence of a tax does not depend on whether the tax is formally placed on consumers or on merchants. This is because a tax can affect demand or supply in the markets for goods, and hence can change equilibrium prices. These price changes can shift the economic burden of a tax away from its formal incidence. Proposition 1 is in fact a general property of payment systems when merchants can surcharge (see Gans and King (2003)). However, note that, as it will become clear in the next section, in the presence of the NSR, EPN's profits will depend on m and r individually.

When merchants are allowed to surcharge card payments, the solution for EPN's problem is

$$m^* - r^* = \frac{2}{3}\left(v - c + \bar{b}_S\right) - \frac{1}{3}\sqrt{4\left(v - c\right)^2 + \bar{b}_S\left(2\left(v - c\right) + \bar{b}_S\right)}. \qquad (49)$$

It is noteworthy that $0 < m^* - r^* < \bar{b}_S$ ensuring the existence of a mass of merchants accepting card payments. In particular, when merchants are allowed to surcharge, the mass of merchants accepting card payments is

$$D^*\left(m^* - r^*\right) = 1 - \frac{\frac{2}{3}\left(v - c + \bar{b}_S\right) - \frac{1}{3}\sqrt{4\left(v - c\right)^2 + \bar{b}_S\left(2\left(v - c\right) + \bar{b}_S\right)}}{\bar{b}_S}.$$

---

[98]A similar result was noted in Schwartz and Vincent (2006) and can be found in previous research work such as Gans and King (2003).

3.5 Equilibrium under the NSR

Under the NSR, each and every merchant accepting card payments will set a single price to all consumers regardless of payment instrument. Hereafter, a variable with an upper indice NSR refers to the model under the NSR.

Merchants that accept only cash will charge $p_c^{NSR} = \frac{v+c}{2}$, while merchants with transactional benefit b_S that accepts both card and cash payments will set a uniform price at $p_e^{NSR}(b_S) = \frac{v+c+\alpha(m+r-b_S)}{2}$.

A merchant will accept to run transactions under the EPN if and only if the option of accepting cash and card is at least as profitable as accepting only cash. Only merchants with a transactional benefit b_S sufficiently high will accept card payments, in particular, under the NSR,

$$b_S \geq b_S^{NSR}(m,r) \equiv \frac{\sqrt{(v-c)^2 + 4(1-\alpha)r(v-c+r)} - (v-c+2r-\alpha(m+r))}{\alpha}. \quad (50)$$

Proposition 2 *Fix m and r at any given positive level. Compared to the case where merchants are allowed to surcharge, the number of merchants accepting card payments is lower under the NSR.*

Proposition 2 can be shown by comparing the expression $b_S^{NSR}(m,r)$ in (50) against $b_S^*(m,r) \equiv m-r$. We get that

$$b_S^{NSR}(m,r) - b_S^*(m,r) > 0 \Leftrightarrow 4\alpha(1-\alpha)r^2 > 0,$$

thus, $b_S^{NSR}(m,r) > b_S^*(m,r)$ and consequently $D^{NSR}(m,r) < D^*(m,r)$. Under the NSR, the mass of merchants accepting card payments is

$$D^{NSR}(m,r) = 1 - \frac{b_S^{NSR}(m,r)}{\bar{b}_S} =$$

$$= 1 - \frac{\sqrt{(v-c)^2 + 4(1-\alpha)r(v-c+r)} - (v-c+2r-\alpha(m+r))}{\alpha\bar{b}_S}.$$

Merchants accepting card payments are clearly worse off through being constrained in the ability to engage in price discrimination (i.e., under the NSR). Therefore, some merchants that choose to accept card payments when allowed to surcharge, will choose not to do so under the imposition of the NSR. When merchants are allowed to surcharge (i.e., engage in price discrimination), they have greater incentives to accept card payments.

EPN's problem is

$$\max_{m,r} \Pi_{EPN}^{NSR}(m,r) = (m-r)T^{NSR}(m,r) \text{ subject to}$$

$$T^{NSR}(m,r) = \frac{\alpha}{\bar{b}_S} \int_{b_S^{NSR}(m,r)}^{\bar{b}_S} \left(\frac{v-c-\alpha(m+r-b_S)}{2} + r \right) db_S,$$

$$b_S^{NSR}(m,r) \leq \bar{b}_S.$$

The optimal solution for the EPN's problem is

$$
\begin{cases}
m^{NSR} = \frac{1}{3}(v-c) + \frac{5}{6}\bar{b}_S - \frac{1}{6}\sqrt{4(v-c)^2 + \bar{b}_S(2(v-c) + \bar{b}_S)} \\
r^{NSR} = -\frac{1}{3}(v-c) + \frac{1}{6}\bar{b}_S + \frac{1}{6}\sqrt{4(v-c)^2 + \bar{b}_S(2(v-c) + \bar{b}_S)}
\end{cases}, \tag{51}
$$

where $m^{NSR} + r^{NSR} = \bar{b}_S$. It is noteworthy that

$$m^{NSR} - r^{NSR} = \frac{2}{3}(v-c+\bar{b}_S) - \frac{1}{3}\sqrt{4(v-c)^2 + \bar{b}_S(2(v-c) + \bar{b}_S)},$$

which is equal to $m^* - r^*$ as set out in (49), i.e., the EPN's profit margin $m-r$ is equal under the two rules (merchant surcharging and NSR). Also, according to Proposition 1, it may be the case that $m^* = m^{NSR}$ and $r^* = r^{NSR}$.

Proposition 3 *Compared to the equilibrium with merchant surcharging and holding* $(m^*, r^*) = (m^{NSR}, r^{NSR})$ *fixed, if the NSR is imposed, then: (i) the volume of cash transactions will fall and the volume of card transactions will rise, while the total volume of transactions will remain unchanged, per merchant that accepts card payments under the NSR; (ii) the total volume of transactions per merchant that accepts card payments under surcharging but not under the NSR will fall; and (iii) the total volume of transactions in the economy as a whole will fall.*

Under surcharging, equilibrium quantities do not depend on how $m^* - r^*$ is divided between m^* and r^*. In particular, the same outcome would arise if we set $(m^*, r^*) = (m^{NSR}, r^{NSR})$, i.e., the optimal values under the NSR. Merchant surcharging leads to a lower retail price charged to cash payers, i.e., $p_c^* = \frac{v+c}{2} \leq \frac{v+c+(\bar{b}_S - b_S)}{2} = p_e^*(b_S)$. However, imposing the NSR while keeping $(m^*, r^*) = (m^{NSR}, r^{NSR})$ constrains the retail pricing for merchants accepting card payments. In particular, under the NSR, merchants accepting card payments will set a single price equal to $p_e^{NSR}(b_S) = \frac{v+c+\alpha(\bar{b}_S - b_S)}{2}$, where $p_c^* \leq p_e^{NSR}(b_S) \leq p_e^*(b_S)$ since $0 < \alpha < 1$ and $b_S \in [0, \bar{b}_S]$ by assumption.[99] The imposition of a single price causes the volume of cash transactions to fall and the volume

---

[99]The condition $\bar{b}_S \leq v - c$ ensures that the group of cash payers will be served when the NSR is in place. Under this condition, $p_e^{NSR}(b_S)$ is lower than cash payers' maximal willingness to pay for a unit of a good, v.

of card transactions to rise, per merchant accepting card payments. The total volume of transactions per merchant remains the same because of the linearity of demands.

In Proposition 2 we discussed that the number of merchants accepting card payments is lower under the NSR than under merchant surcharging. When surcharging is allowed, retail prices are $(p_c^*, p_e^*(b_S))$ and a cardholder purchases a quantity $q_e^*(b_S) = \frac{v-c-(m^*-r^*)+b_S}{2}$, while a cash payer purchases $q_c^* = \frac{v-c}{2} < q_e^*(b_S)$, since $b_S \geq m^* - r^*$. At merchants that accept card payments under surcharging but not under the NSR, cardholders will have to pay cash when the NSR is in place. This means that, at these merchants, cardholders will purchase a lower quantity under the NSR, while cash payers will purchase the same quantity regardless of the NSR implementation.

In a nutshell, if the NSR is implemented, the total volume of transactions will remain unchanged for some merchants,[100] while it will fall for the remaining merchants.[101] We can conclude, then, that in the economy as a whole the total volume of transactions will fall.

3.6 Welfare analysis

In this section, we investigate the welfare variations that the NSR implies on each group of agents in the absence of network effects (Proposition 4). Two different aspects of economic efficiency related to the NSR are discussed: retail price efficiency and efficiency in merchant acceptance. We provide an efficiency justification for the implementation of the NSR (see below the sub-section on "NSR and retail price efficiency for cardholders"). In particular, as a consequence of the NSR implementation, card transactions shift to more cost-effective merchants. Also, we discuss the welfare variations introduced by the NSR in the presence of network effects. If network effects are sufficiently strong, then, with the exception of the EPN, all groups of agents (i.e., cash payers, cardholders and merchants) will be worse off with the NSR implementation (Proposition 5). We conclude that, in the presence of sufficiently strong network effects from merchants to cardholders, the NSR is socially undesirable (Corollary to Proposition 5).[102]

Proposition 4 *Compared to the equilibrium with merchant surcharging, under the NSR: (i) EPN's profit margin per card transaction $(m - r)$ and profit remain unchanged; (ii) cash payers' transactions and the respective consumer surplus are lower; (iii) if α suf-*

---

[100]In particular, the total volume of transactions, per merchant, will remain unchanged for merchants in the range $b_S \in \left[b_S^{NSR}, \bar{b}_S\right]$, and those in the range $b_S \in [0, b_S^*)$ that accept *only* cash, irrespective of the NSR.

[101]In particular, the total volume of transactions, per merchant, will fall for those in the range $b_S \in \left[b_S^*, b_S^{NSR}\right)$ that accept card payments if and only if are allowed to surcharge.

[102]According to Proposition 1, it may be the case that $m^* = m^{NSR}$ and $r^* = r^{NSR}$. Henceforth, we assume for technical simplicity that $(m^*, r^*) = \left(m^{NSR}, r^{NSR}\right)$ will hold.

ficiently high and $b_B = 0$ (i.e., no network effects), cardholders' transactions and the respective consumer surplus will be higher; and (iv) merchants' profits are lower.

Under the NSR, the optimal merchant fee and cardholder reward $\left(m^{NSR}, r^{NSR}\right)$ are determined in (51). The profit margin, per card transaction, $m^{NSR} - r^{NSR}$, is therefore the same as under merchant surcharging. In Proposition 3 we showed that, as a consequence of the NSR implementation, the volume of card transactions per merchant will rise. However, in Proposition 2 we showed that if the NSR is implemented, the number of merchants accepting card payments will decrease. These two effects will offset each other resulting in an unchanged total volume of card transactions in the economy as a whole. Given that neither the profit margin, per card transaction, nor the total volume of card transactions is altered with the NSR implementation, the EPN's profit will be invariant to the NSR.

If the NSR is implemented, cash users will make a lower volume of transactions with merchants that accept both cash and card payments. This is because these merchants mark up their retail prices for all consumers resulting in cardholders being subsidized by cash payers.[103] By "subsidized" we mean that merchant fees are passed on to all consumers in the form of higher retail prices irrespective of the payment instruments that consumers use. Thus, cash payers must pay higher retail prices to cover merchants' fees associated with the payment cards. Given that these merchant fees are used to finance rewards to cardholders, and since cash payers do not receive rewards, cash payers also finance part of the rewards given to cardholders. As a consequence, cash payers' consumer surplus is lower under the NSR.

Despite the fact that the number of card accepting merchants decreases with the NSR (see Proposition 2), cardholders will make the same volume of card transactions. In other words, if the NSR is implemented, cardholders will concentrate the volume of card transactions in a smaller group of merchants (i.e., the group of merchants that accept card payments under the NSR). As previously discussed, under the NSR, cardholders are "subsidized" by cash payers when shopping at merchants that accept both cash and card payments. Therefore, the volume of card transactions, per card accepting merchant, increases with the NSR implementation. Additionally, given that a higher number of merchants will accept only cash under the NSR, cardholders will make more cash transactions (keeping fixed the volume of cash transactions per merchant, as compared to the surcharging equilibrium). In a nutshell, under the NSR, cardholders will make the same volume of card transactions and more cash transactions, as compared to the surcharging equilibrium.

In terms of cardholders' welfare, on the one hand, cardholders' surplus is higher at

---

[103] The result of welfare transfers from cash payers to cardholders was highlighted in previous research work. See, for example, Gans and King (2003), and Schwartz and Vincent (2006). See Schuh, Shy and Stavins (2010) for an empirical application of this result with US data.

merchants that accept card payments under the NSR. This is due to the "subsidy" effect from cash payers. On the other hand, the reduced merchant acceptance of cards triggered by the NSR implementation (see Proposition 2) makes cardholders' surplus to decrease. Cardholders will be better off with the NSR if the "subsidy" effect dominates the "merchant acceptance" effect. It is noteworthy that if the fraction of cardholders, α, is sufficiently high, the "merchant acceptance" effect will be weakened. That is, the number of merchants rejecting card payments because of the NSR will decrease in α, if α is sufficiently high. Hence, in the absence of network effects and for a sufficiently high α the "subsidy" effect dominates the "merchant acceptance" effect and, consequently, cardholders' surplus increases with the NSR implementation.

Merchants, as a group, are clearly worse off through being constrained in their ability to price discriminate, in particular, those merchants that accept card payments. Nonetheless, merchants with a sufficiently low transaction benefit, i.e., merchants in the range $b_S \in [0, b_S^*)$, choosing to accept only cash regardless of the NSR implementation, will be indifferent to the NSR implementation.

From Proposition 4, we can conclude that the NSR may generate opposite welfare variations on different groups of agents. Thus, it is unclear whether the NSR is welfare enhancing for society as a whole. In order to better understand the impacts of the NSR on social welfare, we discuss below two different aspects of economic efficiency related to the NSR: retail price efficiency and efficiency in merchant acceptance (including network effects).

3.6.1   NSR and retail price efficiency for cardholders

A monopolist at the retail level does not efficiently allocate resources. The negative slope of the demand curve means that the price charged by a monopolist is greater than marginal revenue. As a profit-maximizing merchant that equates marginal revenue with marginal cost, the price charged by a monopolist is greater than its marginal cost. The inequality between price and marginal cost is what makes monopoly inefficient. For retail price efficiency, the price of a good should reflect its marginal cost.

In Proposition 3 (i) we set out that the imposition of the NSR causes the volume of card transactions to increase, the volume of cash transactions to fall, while the total volume of transactions, per card accepting merchant, will remain unchanged. This is because $p_c^* \leq p_e^{NSR}(b_S) \leq p_e^*(b_S)$ and so, under the NSR, cardholders will pay a lower price (closer to marginal cost) which improves retail price efficiency for goods sold to cardholders.

An important aspect in our model is that a transfer of a unit of a good from a cash payer to a cardholder (choosing to pay with card, rather than cash) implies a reduction

in the marginal cost of providing that good by an amount equal to $b_S \in \left[b_S^{NSR}, \bar{b}_S\right]$. Thus, a transfer of cash transactions into card transactions, as a consequence of the NSR implementation, will be more cost-effective. Also, it is noteworthy that the total volume of card transactions in the economy as a whole will be the same regardless of whether the NSR is in place (see proof of Proposition 4 (i)). However, under the NSR, card transactions are in the range of merchants with $b_S \in \left[b_S^{NSR}, \bar{b}_S\right]$, rather than $b_S \in \left[b_S^*, \bar{b}_S\right]$ under merchant surcharging, where $b_S^{NSR} > b_S^*$ (see Proposition 2). As a consequence of the NSR implementation, card transactions shift to more cost-effective merchants and may enhance social welfare.[104]

3.6.2 NSR, network effects and (in)efficiency in merchant acceptance

Efficiency in merchant acceptance involves minimizing the costs of supplying a good. In our model, the costs associated with a card transaction are lower compared to cash. Thus, from the social perspective, it is cost-effective that all merchants accept card payments, i.e., $D^{opt} = 1$. Moreover, due to security concerns and/or cardholders' opportunity cost regarding the time required to withdraw cash from banks, cardholders care about the extent of merchant acceptance offered by the EPN. This network effect from merchants to cardholders can be taken into account by setting $b_B > 0$. Below we set out the welfare effects of the NSR when network effects are sufficiently strong.

Proposition 5 *Compared to the equilibrium with merchant surcharging, in the presence of sufficiently strong network effects, i.e., b_B sufficiently high, with the exception of the EPN, all agents (i.e., cash payers, cardholders and merchants) are worse off with the NSR implementation.*

In Proposition 4 we discussed the reasons why cash payers and merchants are worse off with the NSR implementation. In terms of cardholders' surplus, we concluded that in the absence of network effects and for a sufficiently high α, the "subsidy" effect dominates the "merchant acceptance" effect. Thus, cardholders' surplus increases with the NSR implementation. However, cardholders may strongly prefer an EPN with larger acceptance. Given that the number of merchants accepting card payments decreases with the NSR, a sufficiently strong preference for not carrying cash (i.e., sufficiently strong network effect) will make cardholders worse off under the NSR as compared to the surcharging equilibrium.

---

[104]For example, if $\left(v, c, \alpha, \bar{b}_S, b_B\right) = (100, 10, 0.75, 1, 0)$, then $(m^*, r^*) = \left(m^{NSR}, r^{NSR}\right) = (0.75, 0.25)$ and the total surplus in the economy under the NSR will be higher in 2.12 than the one under merchant surcharging. The expressions for the total surplus in the economy can be found in appendix. In particular, see the sub-sections entitled "Welfare analysis with merchant surcharging" and "Welfare analysis under the NSR".

Corollary to Proposition 5 *Compared to the equilibrium with merchant surcharging, in the presence of sufficiently strong network effects, i.e., b_B sufficiently high, total welfare decreases with the NSR implementation.*

In light of the welfare analysis set out above, we concluded that in the presence of sufficiently strong network effects from merchants to cardholders, then, cash payers, cardholders and merchants are worse off under the NSR (see Proposition 5), while the EPN is indifferent to the NSR implementation (see Proposition 4). Thus, it is straightforward that, if network effects are sufficiently strong, total surplus must decrease with the NSR implementation.

3.7   Conclusions

In this chapter, we built a three-party model with consumers (cash payers and cardholders), merchants and an EPN. We consider in a same model: merchant heterogeneity with respect to transactional benefits of accepting cards, network effects from merchants to cardholders and endogenous transaction volumes. Relative to the existing economic literature on the NSR our chapter makes two contributions.

First, we provide an efficiency justification for the implementation of the NSR. We show that the imposition of the NSR causes the volume of card transactions to increase and the volume of cash transactions to fall. Under the NSR, cardholders will pay a lower price (closer to marginal cost) which improves retail price efficiency for goods sold to cardholders. An important aspect in our model is that a transfer of a unit of a good from a cash payer to a cardholder (choosing to pay with card, rather than cash) implies a reduction in the marginal cost of providing that good. Thus, a transfer of cash transactions into card transactions, as a consequence of the NSR implementation, will be more cost-effective.

Second, we discuss the welfare variations introduced by the NSR in the presence of network effects. If network effects are sufficiently strong, then, with the exception of the EPN, all groups of agents (i.e., cash payers, cardholders and merchants) will be worse off with the NSR implementation. We show that the NSR will be socially undesirable if network effects from merchants to cardholders are sufficiently strong. In our model, the NSR implementation reduces card acceptance. Therefore, if network effects on cardholders are sufficiently strong, the NSR destroys value in the cardholder side of the market. This is the case provided that the network size of card acceptance matters to cardholders and under the NSR fewer merchants will accept payment cards.

In our model we assumed monopolistic merchants. However, different market structures co-exist in practice. An extension of our model is to allow for some industries to

be perfectly competitive, while others to be monopolistic. It is noteworthy that the existing literature (e.g., Gans and King (2003), and Wright (2003)) shows that the NSR is irrelevant in perfectly competitive markets since merchants will separate into those that accept cards and those that do not. Therefore, the results with monopolistic merchants as set out in this chapter should carry over to a more general setting with a combination of monopolies and perfectly competitive markets.

In order to focus on how the NSR affects transaction volumes, per consumer, our analysis abstracted away from an endogenous consumers' choice of the payment instruments. A possible extension of this model would include analyzing the welfare effects of the NSR with such endogenous choices. Other direction would be to consider different merchant market structures (e.g., oligopolies, where the "business stealing" effect may play a role).

3.8 References

Baxter, W. F. (1983), *Bank Interchange of Transactional Paper: Legal and Economic Perspectives*, Journal of Law and Economics, 26, pp. 541–88.

Bedre-Defolie, Özlem, and Emilio Calvano (2013), *Pricing Payment Cards*, American Economic Journal: Microeconomics, 5(3), pp. 206-231.

Benoit, Bertrand (2002) *Defiant C&A reignites debate on German shopping laws*, Financial Times, January 9, p. 2.

Bolt, W. and Sujit Chakravorti (2008), *Consumer choice and merchant acceptance of payment media*, Federal Reserve Bank of Chicago Working Paper No.11.

Bolt, W. and Sujit Chakravorti (2008), *Economics of Payment Cards: A Status Report*, DNB Working Paper 193, Netherlands Central Bank, Research Department.

Cabral, Luís (2006), *Market Power and Efficiency in Card Payment Systems: A Comment on Rochet and Tirole*, Review of Network Economics 5, pp. 15–25.

Carlton, Dennis and Alan S. Frankel (1995), *The antitrust economics of credit card networks*, Antitrust Law Journal, Vol. 63, No. 2, pp. 643–668.

Chakravorti, Sujit (2003), *Theory of Credit Card Networks: A Survey of the Literature*, Review of Network Economics, Vol. 2, No. 2, pp. 1–19.

Chakravorti, Sujit and William R. Emmons (2003), *Who pays for credit cards?*, Journal of Consumer Affairs, Vol. 37, No. 2, pp. 208–230.

Chakravorti, Sujit and R. Roson (2006), *Platform Competition in Two-Sided Markets: The Case of Payment Networks*, Review of Network Economics, 5(1), pp. 118–142.

Chakravorti, Sujit and Ted To (2007), *A theory of credit cards*, International Journal of Industrial Organization, Vol. 25, No. 3, June, pp. 583–595.

Chakravorti, Sujit (2010), *Externalities in Payment Card Networks: Theory and Evidence*, Review of Network Economics, Vol. 9, Issue 2, Article 3.

Economides, N. (November–December 1995), *Commentary on Antitrust Economics of Credit Card Networks*, Federal Reserve Bank of St. Louis Review, pp. 60–63.

Economides, N. (2009), *Competition Policy Issues in the Consumer Payments Industry*, in Robert E. Litan and Martin Neil Baily, eds., *Moving Money: The Future of Consumer Payment*, Brookings Institution.

Economides, N. and David Henriques (October 2011), *To Surcharge or Not To Surcharge? A Two-Sided Market Perspective of the No-Surcharge Rule*, European Central Bank Working Paper No. 1388.

Forden, S. and Peter Eichenbaum (Oct 5, 2010), *MasterCard, Visa Settle as Amex Fights U.S. Lawsuit.* Available at `http://www.bloomberg.com`.

Frankel, Alan (1998), *Monopoly and Competition in the Supply and Exchange of Money*, Antitrust Law Journal, Vol. 66, pp. 313–62.

Frankel, Alan (2006), *Towards a Competitive Card Payments Marketplace*, mimeo, p. 32.

Frankel, Alan and Shampine, Allan (2006), *The Economic Effects of Interchange Fees*, Antitrust Law Journal 73, No. 3, pp. 627–673.

Gans, J. and Stephen King (2003), *The Neutrality of Interchange Fees in Payment Systems*, Topics in Economic Analysis and Policy, Vol. 3, No. 1, article 1.

Hayashi, F. (2006), *A Puzzle of Card Payment Pricing: Why Are Merchants Still Accepting Card Payments?*, Review of Network Economics, Vol. 5, Issue 1.

Hunt, Robert (June 2003), *An Introduction to the Economics of Payment Card Networks*, Working Paper No. 03–10, Federal Reserve Bank of Philadelphia.

Jones, Ashby (September 11, 2009), *New York Restaurant Loses Its Appetite for Cash*, The Wall Street Journal, available at `http://online.wsj.com/article/SB125260838282300453.html`.

Katz, Michael (2001), *Reform of Credit Card Schemes in Australia II: Commissioned Report*, Reserve Bank of Australia.

Kitch, Edmund (1990), *The Framing Hypothesis: Is It Supported by Credit Card Issuer Opposition to a Surcharge on a Cash Price?*, Journal of Law, Economics and Organization, Vol. 6, pp. 217–33.

Lowe, Philip (2005), *Payments System Reform: the Australian Experience*, Remarks prepared for the Federal Reserve Bank of Kansas City Conference on *Interchange Fees in Credit and Debit Card Industries: What role for Public Authorities?*

Monnet, Cyril and William Roberds (November 2006), *Credit and the No-Surcharge Rule*, FRB of Atlanta Working Paper No. 2006–25. Available at `http://ssrn.com/abstract=948717`.

Monnet, Cyril and William Roberds (2008), *Optimal pricing of payment services*, Journal of Monetary Economics, Vol. 55(8), pp. 1428–1440.

Murphy, M. and M. Ott (1977), *Retail credit, credit cards and price discrimination,* Southern Economic Journal, 43 (3), pp. 1303–1312.

Prager, R., Manuszak, M. *et al* (2009), *Interchange Fees and Payment Card Networks: Economics, Industry Developments, and Policy Issues*, Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C.

Rafsanjani, Nazanin (October 11, 2006), *Plastic only: Cafe refuses to accept cash*, Morning Edition, National Public Radio, available at `www.npr.org/templates/story/story.php?storyId=6246139`.

Reserve Bank of Australia (2008a), *Reform of Australia's Payments System: Preliminary Conclusions of the 2007/08 Review.*

Rochet, Jean-Charles and Jean Tirole (2002), *Cooperation Among Competitors: Some Economics of Payment Card Associations*, RAND Journal of Economics, Vol. 33, No. 4.

Rochet, Jean-Charles and Jean Tirole (2003), *Platform Competition in Two-Sided Markets*, Journal of the European Economic Association, 1(4), 990–1029.

Rochet, Jean-Charles and Jean Tirole (2006), *Two-sided markets: A progress report*, RAND Journal of Economics, Vol. 37, No. 3, Autumn, pp. 645–667.

Rysman, M. (2007), *An Empirical Analysis of Payment Card Usage*, Journal of Industrial Economics, Blackwell Publishing, Vol. 55(1), pp. 1–36.

Singh, N. and Xavier Vives (1984), *Price and Quantity Competition in a Differentiated Duopoly*, RAND Journal of Economics, Vol. 15, No. 4, pp. 546–554.

Schwartz, M. and Daniel Vincent (2006), *The No Surcharge Rule and Card User Rebates: Vertical Control by a Payment Network*, Review of Network Economics, Vol.5, Issue 1.

Schuh, S., Shy, O. and Joanna Stavins (2010), *Who Gains and Who Loses from Credit Card Payments? Theory and Calibrations*, available at `http://www.bostonfed.org/economic/ppdp/2010/ppdp1003.pdf`.

Shy, O., and Zhu Wang (2011), *Why Do Card Networks Charge Proportional Fees?*, American Economic Review, 101(4), pp. 1575–1590.

The Nilson Report, May 2008; April 2010, issue 946.

Wright, J. (2003), *Optimal card payment systems*, European Economic Review, 47, pp. 587–612.

Wright, J. (2004), *The determinants of optimal interchange fees in payment systems*, Journal of Industrial Economics, Vol. 52, No. 1, pp. 1–26.

Wright, J. (2010), *Why payment card fees are biased against retailers*, available at `http://profile.nus.edu.sg/fass/ecsjkdw`.

3.9 Appendix

3.9.1 Equilibrium with merchant surcharging

Stage V: consumers' choices

Consumers' surplus can be computed from the demands given in (44) and (45). Consumer surplus, per cash payer, for a good, is

$$CS_c \equiv \int_{p_c}^{v} q_c(x)\, dx = \int_{p_c}^{v} (v - x)\, dx = \frac{1}{2}(v - p_c)^2, \tag{52}$$

while consumer surplus, per cardholder paying with card, for a good, is

$$CS_e \equiv \int_{p_e}^{v+r} q_e(x)\, dx = \int_{p_e}^{v+r} (v + r - x)\, dx = \frac{1}{2}(v + r - p_e)^2. \tag{53}$$

At merchants that accept both cash and card payments, cardholders can choose the payment method to use, i.e., the payment instrument that maximizes their surplus at a given merchant.

Stage IV: merchants set prices

If a consumer pays with cash, then a merchant solves

$$\max_{p_c} (p_c - c)(v - p_c) = \frac{1}{4}(v - c)^2$$

$$FOC : v - 2p_c + c = 0 \Leftrightarrow p_c^* = \frac{v + c}{2},$$

$$SOC : -2 < 0.$$

If a consumer pays with card, then a card accepting merchant solves

$$\max_{p_e} (p_e - c - m + b_S)(v + r - p_e) = \frac{1}{4}(v - c - (m - r) + b_S)^2$$

$$FOC : v - 2p_e + r + c + m - b_S = 0 \Leftrightarrow p_e^*(b_S) = \frac{v + c + m + r - b_S}{2},$$

$$SOC : -2 < 0.$$

Stage III: merchants decide whether to join the EPN

A merchant will accept card payments if and only if

$$\frac{1}{4}(v - c - (m - r) + b_S)^2 \geq \frac{1}{4}(v - c)^2 \Leftrightarrow$$

$$b_S \geq b_S^*(m, r) = m - r.$$

Thus, the mass of merchants accepting card payments will be

$$D(m - r) = \int_{m-r}^{\bar{b}_S} \frac{1}{\bar{b}_S} db_S = 1 - \frac{m - r}{\bar{b}_S}.$$

Note that given the prices $(p_c^*, p_e^*) = \left(\frac{v+c}{2}, \frac{v+r+c+m-b_S}{2}\right)$,

$$CS_c^* = \frac{1}{8}(v - c)^2,$$
$$CS_e^* = \frac{1}{8}(v - c - (m - r) + b_S)^2,$$

and a cardholder will choose to pay with card, rather than cash, if and only if

$$CS_c^* \leq CS_e^* \Leftrightarrow b_S \geq b_S^*(m, r) = m - r.$$

Hence, in equilibrium with merchant surcharging, if a merchant accepts card payments, a cardholder will choose to pay with card, rather than cash.

Stage II: EPN's pricing

131

EPN's problem is

$$\max_{m,r} \Pi_{EPN}(m,r) = (m-r)\,T(m,r) \text{ subject to}$$

$$T(m,r) = \frac{\alpha}{\bar{b}_S}\int_{m-r}^{\bar{b}_S}\frac{v-c-m+r+b_S}{2}db_S$$

$$m-r \le \bar{b}_S,$$

where

$$T(m,r) = \frac{\alpha}{\bar{b}_S}\left[\frac{(v-c-(m-r))\,b_S+\frac{b_S^2}{2}}{2}\right]_{m-r}^{\bar{b}_S} =$$

$$= \frac{\alpha}{2\bar{b}_S}\left[(v-c-(m-r))\,(\bar{b}_S-(m-r))+\frac{\bar{b}_S^2}{2}-\frac{(m-r)^2}{2}\right].$$

Platform's problem can be re-written as

$$\max_{m-r\le\bar{b}_S}\Pi_{EPN}(m-r) = \frac{\alpha}{2\bar{b}_S}(m-r)\left((v-c-(m-r))\,(\bar{b}_S-(m-r))+\frac{\bar{b}_S^2}{2}-\frac{(m-r)^2}{2}\right)$$

$$FOC: \frac{\alpha}{2}\left(2\bar{b}_S(v-c)-4(m-r)\bar{b}_S+\bar{b}_S^2-4(v-c)(m-r)+3(m-r)^2\right)=0,$$

$$SOC: \alpha\left(-2(v-c)-2\bar{b}_S+3(m-r)\right)=-\sqrt{4(v-c)^2+\bar{b}_S(2(v-c)+\bar{b}_S)}<0.$$

The solution is[105]

$$m^*-r^* = \frac{2}{3}(v-c+\bar{b}_S)-\frac{1}{3}\sqrt{4(v-c)^2+\bar{b}_S(2(v-c)+\bar{b}_S)}. \tag{54}$$

Note that

$$0 < m^*-r^* < \bar{b}_S, \text{ given}$$

$$0 < \frac{2}{3}(v-c+\bar{b}_S)-\frac{1}{3}\sqrt{4(v-c)^2+\bar{b}_S(2(v-c)+\bar{b}_S)} < \bar{b}_S,$$

ensuring the existence of a mass of merchants accepting card payments. In particular, the mass of merchants accepting card payments is

$$D^*(m^*-r^*) = 1-\frac{\frac{2}{3}(v-c+\bar{b}_S)-\frac{1}{3}\sqrt{4(v-c)^2+\bar{b}_S(2(v-c)+\bar{b}_S)}}{\bar{b}_S}.$$

---

[105]The candidate solution $m-r = \frac{2}{3}(v-c+\bar{b}_S)+\frac{1}{3}\sqrt{4(v-c)^2+\bar{b}_S(2(v-c)+\bar{b}_S)}$ was ruled out because it does not satisfy the SOC of EPN's problem.

3.9.2 Welfare analysis with merchant surcharging

Consumers' surplus

The consumer surplus of the group of cash users, TCS_c, is

$$TCS_c^* = (1 - \alpha) \frac{1}{2} \left(\frac{v - c}{2} \right)^2.$$

The consumer surplus of the group of cardholders, TCS_e, is

$$TCS_e^* = \alpha \left[\frac{1}{2} \left(\frac{v - c}{2} \right)^2 (1 - D^*(m^* - r^*)) + \frac{1}{\bar{b}_S} \int_{m^*-r^*}^{\bar{b}_S} \frac{1}{2} \left(\frac{v - c - (m^* - r^*) + b_S}{2} \right)^2 db_S \right]$$

$$= \alpha \left(\frac{1}{2} \left(\frac{v - c}{2} \right)^2 \frac{m^* - r^*}{\bar{b}_S} + \frac{1}{8\bar{b}_S} \left[\frac{\left(v - c - (m^* - r^*) + \bar{b}_S \right)^3}{3} - \frac{(v - c)^3}{3} \right] \right),$$

where $m^* - r^*$ is determined in (54).

Merchants' profit

The total profit of the group of merchants accepting only cash is

$$\frac{1}{\bar{b}_S} \int_0^{m^*-r^*} \Pi^*(b_S) \, db_S = \left(\frac{v - c}{2} \right)^2 (1 - D^*(m^* - r^*)) = \left(\frac{v - c}{2} \right)^2 \frac{m^* - r^*}{\bar{b}_S},$$

while the total profit of the group of merchants accepting cash and card payments is

$$\frac{1}{\bar{b}_S} \int_{m^*-r^*}^{\bar{b}_S} \Pi^*(b_S) \, db_S = \frac{\alpha}{\bar{b}_S} \int_{m^*-r^*}^{\bar{b}_S} \left(\frac{v - c - (m^* - r^*) + b_S}{2} \right)^2 db_S + (1 - \alpha) \left(\frac{v - c}{2} \right)^2 D^*(m^* - r^*)$$

$$= \frac{\alpha}{4\bar{b}_S} \left(\frac{\left(v - c - (m^* - r^*) + \bar{b}_S \right)^3}{3} - \frac{(v - c)^3}{3} \right) + (1 - \alpha) \left(\frac{v - c}{2} \right)^2 \left(1 - \frac{m^* - r^*}{\bar{b}_S} \right),$$

where $m^* - r^*$ is determined in (54).

EPN's profit

EPN's profit is

$$\Pi_{EPN}^*(m^* - r^*) = \frac{m^* - r^*}{2\bar{b}_S} \left((v - c - (m^* - r^*)) \left(\bar{b}_S - (m^* - r^*) \right) + \frac{\bar{b}_S^2}{2} - \frac{(m^* - r^*)^2}{2} \right),$$

$$(55)$$

where $m^* - r^*$ is determined in (54).

Total surplus

Total surplus derived from all cash transactions is

$$\frac{3}{2}\left(\frac{v-c}{2}\right)^2\left(1-\alpha+\alpha\frac{m^*-r^*}{\bar{b}_S}\right).$$

Total surplus derived from all card transactions is

$$\frac{\alpha\int_{m^*-r^*}^{\bar{b}_S}\left(\frac{v-(c-b_S)+\frac{v+c+m^*-r^*-b_S}{2}-(c-b_S)}{2}\times\frac{v-c-(m^*-r^*)+b_S}{2}\right)db_S}{\bar{b}_S}=$$

$$=\frac{\alpha}{8\bar{b}_S}\int_{m^*-r^*}^{\bar{b}_S}\left(3\left(v-c+b_S\right)+m^*-r^*\right)\left(v-c+b_S-\left(m^*-r^*\right)\right)db_S=$$

$$=\frac{\alpha\left(\bar{b}_S-\left(m^*-r^*\right)\right)\left(3\left(v-c+\bar{b}_S\right)\left(v-c\right)+\left(v-c-\left(m^*-r^*\right)\right)\left(m^*-r^*\right)+\bar{b}_S^2\right)}{8\bar{b}_S}.$$

Total surplus, TS, in the economy is

$$TS^*=\frac{3}{2}\left(\frac{v-c}{2}\right)^2\left(1-\alpha+\alpha\frac{m^*-r^*}{\bar{b}_S}\right)+$$

$$+\frac{\alpha\left(\bar{b}_S-\left(m^*-r^*\right)\right)\left(3\left(v-c+\bar{b}_S\right)\left(v-c\right)+\left(v-c-\left(m^*-r^*\right)\right)\left(m^*-r^*\right)+\bar{b}_S^2\right)}{8\bar{b}_S},$$

where m^*-r^* is determined in (54).

3.9.3 Equilibrium under the NSR

Stage V: consumers' choices

See section "Equilibrium with merchant surcharging" in appendix.

Stage IV: merchants set prices

A merchant that accepts only cash solves

$$\max_{p_c}\left(p_c^{NSR}-c\right)\left(v-p_c^{NSR}\right)=\frac{1}{4}\left(v-c\right)^2$$

$$FOC:v-2p_c+c=0\Leftrightarrow p_c^{NSR}=\frac{v+c}{2},$$

$$SOC:-2<0.$$

A merchant that accepts both card and cash payments solves[106]

$$\max_{p_e^{NSR}} \Pi_e^{NSR}(b_S) = \alpha \left(p_e^{NSR} - c + b_S - m \right) \left(v + r - p_e^{NSR} \right) + (1-\alpha) \left(p_e^{NSR} - c \right) \left(v - p_e^{NSR} \right)$$

$$FOC : v + c - 2p_e^{NSR} + \alpha (m + r - b_S) = 0 \Leftrightarrow p_e^{NSR}(b_S) = \frac{v + c + \alpha (m + r - b_S)}{2},$$

$$SOC : -2 < 0.$$

Hence, under the NSR, the profit of a merchant accepting both card and cash payments is

$$\Pi_e^{NSR}(b_S) = \alpha \left(\frac{v - c + \alpha (m + r - b_S)}{2} + b_S - m \right) \left(\frac{v - c - \alpha (m + r - b_S)}{2} + r \right) +$$

$$+ (1-\alpha) \left(\frac{v - c + \alpha (m + r - b_S)}{2} \right) \left(\frac{v - c - \alpha (m + r - b_S)}{2} \right).$$

Stage III: merchants decide whether to join the EPN

A merchant will accept card payments if and only if

$$\Pi_e^{NSR}(b_S) \geq \frac{1}{4} (v - c)^2.$$

Under the NSR, the marginal merchant accepting card payments is located at[107]

$$b_S^{NSR}(m, r) = \frac{1}{\alpha} \left(\sqrt{(v - c)^2 + 4(1-\alpha) r (v - c + r)} - (v - c + 2r - \alpha (m + r)) \right).$$

$$(56)$$

Note that

$$b_S^{NSR}(m, r) - b_S^*(m, r) > 0 \Leftrightarrow 4\alpha (1-\alpha) r^2 > 0,$$

thus, $b_S^{NSR}(m, r) > b_S^*(m, r)$ and $D^{NSR}(m, r) < D^*(m, r)$. The mass of merchants accepting card payments is

$$D^{NSR}(m, r) = 1 - \frac{b_S^{NSR}(m, r)}{\bar{b}_S} =$$

$$= 1 - \frac{\sqrt{(v - c)^2 + 4(1-\alpha) r (v - c + r)} - (v - c + 2r - \alpha (m + r))}{\alpha \bar{b}_S}.$$

Stage II: EPN's pricing

EPN's problem is

---

[106]Given that the price is the same regardless of the mean of payment, for $r \geq 0$, all cardholders prefer a card payment.

[107]The candidate solution $b_S^{NSR}(m, r) = \frac{-\left(v - c - \alpha m + (2-\alpha) r + \sqrt{4(1-\alpha) r (v - c + r) + (v - c)^2} \right)}{\alpha}$ was ruled out of the analysis since $d\Pi_e^{NSR}(b_S)/db_S$ is negative when evaluated at that candidate solution.

$$\max_{m,r} \Pi_{EPN}^{NSR}(m,r) = (m-r)T^{NSR}(m,r) \text{ subject to}$$

$$T^{NSR}(m,r) = \frac{\alpha}{\bar{b}_S}\int_{b_S^{NSR}(m,r)}^{\bar{b}_S}\left(\frac{v-c-\alpha(m+r-b_S)}{2}+r\right)db_S,$$

$$\bar{b}_S \geq b_S^{NSR}(m,r)$$

where

$$T^{NSR}(m,r) = \frac{\alpha}{\bar{b}_S}\left[\frac{(v-c)b_S - \alpha\left((m+r)b_S - \frac{b_S^2}{2}\right)}{2}+rb_S\right]_{b_S^{NSR}(m,r)}^{\bar{b}_S}.$$

The FOCs for EPN's problem are

$$\begin{cases} \frac{\partial \Pi_{EPN}^{NSR}(m,r)}{\partial m} = 0 \\ \frac{\partial \Pi_{EPN}^{NSR}(m,r)}{\partial r} = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} \alpha\frac{4r\bar{b}_S - 2c\bar{b}_S + 2v\bar{b}_S + 3m^2\alpha - r^2\alpha + \alpha\bar{b}_S^2 + 4cm - 4cr - 8mr - 4mv + 4rv + 4r^2 + 2mr\alpha - 4m\alpha\bar{b}_S}{4\bar{b}_S} = 0 \\ \alpha\frac{2c\bar{b}_S - 8r\bar{b}_S + 4m\bar{b}_S - 2v\bar{b}_S + m^2\alpha - 3r^2\alpha - \alpha\bar{b}_S^2 - 4cm + 4cr + 8mr + 4mv - 4rv - 4m^2 - 2mr\alpha + 4r\alpha\bar{b}_S}{4\bar{b}_S} = 0 \end{cases}.$$

The optimal solution for EPN's problem is[108]

$$\begin{cases} m^{NSR} = \frac{1}{3}(v-c) + \frac{5}{6}\bar{b}_S - \frac{1}{6}\sqrt{\bar{b}_S\left(2(v-c)+\bar{b}_S\right) + 4(v-c)^2} \\ r^{NSR} = -\frac{1}{3}(v-c) + \frac{1}{6}\bar{b}_S + \frac{1}{6}\sqrt{\bar{b}_S\left(2(v-c)+\bar{b}_S\right) + 4(v-c)^2} \end{cases}, \quad (57)$$

note that $m^{NSR} + r^{NSR} = \bar{b}_S$. The Hessian matrix of the EPN's profit function is

$$H(m,r) = \begin{bmatrix} \alpha\frac{\left(3m+r-2\bar{b}_S\right)\alpha - 2(v-c) - 4r}{2\bar{b}_S} & \alpha\frac{2(v-c)-4(m+r)+2\bar{b}_S+(m-r)\alpha}{2\bar{b}_S} \\ \alpha\frac{2(v-c)-4(m+r)+2\bar{b}_S+(m-r)\alpha}{2\bar{b}_S} & \alpha\frac{4\left(m-\bar{b}_S\right)-2(v-c)-\left(m+3r-2\bar{b}_S\right)\alpha}{2\bar{b}_S} \end{bmatrix}.$$

We can verify that at $(m,r) = \left(m^{NSR}, r^{NSR}\right)$

$$|H_1| = \alpha\frac{\left(3m^{NSR}+r^{NSR}-2\bar{b}_S\right)\alpha - 2(v-c) - 4r^{NSR}}{2\bar{b}_S} =$$

$$-\frac{\alpha}{6}\left(\begin{array}{c} 2(1-\alpha)\left(v-c+\bar{b}_S\right) + \\ +(2+\alpha)\sqrt{\bar{b}_S\left(2(v-c)+\bar{b}_S\right) + 4(v-c)^2} \end{array}\right) < 0,$$

---

[108]Three other candidate solutions can be derived from the FOCs. However, those candidates were ruled out of the analysis given that they do not satisfy the SOC.

and

$$|H_2| = -\frac{\alpha^2 (1-\alpha) \left(12 \left(r^{NSR}\right)^2 + \bar{b}_S^2 + 2 \left(2r^{NSR} - \bar{b}_S\right)(v-c) - 8r^{NSR}\bar{b}_S\right)}{\bar{b}_S^2} =$$

$$-\frac{\alpha^2 (1-\alpha) \left(\begin{array}{c} \bar{b}_S \left(2(v-c)+\bar{b}_S\right) + 4(v-c)^2 + \\ -2 \left((v-c)+\bar{b}_S\right)\sqrt{\bar{b}_S \left(2(v-c)+\bar{b}_S\right) + 4(v-c)^2} \end{array} \right)}{3\bar{b}_S^2} > 0,$$

given that

$$\bar{b}_S \left(2(v-c)+\bar{b}_S\right) + 4(v-c)^2 - 2 \left((v-c)+\bar{b}_S\right)\sqrt{\bar{b}_S \left(2(v-c)+\bar{b}_S\right) + 4(v-c)^2} < 0 \Leftrightarrow$$

$$\sqrt{\bar{b}_S \left(2(v-c)+\bar{b}_S\right) + 4(v-c)^2} < 2 \left((v-c)+\bar{b}_S\right) \Leftrightarrow$$

$$\bar{b}_S \left(2(v-c)+\bar{b}_S\right) + 4(v-c)^2 - 4 \left((v-c)+\bar{b}_S\right)^2 = -3\bar{b}_S \left(2(v-c)+\bar{b}_S\right) < 0.$$

3.9.4 Welfare analysis under the NSR

Consumers' surplus

The consumer surplus of a consumer from shopping at a merchant accepting only cash is

$$\frac{\left(v - p_c^{NSR}\right)^2}{2} = \frac{\left(v - \frac{v+c}{2}\right)^2}{2} = \frac{(v-c)^2}{8}.$$

The consumer surplus of a cash payer from shopping at a merchant with a transactional benefit b_S and accepting card payments is

$$\frac{\left(v - p_e^{NSR}(b_S)\right)^2}{2} = \frac{\left(v - \frac{v+c+\alpha\left(m^{NSR}+r^{NSR}-b_S\right)}{2}\right)^2}{2} = \frac{\left(v - c - \alpha\left(\bar{b}_S - b_S\right)\right)^2}{8}.$$

The consumer surplus of a cardholder from shopping at a merchant with transactional benefit b_S and accepting card payments is

$$\frac{1}{2}\left(v + r^{NSR} - p_e^{NSR}(b_S)\right)^2 = \frac{\left(v + r^{NSR} - \frac{v+c+\alpha\left(m^{NSR}+r^{NSR}-b_S\right)}{2}\right)^2}{2} =$$

$$= \frac{\left(v - c + \bar{b}_S - 3\alpha\left(\bar{b}_S - b_S\right) + \sqrt{\bar{b}_S \left(2(v-c)+\bar{b}_S\right) + 4(v-c)^2}\right)^2}{72}.$$

Merchants' profit

137

The profit of a merchant accepting only cash is $\left(\frac{v-c}{2}\right)^2$, while the profit of a merchant with transactional benefit b_S accepting card payments is

$$
\begin{aligned}
\Pi_e^{NSR}(b_S) \;=\;\; & \alpha \left(\frac{v - c + \alpha \left(\bar{b}_S - b_S\right)}{2} + b_S - m^{NSR} \right) \left(\frac{v - c - \alpha \left(\bar{b}_S - b_S\right)}{2} + r^{NSR} \right) + \\
& + (1 - \alpha) \left(\frac{v - c + \alpha \left(\bar{b}_S - b_S\right)}{2} \right) \left(\frac{v - c - \alpha \left(\bar{b}_S - b_S\right)}{2} \right),
\end{aligned}
$$

where $\left(m^{NSR}, r^{NSR}\right)$ as defined in (57).

EPN's profit

For $\left(m^{NSR}, r^{NSR}\right) = (m^*, r^*)$, the EPN's profit is

$$
\Pi_{EPN}^{NSR}\left(m^{NSR} - r^{NSR}\right) \equiv \left(m^{NSR} - r^{NSR}\right) T^{NSR}\left(m^{NSR}, r^{NSR}\right) = \Pi_{EPN}^*\left(m^* - r^*\right),
$$

given that

$$
m^{NSR} - r^{NSR} = m^* - r^* = \frac{2}{3}\left(v - c + \bar{b}_S\right) - \frac{1}{3}\sqrt{4\left(v - c\right)^2 + \bar{b}_S\left(2\left(v - c\right) + \bar{b}_S\right)} \text{ and}
$$
$$
T^{NSR}\left(m^{NSR}, r^{NSR}\right) = T^*\left(m^*, r^*\right).
$$

Total surplus

Total surplus, TS, in the economy is

$$
\begin{aligned}
TS^{NSR} = \;& \frac{1}{\bar{b}_S} \int_0^{b_S^{NSR}\left(m^{NSR}, r^{NSR}\right)} \frac{(v - c) + \left(\frac{v+c}{2} - c\right)}{2} \times \frac{v - c}{2} db_S + \\
& + \frac{(1 - \alpha) \int_{b_S^{NSR}\left(m^{NSR}, r^{NSR}\right)}^{\bar{b}_S} \frac{v - c + \frac{v + c + \alpha\left(\bar{b}_S - b_S\right)}{2} - c}{2} \times \frac{v - c - \alpha\left(\bar{b}_S - b_S\right)}{2} db_S}{\bar{b}_S} + \\
& + \frac{\alpha \int_{b_S^{NSR}\left(m^{NSR}, r^{NSR}\right)}^{\bar{b}_S} \frac{(v - (c - b_S)) + \frac{v + c + \alpha\left(\bar{b}_S - b_S\right)}{2} - r^{NSR} - (c - b_S)}{2} \times \left(\frac{v - c - \alpha\left(\bar{b}_S - b_S\right)}{2} + r^{NSR} \right) db_S}{\bar{b}_S}.
\end{aligned}
$$

3.9.5   Proofs

Proof of Proposition 1 EPN's problem is

$$\max_{m,r} \Pi_{EPN}(m,r) = (m-r)\, T(m,r) \text{ subject to}$$

$$T(m,r) = \frac{\alpha}{\bar{b}_S} \int_{m-r}^{\bar{b}_S} \frac{v - c - (m-r) + b_S}{2}\, db_S$$

$$\bar{b}_S \geq m - r,$$

where

$$T(m,r) = \frac{\alpha}{\bar{b}_S} \left[\frac{(v - c - (m-r))\, b_S + \frac{b_S^2}{2}}{2} \right]_{m-r}^{\bar{b}_S} =$$

$$= \frac{\alpha}{2\bar{b}_S} \left[(v - c - (m-r))\, (\bar{b}_S - (m-r)) + \frac{\bar{b}_S^2}{2} - \frac{(m-r)^2}{2} \right].$$

Replacing $m - r \equiv X$, EPN's problem can be re-written as a single-variable maximization problem

$$\max_{X \leq \bar{b}_S} \Pi_{EPN}(X) = \frac{\alpha X}{2\bar{b}_S} \left((v - c - X)\, (\bar{b}_S - X) + \frac{\bar{b}_S^2}{2} - \frac{X^2}{2} \right).$$

Let X^* denote the solution for the EPN's problem above. If $m^* - r^* = X^*$ and $m' - r' = m^* - r^*$, then it follows that $m' - r' = X^*$. $\square$

Proof of Proposition 2 We want to show that $D^{NSR}(m,r) < D^*(m,r)$, where $D^{NSR}(m,r) \equiv 1 - \frac{b_S^{NSR}(m,r)}{\bar{b}_S}$ and $D^*(m,r) \equiv 1 - \frac{b_S^*(m,r)}{\bar{b}_S}$. Thus, $D^{NSR}(m,r) < D^*(m,r) \Leftrightarrow b_S^{NSR}(m,r) > b_S^*(m,r)$, where $b_S^{NSR}(m,r) \equiv \frac{\sqrt{(v-c)^2 + 4(1-\alpha)r(v-c+r)} - (v-c+2r-\alpha(m+r))}{\alpha}$ and $b_S^*(m,r) \equiv m - r$. We can show that $b_S^{NSR}(m,r) > b_S^*(m,r)$ since

$$\frac{1}{\alpha} \left(\sqrt{(v-c)^2 + 4(1-\alpha)r(v-c+r)} - (v - c + 2r - \alpha(m+r)) \right) > m - r \Leftrightarrow$$

$$\sqrt{(v-c)^2 + 4(1-\alpha)r(v-c+r)} - (v - c + 2r - \alpha(m+r)) > (m-r)\alpha \Leftrightarrow$$

$$\sqrt{(v-c)^2 + 4(1-\alpha)r(v-c+r)} > v - c + 2r(1-\alpha) \Leftrightarrow$$

$$(v-c)^2 + 4(1-\alpha)r(v-c+r) - (v - c + 2r(1-\alpha))^2 > 0 \Leftrightarrow$$

$$4r^2 \alpha(1-\alpha) > 0,$$

since $0 < \alpha < 1$ and $r > 0$ by assumption. $\square$

Proof of Proposition 3 (i) Merchants with transactional benefits in the range $\left[b_S^{NSR}, \bar{b}_S \right]$ will accept card payments under the NSR. The volume of cash transactions, per merchant, in the range $\left[b_S^{NSR}, \bar{b}_S \right]$, is $(1-\alpha)\, q_c^* = (1-\alpha)\, \frac{v-c}{2}$ under surcharging, and $(1-\alpha)\, q_c^{NSR} =$

$(1 - \alpha) \frac{v - c - \alpha(\bar{b}_S - b_S)}{2}$ under the NSR. It is straightforward that $(1 - \alpha) q_c^* \geq (1 - \alpha) q_c^{NSR}$ given that $0 \leq b_S \leq \bar{b}_S$.

The volume of card transactions, per merchant, in the range $[b_S^{NSR}, \bar{b}_S]$, is $\alpha q_e^* = \alpha \frac{v - c - (m^* - r^*) + b_S}{2}$ under surcharging, and $\alpha q_e^{NSR} = \alpha \left(\frac{v - c - \alpha(m^{NSR} + r^{NSR} - b_S)}{2} + r^{NSR} \right)$ under the NSR. We can show that for $(m^*, r^*) = (m^{NSR}, r^{NSR})$, $q_e^* \leq q_e^{NSR}$ since $\frac{v - c - \alpha(m^{NSR} + r^{NSR} - b_S)}{2} + r^{NSR} \geq \frac{v - c + r^{NSR} - m^{NSR} + b_S}{2} \Leftrightarrow m^{NSR} + r^{NSR} \geq b_S \Leftrightarrow \bar{b}_S \geq b_S$ which is true by assumption.

The total volume of transactions, per merchant, in the range $[b_S^{NSR}, \bar{b}_S]$, is

$$Q^* (b_S) \equiv \alpha q_e^* + (1 - \alpha) q_c^* = \frac{v - c - \alpha (m^* - r^* - b_S)}{2}$$

under surcharging, and

$$Q^{NSR} (b_S) \equiv \alpha q_e^{NSR} + (1 - \alpha) q_c^{NSR} = \frac{v - c - \alpha (m^{NSR} - r^{NSR} - b_S)}{2}$$

if the NSR is in place. Hence, if $(m^*, r^*) = (m^{NSR}, r^{NSR})$, it is clear that $Q^* = Q^{NSR}$.

(ii) Merchants with transactional benefits in the range $[b_S^*, b_S^{NSR})$ will accept card payments if surcharging is allowed, otherwise, under the NSR, a merchant will accept only cash. The total volume of transactions, per merchant, in the range $[b_S^*, b_S^{NSR})$, is

$$Q^* = \frac{v - c - \alpha (m^* - r^* - b_S)}{2} \text{ under surcharging, and}$$
$$Q^{NSR} = \frac{v - c}{2}, \text{ under the NSR.}$$

Hence, if $(m^*, r^*) = (m^{NSR}, r^{NSR})$, using the fact that $m^{NSR} - r^{NSR} = \bar{b}_S$ and given that $0 \leq b_S \leq \bar{b}_S$ by assumption, it is clear that $Q^* \leq Q^{NSR}$.

(iii) The total volume of transactions will remain unchanged for merchants in the range $[0, b_S^*)$ as these merchants will always choose to accept only cash irrespective of the NSR. If the NSR is implemented, as shown in part (i) of this proposition, the total volume of transactions will remain unchanged for merchants in the range $[b_S^{NSR}, \bar{b}_S]$. Finally, if the NSR is implemented, as shown in part (ii) of this proposition, the total volume of transactions will fall for merchants in the range $[b_S^*, b_S^{NSR})$. We can conclude, then, that in the economy as a whole the total volume of transactions will fall with the NSR implementation. $\square$

Proof of Proposition 4 (i) The solution to the EPN's problem, when merchants are allowed to surcharge, is given in (54),

$$m^* - r^* = \frac{2}{3} (v - c + \bar{b}_S) - \frac{1}{3} \sqrt{4 (v - c)^2 + \bar{b}_S (2 (v - c) + \bar{b}_S)}.$$

The solution to the EPN's problem, when merchants are under the NSR, is given in (57),

$$
\begin{cases}
m^{NSR} = \frac{1}{3}\left(v - c\right) + \frac{5}{6}\bar{b}_S - \frac{1}{6}\sqrt{\bar{b}_S\left(2\left(v - c\right) + \bar{b}_S\right) + 4\left(v - c\right)^2} \\
r^{NSR} = -\frac{1}{3}\left(v - c\right) + \frac{1}{6}\bar{b}_S + \frac{1}{6}\sqrt{\bar{b}_S\left(2\left(v - c\right) + \bar{b}_S\right) + 4\left(v - c\right)^2}
\end{cases},
$$

thus, it is clear that

$$
m^{NSR} - r^{NSR} = m^* - r^*.
$$

Now, note that when merchants are allowed to surcharge, the volume of card transactions is given by

$$
T^*\left(m^*, r^*\right) = \frac{\alpha}{\bar{b}_S}\int_{m^*-r^*}^{\bar{b}_S} \frac{v - c - \left(m^* - r^*\right) + b_S}{2}\, db_S,
$$

while under the NSR, the volume of card transactions is

$$
T^{NSR}\left(m^{NSR}, r^{NSR}\right) = \frac{\alpha}{\bar{b}_S}\int_{b_S^{NSR}(m^{NSR}, r^{NSR})}^{\bar{b}_S}\left(\frac{v - c - \alpha\left(m^{NSR} + r^{NSR} - b_S\right)}{2} + r^{NSR}\right) db_S,
$$

where $b_S^{NSR}\left(m^{NSR}, r^{NSR}\right)$ is defined in (56). For $m^* = m^{NSR} = m$ and $r^* = r^{NSR} = r$,

$$
T^*\left(m, r\right) - T^{NSR}\left(m, r\right) = \frac{\alpha\left(1 - \alpha\right)\left(\bar{b}_S - \left(m + r\right)\right)^2}{4\bar{b}_S} = 0,
$$

given that, in equilibrium, $m + r = m^{NSR} + r^{NSR} = \bar{b}_S$. Thus, the EPN's profit must be the same regardless of the NSR implementation.

(ii) If merchants are allowed to surcharge, a cash payer will buy $q_c^* = \frac{v-c}{2}$ from each and every merchant. Under the NSR, a cash payer will buy $q_c^* = \frac{v-c}{2}$ from each merchant that accepts only cash, while buying $q_c^{NSR}\left(b_S\right) = \frac{v - c - \alpha\left(\bar{b}_S - b_S\right)}{2}$ from merchants accepting both cash and card payments. Thus, the aggregate volume of transactions made by a cash payer must be lower under the NSR, given that $q_c^{NSR}\left(b_S\right) < q_c^*$ at merchants with transactional benefits $b_S < \bar{b}_S$ accepting card payments.

If merchants are allowed to surcharge, the consumer surplus of a cash user is $\left(v - c\right)^2/8$ at each and every merchant. Under the NSR, the consumer surplus of a cash user is $\left(v - c\right)^2/8$ at merchants accepting only cash, while being $\left(v - c - \alpha\left(\bar{b}_S - b_S\right)\right)^2/8$ at merchants accepting cash and card payments. Thus, the consumer surplus of a cash payer must be lower under the NSR, given that $\left(v - c - \alpha\left(\bar{b}_S - b_S\right)\right)^2/8 < \left(v - c\right)^2/8$ at merchants with transactional benefits $b_S < \bar{b}_S$ accepting card payments.

(iii) If merchants are allowed to surcharge, the total volume of transactions per card-holder is

$$
\frac{v - c}{2}\frac{m^* - r^*}{\bar{b}_S} + \frac{1}{\bar{b}_S}\int_{m^*-r^*}^{\bar{b}_S} \frac{v - c + r - m + b_S}{2}\, db_S, \tag{58}
$$

and the consumer surplus generated by such volume of transactions is

$$\frac{(v-c)^2}{8}\frac{m^*-r^*}{\bar{b}_S}+\frac{1}{\bar{b}_S}\int_{m^*-r^*}^{\bar{b}_S}\frac{1}{2}\left(\frac{v-c-(m-r)+b_S}{2}\right)^2 db_S. \tag{59}$$

Under the NSR, the total volume of transactions per cardholder is

$$\frac{v-c}{2}\frac{b_S^{NSR}\left(m^{NSR},r^{NSR}\right)}{\bar{b}_S}+\frac{1}{\bar{b}_S}\int_{b_S^{NSR}(m^{NSR},r^{NSR})}^{\bar{b}_S}\left(\frac{v-c-\alpha\left(\bar{b}_S-b_S\right)}{2}+r^{NSR}\right) db_S, \tag{60}$$

and the consumer surplus generated by such volume of transactions is

$$\frac{(v-c)^2}{8}\frac{b_S^{NSR}\left(m^{NSR},r^{NSR}\right)}{\bar{b}_S}+\frac{1}{\bar{b}_S}\int_{b_S^{NSR}(m^{NSR},r^{NSR})}^{\bar{b}_S}\frac{1}{2}\left(\frac{v-c-\alpha\left(\bar{b}_S-b_S\right)}{2}+r^{NSR}\right)^2 db_S, \tag{61}$$

where $b_S^{NSR}\left(m^{NSR},r^{NSR}\right)$ is defined in (56).

We can show that the expression in (58) is smaller than the one in (60). Using the fact that $m^{NSR}+r^{NSR}=\bar{b}_S$, the difference between (60) and (58) can be written as

$$\frac{(v-c)\sqrt{4r\left(1-\alpha\right)\left(v-c+r\right)+\left(v-c\right)^2}-2r\left(1-\alpha\right)\left(v-c\right)-\left(v-c\right)^2}{2\alpha\bar{b}_S}>0$$

$$\Leftrightarrow 4\alpha\left(1-\alpha\right)r^2\left(v-c\right)^2>0.$$

We can show that if α is sufficiently high, the expression in (59) will be smaller than the one in (61). Using the fact that $m^{NSR}+r^{NSR}=\bar{b}_S$, the difference between (61) and (59), denoted by ΔTCS_e below, can be written as

$$\Delta TCS_e=\frac{\left(\begin{array}{c}\left(12\left(v-c\right)+8r\right)\left(1-\alpha\right)r^2-2\left(v-c\right)^3+\\+\left(2\left(v-c\right)^2-4r\left(1-\alpha\right)\left(v-c+r\right)\right)\sqrt{4r\left(1-\alpha\right)\left(v-c+r\right)+\left(v-c\right)^2}\end{array}\right)}{24\alpha\bar{b}_S}, \tag{62}$$

where $\lim_{\alpha\to1}\Delta TCS_e=0$, and $\frac{d(\Delta TCS_e)}{d\alpha}=-4r^2\frac{(3(v-c)+2r)\sqrt{4r(1-\alpha)(v-c+r)+(v-c)^2}-6(v-c+r)^2(1-\alpha)}{\sqrt{4r(1-\alpha)(v-c+r)+(v-c)^2}}$ <0, if α sufficiently high. Thus, if α sufficiently high, $\Delta TCS_e>0$.

(iv) Merchants with transactional benefits in the range $[0,b_S^*)$ will only accept cash irrespective of the NSR implementation. Thus, the profit for these merchants must be the same, regardless of the NSR.

Merchants with transactional benefits in the range $\left[b_S^*,b_S^{NSR}\right)$ will accept card payments if surcharging is allowed, otherwise, under the NSR, the merchant will accept only

cash. Within this range, a merchant's profit will be

$$\Pi\left(b_S\right) = \begin{cases} \frac{1}{4}\left(v-c\right)^2 & \text{if NSR} \\ \alpha\left(\frac{v-c+b_S-\left(m^*-r^*\right)}{2}\right)^2 + \left(1-\alpha\right)\left(\frac{v-c}{2}\right)^2 & \text{if surcharging allowed} \end{cases}.$$

We can show that $\alpha\left(\frac{v-c+b_S-\left(m^*-r^*\right)}{2}\right)^2 + \left(1-\alpha\right)\left(\frac{v-c}{2}\right)^2 > \frac{1}{4}\left(v-c\right)^2$, because $\frac{v-c+b_S-\left(m^*-r^*\right)}{2} > \frac{v-c}{2} \Leftrightarrow b_S > b_S^* \equiv m^* - r^*$. Hence, merchants in the range $\left[b_S^*, b_S^{NSR}\right]$ will be worse off with the NSR implementation.

Merchants with transactional benefits in the range $\left[b_S^{NSR}, \bar{b}_S\right]$ will accept card payments regardless of the NSR implementation. Within this range, a merchant's profit will be

$$\Pi^{NSR}\left(b_S\right) = \alpha\left(\frac{v-c+\alpha\left(\bar{b}_S-b_S\right)}{2} + b_S - m^{NSR}\right)\left(\frac{v-c-\alpha\left(\bar{b}_S-b_S\right)}{2} + r^{NSR}\right) + \\ + \left(1-\alpha\right)\left(\frac{v-c+\alpha\left(\bar{b}_S-b_S\right)}{2}\right)\left(\frac{v-c-\alpha\left(\bar{b}_S-b_S\right)}{2}\right)$$

if NSR is in place, and

$$\Pi^*\left(b_S\right) = \alpha\left(\frac{v-c+b_S-\left(m^*-r^*\right)}{2}\right)^2 + \left(1-\alpha\right)\left(\frac{v-c}{2}\right)^2$$

if merchant surcharging is allowed. We can show that $\Pi^{NSR}\left(b_S\right) \leq \Pi^*\left(b_S\right)$, for $b_S \in \left[b_S^{NSR}, \bar{b}_S\right]$. In particular, using the fact that $m^{NSR} + r^{NSR} = \bar{b}_S$, we get

$$\Pi^{NSR}\left(b_S\right) - \Pi^*\left(b_S\right) = -\frac{1}{4}\alpha\left(1-\alpha\right)\left(\bar{b}_S-b_S\right)^2 \leq 0.$$

Hence, merchants in the range $\left[b_S^{NSR}, \bar{b}_S\right]$ will be worse off with the NSR implementation. $\square$

Proof of Proposition 5 In Proposition 4 (i) it is shown that the EPN's profit remains unchanged to the NSR implementation, thus, the EPN will not be worse off with the NSR implementation. In Proposition 4 (ii) and (iv), it is shown, respectively, that cash payers' surplus and merchants' profits are lower under the NSR.

In the presence of network effects from merchants to cardholders, the variation in cardholders' surplus from an equilibrium with surcharging to an equilibrium with the NSR is given by

$$\Delta TCS_e + b_B\left(D^{NSR}\left(m^{NSR}, r^{NSR}\right) - D^*\left(m^*, r^*\right)\right),$$

where ΔTCS_e is defined in (62), and $D^{NSR}\left(m^{NSR}, r^{NSR}\right) - D^*\left(m^*, r^*\right) < 0$ for $\left(m^*, r^*\right) =$

$\left(m^{NSR}, r^{NSR}\right)$ as discussed in Proposition 2. Therefore, the variation in cardholders' surplus will be negative if

$$\Delta TCS_e + b_B \left(D^{NSR}\left(m^{NSR}, r^{NSR}\right) - D^*\left(m^*, r^*\right)\right) < 0 \Leftrightarrow$$

$$b_B > \frac{\Delta TCS_e}{-\left(D^{NSR}\left(m^{NSR}, r^{NSR}\right) - D^*\left(m^*, r^*\right)\right)}. \ \square$$

Proof of Corollary to Proposition 5 In light of the proof of Proposition 5, we can conclude that in the presence of sufficiently strong network effects from merchants to cardholders, then, cash payers, cardholders and merchants are worse off under the NSR, while the EPN is indifferent to the NSR implementation (see Proposition 4). Thus, it is straightforward that, if network effects are sufficiently strong, total surplus must decrease with the NSR implementation. $\square$