Essays on the Regulation of the Telecommunications Industry

by

Jan Yngve Sand

A dissertation submitted for the degree of dr.oecon.

Acknowledgements

There are a number of people that have contributed to this thesis at various stages of the process and in a variety of ways, and I wish to express my gratitude to them. First of all, I am grateful to my supervisor Kåre Petter Hagen for his encouragement and detailed comments and suggestions, without which this thesis would have been of inferior quality. I would also like to thank the other two members of the supervising committee Derek Clark and Nils-Henrik von der Fehr. They have both contributed with very valuable comments to this thesis. In addition, Petter Osmundsen has provided me with valuable advice and comments throughout the process.

Furthermore, I would like to thank colleagues and friends at both the Institute of Economics (NHH) and at the Department of Economics and Management (University of Tromso) for support and encouragement. In particular, I would like to thank Øystein Foros who has been a constant source of ideas through our collaborations, and always available for discussions about topics related to the telecommunications industry (some of which appear in this thesis). Discussions with various people in Telenor (in particular Bjørn Hansen) have been beneficial for my understanding of the industry, and this is highly appreciated.

In the process of writing this thesis I have, on two occasions, been given the opportunity to spend longer periods overseas. In 1998, I spent a year in Toulouse at the Université de Toulouse 1, Gremaq, which was an exciting year. I worked under the supervision of Professor Patrick Rey, who, in addition to several of the other scholars there, provided me with many suggestions for improvement on several of my papers. In the spring of 2001, I was at the Department of Economics and Related Studies at the University of York, during which I had several valuable discussions with Professor Gianni De Fraja. His, and the Department's, hospitality is appreciated.

I would like to thank Telenor, the Foundation for Research in Economics and Business Administration, the Norwegian Central Bank's research fund and Ingegerd and Arne Skaug's legacy for financial support. I am also grateful to the University of Tromso, NFH, for letting me use some of the time that I should have been teaching to work on my dissertation.

Finally, I would like to thank my family - Mariann and Martin - for being patient and supportive.

Tromso, 08 January 2002
Jan Yngve Sand

Contents

Introduction

Part 1: Quality regulation

Essay 1: *Network infrastructure quality regulation*

Essay 2: *Regulation of a vertically differentiated duopoly*

Part 2: Access charge regulation

Essay 3: *Regulation and foreclosure*

Essay 4: *Market structure and regulation*

Essay 5: *A note on first- and second-best access charge regulation*

Part 3: Investment incentives and regulation

Essay 6: *Demand-side spillovers and semi-collusion in the market for mobile communications*

Introduction

The communications markets are typically characterised by their vertical structures, with one or more firms providing essential facilities necessary to produce products sold to end-users. In telecommunications markets, access to the local loop (the last mile of the telecommunications network) is still often considered to constitute an essential facility.

Until the early 1980's, the policy on telecommunications pricing was more guided by distributional goals rather than efficiency concerns, and competition was almost non-existent. Consequently, competition policy and regulatory policy was limited to pricing policy and investment decisions. The industry usually consisted of one single firm that provided both network and end-user services, and this firm could either be a state-owned enterprise or a privately owned utility. Investment decisions were taken as part of the budgetary process of the respective national parliaments. The main economic argument for a monopolised market structure was that the telecommunications market has the characteristics of a natural monopoly. The common understanding now among most economists working on the regulation of the telecommunications industry is that the only part of the market that exhibits the characteristics of natural monopoly (or possibly natural oligopoly) is the local loop.[1]

The ensuing deregulation process in Europe and the USA has put a firmer emphasis on efficiency, but has also raised new regulatory issues that need to be addressed. Some of these issues are:

- Access to essential facilities and the requirements imposed on operators with significant market power (this is referred to as SMP in EU legislation). This relates to both pricing (access pricing) and technical issues (the choice of interface between interconnected firms, the degree to which firms without infrastructure should be given control over parts of the network etc.).[2]
- Should there be sector specific regulators or should the industry simply be subject to general competition rules?

[1] For more about the history of the deregulation in the US and UK, see Woroch (2001) and Armstrong, Cowan and Vickers (1995).

- Should operators be allowed to vertically integrate, or should the regulator impose vertical separation (i.e., separating the infrastructure provision from the production of end-user services)?
- Universal service obligation (USO), or how to ensure that all inhabitants are given access to necessary communications infrastructure at reasonable rates.[3]

If effective competition is the ultimate goal for regulators, it is necessary for firms that want to offer communication services (be it simple voice telephony, or broadband services) to be given access to the telecommunications infrastructure, which is (potentially) an essential facility. Without such access there will be no entry by non-facility based firms (virtual operators). Some regulators have put the emphasis on encouraging facility-based competition, whereas others have focused on determining a regulatory regime that encourages non-facility based entry. The cost associated with duplicating the existing infrastructure is often very high, and the social costs of duplication must be weighed against the benefits of competition. Assuming that there is only a single firm providing local access, there are various methods of ensuring access to the essential facility. These methods range from the simple resale of the local access provider's (LAPs) services (the consumer pays to the reseller, who in turn purchases transmission capacity from the local access provider) to complete local loop unbundling (LLU).[4] Whether such access generates new, or differentiated services, will depend on the degree to which the entrant can control the traffic generated by its own customers. With simple resale, the choice of signalling and switching technology is left entirely to the LAP. With local loop unbundling, the entrant may to a larger extent be able to control crucial technological decisions and thereby exercise greater control over the traffic (depending on the form of LLU). The terms at which firms without infrastructure should be given access is subject to discussions by most regulators and competition policy authorities, and is debated in

---

[2] This can be described as the price and quality of interconnection between networks.

[3] The issue of USO is, however, not discussed in any detail in this thesis. Much of the problem related to USO deals with the financing of the burden of universal service. Currently EU legislation includes voice telephony and fax and voice band data transmission via modems - i.e., access to the Internet (see COM (1999) 539), but as new technology becomes more widespread it is reasonable that more advanced services will be included in the USO.

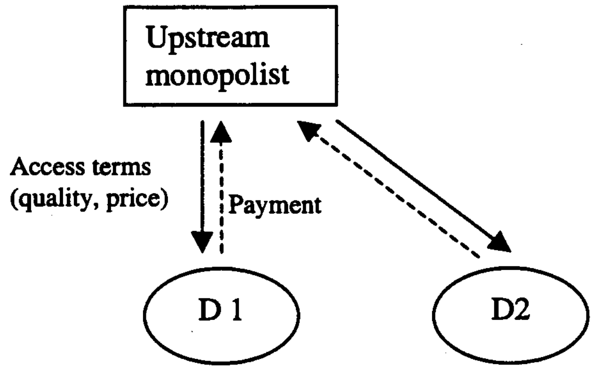
[4] See, e.g., Regulation (EC) No. 2887/2000 of the European Parliament and of the Council, or COM (2000)/237 Unbundled Access to the Local loop.

detail in several of the papers in the present thesis. In the policy debate it is often argued that independent (non-facility based) firms should be given access at the same terms as the network provider's competitive subsidiary. This is highlighted in EU policy documents that call for cost based access charges (when the access provider is in a dominant position, where dominant position is defined as having a market share of more than 50%).[5] The choice of pricing rule should reflect the degree of competition in the market. However, the exact interpretation of what is meant by cost-based access prices is more difficult to ascertain, but it is reasonable to assume that a cost-based access charge should cover some measure of long-run marginal cost (or, alternatively, long-run incremental costs - LRIC). Note that using the LRIC as a pricing rule will imply an arbitrary allocation of fixed costs.

In some circumstances and markets, it is reasonable to assume that there are several firms providing the essential facility. This is in particular true for the market for mobile communications, and it may also be a reasonable description of the market for some types of broadband access. In models with several networks, most notably Laffont, Rey and Tirole (1998a,b) and Armstrong (1998), one of the main issues is the use of the access charge as a collusive device when the access charge is unregulated (or two-way access charges). This is not an issue that is analysed in this thesis. The greater part of this thesis deals with the regulation of the access charge in a situation with a monopolist provider of network services – one way access. This somewhat simplified setting captures most of the relevant issues from a policy perspective, related to, for instance, the socially optimal pricing of the access charge or the socially optimal investment in the quality of the network infrastructure. The simplest set-up that still captures the important aspects consists of one network monopolist upstream, and two downstream firms producing a final output. The upstream monopolist may or may not be vertically integrated with downstream firm D1.

The typical market structure investigated in the present dissertation is illustrated in figure 1 (firms D1 and D2 are two downstream firms producing a final output using the network services of the upstream firm):

[5] See, e.g., Directive 97/33/EC of the European Parliament and of the Council. If the access provider has significant market power (SMP - defined as a market share in excess of 25%), the access charges should normally be determined through commercial negotiations.



The terms at which the producers of the final product gains access to the network monopolist's facilities is composed of a price to be paid for the access service and a quality level attached to that particular service. Regulation often requires that the same price be charged to all firms utilising the network. In the case of a vertically integrated firm such a payment is, of course, only a transfer payment and the magnitude is naturally not affecting the equilibrium output. However, the level of the access charge will affect the market share of independent firms, and the upstream monopolist may, if vertically integrated, have incentives to foreclose (some of) the downstream firms. This is, in particular, true when the access charge is unregulated, in which case independent downstream firms may be foreclosed through setting a sufficiently high access charge. Foreclosure may, however, also take place even in a regulated environment, by the use of non-price discrimination (for some examples of analysis of non-price discrimination see, e.g., Economides, 1998a, Foros, Kind and Sorgard, 2001, and Beard et. al., 2001). The incentives to use non-price discrimination when the access charge is subject to regulation is also the topic of essay 3 in this thesis, and the main results of that analysis will be presented shortly.

Whether there should be restrictions on the network operators' opportunity to integrate vertically into the downstream market is an important question. In the US, the Bell system was broken up in regional Bell operating companies (RBOCs) and AT&T (i.e., the formerly vertically integrated Bell system was vertically separated). The RBOCs offered local access services only (often as regional monopolies) and were not allowed to enter the long-distance market, whereas AT&T were operating as a long-distance operator. The main thrust of the 1996 Telecommunications Act in the US has been to increase the competition in the local access market, and move back towards allowing vertical integration again.[6] The RBOCs will be allowed to enter the long-distance market provided that there is sufficient competition in its own local market. In the EU, vertical separation in the telecommunications industry between network provision and service provision has not been an issue in the market for traditional telephony. However, BT in the UK has been restricted from offering cable-TV services. The network infrastructure has to some extent characteristics that resemble that of a public good, and the issue of access charges may be seen as the pricing of access to a public good. When there is vertical separation, the pricing issue is relatively uncomplicated. One argument against such separation may be that economies of scope between network provision and service provision will not be exploited, which is inefficient. In addition, if non-linear pricing for access cannot be used there will be a problem of double marginalisation when there is vertical separation.[7] With vertical integration the problem is that the integrated firm may have incentives to foreclose rival firms, and regulation of access prices becomes an important issue, which is discussed briefly above (and in more detail in several of the essays below).

The communications industry is an industry where the rapid pace of technological advances means that the issue of investment incentives should be placed high on the agenda. This is also an issue that is closely related to the policy on allowing non-facility based firms access to the infrastructure. There have been a number of cases where small (virtual) operators have complained to national regulators about the lack of access or the terms at which access has been granted.[8] By granting access (for a given level of infrastructure investment), competition may increase which is beneficial for consumers' welfare. However, allowing such access may also deteriorate investment incentives. In addition, allowing access implies that the investments undertaken will have a larger positive external effect on consumers' surplus. The main trade-off (from a regulator's point of view) when determining whether virtual operators should be allowed entry is between the positive effect of competition and the negative effect on investment incentives.

Following the deregulation process that has taken place throughout most of the western economies, there has been a debate about whether there should be sector

[6] See, e.g., Laffont and Tirole (2000) and Economides (1998b).
[7] The classic reference on double marginalisation is Spengler (1950).

specific regulation, or if the industry should be subject to general competition policy oversight only. The main trend until now seems to be that there is a combination of both these aspects of regulation (with the exception of New Zealand).[9] The European Commission suggests in the 1999 Communications Review that there should be a greater reliance on general competition rules when competition becomes more effective. The main argument for using sector-specific regulation is that the competition policy authorities do not have the competence to evaluate the technical side of the industry. It is certainly so that the sector-specific regulator is often concerned with technical issues related to interoperability, but in some countries also the pricing of access to the network infrastructure and other issues related to terms of access. Many of the decisions regarding non-facility based firms' access to the network is not only a technical issue, it is most certainly also an issue that affects the competitive environment in the industry. Thus, such decisions should also involve the competition authorities. Current EU legislation on access to infrastructure is that telecommunications operators with significant market power are required to grant all reasonable requests for access. From a static efficiency point-of-view it may be a sensible idea to allow all operators access, whereas dynamic considerations may call for restricting access to ensure that the incentives to invest in infrastructure are maintained. The issue of sector-specific regulation versus general competition policy rules is, however, not discussed further in this thesis.

What has already been accomplished in the literature?

The literature on access pricing is quite abundant, and I will not attempt to give a complete overview over that literature her. Armstrong (2001) and Laffont and Tirole (2000) provide a more thorough review.

A majority of the literature has chosen to focus on access pricing when the rival firms do not have market power, and more specifically, model rivals as a competitive fringe.[10] The implication of such an assumption is that the fringe firms price their

---

[8] Notably Sense Communication's case against Telenor and Netcom, and Teletopia's case against Telenor (both in Norway).

[9] In the UK, the telecommunications industry is overseen by Oftel (Office of Telecommunications), the FCC (Federal Communications Commission) in the US, and by the Norwegian Post and Telecommunications Authority (PT) in Norway.

[10] Some exceptions are Lewis and Sappington (1999), Vickers (1995) and Armstrong and Vickers (1998).

products equal to perceived marginal costs and earn zero profits. This means that the rival firms' profits play no role in the regulator's welfare function. The general feature of many of these models (and indeed with a more general validity than access charge pricing) is that the socially optimal access price should reflect the cost of providing the service plus the opportunity cost:

Access charge = marginal cost of providing access + opportunity cost

The opportunity cost of selling access to independent firms may, for instance, be the profit foregone in the end-user market by the network provider, which is the main idea of the efficient component pricing rule (ECPR or Baumol-Willig rule; see Willig, 1979). However, the ECPR considers the best access charge for a given price in the final product market, such that the efficiency aspect of the rule depends on the efficiency of the final product market pricing. If there is monopoly pricing in the final product market, the mark-up over costs will be too high from a social point of view that will result in insufficient entry. In Armstrong (2001) a variety of versions (or re-interpretations) of the ECPR are considered, with the main difference between them being the magnitude of the mark-up over costs.[11]

If we were in a first-best situation, with marginal cost pricing in all but the access market (including the final product market), then the socially optimal access charge would imply that access should be priced at marginal cost. In such a case, the ECPR implies pricing access at marginal cost. However, as we have learnt from the theory of second-best, if pricing in some markets diverge from the first-best, then it is not necessarily socially optimal to induce marginal cost pricing of access.[12] Consequently, levelling the playing field by setting access charges equal to marginal cost is not in general the socially optimal policy. Another reason for a departure from marginal cost pricing of the access charge is if there are fixed costs that need to be covered or more generally a budget constraint that must be satisfied. In this case, the access charges may need to be in excess of marginal cost of providing access for Ramsey reasons; i.e., there will be a mark-up over marginal costs to ensure coverage of fixed costs or

[11] In Armstrong, Doyle and Vickers (1996) the ECPR is generalised to account for substitutability between the incumbent's and entrants' final products. In this case the mark-up over the marginal cost is larger the closer substitutes the incumbent's and entrants' products are.
[12] The seminal paper on second-best theory is Lipsey and Lancaster (1956).

to capture the substitutability or complementarity between different products.[13] The latter interpretation is closely related to the analysis in Armstrong, Doyle and Vickers (1996).

Yet another reason for the socially optimal access charge to differ from the marginal cost of providing access is related to non-price discrimination. The relationship between access charges and the incentive to foreclose rival firms is discussed, e.g., in Economides (1998a), Foros, Kind and Sorgard (2001), and Beard et. al. (2001).[14] In Economides (1998a), of which one of the papers of this thesis is an extension, does not (at least explicitly) consider regulated access charges. If the access charge is unregulated and an upstream monopolist is active in a related downstream market, he may have incentives to foreclose his rivals by setting a high access charge (e.g., if his own downstream subsidiary is not too inefficient to rival firms or if there is product differentiation). In the telecommunications industry, the access charge is more often than not subject to regulation, but there may be other means by which the network monopolist may foreclose rivals (e.g., by degrading the quality of access sold to rivals). Thus, it is of interest to examine if and how the socially optimal access charge may differ from marginal cost if the network provider can degrade the quality of access. It seems reasonable that the network provider will be less inclined to refuse (or degrade quality of) access to rivals if the profit margin is high in the access segment. Consequently, if access is priced at marginal cost the network provider will most likely wish to deny rival firms access by degrading quality to such an extent that the rivals do not find it profitable to enter. Thus, another reason for distorting the access charge in excess of marginal cost may be to limit the degree of (potentially) socially costly quality degradation.

When considering the optimality of entry in the competitive segment, it may be necessary to distort access charges from marginal cost and only in the special case of a perfectly competitive downstream market will marginal cost pricing of access be socially optimal. Using the ECPR to price access will under certain conditions ensure that entry will take place only if such entry is efficient. More specifically, entry will take place if the entrant is more efficient than the incumbent in which case entry is socially desirable. In the simplest formulation of the ECPR, it will be the most

[13] See Laffont and Tirole (1994) for details. They develop a model in which the mark-up over marginal cost depends on the super-elasticities, where the super-elasticities incorporate possible complementarities or substitution between products.

efficient firm that provides the final product, and if this firm is the entrant then the network provider will provide network services only. However, pricing access using ECPR will compensate the incumbent for any loss in the competitive segment, which implies that the incumbent will be indifferent between allowing entry and not allowing such entry (i.e., the incumbent does not have any incentives for foreclosure). Consequently, this (access) pricing rule ensures that there is no business stealing effect, which may be the case if pricing access does not take into account the correct opportunity cost measure. However, with a (partially) deregulated downstream segment where the pricing and output decisions of firms is not subject to regulation and where there is imperfect competition, new entry may result in business stealing. This implies that the private incentives for entry may diverge from the social incentives for such entry, and with Cournot competition there may be a tendency towards excess entry if there are no restrictions on entry (see Mankiw and Whinston, 1986).[15]

The main topics of the thesis

The major part of the literature on access pricing problems has focused on situations where the independent firms producing a final product do not possess market power; i.e., they either act as price takers (or a competitive fringe). In this thesis the main idea is to examine, among other things, access pricing in the presence of independent firms with market power. When rival firms do not have market power and if there is marginal cost pricing in the final product market by the incumbent, then I have argued that it is socially optimal to price access at marginal cost (cf. the ECPR). However, if there is imperfect competition in the final product market the pricing will in general diverge from marginal cost. Some of the contribution of this thesis is to capture such situations, and to explain how (and if) the socially optimal access charge should be different from marginal cost of providing access when there is an imperfectly competitive and unregulated final product market.

In essays 3, 4 and 5 presented below, the model is static and there is no explicit consideration of the dynamics between access charges and investments. The cost of providing access may in those models be the (static) equivalent of the long-run

[14] Rey and Tirole (1997) is a more general treatment of the theory of foreclosure.
[15] In their model, Mankiw and Whinston (1986) do not consider vertically related markets.

incremental cost. Cost-based access charges (based on LRIC) may thus be interpreted to mean that the access charge is equal to the marginal cost of providing access. If there is imperfect competition in the downstream market (more specifically, Cournot-competition) the socially optimal access charge is in general below marginal cost of providing access to counter the distortion caused by imperfect competition. However, in these papers, it is shown that the access charge should be in excess of marginal cost if, for instance, the regulated firm can degrade the access quality. It is also shown that if there is entry into the downstream market the access charge should be in excess of marginal cost to eliminate too much costly duplication of fixed costs. It is also shown in this paper that the excess entry result obtained by Mankiw and Whinston (1986) does not necessarily carry over to vertically related markets, and in particular, imposing vertical separation and leaving the access charge unregulated will result in socially optimal entry. In essay 1, the interaction between the level of the access charge and the incentives to invest in infrastructure quality is examined. This analysis is related to both the literature on strategic R&D investments and the literature on (monopoly) quality choice.[16] Essay 1 also discusses how the socially optimal regulation is affected by the choice of market structure. The final essay also discusses issues related to investment incentives and regulation, more specifically between the regulation of roaming quality and investment incentives in the mobile communications market.

Outline of the thesis

In the first essay "Network infrastructure quality regulation", it is assumed that there is one upstream monopolist who may, or may not, be vertically integrated with one (of the two) downstream firms. I study how different vertical arrangements (more specifically, vertical integration versus vertical separation) affect both the private and social incentives to undertake quality investments in the infrastructure. The investments undertaken are either subject to regulation or decided by the upstream monopolist, and it is assumed throughout that the access charge is determined outside the model. A central assumption is that the firms know more than the regulator about the effectiveness of the investments on demand. This implies that the regulator must leave some socially costly information rent to the regulated firm in order to reveal

[16] See D'Aspremont et. al. (1988) on R&D investments, and Spence (1975) for an analysis of the quality choice of a monopolist.

truthfully the relevant information. There are positive external effects due to the investments, which is an argument for regulation of quality. However, due to the asymmetry of information about demand, regulation is socially costly. Consequently, the socially optimal regulation is a trade-off between these effects. The choice of market structure affects the information rent. The value of possessing private information to the upstream monopolist relates to the profitability of operating in the upstream and downstream segments. Under full information, the socially optimal quality level is higher under vertical integration due to the vertically integrated firm's ability to internalise the vertical externality (and produce a higher output in equilibrium). When there is asymmetric information it is shown that the information rent to the regulated firm is higher when vertical integration is allowed, and when access is priced at marginal cost the full information solution can be implemented if there is vertical separation at zero information rent. The socially optimal quality level is, in some circumstances, higher under vertical separation. This is the case if the access charge is sufficiently small, since this means that the vertical externality in the case of vertical separation is small (i.e., the cost disadvantage in the vertical separation case is negligible). Thus, if the access charge is sufficiently close to marginal cost the level of welfare is higher if vertical separation is imposed. Furthermore, the unregulated outcome may be better in terms of welfare compared to regulating the investment decision. This is in particular true when the effect that the investment has on demand is small.

The second essay "Regulation of a vertically differentiated duopoly" considers the regulation of quality when there is asymmetric information about the firms' efficiency levels (or marginal costs in production) and with the assumption that the market is covered. The regulatory mechanism specifies a quality level for each of the two firms, which depend on the efficiency levels of both firms, and a monetary transfer. It is shown that the information rent awarded one firm depends on the choice of quality for the other firm, which implies that there are cross-effects or fiscal externalities in the information rent. This has the effect that the optimal provision of quality cannot be separated from the rent extraction. The choice of quality affects the equilibrium outputs, and the information rent to the firms is based on the output. It is shown that the socially optimal quality levels for the firms are such that the more efficient firm is induced to provide a higher level of quality than the less efficient firm does. The reason for this is that this distorts the market shares in favour of the more efficient firm, which implies that the total output produced is produced at lower cost (note that the total size of the market is exogenously given). The consequence is that, if firms have different efficiency levels, it is socially optimal to induce some degree of vertical differentiation to influence the information rent necessary to induce truthful revelation. The model is extended to capture the effects of cost complementarity between network traffic and the quality of service; i.e., as output increases the marginal cost of providing access is reduced. The presence of cost complementarity affects the information rent. It is less costly to increase the quality provided by the less efficient firm since this increases its output and effectively reduces the marginal cost of providing access. Thus, the presence of cost complementarity partly "finances" the increase in quality.

The third essay, "Regulation and foreclosure", examines the interaction between the access charge and the regulated firm's incentives to use non-price means to foreclose rival firms. Both firms producing the final product possess some degree of market power. If the access charge is unregulated, the upstream monopolist will foreclose rivals by setting a high access charge. When the access charge is subject to regulation, the upstream monopolist may still have incentives to foreclose rivals. This can be achieved by degrading the quality of the input offered to rival firms, which is qualitatively equivalent to raising the cost of rivals. Such foreclosure activities are assumed to be costly for the upstream monopolist. Faced with this problem, the main question in the paper is how the optimal regulation of the access charge should be. In the absence of non-price foreclosure and under full information about the upstream monopolist's cost structure, the socially optimal access charge may in some situations be below marginal cost of providing access to counter the distortion caused by imperfect competition in the downstream market. This, however, creates strong (and socially excessive) incentives to foreclose rivals by non-price means. Since non-price foreclosure is costly, it implies that the regulated firm will require higher transfers (all other things equal) to ensure that the participation constraint is satisfied, and such transfers are socially costly. By distorting the access charge upwards and in excess of the marginal cost of providing access, the foreclosure incentives are weakened.

Essays 4 and 5 are closely related, and focus on many of the same problems. The main difference between these essays is that essay 4 covers a situation with endogenous entry whereas entry is exogenous in essay 5. The fourth essay "Market structure and regulation" investigates the sub-game perfect regulatory policy, where the policy instruments at the regulator's disposal are both structure (vertical integration versus vertical separation) and conduct regulation (access charge regulation). In contrast to the other essays in this thesis, this essay assumes that the independent rival firms enter as long as profit is non-negative. There are no formal restrictions on entry by independent firms. Mankiw and Whinston (1986) show that there is a tendency towards excessive entry with imperfect (Cournot) competition. In my model, there are vertically related markets and one of the main questions examined here is whether the excess entry result holds in this setting. When comparing the degree of entry by independent firms under free entry with the socially optimal degree of entry, I find that the excess entry result does not carry over to all the situations I consider. Free entry may result either in excess entry, insufficient entry, or exactly the socially optimal degree of entry depending on the combination of vertical structure and access charge determination.

The fifth essay "A note on first- and second-best access charge regulation" examines the same problem as the third paper with the difference being that the number of downstream firms is exogenous in contrast to the free entry model of the third paper. The main motivation for this paper is to investigate how access charges should be determined when the regulator can use transfers at no extra (social) cost, and compare this access charge to the access charge when transfers cannot be used. When the regulator can use transfers at no extra (social) cost, the first best outcome can be achieved. This involves setting an access charge below marginal cost to obtain the competitive outcome in terms of quantity.

The final essay "Demand-side spillovers and semi-collusion in the market for mobile communications" (co-authored with Øystein Foros and Bjørn Hansen) considers roaming policy in the mobile market, when there are both facility-based firms and a virtual operator present in the market. I have included this paper in the thesis to provide an analysis of the problem of awarding access to a network from a different perspective from that of the previous papers. This paper analyses the choice of roaming quality in the market for third generation mobile communications (UMTS), and the interaction between regulation and investment incentives. The decision on roaming quality can be seen as a choice of interface between different mobile operators. Higher roaming quality is assumed to result in higher product quality and therefore higher willingness to pay for the final product.[17] One question the paper addresses is whether the investing firms should be allowed to co-ordinate their investments in, and share infrastructure, which is an issue that has become highly topical with the deployment and planning of the next generation mobile communications network. The model presented is a three-stage game, and the choice of roaming quality is decided in the first stage of the game prior to the decision on investments. The regulator may either mandate the roaming decision (in which case the roaming quality is a part of the operators' licenses), or the investing firms may voluntarily choose it. In the final stage of the game, firms compete (non co-operatively) in a Cournot fashion.[18]

We find that in the present model it is welfare optimal to allow firms to co-operate at the investment stage, since this allows some of the effects of the investments to be internalised. By allowing the firms to co-operate, the unregulated firms will choose the socially optimal level of roaming quality, which again induces the socially optimal level of investments. If co-operation is not allowed, an unregulated firm will choose a lower investment level than what is socially optimal to the detriment of the consumers. Furthermore, a welfare maximising regulator will find that allowing the virtual operator access to the mobile infrastructure is better than denying such access.

---

[17] Infrastructure sharing in mobile markets is called *roaming*, and the quality of roaming affects the degree to which investments in infrastructure by one firm benefits the other firms. A high quality of roaming implies that investments are (almost) as beneficial to the non-investing firms as for the investing firm.

[18] This timing of the game may, of course, lead to incentives for re-negotiation after investments are sunk, but we assume that full-commitment contracts can be written. The commitment problem is only present if the virtual operator is not allowed access to the mobile infrastructure.

References

Armstrong, M. (1998), "Network Interconnection in Telecommunications", *Economic Journal* **108**, 545-564

Armstrong, M. (2001), "The theory of access pricing and interconnection", in Cave, Majumdar and Vogelsang (eds.), *Handbook of Telecommunications Economics*, North-Holland, forthcoming 2002

Armstrong, M. and J. Vickers (1998), "The access pricing problem with deregulation: A note", *Journal of Industrial Economics* **46**, 115-121

Armstrong, M., S. Cowan and J. Vickers (1995), *Regulatory reform: Economic analysis and the British experience*, MIT-Press

Beard, T.R., D. Kaserman and J. Mayo (2001), "Regulation, vertical integration and sabotage", *Journal of Industrial Economics*, 319-333

Bergman, et. al. (1998), *Europe's Network Industries: Conflicting Priorities*, Monitoring European Deregulation 1, CEPR, London

COM (1999) 539 Communication from the Commission, "Towards a new framework for electronic communications infrastructure and associated services", European Commission

COM (2000) 237 Communication from the Commission, "Unbundled Access to the Local loop", April 2000, European Commission

D'Aspremont and A. Jacquemin (1988), "Cooperative and non-cooperative R&D in duopoly with spillovers", *American Economic Review* **78**, 1133-1137

Directive 97/33/EC of the European Parliament and of the Council, "On interconnection in Telecommunications with regard to ensuring universal service and interoperability through application of the principles of Open Network Provision (ONP)". Official Journal L 199, 26/07/1997 p. 0032 - 0052

Economides, N. (1998a), "The incentive for non-price discrimination by an input monopolist", *International Journal of Industrial Organization* **16**, 271-284

Economides, N. (1998b), The telecommunications Act of 1996 and its impact", forthcoming, *Japan and the World Economy*

Foros, Ø., H.J. Kind and L. Sorgard (2000), "Access pricing, quality degradation and foreclosure in the Internet", *Journal of Regulatory Economics*, forthcoming

Laffont, J.J. and J. Tirole (1994), "Access pricing and competition", *European Economic Review* **38**, 1673-1710

Laffont, J.J. and J. Tirole (1996), "Creating competition through interconnection: Theory and practice", *Journal of Regulatory Economics* **10**, 227-256

Laffont, J.J. and J. Tirole (2000), *Competition in telecommunications*, MIT-Press

Lewis, T. and D. Sappington (1999), "Access pricing with unregulated downstream competition", *Information Economics and Policy* **11**, 73-100

Lipsey, R. and K. Lancaster (1956), "The general theory of second-best", *Review of Economic Studies* **24**, 11-32

Mankiw, N.G. and M.D. Whinston (1986), "Free entry and social inefficiency", *Rand Journal of Economics* **17**, 48-58

Regulation (EC) No. 2887/2000 of the European Parliament and of the Council of 18 December 2000 "on unbundled access to the local loop"

Rey, P. and J. Tirole (1997), "A primer on foreclosure", in Armstrong and Bresnahan (eds.), *Handbook of Industrial Organization*, vol. 3, North-Holland, forthcoming

Spence, M. (1975), "Monopoly, quality, and regulation", *Bell Journal of Economics* **6**, 417-429

Spengler, J. (1950), "Vertical integration and anti-trust policy", *Journal of Political Economy* **58**, 347-352

Vickers, J. (1995), "Competition and regulation in vertically related markets", *Review of Economic Studies* **62**, 1-17

Willig, R. (1979), "The theory of network access pricing", in *Issues in Public Utility Pricing*, H. Trebing (ed.), Michigan State University Press

Woroch, G. (2001), "Local network competition", in Cave, Majumdar and Vogelsang (eds.), *Handbook of Telecommunications Economics*, North-Holland, forthcoming 2002

Part 1:
Quality regulation

ESSAY 1

Network infrastructure quality regulation

Jan Y. Sand*

University of Tromso, NFH

Department of Economics, N-9037 Tromso,

Norway

Jan.Sand@nfh.uit.no

December 2001

Abstract

The paper considers the optimal regulation of network infrastructure quality when the impact of investments on demand is private information. The socially optimal investment level is contrasted to the unregulated levels under vertical integration and separation. The choice of market structure has an impact on the information rent, and it is shown that the value of the firm's private information is reduced under vertical separation.

JEL Classification: D82, L13, L22, L51

Keywords: infrastructure regulation, vertical relations, duopoly

*I would like to thank Kåre P. Hagen, Derek Clark, Nils-Henrik von der Fehr, Patrick Rey and Jean Tirole for very helpful comments and suggestions on a previous version. Financial support from Telenor, NFR, and Norges Banks Fond til Økonomisk Forskning is gratefully acknowledged.

1 Introduction

A major issue in the communications industry at the present deals with access to broadband networks and services. The exact definition of what broadband changes with time, but it now seems common to define broadband as the capability of supporting bandwidth greater than 2 Mb/s to the consumer. The industry envisions a huge increase in the demand for services which require broadband capacity to operate satisfactorily. Such services may include video-on-demand, high definition TV (HDTV), video-conferencing, games, and interactive information services. Many of these services are likely to be using the network services intensively, with the end-user terminal as an interface. Significant investments are necessary to upgrade the existing networks to be able to offer the broadband capacity. In some of the Scandinavian countries, there have been political discussions regarding whether the government should become involved in building (or financing) broadband networks to guarantee access to high speed network to everyone.

There are two main issues addressed in the present paper: Should the network operator be induced to invest in quality (i.e., should network investments be subject to regulation), and should the network operator be allowed to enter the service market (the content market)? The former issue is closely related to the quality choice problem analysed by, e.g., Spence (1975), where it is shown that the monopolist may have incentives to either over- or under-invest in quality.[1] The latter issue is related to whether there should be limitations on the (potential) entry of the network owner into the markets for end-user services. The paper investigates the effect that changing the vertical structure (i.e., vertical integration versus vertical separation) has on the optimal regulatory policy. The paper addresses questions related to optimal infrastructure quality and market structure issues when the effect of investments on demand is private information to the investing firm. The issue of infrastructure quality is often the concern of a sector-specific regulator, whereas the

[1] The unregulated monopolist's choice of quality reflects the value of that quality to the marginal consumer, whereas the socially optimal quality level is related to the value of quality to the average consumer.

appropriate market structure is a concern for the competition authorities.[2]

In this paper, it is assumed that a non-facility based firm is allowed to enter the services market (the end-user market), and the focus of the regulator is to provide consumers with the appropriate level of quality in the network. The main concerns of many politicians and (potential) users of broadband access and services are whether the market will provide adequate coverage of broadband access, and also the timing of such investments (i.e., will they make these infrastructure upgrades quickly enough). This is, in particular, a major concern among information technology users in less densely populated areas. Obviously, one fiber cable will not be able to serve as many homes in less densely populated areas as in urban areas. Consequently, the investment costs associated with upgrading the network to provide broadband access will necessarily be convex in coverage. The concerns voiced by some people regarding the capability, or rather the willingness to supply broadband access voluntarily to all areas of a region, may well become a real issue.

In the present model I find that an unregulated profit-maximising firm in general chooses a level of infrastructure quality which does not coincide with the socially optimal level. In general, the (unregulated) level of quality can be either higher or lower than the socially optimal level, since the profit-maximising firm does not internalise the effects on consumers' surplus and on rivals' profits. However, as is shown in the paper, the facility-based firm will in the present model choose a lower level of infrastructure quality compared to the socially optimal level when there is full information. Furthermore, the choice of market structure will also have an impact on the value of the private information to the regulated firm. In particular, vertical separation reduces the information rent compared to the vertical integration case. As a consequence, the socially optimal quality level is always lower under

[2]There are many regulatory issues related to broadband, and many of these are essentially the same as those encountered in the market for narrowband networks. Regulatory issues common for both narrowband and broadband could, for instance, be whether the regulatory authorities should encourage facility-based competition or allow non-facility based entrants access to the network. Furthermore, on which terms should they gain access if entrants without their own facilities are allowed to enter (i.e., should the local loop be unbundled, how should access charges be determined)? Many of these issues are discussed in for instance Laffont and Tirole (2000).

vertical integration. Both of these factors are, in isolation, arguments in favour of vertical separation. The value of the regulated firm's private information will depend critically on the level of the access charge.

Throughout the paper I will maintain the assumption that the downstream industry is not subject to price regulation. The number of firms in the industry is exogenously determined, and this may be interpreted as the result of entry restrictions (e.g., through a licensing process). In addition, the level of the access charge is not included in the regulatory contract and is exogenously given, but I will examine how the level of the access charge impacts on the regulatory contract.

Furthermore, it will be assumed throughout the paper that the cost structure of the firms does not involve economies of scope between upstream and downstream activities. This is clearly a simplification. It is often argued that there are significant economies of scope between providing network services (transportation) and the services that run on top of the networks. If we allow for economies of scope between the production of network services and the production of the final product, the costs of producing downstream will be lower if we allow for vertical integration. Any such economies of scope will of course imply that there are welfare gains to vertical integration - welfare gains that are not taken into account in the analysis below.

The rest of the paper is organised as follows: In section 2, the framework for analysis is presented, together with a characterisation of the Cournot equilibrium of the game. In section 3, I consider the socially optimal regulation with full information and the unregulated cases under vertical integration and separation. In section 4, asymmetric information is introduced and the optimal qualities for the cases of vertical integration and separation are characterised, and the effect of the vertical structure on regulatory policies is also investigated. In the final section, some concluding remarks are made.

2 The framework

In the model presented below there are vertically related markets (an upstream and a downstream market), with one firm upstream and two firms downstream. The downstream firms purchase an essential input from the network monopolist. To simplify, I will only consider the case where the number of firms in the downstream industry is determined exogenously, and will not consider issues related to entry and exit here. The regulatory mechanism that is offered consists of a quality level and corresponding transfers necessary to induce truthful revelation of information, and will be based on a signal about the effect of investments on demand that the regulated firm sends to the regulator.[3]

I assume that the game is played in several stages, and the timing of the game is as follows:

Stage 1: The regulator offers a menu of contracts to the regulated firm, which consists of the quality level of the infrastructure and transfer.

Stage 2: The regulated firm reports a demand level (type) to the regulator, and the contract is executed.

Stage 3: Firms compete in quantities in the downstream segment (the final stage of the game is unregulated, and it is assumed that firms move simultaneously).

In addition, the welfare effect of vertical integration versus vertical separation will be discussed. It is assumed that the choice of vertical structure is made prior to the start of the game (i.e., prior to the investment decision).

The regulator must take into account the expected effects the regulation he proposes have on all subsequent stages to be able to internalise the effects of the proposed regulation. The model is solved using backwards induction. We first solve for the outcome of the Cournot competition stage. Using his knowledge of the outcome at this stage, the regulator designs a regulatory mechanism.

[3] A more complete regulatory mechanism would also include the access charge. However, to simplify and maintain focus on investment in infrastructure, I will only discuss how the magnitude of the access charge affects the solution.

2.1 The demand side

In the downstream market, firms face an inverse demand function of the following form: $P = P(Q, \kappa)$, where $\kappa \equiv \beta\theta$ denotes the overall quality of the network, Q is the total production downstream, θ is the quality of the infrastructure, and β denotes the impact of changes in the infrastructure quality has on demand.

Gross consumers' surplus is in this case given by:

$$CS(Q, \theta, \beta) = \int_0^Q P(Q', \beta\theta)\, dQ' \tag{1}$$

The quality level may be subject to regulation, in which case we can think of this as a quality requirement in the incumbent's universal service obligation (USO).[4] It is assumed that any increase in the level of quality upstream translates directly into increases in quality downstream and increased value to consumers. The parameter β can be interpreted as a spillover parameter, and is identical for both firms. One reason for having equal spillovers may, for instance, be that all firms are perfectly interconnected. If firms are symmetric in all other respects, there is no reason that the market expansion should be asymmetric.[5] It is assumed that all firms know the exact value of β whereas the regulator knows only the distribution and support.

Assume $P_\kappa > 0$, $P_{\kappa\kappa} \leq 0$, and $P_{Q\kappa} = 0$. Thus, the higher the spillover, the larger is the effect of increasing the quality of the infrastructure on the willingness to pay. The assumption $P_{Q\kappa} = 0$ rules out the possibility that both the parameters θ and β may affect the slope of the inverse demand function. Thus, quality investments affect the demand downstream through parallel shifts of the (downstream) demand

[4] Incumbent firms subject to USO may be reluctant to introduce new services and technology to a majority of its customers, fearing that regulators may be tempted to include these in the USO. For instance, normally a USO does not incorporate ISDN or ADSL. However, if these technologies become widespread, it is reasonable to suspect that regulators may want to include these in the USO.

[5] In the R&D literature, these spillovers can typically take on different values for different firms, with a higher level of spillover for the investing firm (the regulated firm in the present model). In the context of R&D investments, it is often reasonable that it is more difficult to transfer knowledge between firms unless they set up joint-venture R&D facilities, with perfect information sharing. See, e.g., d'Aspremont and Jacquemin (1988).

function, and the degree of spillover simply adjust the magnitude of this shift. This implies that changes in the quality level affects consumers' surplus only indirectly through a market expansion effect.[6] I assume furthermore that the inverse demand function is decreasing and concave in Q ($P_Q < 0$, $P_{QQ} \leq 0$).[7] When comparing quality levels and information rents, I will for simplicity use a linear inverse demand function of the form: $P(Q, \beta, \theta) = a + \beta\theta - Q$.

2.2 The supply side

The three main elements to the profit function of the vertically integrated firm are; the profit from its downstream activities, the profit from upstream sales, and the investment cost function (for investments in infrastructure quality). The cost of investment is given by the function $K(\theta)$, which is assumed to be an increasing and convex function. It is assumed that the network provider (i.e., the upstream firm) undertakes the investments in network infrastructure quality. The profit function of

[6] The implication of assuming that changes in quality does not affect the slope of the inverse demand function is that the willingness to pay for quality is identical for the average and marginal consumer. However, in contrast to the literature on the monopolist's quality choice (Spence, 1975), the present paper considers a situation with vertically related markets and imperfect competition in which the investing monopolist cannot capture the entire surplus of the investments in quality. Consequently, even if the average and marginal consumer value quality equally, which in the pure monopoly case would imply that the monopolist chooses the socially optimal level of quality, this is not the case in the present paper.

[7] The assumption of a concave inverse demand is not necessary to ensure stability of the equilibrium. Normally the following regularity conditions are imposed on models of Cournot competition: i) $P_Q + P_{QQ}q < 0$ (where q is the downstream output of an individual firm), and ii) $3P_Q + P_{QQ}Q < 0$ (where Q is total downstream output). These regularity conditions imply that the demand function cannot be too convex in Q, and implies that the reaction curves are downward sloping. Condition (i) is identified by Hahn (1962) as the condition required to ensure (local) stability of the equilibrium. Hahn allows for non-linearities in costs and includes the condition $c_{qq} - P_q > 0$ in his requirements, where c_{qq} is the second-derivative of production costs. These two conditions ensure that the own effect on marginal profitability of changing own quantity dominates the (sum of) cross-effects of changing other firms' quantities, and put together they yield condition (ii) in the present model. In the present model, these conditions are trivially satisfied since demand is assumed to be concave.

the vertically integrated firm is given by:

$$\Pi^v = (P - w)\, q^v + (w - c)\, Q^v - K(\theta) + t \tag{2}$$

where q^v is the downstream production of the vertically integrated firm, Q^v is the total production downstream if there is vertical integration, and t is the transfer from the regulator. The marginal cost of purchasing all other inputs besides local access is assumed to be identical for both downstream firms, and is normalised to 0. The variable w is the (exogenously determined) price all downstream firms pay per unit for the inputs purchased from the upstream firm. Below, I will discuss the effects of varying the level of the access charge w. I will assume the following about the level of the access charge:

Assumption 1

The access charge covers the marginal costs of providing access; i.e., $w \geq c$.

The main reason for assuming that the access charge must cover marginal cost of providing access is that regulators often argue (and require) that the network providers should be able to provide their services without relying on subsidies from the government.[8]

The profit function for upstream operations only, which is also the profit earned by a vertically separated firm, is given by:

$$\Pi^u = (w - c)\, Q - K(\theta) + t \tag{3}$$

The profit function of the (independent) firms competing downstream is given by:

$$\Pi^i = (P - w)\, q^i \tag{4}$$

[8] In the present paper I am only considering linear access charges. If we allow for non-linear access tariffs, the per unit access charge does not necessarily have to be in excess of the marginal cost of providing access.

where q^i is the production level of firm $i = 1, 2$. It is assumed throughout that there is always two firms in the downstream market. Consequently, if there is vertical separation, I assume that there are two independent firms competing.

2.3 Welfare

The regulator maximises the utilitarian welfare function given by (for vertical integration and vertical separation):

$$W = \begin{cases} CS - PQ^v + \Pi^v + \Pi^i - (1+\lambda)\, t \\ CS - PQ^s + \Pi^u + \sum_{i=1}^{2} \Pi^i - (1+\lambda)\, t \end{cases} \tag{5}$$

where $\lambda > 0$ is the shadow cost of public funds, which is assumed to be strictly positive due to efficiency losses from distortionary taxation. The welfare function is assumed to be concave in θ. The regulator has imperfect knowledge of the degree of spillover, and therefore only imperfect knowledge about the exact effect investments in the infrastructure quality has on demand. The regulator knows only the distribution, $G(\beta)$, with the strictly positive density function $g(\beta) > 0$ and the support of the level of spillovers, $\beta \in [\underline{\beta}, \overline{\beta}]$. The case of $\underline{\beta} = 0$ corresponds to a case where investing in the quality of the network infrastructure has no effect on demand. To avoid such a trivial case, I will assume that spillovers are always strictly positive; i.e., $\underline{\beta} > 0$.

2.4 Cournot competition

Firms compete in quantities in the final product market. The vertically integrated firm v chooses the level of quantity to maximise Π^v, where Π^v is defined by eqn. (2). The independent firm i solves a similar maximisation problem, maximising (4) with respect to q^i. The equilibrium quantities for firm v and firm i are defined as q^{*v} and q^{*i}. Total equilibrium quantity when there is vertical integration is defined as $Q^* = q^{*v} + q^{*i}$. When there is vertical separation, $i = 1, 2$, and $Q_S^* = q_1^* + q_2^*$, where subscript S denotes vertical separation.

The vertically integrated firm's optimal quantity choice q^{*v} is given by:

$$P_Q q^{*v} + (P(Q^*, \beta\theta) - c) = 0 \tag{6}$$

For an independent firm i optimal quantity q^{*i} is implicitly given by:

$$P_Q q^{*i} + (P(Q^*, \beta\theta) - w) = 0 \tag{7}$$

To ensure that the independent firm(s) produces a positive level of output in equilibrium, we need to assume that $P > w$.

Vertical integration I will first consider the Cournot equilibrium in the case where the network monopolist is allowed to integrated vertically into the downstream market. Using the system of first-order conditions defined by equations (6) and (7), we can solve for the effects that changing exogenous variables have on q^{*v} and q^{*i}.

Whereas the equilibrium output of the firm i (when there is vertical integration) is increasing in θ and β, both the effect of changing the infrastructure quality θ, and of changing the spillover level β, on firm v's equilibrium output are ambiguous. I will however, make the following assumption to ensure that firm v's output, and therefore both firms' output, is increasing in both θ and β:

Assumption 2

Provided that $P_{QQ} \leq 0$, the following condition ensures that firm v's output is increasing in equilibrium: $P_Q/P_{QQ} \geq (q^{*v} - q^{*i})$.

If inverse demand is strictly *concave* in Q, then necessary conditions for both firms' outputs to increase as θ and β increase are: $P_Q/P_{QQ} \geq (q^{*v} - q^{*i}) \geq -P_Q/P_{QQ}$.[9] The first inequality ensures that firm v's output increases, whereas the second inequality ensures that firm i's output increases. Note that a necessary condition for

[9] If , for instance, we consider the strictly concave inverse demand function $P(Q, \theta) = a + \beta\theta - bQ^2/2$, the set of inequalities are satisfied (the requirement is that the rival firms has a non-negative market share, which is obviously satisfied).

the latter inequality to be satisfied is that firm v has a larger market share than firm i, which is ensured by assumption 1. The first inequality is satisfied if, for instance, the inverse demand function is not too concave. If the inverse demand is sufficiently concave, an increase in output from increased quality (or spillover) results in a large reduction in price. This implies a significant loss on inframarginal units of the final product for the firm with the larger market share (i.e., firm v). Consequently, it may be desirable for a dominant firm to reduce its own quantity when infrastructure quality (or spillover) increases to maintain a relatively high price on the final product.

Under assumption 2 and the assumptions made on the inverse demand function, the profit-maximising quantity choices made by the firms have the following properties:[10]

Lemma 1 *(Vertical integration)*

Firm v's supply has these properties: $\partial q^{*v}/\partial w > 0$, $\partial q^{*v}/\partial c < 0$, $\partial q^{*v}/\partial \beta > 0$, and $\partial q^{*v}/\partial \theta > 0$.

Firm i's supply has these properties: $\partial q^{*i}/\partial w < 0$, $\partial q^{*i}/\partial c > 0$, $\partial q^{*i}/\partial \beta > 0$ and $\partial q^{*i}/\partial \theta > 0$.

The effect on total downstream demand in equilibrium, Q^*, is simply the sum of the effects on q^{*v} and q^{*i}, and is summarised in lemma 2.

Lemma 2 *(Vertical integration)*

Total downstream production has the following properties: $\partial Q^*/\partial \beta > 0$, $\partial Q^*/\partial \theta > 0$, $\partial Q^*/\partial c < 0$, $\partial Q^*/\partial w < 0$.

In equilibrium, the total quantity depends only on the sum of the (perceived) marginal costs of the two firms, with the perceived marginal cost equal to c and w for firms v and i, respectively. As is expected, the total equilibrium quantity is decreasing in the cost parameters. Furthermore, the total quantity is increasing in the infrastructure quality and the spillover parameter.

[10]See Appendix 1 for details.

Vertical separation When there is vertical separation, the downstream firms are symmetric and each independent firm's Cournot quantity is in equilibrium given implicitly by eqn. (7). It can then be shown that the properties of the equilibrium are qualitatively the same as for the independent firm i in the vertical integration case. Note that Assumption 2 is not necessary in the vertical separation case.

3 The benchmark cases

The main issue of the present paper is to investigate how different vertical market structures affects the socially optimal level of infrastructure quality and how this relates to the unregulated solution. Furthermore, I will examine how the presence of asymmetry of information about the effect of changes in infrastructure quality impacts on the regulator's socially optimal investment strategies. In order to evaluate the effect of asymmetry of information, the socially optimal solution under full information must be used as the benchmark case. The socially optimal solutions will also be compared to the unregulated solutions in order to identify the situations in which a regulator should avoid interference in the market place.[11]

3.1 Infrastructure quality regulation

When there is vertical integration, the regulator solves the maximisation problem given by eqn. (5) with respect to θ, subject to Cournot equilbrium outputs and the participation constraint for the regulated firm (the participation constraint will be binding for all types in the full information case since the shadow costs of public funds is positive):[12]

[11] It should, however, be noted that since the downstream market is assumed to be unregulated and involving firms with significant market power, the benchmark case is *not a* first-best scenario.

[12] The welfare function can be rewritten as follows (taking into account that the participation constraint for firm v is binding): $W = \int_0^{Q^*} P(Q', \theta\beta)\, dQ' - cQ^* - K(\theta)$.

$$
\begin{aligned}
& \beta P_{\kappa} q^{*v} + \left(P\left(Q^{*v}, \beta\theta\right) - w\right) \frac{\partial q^{*v}}{\partial \theta} + \left[(w-c) + P_Q q^{*v}\right] \frac{\partial Q^{*v}}{\partial \theta} \\
= \; & K^{'}(\theta) - \frac{d\Pi^i}{d\theta} - \frac{dCS^n}{d\theta}
\end{aligned}
\tag{8}
$$

The socially optimal infrastructure quality under full information, θ^{FI}, is defined implicitly by eqn. (8). CS^n is defined as net surplus to consumers, or $CS - PQ^{*v}$, where Q^{*v} is the total downstream output in equilibrium with vertical integration. It is assumed that net surplus to consumers is increasing in network infrastructure quality.[13] Since it is assumed that quality does not affect the slope of the inverse demand function, the only effect on consumers' surplus of changing quality is indirectly through an increase in the output in equilibrium.

From condition (8), we observe that optimal policy under full information with respect to the vertically integrated firm's upstream efficiency level, depends in part on the effect of infrastructure quality on the competitor's profit, $d\Pi^i/d\theta$. If the sign is positive (negative), this means that the investment imposes a positive (negative) externality on the competitor. This is analogous to the industrial organisation literature on strategic investments where an action taken in the first stage of a game may affect the intensity of the second-period competition (see Tirole, 1988, or Fudenberg and Tirole, 1984). The sign of $d\Pi^i/d\theta$ is determined by the following relationship:[14]

$$
\frac{d\Pi^i}{d\theta} = \beta P_{\kappa} q^{*i} + q^{*i} P_Q \frac{\partial q^{*v}}{\partial \theta}
$$

Investing in infrastructure quality has an unambiguously positive *direct* effect on firm i's profit, through its effect on demand (increasing the quality in effect increases the size of the market; i.e., it shifts the demand curve outwards). This is the term $\beta P_{\kappa} q^{*i}$. However, there is another effect which is related to the strategic interaction

[13]It can be shown that in equilibrium $\frac{\partial CS}{\partial \theta} = 2\beta\theta P_{\kappa} P_Q Q^{*} / \left(3P_Q + P_{QQ} Q^{*}\right) \geq 0$ and $\frac{\partial^2 CS}{\partial \theta^2} = 2\beta P_Q / \left(3P_Q + P_{QQ} Q^{*}\right) \left(\beta P_{\kappa\kappa} Q^{*} + 3 P_Q P_{\kappa} \frac{\partial Q^{*}}{\partial \theta} / \left(3P_Q + P_{QQ} Q^{*}\right)\right)$. The latter expression is positive if $P_{\kappa\kappa} = 0$ or sufficiently small.

[14]We can write this in more general terms as (assuming that firm i maximises profits with respect to its own quantity): $\frac{d\Pi^{*i}}{d\theta} = \frac{\partial \Pi^{*i}}{\partial \theta} + \frac{\partial \Pi^{*i}}{\partial q^{v}} \frac{dq^{*v}}{d\theta}$.

and this effect is negative. The *strategic* effect is the impact on prices given by the term $q^{*i}P_Q\frac{\partial q^{*v}}{\partial\theta}$. For a given level of production for firm i, increasing θ implies increased production by firm v, which again implies reduced price. This is a negative effect on firm i's profit.[15] Under the assumption of a concave demand function and perfect spillover, the externality is indeed always positive. Consequently, we only need to consider the case of a positive external effect.[16]

Assuming that welfare, W, is concave in θ.[17] Since an increase in infrastructure quality has a *positive* external effect on firm i's profit (under our quite general assumptions), then optimal quality is *higher* than it would be in the absence of the effect on firm i.[18]

The welfare maximising regulator takes into account the effects that changes in the quality level of the network infrastructure has on downstream competition and consumers' surplus. The effect on consumers' surplus is a positive externality, which the regulator internalises. If it is appropriate to include the rival's profit in the welfare function, there is an additional positive external effect imposed on competitors by increasing quality, represented by the expression $d\Pi^i/d\theta$. Since both these externalities are positive, this implies that the marginal social cost of increasing infrastructure quality is less than the marginal private cost. A welfare maximising

[15] We observe that if $\left(\beta P_\kappa + P_Q\frac{\partial q^{*v}}{\partial\theta}\right) > 0$, then quality investment makes the regulated firm "soft"; i.e., it increases firm i's profit. If, on the other hand, the expression is negative, then investments make the regulated firm "tough", to use the terminology of Fudenberg and Tirole (1984). For our situation it is perhaps more appropriate to use the term negative (positive) externality instead of "toughness" ("softness").

[16] $d\Pi^i/d\theta = 2\beta P_\kappa q^{*i}\left(P_Q + P_{QQ}q^{*v}\right)/\left(3P_Q + P_{QQ}Q^*\right)$. Thus, the concavity of demand ensures that increasing θ imposes a positive external effect on the rival firm. Note that if P is linear in θ, then Π^i is increasing and convex in θ.

[17] The second-order condition on welfare with respect to θ is given as (assuming that $P_{\kappa\kappa} = 0$): $\frac{\partial^2 W}{\partial\theta^2} = -\left(\frac{2\beta P_\kappa}{3P_Q+P_{QQ}Q^*}\right)^2\left(\frac{(P(Q^*,\theta\beta)-c)P_{QQ}}{3P_Q+P_{QQ}Q^*} + 2P_Q + P_{QQ}Q^*\right) - K''(\theta)$. Concavity of W in θ implies that the investment cost function is sufficiently convex in θ.

[18] If, for instance, the rival firm is a subsidiary of a foreign multinational corporation it is reasonable that the effect on firm i's profit is excluded from the welfare function. In such a case, most of the profits earned by the rival is transferred overseas and, consequently, ignored by the (domestic) regulator.

regulator will internalise the gain to the rival's profit and to consumers' surplus of an increase in the infrastructure quality.

The first-order condition (8) which determines the socially optimal level of infrastructure quality under full information can be rewritten as:

$$\left(P\left(Q_k^*, \theta\beta\right) - c\right) \frac{\partial Q_k^*}{\partial \theta} + \beta P_\kappa Q_k^* - K'(\theta) = 0 \tag{9}$$

where $k = \{\text{vertical integration } I, \text{ vertical separation } S\}$. From the rewritten first-order condition (9) the socially optimal infrastructure quality under both vertical integration and separation is implicitly given, with the difference in optimal quality being only due to differences in the total equilibrium output. Eqn. (9) defines implicitly θ_k^{FI}, where $k = (I, S)$. In the rewritten form, the various real economic factors are more easily noticed. First, the two last elements, $\beta P_\kappa Q_k^* - K'(\theta)$, are the effects on welfare through raising the willingness to pay for the final product for a given quantity supplied and the costs associated with increasing quality, respectively. The first element, $\left(P\left(Q^*, \theta\beta\right) - c\right) \frac{\partial Q_k^*}{\partial \theta}$, indicates the value of raising quality through its effect on quantity, where the difference between price and cost $(P - c)$ represents the value to consumers of each extra unit they consume.

In general, the socially optimal infrastructure quality differs depending on the choice of market structure (i.e., whether there is vertical integration or vertical separation), since equilibrium quantity will be different when $w > c$. However, the socially optimal level of network infrastructure quality is independent of the market structure (i.e., independent of whether there is vertical integration or vertical separation) when the regulator has full information and the access charge is equal to marginal cost of providing access (i.e., if $w = c$). In that case, there are no vertical externalities that the vertically integrated firm can internalise to create a cost advantage over the vertically separated situation. However, as will be shown below, the choice of market structure does play a role in determining the optimal level of quality when there is asymmetric information even when there is marginal cost pricing of access. This is due to the effect that changing the market structure has on the information rent to the regulated firm.

3.2 Optimal versus unregulated level of quality

Will there in general be a need for governmental intervention in the market? Or, put differently, what is the relationship between the *socially* optimal level of quality and the *unregulated*, profit maximising level of quality? The answer to this question obviously depends on how the investment decision taken by the network monopolist affects consumers and independent firm(s), and whether the firm providing network services is vertically integrated or not. The case of vertical integration will be discussed first.

3.2.1 Vertical integration

Let us assume that the upstream monopolist is allowed to enter the downstream market. Then, firm v will choose a quality level to maximise its profit, given by eqn. (2) while taking into account the Cournot outputs in the final stage of the game. The result can be summarised in the following proposition:

Proposition 1 *Assume a concave welfare function. Then, an unregulated, profit-maximising choice of quality for a vertically integrated firm results in a socially suboptimal level;* $\theta_I^{FI} \geq \theta_I^{ur}$. *Furthermore, the unregulated firm will choose not to invest at all (i.e.,* $\theta_I^{ur} = 0$*) if the degree of spillover is sufficiently small.*

Proof. Maximisation of (2), subject to $q^v = q^{*v}$ and $q^i = q^{*i}$, yields a quality level θ_I^{ur} implicitly defined by (assuming an interior solution):[19]

$$\frac{d\Pi^{*v}}{d\theta} = \beta P_\kappa q^{*v} - K'(\theta) + (w - c + P_Q q^{*v}) \frac{\partial q^{*i}}{\partial \theta} = 0 \tag{10}$$

The first two elements of eqn. (10) are the *direct* effects of investments in quality, and the latter element is the *strategic* effect which is negative.[20] Observe that the left-hand side of (10) is smaller than the left-hand side of (9), which together with

[19]The total derivative of firm v's profit with respect to θ is given by: $\frac{d\Pi^{*v}}{d\theta} = \frac{\partial\Pi^{*v}}{\partial\theta} + \frac{\partial\Pi^{*v}}{\partial q^i}\frac{\partial q^{*i}}{\partial\theta} + \frac{\partial\Pi^{*v}}{\partial q^v}\frac{\partial q^{*v}}{\partial\theta}$. By utilising the envelope theorem, we can eliminate $\frac{\partial\Pi^{*v}}{\partial q^v}\frac{\partial q^{*v}}{\partial\theta}$.

[20]This can be derived from the firms' profit maximisation problems. Assume $[(w-\beta) + P_Q q^{*v}] > 0$. The first-order condition for the profit-maximising quantity choice for the vertically integrated firm is given by: (*) $[P_Q q^{*v} + (w - \beta)] + P - c - w = 0$. Then, the

$K'' > 0$ implies that $\theta_I^{FI} \geq \theta_I^{ur}$. Observe that if β is sufficiently small $\frac{d\Pi^{*v}}{d\theta} < 0$, which implies that $\theta_I^{ur} = 0$. ■

There are positive external effects on both consumers' surplus and the rival firm's profit associated with the investments in infrastructure quality. The profit of the downstream subsidiary of firm v will also be increasing in the level of quality. An unregulated firm will, in general, choose a level of network infrastructure quality which does not coincide with the welfare maximising level, since the external effects imposed on the rival firm's profit and on consumers' surplus cannot be internalised by firm v. However, a vertically integrated firm will be able to internalise parts of the external effect; i.e., the effects changes in infrastructure quality has on the profit of its own downstream subsidiary.

The negative sign of the strategic effect seems contrary to results obtained in the literature on strategic R&D investments, but can be explained as follows. In many of the models analysing strategic R&D investments, firms may have an incentive to over-invest for strategic reasons. Investments in the quality of the infrastructure have very similar features to investments in research and development (R&D), in that there are spillover effects associated with the investments. In the present model there are perfect spillovers, meaning that any investment undertaken by the facility-based firm has equal impacts on the demand faced by all firms. Perfect spillovers may be due to perfect interconnection of networks, which may be a reasonable assumption if, for instance, there is local loop unbundling (LLU) and the non-facility based firm is entirely in control over the choice of signaling technology used in the local loop. In R&D models, however, the general feature is that the degree of spillover often is less than perfect, partly due to difficulties in obtaining perfect information sharing between firms in an industry.[21] Since spillovers are perfect in the present model, an investing firm will not be able to gain from a product differentiation effect, and

condition $[P_Q q^{*v} + (w - \beta)] > 0$ implies $P - w < c$. The first-order condition for the rival firm is given by: (**) $P_Q q^{*i} + P - c - w = 0$. Consequently, $[P_Q q^{*v} + (w - \beta)] > 0$ is not consistent with profit maximisation in the duopoly case.

[21] See e.g., D'Aspremont and Jacquemin (1988). In the literature on strategic R&D investments there is typically more than one firm investing.

the non-facility based firm will free-ride on the investment. With less than perfect spillovers, the (strategic) under-investment result above will be modified, since there will be a product differentiation effect which generates additional returns on the investment, and the free-riding effect becomes less pronounced.[22]

3.2.2 Vertical separation

Let us now briefly examine the case of unregulated investments if there is vertical separation. When there is vertical separation, it is assumed that there are two independent firms competing in the downstream market. The socially optimal level of quality is defined as θ_S^{FI} if there is vertical separation. The result of comparing the unregulated and socially optimal quality choice with vertical separation is summarised in the following porposition:

Proposition 2 *Assume a concave welfare function in quality. When there is vertical separation, we find that the socially optimal level of quality is in excess of the unregulated choice of quality. No investment will take place if the access charge is equal to marginal cost of providing access.*

Proof. The vertically separated firm, which we denote with superscript u, maximises the profit function (3) with respect to θ, which defines θ_S^{ur}:[23]

$$(w-c)\frac{\partial Q_S^*}{\partial \theta} - K'(\theta_S) = 0 \tag{11}$$

where subscript S denotes vertical separation, and where $\partial Q_S^*/\partial\theta > 0$.[24] Compare eqns. (9) and (11). Since $P > w$ and $\beta P_\kappa Q_k^* > 0$, together with $K'' > 0$, we must

[22] Another factor that may modify the no-investment result may be modified is if the downstream firms have different marginal costs. If the downstream subsidiary of firm v is sufficiently inefficient relative to firm i, we may have $(w - c + P_Q q^{*v}) > 0$, in which case firm v may choose a positive level of investment in quality (i.e., differences in marginal costs and imperfect spillovers yield equivalent results).

[23] The second-order condition for firm u is always satisfied under the assumptions in the present model, and is given as: $\frac{\partial^2 \Pi^u}{\partial \theta^2} = (w-c)\frac{\partial^2 Q^*}{\partial \theta^2} - K''(\theta) < 0$.

[24] By comparing eqns. (11) and (9) when there is vertical separation, we will be able to determine whether an unregulated, vertically separated network provider chooses to invest more or less in quality than a welfare maximising regulator with the same market structure.

have $\theta_S^{FI} \geq \theta_S^{ur}$. ■

The first element $(w-c)\frac{\partial Q_S^*}{\partial \theta}$ in eqn. (11) measures the marginal gain to gross profit from expanding the market (i.e., from investing in infrastructure quality), whereas the latter element $K'(\theta_S)$ gives the marginal cost of the investments.

First, the unregulated choice of network quality will never be identical to the socially optimal level of quality, except for the case where both are equal to zero. The reason for this is that the regulator will be able to internalise a larger part of the marginal gain due to investments, and in addition, will be able to internalise the increased willingness to pay for the inframarginal units (i.e., the direct increase in consumers' surplus). This implies that $\theta_S^{ur} \leq \theta_S^{FI}$.

Furthermore, we observe from eqn. (11) that if the access charge is set equal to marginal cost of providing access (i.e., $w = c$), the first-order condition for the vertically separated firm is always negative. Thus, the unregulated choice of infrastructure quality will be to set θ as low as possible, where the minimum quality level is normalised to 0. This observation is quite intuitive. Since the network provider's only income is from the sale of access, he has no incentives to improve the quality of the infrastructure if he is not allowed a return on his investments.

3.2.3 Quality comparisons with a linear inverse demand function

Assume that the demand function is given by $P(Q, \beta, \theta) = a + \beta\theta - Q$. In this case, changes in the infrastructure quality will affect the elasticity of demand, making demand less elastic when the quality level is high. Consequently, investment in infrastructure is beneficial (from the point of view of the firms) for two reasons; it raises consumers' willingness-to-pay for any given level of quantity, and demand becomes less elastic.[25] Furthermore, let us normalise the marginal cost of providing access to zero (i.e., $c \equiv 0$) without loss of generality, and let the cost of investing in quality be given by the quadratic cost function $K(\theta) = \varphi\frac{\theta^2}{2}$. I will restrict my attention to parameter values such that both downstream firms are active. To ensure this, the following must be satisfied: $a > w$.

[25] The elasiticity of demand is in this case given by: $\varepsilon = -P/(a + \beta\theta - P)$.

With these specifications, the model gives us the following explicit solutions for infrastructure quality investments in the three cases discussed above:[26]

$$\theta_I^{FI} = \frac{(8a - w)\,\beta}{9\varphi - 8\beta^2} \tag{12}$$

$$\theta_S^{FI} = \frac{(8a - 2w)\,\beta}{9\varphi - 8\beta^2} \tag{13}$$

$$\theta_I^{ur} = \frac{(2a + 5w)\,\beta}{9\varphi - 2\beta^2} \tag{14}$$

$$\theta_S^{ur} = \frac{2w\beta}{3\varphi} \tag{15}$$

Observe that contrary to case with a generalised demand function, the quality level is strictly positive if $w > 0$ (i.e., if the access charge is in excess of the marginal cost of providing access) in all cases.[27] Note also that an unregulated, vertically separated firm will choose not to invest at all if $w = 0$, since it has no other way of internalising the increased value of the higher quality level than through the price of access. Furthermore, we observe that $\theta_S^{FI} = \theta_I^{FI}$ when $w = 0$.

By comparing eqns. (12), (14) and (15), we can rank the investment levels. This ranking is summarised in the following lemma:[28]

Lemma 3 *When demand is linear, investment cost quadratic and there is full information about the spillover level, the socially optimal quality level is always higher than the unregulated levels. Furthermore, if firms are unregulated, the level of quality provided by a vertically integrated firm is always higher than the corresponding choice of a vertically separated firm (that is, $\theta_I^{FI} > \theta_S^{FI} > \theta_I^{ur} > \theta_S^{ur}$).*

The reason that the socially optimal solution is always strictly higher than the unregulated cases is discussed above, and is due to the fact that a benevolent regulator is able to internalise the positive external effects on consumers' surplus and

[26] I assume that all the relevant second-order conditions are met, which amounts to requiring that the investment costs are sufficiently convex: $9\varphi - 8\beta^2 > 0$. This condition is sufficient for all three cases under full information.

[27] Note that $\beta > 0$ by assumption.

[28] Note that a sufficient, but not necessary condition for the second inequality is $a \geq \frac{63}{54}w$.

the rival firm's profit of increasing the network quality. An unregulated and vertically integrated firm can internalise the benefit increasing the quality has on its own downstream subsidiary's profit. For an unregulated and vertically separated firm, investing in infrastructure quality has only the effect of increasing the total sales of access to independent downstream firms, and such a firm will only be able to internalise the gain of investing in the infrastructure through the price of access. Consequently, if the market is unregulated, a vertically integrated firm will choose to invest more in quality than a vertically separated firm.

4 Asymmetric Information

A realistic assumption is that the regulator does not know how investments affect demand perfectly. We assume that only the distribution and support of the demand scaling parameter β is known to the regulator.[29] He must therefore offer a regulatory contract which induces the regulated firm to reveal its true demand. The regulator will then maximising welfare, subject to incentive and participation constraints. As we will see, the optimal policy with respect to infrastructure quality depends, contrary to what is the case when there is full information, on the choice of market structure. The reason for this is that the information rent will be affected by the choice of market structure, and consequently, the real cost to the regulator of inducing a specific level of quality will be different.

The regulator offers the incentive compatible contract of the form $\left(\theta\left(\widehat{\beta}\right), t\left(\widehat{\beta}\right)\right)$ to the upstream firm to maximise the welfare function (5), using his knowledge about the distribution and support of the unknown parameter and the structure of the game.[30] The contract specifies infrastructure quality, θ, and transfer to the regulated firm, t.

The regulator maximises the expected value of the welfare function, with expectations taken over the level of spillovers, β, subject to participation and incentive

[29] Lewis and Sappington (1988) considers the regulation of a monopolist when demand is uncertain.

[30] $\widehat{\beta}$ is the report from the regulated firm to the regulator.

constraints.[31] In addition, the regulator knows that if the network owner is also present downstream, truthful revelation must be based on the joint profit function for the vertically integrated firm, i.e., equation (2). The reason for this is that the report of demand made to the regulator will internalise any effects that the report (and resulting infrastructure quality) has on its own downstream profits. If, on the other hand, there is vertical separation the relevant expression for the incentive constraint would be equation (3).

Lemma 4 *When there is vertical integration, (first-order) incentive compatibility requires the following:*

$$\frac{d\Pi^v}{d\beta} = \left(\theta P_\kappa + P_Q \frac{\partial q^{*i}}{\partial \beta}\right) q^{*v} + (w - c) \frac{\partial q^{*i}}{\partial \beta}$$

Proof. Apply the envelope theorem to (2), given that q^v and q^i are chosen optimally at stage 3. More generally, we can write the first-order incentive constraint as: $\frac{d\Pi^v}{d\beta^u} = \frac{\partial \Pi^v}{\partial \beta} + \frac{\partial \Pi^v}{\partial q^i}\frac{\partial q^{*i}}{\partial \beta} + \frac{\partial \Pi^v}{\partial q^v}\frac{\partial q^{*v}}{\partial \beta}$. However, using the first-order condition determining q^{*v}, this simplifies to $\frac{d\Pi^v}{d\beta} = \frac{\partial \Pi^v}{\partial \beta} + \frac{\partial \Pi^v}{\partial q^i}\frac{\partial q^{*i}}{\partial \beta}$. ■

We can distinguish between three different effects in the incentive constraint with vertical integration. The first component of the incentive constraint, $\theta P_\kappa q^{*v}$, is the *direct* effect, which would also be present in the absence of competition. The effect of increasing the degree of spillover implies that the value to the marginal consumer of a given level of infrastructure quality is higher and the downstream subsidiary of firm v can, all other things held constant, sell its output q^{*v} at a higher price. Furthermore, increasing the degree of spillover implies (under the assumptions I have made) that the rival's output q^{*i} increases, which raises firm v's profit in the local access segment (this is $(w-c)\frac{\partial q^{*i}}{\partial \beta}$). The expression $P_Q q^{*v} \frac{\partial q^{*i}}{\partial \beta}$ in the incentive constraint deals with the indirect effect on downstream profit of increasing the degree of spillover. Since increasing spillover implies a higher total output in equilibrium, there is a loss on all inframarginal units. It is easily shown that the sign of the

[31] The degree of spillover is common knowledge among all firms, and is revealed prior to the final stage of the game.

incentive constraint is greater or equal to zero, when demand is concave (or more generally, not too convex), for the whole support for β.

In the case of vertical separation, the incentive constraint is positive for all $\beta \in [\underline{\beta}, \overline{\beta}]$ provided that $w \geq c$, but takes on a simpler form as the provider of local access does not internalise the effects on downstream profits:

Lemma 5 *When there is vertical separation, (first-order) incentive compatibility requires the following:*

$$\frac{d\Pi^u}{d\beta} = -(w-c)\frac{2\theta P_\kappa}{3P_Q + P_{QQ}Q_S^*}$$

Proof. More generally, we can write the first-order incentive constraint as: $\frac{d\Pi^u}{d\beta} = \frac{\partial \Pi^u}{\partial \beta} + \frac{\partial \Pi^u}{\partial q_1}\frac{\partial q_1^*}{\partial \beta} + \frac{\partial \Pi^u}{\partial q_2}\frac{\partial q_2^*}{\partial \beta}$. Contrary to the case of vertical integration, the envelope theorem is not utilised here. Note that $\frac{\partial \Pi^u}{\partial q_1} = \frac{\partial \Pi^u}{\partial q_2} = (w-c)$. The proof makes use of the properties of the equilibrium quantities (proven in appendix 1; insert for the symmetric equilibrium quantities in the vertically separated case) to determine that $\frac{\partial q_1^*}{\partial \beta} + \frac{\partial q_2^*}{\partial \beta} = -\frac{2\theta P_\kappa}{3P_Q + P_{QQ}Q_S^*} \geq 0$. ■

We observe that $d\Pi^u/d\beta \geq 0$ for all $\beta \in [\underline{\beta}, \overline{\beta}]$, with equality only for $\theta = 0$ and/or $w = c$. When there is vertical separation, asymmetry of information benefits the regulated firm since changes in the degree of spillover affects the total size of the access market. The equilibrium profit of the vertically separated firm is increasing in the degree of spillover, which must be taken into account by an incentive compatible mechanism (i.e., the transfers will have to be increasing in the degree of spillover). A high degree of spillover implies that investing in infrastructure quality increases the demand for the final product, and consequently, the demand for access substantially. However, the value of the increase in demand for access naturally depends on the profit margin in the access segment. A vertically separated firm can only recoup the benefit from the investments by capturing some of the increased value of the downstream output through a positive profit margin on access (as opposed to the vertical integration case, where the regulated firm is able to internalise some of the value creation of the investment through its downstream subsidiary). Consequently, the value of the private information vanishes if the regulated firm does not make positive profits on the sale of access. A regulated firm will not be able to gain anything from pretending to have a different level of spillover than the realised level, since the access revenue is identical and equal to zero no matter what is reported.[32] In such a case, the transfer from the regulator will only have to cover the investment costs.

If the mechanism is not designed to achieve truthful revelation the regulated firm will, by imitating a firm with lower degree of spillover, be able to get a higher transfer combined with the profit of a high spillover firm. This implies that the firm is able to capture some additional profits. If the informational asymmetry is not taken into account when designing the regulation, a firm reporting a high degree of spillover will end up with a smaller transfer (information rent) than a firm (mis)reporting a lower degree of spillover. Thus, to be able to reveal the degree of spillover truthfully, the transfer must take this opportunity of misreporting into account.

4.1 Optimal policies

The regulator's optimisation problem is given by:

$$\max_{\theta} \int_{\underline{\beta}}^{\overline{\beta}} \left[CS(\beta,\theta) - cQ^*(\beta,\theta) - K(\theta) - \lambda \Pi^j(\beta,\theta) \right] dG(\beta) \tag{16}$$

where $j = (v, u)$ subject to either

$$\frac{d\Pi^v}{d\beta} = \theta P_\kappa q^{*v} + \left[(w-c) + P_Q q^{*v}\right] \frac{\partial q^{*i}}{\partial \beta} \tag{ICVI}$$

$$\Pi^v(\beta) \geq 0 \tag{PCVI}$$

[32]This result is similar to what Lewis and Sappington (1988) finds for a monopolist when demand is uncertain. In their model, the monopolist earns no information rent and has no incentives to report a demand different from the true demand.

if there is vertical integration, or

$$\frac{d\Pi^u}{d\beta} = -(w-c)\frac{2\theta P_\kappa}{3P_Q + P_{QQ}Q_S^*} \tag{ICVS}$$

$$\Pi^u(\beta) \geq 0 \tag{PCVS}$$

if there is vertical separation. In addition, the following constraint must be taken into account in both cases:

$$q^{*v} \geq 0,\ q^{*i} \geq 0 \tag{Cournot}$$

The last constraint is simply the Cournot equilibrium from the final stage of the game. Obviously, in the vertical separation case the equilibrium outputs are symmetric. The participation constraint ensures non-negativity of profits and must be satisfied. In addition to the constraints set out above, we also need to ensure that the optimal solution satisfies the second-order incentive constraint.[33]

Integrating the incentive constraint by parts, assuming that [PC] binds for the lowest level of spillover, $\underline{\beta}$, we get an expression for the regulator's virtual surplus (i.e., welfare adjusted to take account of the informational costs of inducing truthful revelation).

In the case of vertical integration, the virtual surplus is given by:

$$\begin{aligned} VS_{int} \;=\; & \int_{\underline{\beta}}^{\overline{\beta}} \left\{ \int_0^Q P(Q',\theta)\,dQ' - cQ_I^* - K(\theta) \right. \\ & \left. +\lambda \frac{1-G(\beta)}{g(\beta)} \left[\left((w-c) + P_Q q^{*v}\right)\frac{\partial q^{*i}}{\partial \beta} + \theta P_\kappa q^{*v} \right] \right\} dG(\beta) \end{aligned} \tag{17}$$

[33]Provided that $d\Pi^k/d\beta$ is increasing in θ, where $k = v, u$, then the second-order incentive constraint requires that the optimal solution for θ must be increasing in β. The proof of this is standard and is omitted. For details, consult Guesnerie and Laffont (1984).

If there is vertical separation, virtual surplus is:

$$\begin{aligned} VS_{sep} &= \int_{\underline{\beta}}^{\overline{\beta}} \left\{ \int_{0}^{Q} P(Q', \theta)\, dQ' - cQ_S^* - K(\theta) \right. \\ &\quad \left. -\lambda \frac{1-G(\beta)}{g(\beta)} \left[(w-c) \frac{2\theta P_\kappa}{3P_Q + P_{QQ}Q_S^*} \right] \right\} dG(\beta) \end{aligned} \tag{18}$$

4.1.1 Optimal infrastructure quality

Maximising the integrand of expression (17) and (18) pointwise with respect to θ, we obtain the following first-order condition for optimal quality under asymmetric information, θ_j^{AI}, for $j = I, S$:[34]

$$\left(P\left(Q_j^*, \beta\theta_j^{AI}\right) - c\right) \frac{\partial Q_j^*\left(\theta_j^{AI}\right)}{\partial \theta} + \beta P_\kappa Q_j^*\left(\theta_j^{AI}\right) = K'\left(\theta_j^{AI}\right) + ICT_j \tag{19}$$

The only difference from the full information solution is the information correction term, ICT_j. The exact characteristics of the incentive correction terms is summarised in Lemma 6:

Lemma 6

a) If there is vertical integration, the incentive correction term is given by:

$$\begin{aligned} ICT_I = \lambda \frac{1-G(\beta^u)}{g(\beta^u)} &\left[\left(w - c + P_Q q^{*v}\right) \frac{\partial^2 q^{*i}}{\partial\theta\partial\beta} \right. \\ + \left(P_Q \frac{\partial q^{*v}}{\partial\theta} + P_{QQ} q^{*v} \frac{\partial Q_I^*}{\partial \theta} \right) &\left. \frac{\partial q^{*i}}{\partial \beta} + \left(P_\kappa + \beta\theta P_{\kappa\kappa}\right) q^{*v} + \theta P_\kappa \frac{\partial q^{*v}}{\partial \theta} \right] \end{aligned} \tag{20}$$

b) If there is vertical separation, the incentive correction term is given by:

$$ICT_S = -\lambda \frac{1-G(\beta)}{g(\beta)} \left[(w-c) \frac{2P_\kappa \left(3P_Q + P_{QQ}Q_S^*\right) - 2\theta P_\kappa P_{QQ}\left(\partial Q_S^* / \partial \theta\right)}{\left(3P_Q + P_{QQ}Q_S^*\right)^2} \right] \tag{21}$$

The information correction term ICT_I is positive if the information rent increases in θ, or equivalently, $\partial^2 \Pi^{*v} / \partial\theta\partial\beta > 0$.[35] If the incentive correction term is positive,

[34]The second-order incentive constraint is assumed to be satisfied. The following conditions are sufficient, but not necessary, to ensure this (subscripts denote partial derivatives): i) $\frac{d}{d\beta}\left(\frac{1-G(\beta)}{g(\beta)}\right) \leq 0$, ii) $\Pi_{\beta\theta} \geq 0$, iii) $\Pi_{\beta\beta\theta} \leq 0$, iv) $Q_{\beta\theta} \leq 0$, and v) $CS_{\beta\theta} \geq 0$. It can be shown that the necessary conditions for second-order incentive compatibility are satisfied in the linear inverse demand case.

[35]This is the single-crossing condition on equilibrium profit.

this has the effect of increasing the regulator's costs, which implies that optimal quality should be lower than in the full information case. In addition, there is "no distortion at the top". These are standard results in many models of asymmetric information. The term correcting for the asymmetry of information is, however, not unambiguously positive and the sign is difficult to determine in general.

If ICT_S is positive, then the asymmetry of information dictates a lower level of quality compared to the full information solution. If $w > c$, then $ICT_S > 0$ if the following inequality is satisfied: $(3P_Q + P_{QQ}Q_S^*)^2 + 2\beta\theta P_\kappa P_{QQ} \geq 0$.[36] This is equivalent to saying that the rate at which equilibrium output increases with the spillover level is increasing in the level of quality. This inequality will be satisfied for, e.g., linear and quadratic inverse demand functions.

Observe that if $w = c$ when there is vertical separation, the term correcting for the asymmetry of information vanishes. In such a case, we have seen that the value of the regulated firm's private information vanishes. Consequently, the socially optimal policy on quality investment will be unaffected by the asymmetry of information if access is priced at the marginal cost of providing the access service, and the socially optimal solutions under full and asymmetric information coincides. The regulator will, if $w = c$, simply set a transfer equal to the cost of investing in the infrastructure and leave the regulated firm zero economic profit.

4.1.2 Optimal quality levels with linear demand

Let us now examine the explicit solution when the demand side and costs are specified as above. I assume here that the degree of spillover is uniformly distributed on the interval $\left[\underline{\beta}, \overline{\beta}\right]$. Solving first the regulator's maximisation problem in the vertical integration case when there is asymmetric information about the level of spillover, we find that the socially optimal level of quality is given by:[37]

$$\theta_I^{AI} = \frac{(8a - w)\,\beta + (2a - w)\,\lambda\left(\beta - \overline{\beta}\right)}{9\varphi - 4\beta\left(2\beta + \lambda\left(\beta - \overline{\beta}\right)\right)} \tag{22}$$

[36]This inequality is obtained by inserting for $\partial Q_S^*/\partial\theta$ in ICT_S.

[37]I assume that the second-order conditions are met, which ensure that the denominators in the expressions for optimal quality are positive.

If the upstream firm is not allowed to compete downstream, but must produce access for two independent downstream firms, the socially optimal level of quality is given by:

$$\theta_S^{AI} = \frac{(8a - 2w)\beta + 6w\lambda\left(\beta - \overline{\beta}\right)}{9\varphi - 8\beta^2} \tag{23}$$

Observe that when there is asymmetric information about the degree of spillover and the regulator must pay the regulated firm an informational rent, it may happen that the optimal quality levels as defined by eqns. (22) and (23) are negative for some parameter combinations, which cannot happen when there is full information. This may, in particular, happen if the degree of spillover is sufficiently small and the shadow cost of public funds is sufficiently large. I assume that quality investments cannot be negative, so in those cases, the regulator will induce the regulated firm not to invest at all. In the linear demand case, we obtain the standard result that the quality for all but the highest spillover type is distorted downwards; i.e., the socially optimal level of quality is always higher under full information relative to the asymmetric information solution when there is vertical integration and separation, respectively (i.e., $\theta_j^{FI} \geq \theta_j^{AI}$ for $j = I, S$).

Note that from eqn. (23) we see that the optimal level of infrastructure quality is strictly positive when $w = 0$, and that this level corresponds to the socially optimal level under full information given by eqn. (12). The reason for this has been discussed above, and relates to the fact that the regulated firm does not possess valuable private information when the access charge is equal to marginal cost. However, if there is vertical integration and asymmetric information, the regulator will want to choose a level of quality which is different to the full information case even if $w = 0$, except for the highest level of spillover $\beta = \overline{\beta}$ (i.e., no distortion at the top). With vertical integration the cost of inducing truthful revelation is, for $\beta \neq \overline{\beta}$, never zero. Thus, the presence of asymmetry of information will influence the solution even with marginal cost pricing of access. By distorting the quality level downward relative to the full information solution it becomes less profitable to imitate firms with lower degrees of spillover, which lowers the information rent. A lower level of quality implies that the value of the final product is lower at the margin, and furthermore, it lowers the profit from the access segment due to a lower total equilibrium output. Both these effects reduce the incentives to imitate lower spillover types. However, lower quantity implies higher prices for the inframarginal consumers in the downstream market, but this effect is dominated by the two former effects.

4.2 The effect of the vertical structure on regulatory policies

In this section I will discuss how changes in the choice of the vertical structure affects the choice of regulatory policy. As we have seen, the vertical structure has an impact on the incentive compatibility constraint, and consequently, changes in the vertical structure may affect the level of information rents necessary to induce truthful revelation. This may have an impact on the socially optimal quality levels under vertical separation and integration.

Before the formal analysis is undertaken, I will briefly discuss the intuition behind the results. The positive external effects on the downstream market of investments is an argument for regulating the infrastructure quality. We have seen above that the socially optimal level of quality is higher with vertical integration than with vertical separation when there is full information. This is due to the vertically integrated firm's ability to internalise the vertical externality, and produce at lower cost than when there is vertical separation. When there is asymmetric information about the spillover levels, however, the informational cost is in itself an argument against regulation of quality. The cost of regulation in terms of the information rent must then be weighed against the benefit of regulation (internalising the positive external effect) to determine whether regulation should take place. The magnitude of the information rent will naturally be important in this trade-off, and as we will see below, the choice of market structure is important in determining the level of the information rent. In the asymmetric information case, we find that the socially optimal quality in some cases is higher under vertical separation than under vertical integration due to the effect that differences in the vertical structure has on the information rent.

4.2.1 Comparison of the information rents

The *information rent* in the case of vertical separation may be written as (assuming that the participation constraint binds for the type with the lowest level of spillover, or the least efficient type):[38]

$$\Pi_S^u(\beta) = \int_{\underline{\beta}}^{\beta} \left[(w-c) \frac{\partial Q_S^*}{\partial \beta} \right] d\tau \tag{24}$$

and for vertical integration:

$$\Pi_I^v(\beta) = \int_{\underline{\beta}}^{\beta} \left[\theta_I P_\kappa q_I^{*v}(\theta_I, \tau) + \left[(w-c) + P_Q q_I^{*v}(\theta_I, \tau) \right] \frac{\partial q_I^{*i}}{\partial \beta} \right] d\tau \tag{25}$$

Subscript S and I refer to vertical separation and vertical integration, respectively. The difference in information rents is defined as: $\Delta R = \Pi_I^v(\beta) - \Pi_S^u(\beta)$. The sign of ΔR clearly depends on optimal infrastructure quality (under vertical separation and integration), among other things, and is determined by the following expression:

$$\Delta R = \int_{\underline{\beta}}^{\beta} \left[\theta P_\kappa q_I^{*v}(\theta_S, \tau) - (w-c) \left(\frac{\partial Q_S^*}{\partial \beta} - \frac{\partial q_I^{*i}}{\partial \beta} \right) + P_Q q^{*v}(\theta_I, \tau) \frac{\partial q_I^{*i}}{\partial \beta} \right] d\tau \tag{26}$$

Proposition 3 *The information rent to the regulated firm is higher under vertical integration compared to vertical separation.*

Sufficient conditions for the expression (26) to be non-negative; i.e., $\Delta R \geq 0$, is that the inverse demand function is concave and that assumption 1 is satisfied.[39] Consequently, under the assumptions in this paper, the information rent to a vertically integrated and regulated firm will always be at least as large as the information rent if there is vertical separation.

[38] To save on notation, I only include the variables related to upstream costs and the infrastructure quality. Subscripts S and I refer to the cases of vertical separation and vertical integration, respectively. Note that the derivatives of the equilibrium downstream quantities are functions of the parameter β.

[39] The result is maintained even if the inverse demand function is convex, provided that demand is not too convex (see, e.g., Hahn, 1962).

4.2.2 Comparison of optimal quality

When comparing the optimal quality under vertical separation and integration, respectively, we need to compare the first-order condition (19) for $j = I, S$. We have seen that when $w = c$, the full information solution is identical both under vertical integration and separation. Thus, the only difference between these two conditions when $w = c$ will be due to differences in how the asymmetry of information affects the solution, and in that special case we need only compare the information correction terms, ICT_S and ICT_I, given in Lemma 6.

Let us first examine the difference in the incentive correction terms, before examining the difference in optimal quality in the two cases (the comparison of quality will only be undertaken in the simplified case of linear inverse demand and convex investment costs).

Define $\Delta ICT \equiv ICT_S - ICT_I$, which is given by:

$$\begin{aligned}\Delta ICT &= -\left[(w-c)\left(\frac{2P_\kappa\left(3P_Q + P_{QQ}Q_S^*\right) - 2\theta P_\kappa P_{QQ}\left(\partial Q_S^*/\partial\theta\right)}{\left(3P_Q + P_{QQ}Q_S^*\right)^2} + \frac{\partial^2 q_I^{*i}}{\partial\theta\partial\beta}\right)\right. \\ &\quad \left. + \left(P_Q\frac{\partial q_I^{*v}}{\partial\theta} + P_{QQ}\frac{\partial Q_I^*}{\partial\theta}\right)\frac{\partial q_I^{*i}}{\partial\beta} + \left(P_\kappa + \beta\theta P_{\kappa\kappa}\right)q_I^{*v} + \theta P_\kappa\frac{\partial q_I^{*v}}{\partial\theta}\right]\lambda\frac{1-G(\beta)}{g(\beta)}\end{aligned} \tag{27}$$

The sign of the expression (27) is difficult to establish with a general demand function. I will therefore only consider the case with a linear inverse demand, with $P(Q,\theta) = a + \beta\theta - Q$. The marginal cost of access provision is normalised to zero; i.e., $c = 0$. In this case, the difference in the information correction terms can be rewritten as:

$$\Delta ICT = -\lambda\frac{1-G(\beta)}{g(\beta)}\left(\frac{3a + 5\beta\theta_I}{9}\right) \tag{28}$$

From eqn. (28) we observe that the sign on ΔICT will be negative, and strictly negative for $\beta > \underline{\beta}$, which means that the information correction term under vertical integration is always strictly larger than in the case of vertical separation. This tells us that the marginal information rent is larger if there is vertical integration, which in isolation tends towards lower quality under vertical integration than with vertical separation. However, with vertical integration some of the vertical externalities between the upstream and downstream markets can be internalised which implies that the cost of providing a certain output downstream is lower with vertical integration. This tends towards higher quality when there is vertical integration than when there is vertical separation. Consequently, the only reason for the socially optimal level of quality to be higher under vertical integration when there is asymmetric information is that the cost advantage due to internalisation of the vertical externality becomes sufficiently large. When this happens, the benefit of a lower marginal information rent with vertical separation is dominated by the cost advantage of producing with vertical integration.

Since the marginal information rent (or information correction term) always is higher under vertical integration, we know that the socially optimal level of infrastructure quality is strictly higher under vertical separation when there is marginal cost pricing of access.

More generally (i.e., when $w > c$) it can be shown that the regulated access charge with vertical separation is larger than under vertical integration if the following inequality is satisfied:

$$\begin{aligned}&\left(9\varphi - 8\beta^2\right)\left(-w\beta - \lambda\left(\beta - \overline{\beta}\right)(2a - 7w)\right) \\ &-4\beta\lambda\left(\beta - \overline{\beta}\right)\left((8a - 2w)\beta + 6w\lambda\left(\beta - \overline{\beta}\right)\right) \geq 0\end{aligned} \tag{29}$$

We observe that for $w = 0$, the inequality is indeed satisfied, which implies that marginal cost pricing of access results in a higher quality level with vertical separation than with vertical integration. By continuity, this also holds for small access charges in excess of marginal cost. When the access charge becomes sufficiently large, the inequality (29) switches sign. The reason for this is that when there is vertical separation and the access charge becomes sufficiently large, the cost disadvantage imposed by the vertical externality of the vertical separation case will come to dominate the difference in the incentive correction terms. This is summarised in the following proposition:

Proposition 4 *Let us assume that the inverse demand function is linear.*

a) If the access charge is sufficiently low (where $w_{crit} \geq w \geq 0$), then the socially optimal level of infrastructure quality is strictly larger under vertical separation compared to the case of vertical integration; i.e., $\theta_S^{AI} \geq \theta_I^{AI}$.

b) If the access charge is sufficiently high (i.e., $w > w_{crit}$), then the socially optimal quality level is larger under vertical integration; i.e., $\theta_I^{AI} > \theta_S^{AI}$.

c) For the highest spillover level (i.e., $\beta = \overline{\beta}$), the socially optimal quality level is highest under vertical integration.

Proof. Solving eqn. (29) for w, we find a critical value for the access charge $w_{crit} \equiv -\frac{\left(32a\beta^2\lambda\left(\beta-\overline{\beta}\right)-2a\lambda\left(\beta-\overline{\beta}\right)\left(9\varphi-8\beta^2\right)\right)}{\left(9\varphi-8\beta^2\right)\left(7\lambda\left(\beta-\overline{\beta}\right)-\beta\right)-4\beta\lambda\left(\beta-\overline{\beta}\right)\left(6\lambda\left(\beta-\overline{\beta}\right)-2\beta\right)}$. If $w > w_{crit}$, then $\theta_I^{AI} > \theta_S^{AI}$. If $w \leq w_{crit}$, then $\theta_I^{AI} \leq \theta_S^{AI}$. For $\beta = \overline{\beta}$ we observe that $w_{crit} = 0$, and consequently, since $w \geq 0$ per assumption, then $\theta_I^{AI} \geq \theta_S^{AI}$. ■

The intuition behind this result can be explained by examining how the (first-order) incentive constraints are affected by changes in the regulated quality, given by lemmas 4 and 5. The profit of the regulated firm at the margin increases faster when the degree of spillover increases if the firm is vertically integrated, which results in a higher marginal information rent. The value to the regulated firm of having superior information depends on whether there are restrictions on which segments the firm can operate in, and the value of private information comes from both the upstream and downstream segments. Restricting the regulated firm's opportunity to operate in both markets will also restrict the earning potential of the firm (i.e., it will limit the value of the information, and thereby the level of the information rent).

When the degree of spillover is equal to the upper support (i.e., $\beta = \overline{\beta}$), we observe that the socially optimal quality is higher if there is vertical integration compared to the vertical separation case. This is equivalent to the situation under full information. This is, of course, due to the fact that there is "no distortion at the top"; the regulator chooses to distort all types' allocations except for the highest spillover type.

4.2.3 Optimal versus unregulated quality under asymmetric information

Is it necessary for the regulator to regulate infrastructure quality? In this section I will examine how the socially optimal solutions under vertical integration and separation correspond to the unregulated outcomes. The main point is then to characterise when the the industry should be subject to regulation of the infrastructure quality, and when should the regulator leave the market to itself? To say something about this, we need to compare the socially optimal outcomes and the unregulated outcomes under vertical integration and separation. To simplify the analysis, I will only examine the case with linear inverse demand and quadratic investment costs (as described above).

Vertical integration Assume that there are no vertical restrictions on the upstream monopolist, and that this firm finds it profitable to enter the downstream market. If the firm is subject to regulation, we know from eqn. (25) that the regulator must pay the vertically integrated firm a strictly positive information rent, which is socially costly. Furthermore, if the quality level in the regulated case is lower than the unregulated outcome, we get a lower level of quality at a higher cost. Thus, it is in society's interest to avoid regulation if this is the case.

Define $\Delta_I \equiv \theta_I^{AI} - \theta_I^{ur}$, where θ_I^{AI} is defined by eqn. (22) and θ_I^{ur} is defined by eqn. (14). Then, we have the following:

$$\begin{aligned} \Delta_I &= 2\beta^2 \left[21w\beta + (2a + 11w)\lambda\left(\beta - \overline{\beta}\right)\right] \\ &\quad +9\varphi\left[6(a-w)\beta + (2a-w)\lambda\left(\beta - \overline{\beta}\right)\right] \end{aligned} \tag{30}$$

It is difficult to determine the sign of eqn. (30), as there are opposing effects, but some observations can be made. First, we see that the magnitude of the spillover parameter β plays an important role. For $\beta = \overline{\beta}$, we find that $\Delta_I > 0$, which implies that an unregulated, vertically integrated firm will choose a socially suboptimal level of infrastructure quality when the spillover is at its highest. However, Δ_I becomes smaller when the degree of spillover is reduced, and for a sufficiently low level of spillover the sign on Δ_I changes and becomes negative.[40] Furthermore, these

[40]Numerical calculations suggest that switching level for the spillover parameter may be as low

opposing effects are present even when the access charge is set equal to zero (i.e., equal to the marginal cost).

For a sufficiently low degree of spillover the unregulated firm will choose a higher level of quality than the regulated quality level, and we know that an increase in the level of quality has a beneficial effect on welfare (*ceteris paribus*). Furthermore, no regulation implies that the regulator does not have to pay socially costly information rents, which is also a benefit. Consequently, if the degree of spillover is sufficiently low and there is vertical integration, then there should be no regulation.[41]

Vertical separation Let us now assume that the upstream monopolist is prohibited from entering the downstream market. In this case, the information rent payable to the regulated firm is given by eqn. (24), and we have already seen that the level of information rent approaches zero when the access charge approaches the marginal cost of providing access.

Define $\Delta_S = \theta_S^{AI} - \theta_S^{ur}$, where θ_S^{AI} is defined by eqn. (15) and θ_S^{ur} is defined by eqn. (15). Then we have the following:

$$\Delta_S = \left(24\,(a-w)\,\beta + 18w\lambda\left(\beta - \overline{\beta}\right)\right)\varphi + 16w\beta^3 \tag{31}$$

Again, as in the vertical integration case, there are opposing effects which make it difficult to determine the sign on eqn. (31). However, some observations can be made. First, if either $w = 0$ or $\beta = \overline{\beta}$, then we find that Δ_S is strictly positive (provided, of course, that $\beta > 0$). In this case, we know from eqn. (15) that an unregulated firm will invest nothing in quality. Thus, marginal cost pricing of access together with vertical separation implies that regulated level of quality is strictly larger than the unregulated level for all positive degrees of spillover (contrary to the vertical integration case). Furthermore, we know from eqn. (24) that the

as $\beta \approx 0.29$, but the switching level changes when both the level of the access charge and the shadow cost of public funds change. More specifically, the switching level is increasing in both w and λ.

[41] This essentially tells us that a socially optimal regulatory mechanism should specify: 1) a contract which depends on the level of spillover reported only if β is larger than some critical level, and 2) no regulation below this level.

information rent to the regulated firm is zero in this case. Consequently, with marginal cost pricing of access and vertical separation, the regulator should regulate the quality level, provided that the social cost of investments (which is transferred to the regulated firm) is smaller than the welfare loss due to the lower unregulated quality level. However, for an access charge in excess of marginal cost, $w > c$, we have a similar story to the vertical integration case above. Then, for a sufficiently low level of the spillover parameter β, the sign on Δ_S changes from positive to negative.[42]

4.2.4 Welfare considerations

From eqn. (21) we can deduce that if the access charge is cost-based and, more specifically, equal to the marginal cost of providing access, $w = c$, then vertical separation implies that the full information solution can be implemented.[43] The only transfer necessary is to cover the investment costs. Thus, if the regulator invokes vertical separation prior to the start of the game, he will in the present model be able to implement the best possible outcome at no extra cost when there is asymmetric information about the effect of investments on demand. Under vertical integration the socially optimal quality level will be strictly less than the full information solution, and the regulator must pay the regulated firm a strictly positive information rent in addition to the cost of investment. This implies that a welfare maximising regulator will be able to generate a higher level of welfare in the vertical separation case relative to the vertical integration case if the infrastructure quality is subject to regulation and if the regulated quality level is higher under vertical separation. As we have seen above, the latter will be satisfied for a sufficiently low level of the access charge and is certainly satisfied for marginal cost pricing of access.

If, however, infrastructure quality remains unregulated no information rent is necessary, which is beneficial in terms of overall welfare. On the other hand, in some

[42]Numerical calculations suggest that, for an access charge in excess of marginal cost, the switching level on the spillover parameter is increasing in λ (the shadow cost of public funds).

[43]If $w > c$ then the solution under asymmetric information and vertical separation will, of course, be closer to the full information solution the closer w is to c.

cases the socially optimal level of quality (the regulated level of quality) is higher than the level an unregulated, profit-maximising firm would choose. This is the case if, e.g., the degree of spillover is sufficiently high. Since quality increases overall welfare, a suboptimal level of quality is detrimental to welfare. If the regulated quality level of higher than the unregulated level, the gain in terms of welfare of the higher quality level must be weighed against the social cost of awarding the regulated firm an informational rent.

5 Concluding remarks

This paper has analysed the regulation of infrastructure quality in vertically related markets in the presence of asymmetric information about the effect of these investments on the demand for the downstream product. Furthermore, the effects of differences in market structure - vertical integration versus vertical separation - are investigated. There are some external effects associated with the investments, both on the consumers' surplus and on the profit of rival firm(s), which can be internalised through regulation. This is, in isolation, an argument in favour of regulating the infrastructure quality investment decision. The vertical structure of the market, and the assumptions about the firms' cost structures in the present model, also imply that the socially optimal quality is higher under vertical integration because the vertical externality is avoided. However, the presence of asymmetric information about the effect of investments on demand has an impact on the information rent, which may reverse the result.

The results in the paper also suggest that regulating the infrastructure quality is not necessarily the socially optimal policy, even if the regulated quality is shown to be higher than the unregulated quality levels. When the government regulates quality, truthful revelation necessitates the payment of an information rent and transfer to cover (in part or in full) the investment costs, which is assumed to be socially costly due to distortionary taxation. Provided that the effect of investments on demand is sufficiently low, it is better to let the quality investment decision be taken by the investing firm without interference, since no transfers are necessary in the unregulated case.

The cost structure of the present model does not incorporate economies of scope between network and end-user services. This is obviously a simplification, and may not be a reasonable assumption for some potential applications of the model. The presence of economies of scope may modify some of the results. With regulation I show that the information rent is lower under vertical separation relative to vertical integration, which in isolation is an argument in favour of invoking vertical separation. I argue that welfare is higher with vertical separation in the regulated case at least if the access charge is sufficiently low. However, with economies of scope the result may be reversed for a given level of the access charge. Put differently, for vertical separation to yield higher welfare than vertical integration in the presence of economies of scope it is necessary to reduce the access charge compared to the scenario without economies of scope.

6 References

Baron, D.P. and R. Myerson (1982), "Regulating a monopolist with unknown cost", *American Economic Review* **50**, 911-930

Clark, D. (1999), "High-Speed Data Races Home", *Scientific American*, October. Available at: http://www.sciam.com/1999/1099issue/1099clark.htm

D'Aspremont and A. Jacquemin (1988), "Cooperative and non-cooperative R&D in duopoly with spillovers", *American Economic Review* **78**, 1133-1137

Fudenberg, D., and J. Tirole (1984), "The Fat Cat Effect, the Puppy Dog Ploy and the Lean and Hungry Look", *American Economic Review, Papers and Proceedings* **74**, 361-368

Guesnerie, R. and J.J. Laffont (1984), "A complete solution to a class of principal-agent problems with an application to the control of a self-managed firm", *Journal of Public Economics* **25**, 329-369

Hahn, F. (1962), "The stability of the Cournot oligopoly solution concept", *Review of Economic Studies* **29**, 329-331

Laffont, J.J. and J. Tirole (1993), *A theory of incentives in procurement and* *regulation*, Cambridge, MA: MIT-Press

Laffont, J.J. and J. Tirole (2000), *Competition in telecommunications*, Cambridge, MA: MIT-Press

Lewis, T. and D. Sappington (1988), "Regulating a monopolist with unknown demand", American Economic Review 78, 986-998

Spence, M.A. (1975), "Monopoly, quality, and regulation", *Bell Journal of Economics* **6**, 417-429

Tirole, J. (1988), *The theory of industrial organization*, Cambridge, MA: MIT-Press

7 Appendix

Appendix 1

Proof of lemmas 1 and 2

In this appendix we will show the properties of the equilibrium quantities of the final stage only. The system of first-order conditions is given by:

$$\frac{\partial \Pi^v}{\partial q^v} = P_Q q^{*v} + (P(Q^*, \beta\theta) - c) = 0 \tag{32}$$

and

$$\frac{\partial \Pi^i}{\partial q^i} = P_Q q^{*i} + (P(Q^*, \beta\theta) - w) = 0 \tag{33}$$

Equations (32) and (33) defines a system of first-order conditions, which can be used to determine the effects on equilibrium quantities of changing exogenous variables. The exogenous variables at the final stage of the game are: w (access charge), θ (infrastructure quality), β (spillover), and c (upstream marginal cost).

To save on notation, let us define x to be the vector of exogenous variables at the final stage of the game: $x = (w, \theta, c, \beta)$. Total differentiation of the system of equations yields:

$$\begin{aligned}
\frac{\partial^2 \Pi^v}{\partial x \partial q^v} + \frac{\partial^2 \Pi^v}{\partial^2 q^v}\frac{\partial q^{*v}}{\partial x} + \frac{\partial^2 \Pi^v}{\partial q^i \partial q^v}\frac{\partial q^{*i}}{\partial x} &= 0 \\
\frac{\partial^2 \Pi^i}{\partial x \partial q^i} + \frac{\partial^2 \Pi^i}{\partial q^v \partial q^i}\frac{\partial q^{*v}}{\partial x} + \frac{\partial^2 \Pi^i}{\partial^2 q^i}\frac{\partial q^{*i}}{\partial x} &= 0
\end{aligned}$$

or by rearranging:

$$\begin{pmatrix} \frac{\partial^2 \Pi^v}{\partial^2 q^v} & \frac{\partial^2 \Pi^v}{\partial q^i \partial q^v} \\ \frac{\partial^2 \Pi^i}{\partial q^v \partial q^i} & \frac{\partial^2 \Pi^i}{\partial^2 q^i} \end{pmatrix} \begin{pmatrix} \frac{\partial q^{*v}}{\partial x} \\ \frac{\partial q^{*i}}{\partial x} \end{pmatrix} = - \begin{pmatrix} \frac{\partial^2 \Pi^v}{\partial x \partial q^v} \\ \frac{\partial^2 \Pi^i}{\partial x \partial q^i} \end{pmatrix}$$

Let us call the 2×2 matrix A. This system can be solved using Cramer's rule. The determinant to the matrix A, $|A| = [P_Q (3P_Q + P_{QQ}Q)]$, is positive under assumption 2. This yields the following:

$$\begin{array}{ll}
\frac{\partial q^{*v}}{\partial \theta} = \frac{\beta P_\kappa \left(P_{QQ}\left(q^{*v}-q^{*i}\right)-P_Q\right)}{|A|} & \frac{\partial q^{*i}}{\partial \theta} = \frac{\beta P_\kappa \left(P_{QQ}\left(q^{*i}-q^{*v}\right)-P_Q\right)}{|A|} \\
\frac{\partial q^{*v}}{\partial \beta} = \frac{\theta P_\kappa \left(P_{QQ}\left(q^{*v}-q^{*i}\right)-P_Q\right)}{|A|} & \frac{\partial q^{*i}}{\partial \beta} = \frac{\theta P_\kappa \left(P_{QQ}\left(q^{*i}-q^{*v}\right)-P_Q\right)}{|A|} \\
\frac{\partial q^{*v}}{\partial w} = -\frac{P_Q + P_{QQ}q^{*v}}{|A|} > 0 & \frac{\partial q^{*i}}{\partial w} = \frac{2P_Q + P_{QQ}q^{*v}}{|A|} < 0 \\
\frac{\partial q^{*v}}{\partial c} = \frac{2P_Q + P_{QQ}q^{*i}}{|A|} < 0 & \frac{\partial q^{*i}}{\partial c} = -\frac{P_Q + P_{QQ}q^{*i}}{|A|} > 0
\end{array}$$

The properties of total downstream demand is simply the sum of the effects on the two firms' demands: $\partial Q^*/\partial \beta = -2\theta P_\kappa / (3P_Q + P_{QQ}Q^*) > 0$, $\partial Q^*/\partial \theta = -2\beta P_\kappa / (3P_Q + P_{QQ}Q^*) > 0$, $\partial Q^*/\partial c = \partial Q^*/\partial w = 1/(3P_Q + P_{QQ}Q^*) < 0$.

ESSAY 2

Regulation of a vertically differentiated duopoly*

Jan Yngve Sand†

April 2001

Abstract

This paper focuses on the optimal quality regulation of vertically differentiated duopolies in the presence of asymmetric information. In the model presented there are cross-effects on the information rent. Contrary to standard single-agent models, the production levels are distorted in favour of the most efficient firm, whose production level is increased under asymmetric information relative to full information. The first-best outcomes can only be achieved if both firms are of the most efficient types. The optimal degree of vertical differentiation is also discussed. Furthermore, some extensions to the model are examined (the presence of cost complementarity, quality as complements etc.).

JEL classification: D82, L51, L96

Keywords: Vertical differentiation, asymmetric information,
duopoly, regulation, quality

---

*I would like to thank Kare P. Hagen, Derek Clark, Petter Osmundsen, Patrick Rey, Emmanuelle Auriol, and Nils-Henrik von der Fehr for valuable comments on an earlier draft.

†University of Tromso, NFH, Department of Economics, N-9037 Tromso, Norway, tel. +47 77 64 60 00, fax. + 47 77 64 60 20, e-mail: Jan.Sand@nfh.uit.no. Financial support from Telenor and Norges Banks Fond til Økonomisk Forskning is gratefully acknowledged.

1 Introduction

In a situation of increasing competition, many firms will attempt to differentiate their products from those of their competitors to earn higher levels of profits. In the market for (tele)communications services, one way of doing this would be to vary the degree of functionality (or quality) of a service. With the introduction of digital technology, providers of telecommunications services have a wide range of options on how to do this. The old copper wires previously used for analogue telephony only, can now be used to supply, e.g., digital telephony, ISDN (Integrated Service Digital Network) or ADSL (Asymmetric Digital Subscriber Line). All of which no doubt will increase the value of the subscription to consumers.

It is reasonable to assume that some kind of regulation will remain in this market for some time still as efficient competition establishes itself. One type of regulation which would be reasonable to investigate more closely is the regulation of quality. A frequent (unwanted) consequence of competition is the degradation of quality. This fear is certainly well-founded if the industry is regulated by price-caps, as price-caps normally do not give benefits to firms for increasing the quality of the product they sell.[1] Quality of access to services such as telecommunications is an important aspect of the communications industry, and is often emphasised in the political debate. It is therefore not unlikely that some sort of regulation of quality will be considered for this sector, to ensure that the benefits of competition does not come at the expense of the quality offered.

Here I focus on the case where a regulatory body induces the two firms to choose certain levels of quality, using a direct revelation mechanism, which implies that the extent of vertical differentiation is in fact induced through the regulation mechanism. I show that the regulator may find it welfare improving to induce the firms to supply different levels of quality. In a (tele)communications setting, this could amount to supplying different speeds of connection, or different choices of technology.[2] These

[1]Following the deregulation of the UK telecommunications industry quality did not fall, but did not increase as much as would be expected taking into account the technical advances which were made (Armstrong, Cowan and Vickers, 1994).

[2]We will see that the most efficient firm (with the lowest unit cost of producing *quantity*) is

issues are closely related to the literature on (strategic) investments in R&D, but the majority of this literature considers situations with full information.[3] In the R&D literature the external effect imposed on rival firms is (usually) positive provided that the degree of spillover between firms is sufficiently high.

An alternative interpretation of the model could be the *regulation of advertising.* In such a case, one could assume that the firms produce goods which are identical physically, and the firms use advertising to attempt to distinguish their product from that of the competitor and thereby capturing market shares. Such advertising expenditure is socially wasteful. One concern could be that opening up a sector to free competition results in, at best, no positive effect on welfare (or even a loss of welfare) due to excess spending on advertising, and a regulator may want to limit such behaviour. In such a setting the result that the most efficient firm is allowed a higher level of the regulated variable (here, advertising expenditure) is interpreted as a measure to allow this firm to capture a larger share of the market.

The main goal of this paper is to provide a starting point for further analysis of the regulation of multi-agent settings, where the agents (firms) produce vertically differentiated products. The level of quality is the means for differentiating the products. An analysis of multiproduct monopolies would have similar features as the model presented below, but there is an important difference. In a model of the optimal regulation of a multiproduct monopoly the incentive correction has the standard property that allocations are not distorted for the most efficient type, commonly known as "no distortion at the top". However, in the model of the optimal regulation of single-product duopolies, this is not generally the case. The result of "no distortion at the top" requires either that both firms are of the most efficient type, or that firms are equally inefficient. Thus, the standard properties of single-agent models does not necessarily carry over to multi-agent settings.

Optimal regulation of duopolies under asymmetric information is not very well

induced to supply high quality access (e.g., broadband, and possibly wireless access). The other firm, which has a higher unit cost, is induced to supply access of lower quality (access with lower capacity and speed).

[3]D'Aspremont and Jacquemin (1988) and Kamien, Muller and Zang (1992) consider the strategic incentives to invest in R&D.

covered in the literature. Furthermore, the literature on regulation and quality concerns is relatively scarce. Other examples of regulation and quality can be found in Laffont and Tirole (1993: chapter 4), Lewis and Sappington (1991), Auriol (1992, 1998), and Wolinsky (1997). Both Laffont and Tirole (1993) and Lewis and Sappington (1991) considers the regulation of a monopoly. Wolinsky (1997) presents a model where the quality choices of the firms cannot be controlled by the regulator, and considers the regulation of prices and market shares. The results he obtains are similar to some of the results in this paper (those obtained in the absence of cost complementarity and for a simple demand and cost structure specification). My work is closer to the paper by Auriol (1998), but there are some notable differences.

My model considers the regulation of two firms who produce a single product each, in the absence of strategic interaction, whose cost characteristics are independently distributed. The regulator designs a menu of contracts for the firms, specifying quality and transfers as functions of reported types. The qualities of the two firms are assumed to be substitutes, which corresponds to the case where firms invests in quality to differentiate their products. The two firms' products are thus assumed to be vertically differentiated. The regulator is unable to observe total costs, and each firm has private information about its level of efficiency. Auriol (1998) considers a setting which is similar to mine, but where qualities are complements. The products are supplied through a common network by two firms. The only things that matters in the gross surplus function is total quality and whether both firms produce at the optimum (diversity increases consumer surplus). This implies that there is a public good aspect to quality; if one of the firms invests in quality, both the investing firm's customers and the customers served by the other firm benefit from the increased quality. This, together with unverifiable quality, introduces a problem of free-riding in quality provision which is absent in my initial model.[4] This point is, however, elaborated in the section on extensions to the basic

[4]The reason for this is of course that quality is observable, or verifiable, to the regulator in my model. If quality is unverifiable to the regulator, the solution would be interior in the general case. That is, both firms supply positive levels of quality. Furthermore, by rephrasing the model to analyse qualities as complements, free-riding is a solution only if there is cost complementarity

model. Furthermore, in Auriol's model quantity is regulated directly, whereas I consider the regulation of qualities. The choice of quality levels affect the production levels, which indicates that the regulation in my model is an indirect regulation of quantities. Contrary to the model by Auriol, the quality provision and information rent are not separated. In Auriol (1998), quantity is regulated and is independent of the level of quality provided. This is not so in my model. The information rent depends on the production levels, and the production levels are affected by the (regulated) quality levels. Thus, the level of quality affects the information rent. The implication is that the regulator may choose to distort quality levels to affect the firms' incentives. Furthermore, I obtain conditions which characterise the optimal degree of vertical differentiation. I consider a more general cost function than does Auriol (1992, 1998) which allows for cost complementarity, and I show that cost complementarity may, or may not, decrease the marginal information cost.

Olsen (1993) considers a model of multiagent regulation, where the (R&D) actions of the firms are substitute activites. He obtains the result that only if one firm is close to being maximally efficient and the other firm is less efficient, then the maximally efficient firm produces more than its full information level. If firms' efficiency parameters are identically distributed, but not equally inefficient, nothing can be said about the relations between full and asymmetric information production levels. However, in my model, production levels are distorted in favour of the relatively more efficient firm even for ex ante identical firms when total demand is exogenous.

There may be several reasons for there being more than one operator in a market that previously has been characterised as a natural monopoly (involving large fixed costs). First of all, there may be a yardstick effect, in which competition among firms with correlated costs reduces the information rent necessary to induce truthful revelation.[5] Secondly, the natural monopoly argument may lose some of its validity in some sectors (due to technological innovations). The telecommunications sector

between quantity and quality. This is briefly discussed in the extensions to the basic model.

[5]Auriol and Laffont (1993) consider, among other things, the yardstick effect as a reason for socially valuable duplication of fixed costs.

is one in which possibly only the local access network is still considered a natural monopoly.[6] Thirdly, and possibly most importantly, by introducing competition one may, at a later stage, be able to abolish regulations altogether. Thus, competition may be seen as a substitute for regulation. However, even if the main reason for allowing new firms into a market is to be able to use competition as a substitute for regulation, there may still be scope for some sort of regulation. The reason for this is that unregulated competition does not generally result in socially optimal outcomes.

The paper is organised as follows: In section 2, the model is set out. In section 3, the optimal solution under full information is analysed, section 4 examines the requirements for implementation, and section 5 analyses the optimal solution under asymmetric information. In section 6, some extensions to the basic model is considered.

2 The model

The model consists of three basic elements: (1) The demand side and consumer welfare, (2) the firms' choice of quality, or functionality, and (3) the welfare maximising regulator. The market is of a fixed size, and regulating the firms affects the distribution of market shares. Prices are exogenous to the model, whereas the level of quality is endogenous. The assumption of an exogenous price is obviously a simplification. A more complete regulatory mechanism where prices, in addition to the quality levels, are included would be preferred. However, including prices in the regulatory mechanism would not change the basic results of the model (and will in some situations even strengthen the results).

The quality variable could be interpreted as the quality of access. For instance, whether the line is analog, or digital (the speed of the connection), or simply what type of digital technology should be adopted. At the top-end, one could think

[6]Hansen (1996) finds evidence of natural monopoly when analysing the cost structure of Telenor. In his model, if competition somehow reduces slack (by only a small amount), thus making the firms produce more efficiently, competition may be welfare enhancing.

of wireless broadband links, while at the bottom of the scale we would find fixed analog connections. Firms have private information with respect to some efficiency parameter, β. The regulator is assumed to be a benevolent maximiser of welfare; that is, he has no agenda of his own, and is purely interested in maximising a utilitarian welfare function.

2.1 Consumers

The demand functions are in general given by $q_i = q_i\left(s, \overline{\theta}\right)$, for $i = 1, 2$, where $s = (s_i, s_j)$ is the vector of qualities. The vector of qualities could for instance be interpreted as the quality offered for access, as offered by the two firms operating in the market. The market is of fixed size $\overline{\theta}$, which implies that $q_i = \overline{\theta} - q_j$. The demand functions have the following properties:

Assumption 1 *Demand is increasing in own quality, decreasing in the other firm's quality, and the marginal effect of quality changes on demand is decreasing in the other firm's quality.*

Mathematically, $\partial q_i/\partial s_i > 0$, $\partial q_i/\partial s_j < 0$, $\partial^2 q_i/\partial s_i \partial s_j < 0$ for all (s_i, s_j). In addition, demand increases with the size of the market.

Consumers' net surplus is in general given by: $S = S\left(s, \overline{\theta}\right)$, which implies that net surplus is a function of qualities and the size of the market. Consumers' net surplus has the following characteristics:

Assumption 2 *Consumers' net surplus is increasing and concave with respect to quality.*

The conditions$\frac{\partial S}{\partial s_i} > 0$, $\frac{\partial S}{\partial s_j} > 0$, $\frac{\partial^2 S}{\partial s_i^2} < 0$, and $\frac{\partial^2 S}{\partial s_i \partial s_j} < 0$, together with $\frac{\partial^2 S}{\partial s_1^2}\frac{\partial^2 S}{\partial s_2^2} - \left[\frac{\partial^2 S}{\partial s_1 \partial s_2}\right]^2 \geq 0$, secures that net surplus is an increasing and concave function on (s_1, s_2).

The assumption $\frac{\partial^2 S}{\partial s_i \partial s_j} < 0$ together with the assumption $\frac{\partial^2 q_i}{\partial s_i \partial s_j} < 0$, ensures that the marginal value of s_i to the regulator decreases in s_j if $|\beta_i - \beta_j|$ is sufficiently small. In the definition of consumers' net surplus, S, the consumers' expenditures for purchasing the products supplied by firms 1 and 2 are taken into account. Thus, the fact that net surplus increases in quality indicates that the consumers prefer higher levels of quality for given prices.

As noted, prices are exogenous and equal for both firms even if quality offered by the two firms are different. An explanation for why some consumers still buy from the firm providing the product with the lowest quality could be that there is an unmodelled horisontal differentiation aspect. Thus, a consumer may prefer the product from the low quality provider because it is closest to his preferred product. An alternative explanation may be that some consumers for some reason are locked-in (at least in the short run), and may continue to buy a product of inferior quality even if there is a higher quality product available at an identical price.

2.2 Firms

The firms operate as profit maximising entities, and provide a single quality product each. Prices are exogenous, and given as $p_i = p > \beta_i \; \forall i$ (see note above). Firm i provides functionality level s_i, where $s_i \geq \underline{s} \geq 0$, for $i = 1, 2$. The lower bound on the level of quality can be thought of as a minimum quality standard. The firms are said to be vertically differentiated if $s_i \neq s_j$, for $i \neq j$.

As for firms' information about each other's costs, it is assumed that each firm knows only its own efficiency level. The firms' efficiency parameters are perceived to be (independently) distributed according to common knowledge distributions by the principal/regulator and the (other) firms/agents.[7]

[7] The assumption of "ignorance" makes the analysis non-trivial. If firms have complete information about each other's efficiency parameters, the regulator could potentially force each firm to report their true type by constructing a mechanism which results in an infinitely high penalty to the firms if their reports do not coincide. Each firm is in such a mechanism assumed to report the other firm's efficiency parameter in addition to its own. See Moore (1992) for more on this. Of course, the realism of such a regulatory mechanism could be questioned. The assumption that firms' efficiency parameters are *independently* distributed means that yardstick competition is not

The firm maximises the following objective function:

$$\max_{\widehat{\beta}_i} \pi_i = E_{\beta_j}\left[(p-\beta_i)\, q_i\left(s\left(\widehat{\beta}_i,\beta_j\right),\bar{\theta}\right)+t_i\left(\widehat{\beta}_i,\beta_j\right)-\psi^i\left(s_i\left(\widehat{\beta}_i,\beta_j\right)\right)\right] \qquad (1)$$

where $i=1,2$. E denotes the expectations operator. Expectations are taken over firm j's efficiency parameter, since firm i does not know firm j's type. The transfer from the regulator to firm i is given by t_i, and β_i denotes unit (and marginal) costs. β_i and β_j are assumed to be independently distributed. This would be the case if the two firms utilise different technologies (e.g., wireless vs. fixed-line technologies), which are subjected to independent technology shocks. Marginal cost of producing one unit of good i is initially assumed to be independent of the level of quality.[8] $\psi^i(s_i)$ defines firm i's costs of undertaking quality enhancing investments. The cost function is a cost element that is independent of quantity produced, or can be interpreted as investment costs, and is an increasing and convex function.

Total costs for firm i are thus defined as[9]

$$TC_i = \beta_i E_{\beta_j} q_i\left(s,\bar{\theta}\right)+E_{\beta_j}\psi^i(s_i)$$

Assumption 3 *Marginal cost of quantity increases with inefficiency.*

Assumption 3 is the familiar single-crossing condition, which in this case is $-\frac{\partial^2 TC_i}{\partial\beta_i\partial q_i}=-1<0$, for $i=1,2$. Total costs are assumed to be unobservable to the regulator.

an issue.

[8]See section 6 for the analysis of a case with a more general cost function.

[9]This cost structure bears resemblance to the model presented by Lewis and Sappington (1991).

2.3 The regulator

The regulator is assumed to be benevolent (that is, he has no agenda of his own), and maximises a utilitarian welfare function. By assuming that there are positive shadow costs of public funds (the approach used in Laffont and Tirole, 1993), there is a trade-off between information rents to the firms and economic efficiency.

The regulator's objective function is given as:

$$W = S + \pi_1 + \pi_2 - (1+\lambda)(t_1 + t_2)$$

The regulator maximises a utilitarian welfare function, but takes into account the welfare loss of transfers due to shadow costs of public funds. By inserting for transfers and simplifying (using the fact that $q_1 + q_2 = \overline{\theta}$), we get the welfare function (2):

$$\begin{aligned} &W\left(p, s, \lambda, \pi_1, \pi_2, \overline{\theta}\right) = S\left(s, \overline{\theta}\right) - (1+\lambda)\left[\psi^1(s_1) + \psi^2(s_2)\right] \\ &+ (1+\lambda)\left[p\overline{\theta} - \beta_1 q_1\left(s, \overline{\theta}\right) - \beta_2 q_2\left(s, \overline{\theta}\right)\right] - \lambda\pi_1 - \lambda\pi_2 \end{aligned} \tag{2}$$

where s is the vector $\left(s_1(\beta_1, \beta_2), s_2(\beta_1, \beta_2)\right)$.

Assumption 4 *The welfare function is increasing and concave on* (s_1, s_2).

In the full information case, the regulator maximises W subject to the participation constraints only. In the case of asymmetric information, the regulator knows only the support and distribution of the efficiency parameters, β_1 and β_2, and will thus have to maximise expected welfare, with expectations taken over β_1 and β_2. These (for the regulator) random variables have the (cumulative) distribution $F(\beta_i)$, with corresponding density function, which is assumed to be strictly positive over the relevant range, with $\beta_i \in [\underline{\beta}, \overline{\beta}]$. Furthermore, $F\left(\underline{\beta}_i\right) = 0$ and $F\left(\overline{\beta}_i\right) = 1$, for $i = 1, 2$. The efficiency parameters are assumed to be independently distributed, and the monotone hazard rate property is assumed to hold (assumption 5):

Assumption 5 $F(\beta_i)/f(\beta_i)$ *is non-decreasing.*

The regulator maximises the welfare function under the restrictions that the firms report their true types (the incentive compatibility constraints), and that firms choose to participate voluntarily (the participation constraints).

3 Optimal regulation under symmetric information

The main purpose of this section is to provide the full information benchmark of duopoly regulation of qualities. Below, I will compare optimal policies under symmetric and asymmetric information.

In the case of full information, the regulator can instruct the firms to implement whatever qualities he finds to be optimal. He has, however, to consider the fact that the firms may not wish to participate. By maximising the welfare function (eqn. (2)), subject to the participation constraints that each firm is secured a non-negative level of profit in expectation terms, $E_{\beta_j}\pi_i(\beta_i,\beta_j) \geq 0$, or $E_{\beta_j}\pi_i(\bar{\beta}_i,\beta_j) \geq 0$,[10] we obtain the optimal quality for firm i, s_i^{FI}, as defined by equation (3):

$$\frac{\partial\psi^i}{\partial s_i} + (\beta_i - \beta_j)\frac{\partial q_i}{\partial s_i} = \frac{1}{1+\lambda}\frac{\partial S}{\partial s_i} \tag{3}$$

for $i = 1, 2$, and $i \neq j$ (by using the fact that the market size is fixed). Quality is set such that the marginal cost of providing quality is equal to the sum of the weighted marginal (net) consumer surplus and price-cost margins (weighted by the marginal effect on demand from a change in quality). An unregulated firm i would ignore the effect of a change in the level of quality on consumers' net surplus and

[10] A note to the formulation of the participation constraints: Since the incentive compatibility condition for firm i is strictly decreasing in the (in)efficiency parameter, β_i, for any given β_j, it suffices to require that the participation constraint for the least efficient type is satisfied. Then, the participation constraint is satisfied for any type. In order for this to be true, the regulator needs to design a mechanism such that the least efficient type gets a profit of zero for any report made by the other firm, since firm i's profit is strictly increasing in the other firm's (in)efficiency parameter.

the cross-effect on firm j's profit.[11] Since the outcome of an unregulated duopoly competing in qualities is not generally identical to the socially optimal quality levels, there may be some scope for the regulation of qualities.[12] There are external effects due to investments that the unregulated firms cannot internalise. There are positive external effects on consumers' surplus from investing in quality. In addition, the firm that invests in quality imposes a negative external effect on the other firm from the business stealing effect of the investments. Due to the opposing signs of the external effects, the unregulated quality levels may be both too high or too low compared to the socially optimal levels.

If firms' investment costs are identical, an exogenously given market size, and a fully covered market (such that $\partial q_i/\partial s_i = -\partial q_j/\partial s_i$), the most efficient firm always provides the higher level of quality. Thus, we have the following result:

Lemma 1 *If $\psi^i(s_i) = \psi^j(s_j)$ and $\partial q_i/\partial s_i = -\partial q_j/\partial s_i$ for $\forall i, j$, $i \neq j$, then for $\beta_i < \beta_j$ we have $s_i^{FI} > s_j^{FI}$.*

Proof. Compare eqn.(3) for firms i and j. Using the assumptions on ψ, S, and q_i, we observe that $s_i^{FI} = s_j^{FI}$ only if $\beta_i = \beta_j$. It can be shown that in order to satisfy the first-order conditions defined by eqn. (3), $\beta_i < \beta_j$ must imply $s_i^{FI} > s_j^{FI}$. If this is not the case, the first-order conditions with respect to s_i and s_j cannot hold. ■

4 Implementation

A set of contracts is incentive compatible if a set of first- and second-order conditions are met. Local incentive compatibility (by using the first-order condition) requires that the state variables (here, profits) vary in a certain manner with the efficiency parameter. The second-order conditions ensure that the local optimum is also a global optimum.

[11] An unregulated profit maximising firm would choose its level of quality after the following rule: $\frac{\partial \psi^i}{\partial s_i} = (p - \beta_i)\frac{\partial q_i}{\partial s_i}$, which generally differs from the socially optimal quality level.

[12] If social costs of public funds approaches infinity, optimal quality for the least efficient of the two firms is the minimum quality whereas the more efficient firm produces a positive level of quality.

I assume that the regulator designs a (direct) revelation contract of the form (utilising the Revelation Principle); $M_i = \left\{s_i(\hat{\beta}_i, \hat{\beta}_j), t_i(\hat{\beta}_i, \hat{\beta}_j)\right\}$, where $\hat{\beta}_i$ is firm i's report of its efficiency parameter to the regulator. Let $\pi_i(\beta_i) \equiv E_{\beta_j}[\pi_i(\beta_i, \beta_j)]$, where E is the expectations operator.[13] In order for the regulator to maximise welfare under asymmetric information, the following conditions must be met:

(IC) $E_{\beta_j}[\pi_i(\beta_i, \beta_j)] \geq E_{\beta_j}\left[\pi_i\left(\hat{\beta}_i, \beta_j\right)\right]$, for all $\left(\beta_i, \hat{\beta}_i\right), \forall i$,
and

(PC) $E_{\beta_j}\pi_i(\beta_i, \beta_j) \geq 0$, for all $\beta_i, \forall i$.

The requirements for implementation of incentive compatible contracts are summarised in lemma 2 (local incentive compatibility) and proposition 3 (second-order conditions):

Lemma 2 *When quality is verifiable to the regulator, local incentive compatibility requires (using the envelope theorem):*

$$\frac{d\pi_i}{d\beta_i} = \frac{\partial \pi_i}{\partial \beta_i} = -E_{\beta_j}\left[q_i\left(s_i(\beta_i, \beta_j), s_j(\beta_i, \beta_j)\right)\right] < 0 \tag{4}$$

Both firms will earn information rents, except for firms of the least efficient types, $\overline{\beta}$.[14] The reason that both firm 1 and 2 earn rents is that the firms' efficiency parameters are stochastically independent. Any information the regulator may have on either firm is useless for the purpose of rent extraction. The information rent for firm i is given by equation (5):

$$\pi_i(\beta_i) = \int_{\beta_i}^{\overline{\beta}} E_{\beta_j} q_i\left(s_i\left(\tilde{\beta}_i, \beta_j\right), s_j\left(\tilde{\beta}_i, \beta_j\right)\right) d\tilde{\beta}_i + E_{\beta_j}\pi_i\left(\overline{\beta}_i, \beta_j\right) \tag{5}$$

for $i, j = 1, 2$.

[13] Let $E_{\beta_i}[x] \equiv \int_{\underline{\beta}}^{\overline{\beta}} x f(\beta_i)\, d\beta_i$.

[14] The regulator designs the contracts such that if a firm reports the highest β-value, he gets a profit of zero no matter what the other firm reports.

The first element on the right-hand side is the information rent a firm of type (efficiency level) β_i earns (with expectations taken over firm j types). The information rent is, of course, positive since the integrand is positive. Since rents to firms are costly, the profit to firm i of type $\overline{\beta}$ is set equal to zero; i.e., $E_{\beta_j}\pi_i\left(\overline{\beta}_i, \beta_j\right) = 0$.

The information rent is increasing in a firm's own quality level, but decreasing in the other firm's quality level. This can be seen by differentiating the information rent expression with respect to s_i and s_j, respectively:

$$\frac{\partial \pi_i}{\partial s_i} = \left[\int_{\beta_i}^{\overline{\beta}} E_{\beta_j} \frac{\partial q_i}{\partial s_i} d\widetilde{\beta}_i\right] > 0$$

$$\frac{\partial \pi_i}{\partial s_j} = \left[\int_{\beta_i}^{\overline{\beta}} E_{\beta_j} \frac{\partial q_i}{\partial s_j} d\widetilde{\beta}_i\right] < 0$$

From assumption 1 we know that the partial derivative of q_i with respect to s_i is positive, and with respect to s_j negative. All other things equal, distorting firm i's quality downwards and firm j's quality upwards, reduces the information rent to firm i since it reduces the quantity firm i produces. This implies that it is less tempting for firm i (of any given level of efficiency) to imitate less efficient types. Thus, the revelation process is made cheaper for the regulator.[15] Note that there are two different aspects to changing the level of information rent a firm earns. First, there is the issue of distorting the production levels for less efficient types for a given firm (the standard result in single-agent models). Second, we must consider incentives between the two firms. This amounts to awarding the relatively more efficient of the two firms a higher production level. Given that firm j is of the most efficient type, distorting downwards the quality levels of all but the most efficient type of firm i implies that firm i's incentives to imitate less efficient types is weakened. This is so for two reasons: First (ignoring the other firm), reducing the quality and thereby the quantity of less efficient types of firm i makes it less profitable for a more efficient type to imitate less efficient type. Second, by imitating a less

[15]However, since qualities (and hence quantities) are substitutes, firm j's production level, and hence information rent, is increased.

efficient type, his level of quality and thereby quantity, is further reduced because of the comparison between the two firms. Thus, because there are two firms in the market, the incentives to portray oneself as less efficient are weakened further relative to a monopoly situation.

Note that a more efficient firm is a firm which produces *quantity, and* not necessarily quality, at a lower cost. Since the total size of the market is assumed to be fixed, and fully covered, the fact that the distortion in quality levels (and market share) in favour of the more efficient firm implies that total production is made at a lower cost.

We have seen that the information rent of any given firm depends on the level of quality of both firms. These cross-effects which affect the information rent may be termed *fiscal externalities.* The provision of quality is therefore, unlike the result in Auriol (1998), not separated from the rent extraction. The reason for this is due to the difference in the regulation mechanism. In my model the production level of each firm depends on the quality levels since quantity is not regulated directly. Quantity is regulated indirectly through qualities, whereas Auriol (1998) considers direct regulation of quantities.

The second-order sufficient conditions for incentive compatibility are given in proposition 1:

Proposition 1 *Given the assumption on single-crossing of cost curves, A.3, sufficient conditions for incentive compatibility on the vector of control variables, s, are: (i)* $\frac{\partial s_i}{\partial \beta_i} \leq 0$*, and (ii)* $\frac{\partial s_i}{\partial \beta_j} \geq 0$ *for* $i, j = 1, 2$*, and* $i \neq j$.

The proof of Proposition 1 is in appendix 1.

This implies that quality, s, must be non-increasing in the firm's own inefficiency parameter, β, and must be non-decreasing in the other firm's (in)efficiency parameter. This is satisfied under assumptions 1, 5, and if the virtual surplus function (expression (2) inserted for the informational costs) is concave in quality.

5 Optimal regulation under asymmetric information

In the case of asymmetric information, I assume that the regulator knows only the distribution and support of the random variables β_1 and β_2 , and therefore maximises expected welfare. Furthermore, the regulator needs to induce the firms to reveal whatever private information they may have. By utilising the Revelation Principle (e.g., Myerson, 1979), the regulator may formulate a direct mechanism in which each firm will choose to reveal their types, provided that the transfer function is constructed in such a manner that the mechanism is incentive compatible.

5.1 Optimal policies

The regulator's maximisation problem is the following:

$$\max_{\{s_1,s_2\}} \int_{\beta_1} \int_{\beta_2} W\left(s_1\left(\beta_1,\beta_2\right), s_2\left(\beta_1,\beta_2\right), \lambda, p, \pi_1, \pi_2\right) dF\left(\beta_1\right) dF\left(\beta_2\right) \tag{6}$$

subject to

$$\frac{d\pi_1}{d\beta_1} = -E_{\beta_2}\left[q_1(s_1\left(\beta_1,\beta_2\right), s_2\left(\beta_1,\beta_2\right), \bar{\theta})\right] \tag{IC1}$$

$$\frac{d\pi_2}{d\beta_2} = -E_{\beta_1}\left[q_2(s_1\left(\beta_1,\beta_2\right), s_2\left(\beta_1,\beta_2\right), \bar{\theta})\right] \tag{IC2}$$

$$E_{\beta_2}\pi_1\left(\overline{\beta}_1, \beta_j\right) \geq 0 \tag{PC1}$$

$$E_{\beta_1}\pi_2\left(\overline{\beta}_2, \beta_j\right) \geq 0 \tag{PC2}$$

Since transfers are costly, we can safely assume binding participation constraints for the least efficient firms (for both firms 1 and 2). Integrating by parts the incentive constraints, taking into account that the participation constraints bind at type $\overline{\beta}$ for both firms, and inserting into the welfare function (2), we obtain the virtual surplus function. Let s denote the vector $\left(s_1\left(\beta_1,\beta_2\right), s_2\left(\beta_1,\beta_2\right)\right)$:

$$\begin{aligned} VS \;=\; & \int_{\beta_1}\int_{\beta_2}\{S(s,\bar{\theta})+(1+\lambda)\left[p\bar{\theta}-\beta_1 q_1(s,\bar{\theta})-\beta_2 q_2(s,\bar{\theta})\right] \\ & -(1+\lambda)\left[\psi^1(s_1)+\psi^2(s_2)\right] \\ & -\lambda\left[q_1\left(s,\overline{\theta}\right)\frac{F\left(\beta_1\right)}{f\left(\beta_1\right)}+q_2\left(s,\overline{\theta}\right)\frac{F\left(\beta_2\right)}{f\left(\beta_2\right)}\right]\}dF\left(\beta_1\right)dF\left(\beta_2\right) \end{aligned} \tag{7}$$

Maximising the expression (7) with respect to s_i, we obtain the formula for optimal quality, s_i^{AI}, for $i=1,2$, and $i\neq j$, given by:

$$\frac{\partial\psi^i}{\partial s_i}+\left(\beta_i-\beta_j\right)\frac{\partial q_i}{\partial s_i}=\frac{1}{1+\lambda}\frac{\partial S}{\partial s_i}-\frac{\lambda}{1+\lambda}\left[\frac{\partial q_i}{\partial s_i}\frac{F\left(\beta_i\right)}{f\left(\beta_i\right)}+\frac{\partial q_j}{\partial s_i}\frac{F\left(\beta_j\right)}{f\left(\beta_j\right)}\right] \tag{8}$$

Note here that the formulas for optimal quality for the asymmetric information case are identical to the full information formulas, except for the incentive correction component. The right-hand side of the formula for optimal s_1 is reduced by the (incentive correction) term $\frac{\lambda}{1+\lambda}\left[\frac{\partial q_1}{\partial s_1}\frac{F(\beta_1)}{f(\beta_1)}+\frac{\partial q_2}{\partial s_1}\frac{F(\beta_2)}{f(\beta_2)}\right]$. For firm 1, we have that the quality level, s_1 , should be distorted downwards if this expression is positive, since the quality investment function is increasing and convex.

By assuming that the whole market is covered, we have $\frac{\partial q_i}{\partial s_i}=-\frac{\partial q_j}{\partial s_i}$. This implies that the increase in demand for the investing firm is fully offset by the reduction in the other firm's demand. Then the incentive correction terms for s_1^{AI} and s_2^{AI} are reduced to expressions (9) and (10), respectively:

$$-\frac{\lambda}{1+\lambda}\left[\frac{F\left(\beta_1\right)}{f\left(\beta_1\right)}-\frac{F\left(\beta_2\right)}{f\left(\beta_2\right)}\right]\frac{\partial q_1}{\partial s_1} \tag{9}$$

$$\frac{\lambda}{1+\lambda}\left[\frac{F\left(\beta_1\right)}{f\left(\beta_1\right)}-\frac{F\left(\beta_2\right)}{f\left(\beta_2\right)}\right]\frac{\partial q_2}{\partial s_2} \tag{10}$$

If the sign of the bracketed term is assumed to be positive,[16] then because of the convexity of the quality investment functions, optimal quality for firm 1 is reduced

[16] The sign is positive if $\beta_1\geq\beta_2$ when distributions are identical using assumption 5.

under asymmetric information. For firm 2 the exact opposite result holds; under asymmetric information; optimal s_2 is increased.

A note to the relevance of actual efficiency levels: When the principal is to implement the optimal asymmetric information policies, he knows the true value of the efficiency parameters. The reason is that he has already devised an incentive compatible scheme, which induces firms to reveal their true types.

For the model presented here (with full market coverage), we have the following results:

Proposition 2 *If firm i is less efficient than firm j, i.e., $\beta_i > \beta_j$, then we have the following relationships: (1) $s_i^{FI} \geq s_i^{AI}$, and (2) $s_j^{FI} \leq s_j^{AI}$.*

Corollary *Assume that $\beta_i \neq \beta_j$. Since the most efficient firm always provides a higher level of quality under full information, the optimal policy under asymmetric information leads to more differentiated products than in the full information case.*

This result is quite obvious. From lemma 1 we know that the most efficient firm always provides the higher level of quality under full information. The economic intuition behind this result is that an increase in the degree of differentiation distorts the division of the market in favour of the most efficient producer. Increasing the quality level of the most efficient of the two firms and reducing it for the other firm increases the production level of the most efficient and reduces the production level of the other firm (see assumption 1).

Reducing the quality level of the less efficient firm and increasing it for the more efficient firm implies that the most efficient firm's production level increases (and the production level of the other firm decreases). Thus, contrary to standard one-agent models, production levels under asymmetric information is higher for the most efficient firm relative to the full information solution.[17] The explanation for this is that a departure from the full information solutions distorts the market shares in favour of the firm with the highest level of efficiency. Increasing the most efficient firm's

[17]In standard one-agent models, production is distorted downwards to reduce information rent payments. This, however, is not the case here. The reason is that it is the relative efficiency levels (between the two firms) which matters.

market share implies that the total production is made more cheaply, since efficiency is related to the production of quantity (and not necessarily the cost of producing quality). This result correspond to Olsen (1993), in which the most efficient agent is required to have a higher R&D output in the asymmetric information case. The reason for this result is, since actions are substitutes in his paper, "..the principal's need to balance the gains from coordination of agents' outputs against the costs associated with giving the agents rents.." (Olsen, 1993:p.535). In my paper it is the fact that qualities are substitutes which gives rise to the similarity. This implies that quantities/outputs are substitutes, and thus the analogy to the R&D result is clear. However, there is a difference in the results. In my model, contrary to Olsen (1993), the result that the production is distorted in favour of the more efficient firm is obtained when firms' probability distributions are identical.[18] This result is obtained by assuming that total demand is given, and that the whole market is served (or more precisely, that any increase in the demand for the investing firm's product is exactly offset by the reduction in demand for the other firm's product). Thus, if production is distorted in favour of the most efficient firm in the asymmetric information case, this firm must also increase its production relative to the full information case.

Since qualities are substitutes in the regulator's welfare function, the increase in the necessary total information rents to the firms is balanced by the increase in welfare. Optimal policies call for increasing the quality of the most efficient of the two firms and reducing the quality of the other firm - thus the most efficient of the two increases production and the less efficient reduces production, with the result that information rents to the less efficient of the two is reduced, whereas the information rents to the more efficient is increased. On the other hand, if qualities are complements the regulator would attempt to increase both firms' qualities to increase welfare, but this would also result in an increase in both firms' productions and information rents. Thus, the optimal policy in such a case would call for reductions

[18]In my model there is also the result that first-best levels are obtained if firms' distributions are identical and firms are equally inefficient. In this case, the incentive correction terms vanish. This is also different from Olsen (1993: p.536).

in production under asymmetric information (see McAfee and McMillan, 1991).[19]

If we consider the case where $\beta_1 < \beta_2$, then the optimal quality for firm 1 is increased, while firm 2's quality is decreased (again compared to the full information solution). This is in accordance with the monotonicity condition, which requires that quality be non-increasing in the parameter β. This implies that market shares are distorted in favour of firm 1; i.e., the firm with the highest actual level of efficiency. This is an illustration rather than a rigorous proof of how the optimal policies under asymmetric information in fact comply with the monotonicity conditions.

5.2 First best market shares under asymmetric information

From the analysis above we have the subsequent result:

Proposition 3 *Contrary to standard results in single-agent regulation models, first-best solutions (for market shares) are only obtained if both firms are of the most efficient types, or if firms are equally inefficient.*

The incentive correction terms for optimal qualities under asymmetric information, eqns. (9) and (10), will vanish only if the firms' efficiency parameters are identical.[20] In such a case, even if the optimally regulated qualities (and indirectly, market shares) results in first-best levels, the outcome is still second-best because of the shadow costs of transfers. Generally, the specifics of a first-best solution depends on the distributions, the specification of the cost function, and the demand structure. A sufficient condition for a first-best solution (in the general case where we allow for firms' efficiency parameters being drawn from different distributions) is that both firms are of the most efficient type.

[19] An increase in quantity may, in the case of complements, be a result of increased qualities, ceteris paribus. Such an increase in quantity results in increased information rent payments to both firms (since both firms' production levels are increased). Thus the cost of providing quality to consumers is increased. E.g., for a given level of quality for firm 2, and increase in the quality level of firm 1 would increase the production levels of both firm 1 and firm 2. This would increase the informational costs of production. Thus optimal policy would be to reduce both firms' qualities, and thus production levels.

[20] For a similar result in a different setting, I refer the reader to Osmundsen (1997).

If firm i is of the least efficient type, $\beta_i = \overline{\beta}_i$, and firm j's efficiency parameter is $\beta_j \in [\underline{\beta}, \overline{\beta})$, then the market shares are distorted in favour of firm j. Observe that in accordance with conventional wisdom, the most efficient of the two firms earns a higher rent, for identical distributions. However, if we allow for non-identical distributions, firms of equal efficiency levels may earn different information rents.[21] This is due to the fact that the inverse hazard rates are different for a given level of efficiency, which again leads to a distortion in the market shares relative to the full information solution.

6 Extensions to the basic model

6.1 Cost complementarity and optimal regulation of quality

In some cases, it may be natural to assume that there is cost complementarity between, for instance, network traffic and quality of a service (where quality may be interpreted as content, or functionality). For the case of telecommunications, it may be reasonable to assume that the owner of a network has lower costs for the provision of content, or functionality.[22] It has also been argued that there is a link between the provision of high-speed internet connections and size of the network. Higher-speed is here interpreted as higher quality. Having a large network and a large customer base implies that more of the traffic can move via fewer network links, and thus the speed of the traffic can increase. Thus, the variant of the model presented in this section could be interpreted as analysing the regulation of quality in the internet infrastructure. However, it should be noted that the Internet, as such, remains an unregulated industry, so the regulation is here interpreted purely as the regulation of the underlying infrastructure (much of the Internet traffic is transported over telecommunications networks, which currently are subjected to regulation).

Let us assume that each firm has its own network, with traffic level indicated by

[21] The exception is if both firms are of the least efficient types, in which case none of them earn any information rents.

[22] A similar situation can be found in the software industry, where a supplier of both the operating system and applications may have a cost advantage over a firm which only supplies applications.

the size of q.[23] The firm with the highest production level then has a cost advantage in the provision of quality (interpreted as content or functionality). Below, I will examine the effects of economies of scope on the optimal regulatory policy for two alternative specifications of the cost function.

The model is augmented only on the firm's cost side to incorporate the concept of economies of scope. One possible cost structure which allows for the idea of cost complementarity is the following:

$$TC_i = c_i(\beta_i, s_i)\, q_i\left(s, \bar{\theta}\right) + \psi^i(s_i) \tag{11}$$

where

$$\frac{\partial c_i}{\partial \beta_i} > 0, \frac{\partial c_i}{\partial s_i} < 0, \frac{\partial^2 c_i}{\partial \beta_i \partial s_i} \leq 0$$

Note that the only difference to the model presented in section 2.2 is with respect to marginal costs. Here, marginal cost is a function of both the efficiency parameter, β, and the level of quality. To have cost complementarity, we need marginal production costs to be decreasing in the level of quality, or equivalently, marginal costs of quality to be decreasing in the production level.[24] This could for instance be the case if we assume that there are learning effects in production, such that higher production levels lowers the marginal costs of providing quality. Another interpretation could be that a large customer base, or equivalently a high level of traffic, makes it possible for the firm to maintain a R&D department, which again may be able to reduce the cost of functionality enhancing activities.

6.1.1 Symmetric information

For optimal qualities, we need to take into account that the marginal cost of producing depends on the level of quality (still assuming a fixed market size):

$$\frac{\partial \psi^i}{\partial s_i} = \frac{1}{1+\lambda}\frac{\partial S}{\partial s_i} + (c_j - c_i)\frac{\partial q_i}{\partial s_i} - \frac{\partial c_i}{\partial s_i} q_i\left(s, \bar{\theta}\right) \tag{12}$$

for $i, j = 1, 2$ and $i \neq j$. Compared to the full information qualities in section 3, the only difference is the additional term $-\left(\frac{\partial c_i}{\partial s_i}\right) q_i$, which is positive if there is

[23]I assume that the networks are interconnected.

[24]By assuming $\frac{\partial c_i}{\partial s_i} < 0$, there is cost complementarity between q and s.

cost complementarity between quantity and quality.[25] Thus, all other things equal, cost complementarity increases the optimal quality under full information.

6.1.2 Asymmetric information

The presence of cost complementarity affects the information rent expression. Therefore, the incentive correction term is affected. The formula for optimal quality changes somewhat, and is given by:

$$\begin{array}{l} \frac{\partial \psi^i}{\partial s_i} = \frac{1}{1+\lambda}\frac{\partial S}{\partial s_i} + (c_j - c_i)\frac{\partial q_i}{\partial s_i} - \frac{\partial c_i}{\partial s_i} q_i \\ -\frac{\lambda}{1+\lambda}\left[\left(\frac{\partial c_i}{\partial \beta_i}\frac{\partial q_i}{\partial s_i} + \frac{\partial^2 c_i}{\partial \beta_i \partial s_i} q_i\right)\frac{F(\beta_i)}{f(\beta_i)} + \frac{\partial c_j}{\partial \beta_j}\frac{\partial q_j}{\partial s_i}\frac{F(\beta_j)}{f(\beta_j)}\right] \end{array} \tag{13}$$

for $i, j = 1, 2$ and $i \neq j$. The incentive correction term has an additional element, $\left(\frac{\partial^2 c_i}{\partial \beta_i \partial s_i}\right) q_i$, which is assumed to be less than, or equal to zero.[26] This implies that the marginal information cost (the incentive correction term) of providing quality is reduced when cost complementarity is introduced. Thus, optimal quality is increased (or decreased less) compared to the solution in section 5. The intuition behind this result is that cost complementarity, and thus the quality levels, affects the value of having private information. It is the private information about the cost difference which determines the information rent given to the firms. For the case where firm i is less efficient than firm j (i.e., $\beta_i > \beta_j$), an increase in firm i's quality level in effect reduces the real cost difference in providing quality.[27] By increasing firm i's optimal level of quality, the regulator is able to reduce the value of private information to firm i, and therefore reduce the information rent payment to this firm. This, of course, reduces the (virtual) cost of providing quality, and optimal quality can be increased relative to the solution in section 5.

[25] Obviously, marginal costs are also affected. If a firm supplies a positive level of quality, then, if there is cost complementarity, marginal cost is reduced.

[26] If c is linear in β and s, e.g., $c = \beta - s$, this term is equal to zero.

[27] The regulator is assumed to know the structure of the cost function, and thus he knows whether there is cost complementarity or not.

6.2 Quality as complements

The model analysed above can easily be transformed into a model where the quality of access mandated by the regulator (or chosen by the firms) are complements. We need to modify the assumptions on the demand functions and consumers' surplus in the following manner:

Assumption 1'

Quantity is increasing in both firms' quality, and marginal effect on demand of increasing a firm's own quality is increasing in the other firm's quality.

Assumption 2'

The marginal effect of increasing quality of one firm on consumers' net surplus is increasing in the other firm's quality.

In such a model, the result with respect to optimal quality is qualitatively different from the substitutes case. The following proposition summarises the result in the complements case:

Proposition 4

When individual quality contributions are verifiable to the regulator and in the absence of cost complementarity, the optimal level of quality is reduced under asymmetric information relative to the full information case. Cost complementarity makes the result more ambiguous.

The proof is omitted, but can be obtained from the author.

The intuition behind this result is that increases in quality for a given firm raises the (marginal) consumers' net surplus (with respect to the other firm's quality), but it also increases the information rent necessary for truthful revelation to both firms. An increase in quality by firm i raises firm i's demand and also has positive spill-over effects on firm j's demand. Thus, since a firm's information rent is increasing in quantity, this leads to a higher cost of quality provision. Thus, the socially optimal level of quality is reduced. However, if there is cost complementarity and this effect is sufficiently strong, this may "finance" the additional information rents to firms 1 and 2 and make it optimal to increase quality.

6.3 Unverifiable quality

The unverifiability of qualities may have several justifications. There may be many aspects of quality which is not readily observed by anyone but the user, and these may be aspects which are difficult or expensive to ascertain for the regulator (see the discussion in Laffont and Tirole, 1993: chapter 4). In this respect, we may see the complete quality as a function of both observable and unobservable features. Based on the observable features, the regulator is assumed to be able to set a minimum quality standard. This would mean that the quality investments made by the firms affect the network quality with respect to the unobservable features. Since it is assumed that quality is the only regulatory instrument, the case of unverifiable quality is in reality the unregulated case (with the restriction that prices are exogenously given). I will analyse both qualities as substitutes and as complements in separate subsections below.

The problem of optimal contracts for teams, when the action variables (here, qualities) are complements has recently been analysed by Auriol (1998) and McAfee and McMillan (1991).[28] In Auriol's model, unverifiability results in a free-rider problem in quality provision, whereas in McAfee and McMillan there is no such problem.[29] The model by McAfee and McMillan suggests "that the source of team problem is not the unobservability of team members' efforts or abilities per se" (McAfee and McMillan, 1991: p.571). They suggest that features such as risk aversion, or collusion may be the source of such inefficiencies.

In my model, there are two main results when qualities are complements and verifiable. First, optimal quality is distorted downwards under asymmetric information (relative to the full information case) for both firms, as both firms' incentive correction terms are unambiguously negative - a result which differs from Auriol (1998).

[28]Holmstrom (1982) is the seminal paper on the problem of team production.

[29]It should be noted that McAfee and McMillan does not consider quality provision, but their models is a more general team model where the actions of the team members are complements.

In her model, optimal quality is the same under both symmetric and asymmetric information for verifiable quality: The quantity level determines the information rent in both the present model and her model, but since quantity (in addition to quality) is regulated directly in Auriol (1998) this implies that there is no effect on the information rent of distorting the quality levels in her model. Second, in the absence of cost complementarity and if prices are insensitive to the quality levels, there is no problem of free-riding in quality provision. However, by assuming that prices do change with the level of quality, free-riding is a problem - similar to the result of Auriol (1998). The result of no free-riding corresponds to the result of McAfee and McMillan (1991). Thus, in my model the unobservability of firms' actions is not necessarily the source of the problem of free-riding. Free-riding is, however, a problem in my model if either cost complementarity is present, and/or prices are sensitive to the quality level. Either of these two factors in effect introduce an advantage of having a large market share, or asymmetry between firms. Thus, the reason for free-riding being a problem in my model is that firms' payoffs for identical levels of quality are different if their market shares differ. Cost complementarity implies that a firm enjoys a cost advantage in the provision of quality if it has a larger market share than its competitor. Price sensitivity to the quality level has an impact on the marginal revenue of changing quality. The more sensitive prices are to the level of quality, the higher is the potential for increasing profits by undertaking quality enhancing investments. Marginal profit of quality is higher for the firm with the larger market share. The analysis of optimal quality choice by the firms resembles the traditional analysis of monopoly pricing. First of all we have the effect of changing the quality level through the effect on demand which in reality is a second-order effect. Increased quality raises demand, and at given prices and costs, revenue is increased. This effect would disappear if prices are determined optimally. Then there is the *direct* effect which works through the price and cost effects. Increased quality raises the price the product can be sold for, and similarly, in the presence of cost complementarity, reduces the marginal cost of providing the product at a given level of quality.

To simplify the representation, I assume that the efficiency parameters are common knowledge among the firms and symmetric (unless otherwise stated). I assume that the firms maximise their profit functions given by equation (6), and choose their respective levels of quality simultaneously. We consequently look for Nash equilibria in the quality game.

6.3.1 Qualities as complements

In the substitutes case there is no problem of free-riding in quality investments, whereas when qualities are complements this may be a problem. This analysis is summarised in propositions 5 and 6. Propositions 5 and 6 assume that prices are exogenously given. This implies that quality investments has no effect on the firms' revenue except through the quantity effect. It may also affect their costs due if there is cost complementarities. To focus on the effect of cost complementarities, I choose to ignore asymmetries due to differences in the efficiency levels (i.e., I assume that $\beta_i = \beta_j$ except where stated otherwise).

One may then ask why firms choose to provide additional quality if such an investment has no effect on prices and which increases costs? Increasing quality does increase the quantity of the investing firm, but whether the firm will choose to invest in quality will depend on the cost structure and the profit-margin. If the firm chooses to invest when the price is fixed, it will certainly do so if the price is increasing in quality.[30]

Let us assume that the first-order condition is binding for some interior value of s. Let $\widetilde{s}_i$ be defined by equation (14) when there is no cost complementarity:

$$(p_i - \beta_i) \frac{\partial q_i (p, s_i + s_j)}{\partial s_i} - \frac{\partial \psi}{\partial s_i} = 0 \tag{14}$$

for some $\widetilde{s}_i \in (\underline{s}, \overline{s})$ for a firm with marginal cost β_i. If the price-cost margin is sufficiently high relative to the marginal investment costs, an interior solution will

[30]It is reasonable to expect that prices are increasing in quality in the absence of any restrictions on prices.

exist. If there exists such a $\widetilde{s}_i$, then both firms provide positive levels of quality.

If there is cost complementarity (of the particular type considered here), the profit maximising choice of quality is given by s_i^+ and is defined by equation (15):

$$(p_i - c_i(\beta_i, s_i))\frac{\partial q_i}{\partial s_i} - \frac{\partial c_i}{\partial s_i}q_i - \frac{\partial \psi}{\partial s_i} = 0 \tag{15}$$

where $\partial c_i / \partial s_i < 0$. Observe that the (symmetric) level of quality provided when there is no cost complementarity will be lower than the (symmetric) level provided if there is cost complementarity; i.e., $\widetilde{s}_i \leq s_i^+$.

Proposition 5

Assume that individual quality contributions are unverifiable and firms are equally efficient ex ante (i.e., $\beta_1 = \beta_2$). If prices are exogenous and there are no cost complementarities, then there is no problem of free-riding in quality provision. Both firms provide a positive level of quality unless the price is sufficiently low, or the investment cost sufficiently convex.

Proposition 6

If there is cost complementarity and quality is unregulated, then the problem of free-riding in quality provision is introduced. The firm with the largest market share will be the sole provider of quality.

The proof is found in appendix 2.

The intuition behind the result of Proposition 6 is that cost complementarity between quantity (or network capacity) and quality gives rise to a cost advantage for the firm with the largest market share.[31] This factor indicates that there is an advantage in having a large market share. Optimal quality is then increasing in the market share, and it is the firm with the largest market share which invests in

[31] In addition, if prices are sensitive to quality changes the cost of increasing quality is partly offset by the increase in price.

quality, whereas the firm with the smallest market share free-rides on the investment made by the other firm.

The result in proposition 5 can be explained by observing that the absence of cost complementarity makes firms symmetric. The level of optimal total quality is the same for both firms, since there are no cost advantage of having high production levels. The quality levels are then determined by the outcome of the quality game between the two firms, an equilibrium which is symmetric and in which both firms provide half of the optimal total quality each. If prices are subjected to price-cap regulation, we observe that we may have investments in quality even though firms face a binding price-cap. To see how quality affects a firm's marginal profitability, observe that the marginal profit (with respect to quality) increases with the quantity produced, since $\frac{\partial^2 \pi_i}{\partial q_i \partial s_i} = -\frac{\partial c_i}{\partial s_i} > 0$. Thus, for each extra unit produced by firm i, marginal profit with respect to quality is increased by $-\partial c_i / \partial s_i$. Thus, cost complementarity implies that for a given level of quality, firm i enjoys higher marginal profit (at this level of quality) than firm j if firm i has a higher market share than firm j; that is, if $q_i > q_j$. Firm i will produce a higher output in equilibrium (i.e., $q_i > q_j$) if, for instance, firm i is more efficient than firm j ($\beta_i < \beta_j$).

What about the effect on optimal (total) quality of changes in the efficiency parameter, β? If firms' efficiency parameters are different, this also affects the optimal level of quality in a similar way as both cost complementarity and quality sensitive prices, since different values of β_i and β_j introduces asymmetry between the firms. If $\frac{ds_i^*}{d\beta_i} < 0$, more inefficient firms has a lower optimal (total) quality. This inequality holds in the absence of cost complementarity. When there is cost complementarity, the inequality may be reversed, since higher level of quality reduces the marginal cost of producing the product. A larger market share increases the potential for having $\frac{ds_i^*}{d\beta_i} > 0$.[32] Thus, for a large enough market share for the most inefficient firm, the effect may be that it is the provider of the highest level of quality whereas the most efficient firm free-rides on the less efficient firm's investment in quality. However, if firms are symmetric the equilibrium of the game will be

[32]The sign of the inequality is determined by the sign of the following expression: $\frac{\partial^2 \pi_i}{\partial \beta_i \partial s_i} = -\frac{\partial c_i}{\partial \beta_i}\left(\frac{\partial q_i}{\partial s_i} + \frac{\partial q_i}{\partial p_i} p_i'\right) - \frac{\partial^2 c_i}{\partial \beta_i \partial s_i} q_i \lesseqgtr 0$.

symmetric in terms of quantities. Consequently, to justify the more inefficient firm having a larger market share than the more efficient firm we need to make some *ad hoc* assumptions about, e.g., one firm being an incumbent with an installed base.

It could also happen in the case of observable quality that the less efficient firm provides the highest level of quality if, for instance, the market share of the less efficient firm is sufficiently large. In the observable case the socially optimal level of quality is still higher than the outcome of the unregulated game. Furthermore, it is generally the case that both firms provide positive levels of quality when the regulator can design a contract contingent upon the quality levels, unless the marginal (quality) investment cost is very large.

6.3.2 Qualities as substitutes

The case of verifiable quality in the substitute case is considered in detail in the analysis of the basic model. We have seen that with *ex ante* symmetric firms (i.e., the efficiency parameters are drawn from identical distributions), the most efficient firm provides a higher level of quality. The firms provide positive levels of quality (or rather, is instructed to provide positive levels of quality), if the marginal cost of investing in quality and the shadow cost of public funds is not too large.

For the case of unverifiable quality, however, the solution may depend on whether there is cost complementarity.[33] In the absence of cost complementarity, it may be the case that both firms provide the same (positive) level of quality, or that both firms decides on no quality investment. In any case, the absence of cost complementarity leads to symmetric Nash equilibria.[34] Thus, we have the following proposition:

Proposition 7

Assume unverifiable quality and equally efficient firms. When qualities are substitutes, and in the absence of cost complementarity, then there is no vertical dif-

[33]The price sensitivity to quality changes and differences in efficiency may also affect the solution. However, in the absence of these two effects, the solution depends critically on whether there is cost complementarity.

[34]This should be no surprise, since firms (at a given level of efficiency) are symmetric in the absence of cost complementarity and if prices are insensitive to quality changes.

ferentiation between the networks. Either both firms provide the same positive level of quality, or both make no quality investments. There is therefore no problem of free-riding on the other's investments.

The proof is provided in appendix 3.

When allowing for cost complementarity of the particular type considered here, the optimal level of quality for any given firm is increasing in the market share. A larger market share (corresponds to a large q) implies a cost advantage in quality provision, which again may lead the firm to increasing its level of quality whereas the firm with the smaller market share reduces his. Thus, in the presence of cost complementarity (and/or price sensitivity to quality changes) the effect is that the degree of vertical differentiation is increased. This is similar to the effect of cost complementarity in the case of observable quality.

Furthermore, the unregulated choice of quality is in general different from the socially optimal quality level, but the bias of the distortion is difficult to ascertain in general. The reason for this is that the unregulated firm cannot internalise all the external effects from the investments. The investing firm is able to internalise is the demand effect (on its own demand), but is not able to internalise the (negative) external effect on the other firm nor the (positive) external effect on consumers' surplus.

7 Summary

The situation considered in this paper is the situation where a benevolent regulatory agency is able to affect (through the regulation mechanism) the access quality which the two firms offer to end-users. Assuming that firms' efficiency levels are different, we have seen that the regulator should induce firms to produce access of different quality; i.e., that some degree of (vertical) differentiation is indeed socially optimal in the setting of this model. The presence of asymmetric information necessitates even more differentiation in order to sufficiently distort the market shares in favour of the most efficient of the two firms, which implies that the more efficient firm produces more under asymmetric information. The primary reason for the optimal degree of differentiation to depend on the information which is available to the regulator is related to the cross-effects on information rents. Another reason is that both firms' efficiency parameters are drawn from identical distributions. Allowing for different distributions may modify the results. It is then not necessarily the case that the term correcting for asymmetric information unambiguously distorts optimal quality in favour of the most efficient of the two firms.

It is, of course, important for the regulator to know both the cost structure and how the level of quality impacts on the demand side. In the extensions to the basic model, I allow for variations in the model. The results there highlight the importance of knowing the structure of both demand and costs for optimal regulatory policy.

8 Appendices

Appendix 1

Proof of Proposition 1

I follow the standard approach of, e.g., Guesnerie and Laffont (1984) when clarifying the requirements for implementation under asymmetric information. However, there are some changes since expectations must be taken into account.

Firm i's profit function, when taking into account that firms have private information, and for a more general cost function, is given by (taking expectations over the other firm's efficiency parameter)

$$\begin{aligned} E_{\beta_j}\left[\pi_i\left(\hat{\beta}_i,\beta_i,\beta_j\right)\right] &= E_{\beta_j}\left[(p-\beta_i)\,q_i\left(s_i\left(\widehat{\beta}_i,\beta_j\right),s_j\left(\widehat{\beta}_i,\beta_j\right)\right)\right. \\ &\quad \left. +t_i(\hat{\beta}_i,\beta_j)-\psi^i\left(s_i(\hat{\beta}_i,\beta_j)\right)\right] \end{aligned}$$

for $i,j=1,2$, and $i\neq j$.

Firm i chooses its report to maximise profits, and the report must satisfy the following first- and second-order conditions:

(1) $\frac{\partial E_{\beta_j}\left[\pi_i\left(\left(\hat{\beta}_i,\beta_i,\beta_j\right)\right)\right]}{\partial\hat{\beta}_i}=0$

(2) $\frac{\partial^2 E_{\beta_j}\left[\pi_i\left(\left(\hat{\beta}_i,\beta_i,\beta_j\right)\right)\right]}{\partial \hat{\beta}_i^2} \leq 0$

Differentiating (1) with respect to the true efficiency parameter, β_i, and noting that at the optimum the reported type is equal to actual type (this is the property of incentive compatibility), the second-order condition can be rewritten as:

$$\frac{\partial^2 E_{\beta_j}\left[\pi_i\left(\left(\beta_i,\beta_j\right)\right)\right]}{\partial s_i \partial \beta_i}\frac{\partial s_i}{\partial \beta_i} + \frac{\partial^2 E_{\beta_j}\left[\pi_i\left(\left(\beta_i,\beta_j\right)\right)\right]}{\partial s_j \partial \beta_i}\frac{\partial s_j}{\partial \beta_i} \geq 0 \qquad (*)$$

We can show that $\frac{\partial^2 E_{\beta_j}[\pi_i((\beta_i,\beta_j))]}{\partial s_i \partial \beta_i} = -E_{\beta_j}\left[\frac{\partial q_i(\beta_i,\beta_j)}{\partial s_i}\right]$, which is assumed to be negative, and $\frac{\partial^2 E_{\beta_j}[\pi_i((\beta_i,\beta_j))]}{\partial s_j \partial \beta_i} = -E_{\beta_j}\left[\frac{\partial q_i(\beta_i,\beta_j)}{\partial s_j}\right]$, which is assume to be positive. Furthermore, by symmetry we have that $\frac{\partial s_j}{\partial \beta_i} = \frac{\partial s_i}{\partial \beta_j}$. Thus, the inequality (*) is satisfied for the following conditions:

(i) $\frac{\partial s_i}{\partial \beta_i} \leq 0$

(ii) $\frac{\partial s_i}{\partial \beta_j} \geq 0$

Thus, inequalities (i) and (ii) are sufficient conditions for implementation. These inequalities will be utilised as (the equivalent of the) second order conditions of incentive compatibility in the section on optimal regulation under asymmetric information. The standard procedure is to ignore these constraints ex ante, and checking whether they are satisfied ex post to avoid unnecessary messy calculations. If the monotonicity conditions are not satisfied, optimal policies may be characterised by partial pooling contracts.

Appendix 2

(proof of Propositions 5 and 6. The proof is adapted from Auriol, 1998)

The firms chooses a level of quality which maximises profits. Optimal (total) quality for firm i, s_i^* , is given by the first-order condition: $\frac{\partial \pi_i}{\partial s_i} = 0$, or $(p_i - c_i)\left(\frac{\partial q_i(p,s_i+s_j)}{\partial s_i} + \frac{\partial q_i}{\partial p_i}p_i'\right) + \left(p_i' - \frac{\partial c_i}{\partial s_i}\right) q_i - \partial \psi / \partial s_i = 0$. Qualities are complements, and what matters is total quality, not which firm supplies the quality. s_i^* defines the level of total quality which maximises firm i's profit, and can be supplied by firm i alone, by firm j alone, or as a team effort by both firms. Firm i's best response is given by: $s_i(s_j) = s_i^* - s_j$, for $s_j \leq s_i^*$, and $s_i(s_j) = 0$ otherwise. Implicit differentiation of the first-order condition with respect to q_i yields:

$\frac{ds_i^*}{dq_i} = -\left(\frac{\partial^2\pi_i}{\partial q_i \partial s_i}\right) / \left(\frac{\partial^2\pi_i}{\partial s_i^2}\right)$. The denominator (the second-order condition) is assumed to be negative. The sign of $\frac{ds_i^*}{dq_i}$ is determined by the sign of the numerator: $\frac{\partial^2\pi_i}{\partial q_i \partial s_i} = p_i' - \frac{\partial c_i}{\partial s_i} \geq 0$. In the absence of cost complementarity (CC) and if prices are insensitive to quality changes (that is, if $p_i' = 0$), $\frac{ds_i^*}{dq_i} = 0$, and for CC, or if prices respond to quality changes, $\frac{ds_i^*}{dq_i} > 0$. Thus, if there is CC, the optimal (total) quality is increasing in the market share and otherwise not (assuming $p_i' = 0$). For CC or for quality sensitive prices, $q_i > q_j$ implies $s_i^* > s_j^*$. Thus, the unique Nash equilibrium in the quality game corresponds to Auriol (1998): For $q_i > q_j$, $s_i = s_i^*$ and $s_j = 0$. Thus, free-riding in quality provision is present. In the absence of CC (and for $p_i' = 0$), optimal quality for both firms is independent of the market shares; $s_i^* = s_j^* = s^*$. Best responses for firm i and j are given by: $s_i(s_j) = s^* - s_j$ and $s_j(s_i) = s^* - s_i$. Thus, the equilibrium strategy is to have $s_i = s_j$, which amounts to: $s_i(s_j) = \frac{1}{2}s^*$ for i, j.

Appendix 3

(proof of proposition 7)

Let $s_i^+ = \arg\max\left[(p_i(s_i) - c_i(\beta_i, s_i))\, q_i - \psi(s_i)\right]$ define the profit maximising quality choice for firm i. In the absence of cost complementarity and with unresponsive prices, firms are symmetric if $\beta_i = \beta_j$. This implies that at the symmetric Nash equilibrium $s_i^+ = s_j^+ = s^+$; that is, the optimal level of quality is identical for both firms $i = 1, 2$. Whether the firms provide additional quality (in excess of the minimum quality standard) depends on the sign of the first-order condition at $\underline{s}$: $s_i = s_j = \underline{s}$ if $\frac{\partial \pi_i}{\partial s_i} \leq 0$ at $s_i = \underline{s}$, and $s_i = s_j > \underline{s}$ if $\frac{\partial \pi_i}{\partial s_i} > 0$ at $s_i = \underline{s}$. For a price p sufficiently close to marginal cost β_i, or for sufficiently high marginal investment cost, s_i^+ is equal to zero for $i = 1, 2$.

9 References

Armstrong, M., S. Cowan and J. Vickers (1994), *Regulatory reform: Economic analysis and the British experience*, Cambridge,MA: MIT-Press

Auriol, Emmanuelle (1992), Reglementation du duopole naturel, These pour le

Doctorat en Sciences Economiques, Universite des Sciences Sociales de Toulouse

Auriol, Emmanuelle (1998), “Deregulation and quality”, *International Journal of Industrial Organization*, 16, 169-194

Auriol, E. and J.J. Laffont (1993), "Regulation by duopoly", *Journal of Economics & Management Strategy*, 507-533

D'Aspremont and A. Jacquemin (1988), "Cooperative and non-cooperative R&D in duopoly with spillovers", *American Economic Review* **78**, 1133-1137

Hansen, Bjorn (1996), "Regulation when output is exogenous", mimeo, University of Oslo

Holmstrom, B. (1982), "Moral hazard in teams", *Bell Journal of Economics* **13**, 324-340

Kamien, M., E. Muller and I. Zang (1992), "Research joint-ventures and R&D cartels", *American Economic Review* **82**, 1293-1306

Laffont, J.J. and J. Tirole (1993), *A theory of incentives in procurement and regulation*, Cambridge, MA: MIT-Press

Lewis, T. and D. Sappington (1991), “Incentives for monitoring quality”, *RAND Journal of Economics*, Vol.22, No.3, Autumn

McAfee, R.P. and J. McMillan (1991), “Optimal contracts for teams”, *International Economic Review*, 32, 561-577

McGuire, T. and M. Riordan (1995) “Incomplete information and Optimal Market Structure: Public Purchases from Private providers”, *Journal of Public Economics*, 56, 125-141

Moore, J. (1992), "Implementation, contracts, and renegotiation in environments with complete information", in Laffont, J. J. (ed.), *Advances in Economic Theory*, vol. I, Cambridge, UK: Cambridge University Press

Myerson, R. (1979), “Incentive compatibility and the bargaining problem”, *Econometrica*, 47, 61-73

Osmundsen, P. (1997), “Regulation of industry-specific externalities under private information”, Revised version of Discussion Paper 01/97, Institute of Economics, Norwegian School of Economics and Business Administration

Olsen, T.E. (1993), “Regulation of multiagent research and development”, *RAND*

Journal of Economics, 24, 529-541

Tirole, J. (1988), *The theory of industrial organization*, Cambridge, MA: MIT-Press

Wolinsky, A. (1997), "Regulation of duopoly: Managed competition vs regulated monopolies", *Journal of Economics and Management Strategy*, **6**(4), pp. 821-47

Part 2:
Access charge regulation

ESSAY 3

Regulation and foreclosure

Jan Y. Sand*

University of Tromso, Department of Economics,

N-9037 Tromso, Norway

Jan.Sand@nfh.uit.no

October 2000

Abstract

The paper considers the optimal regulation of access charges, and the effect such regulation has on incentives to foreclose downstream rival firms. I show that when a vertically integrated firm is able to discriminate against rivals by means of non-price measures, optimal access charges must be set higher than in the case when no discrimination is possible and will always provide a positive access margin. The reason is that the level of the access charge affects incentives to practice foreclosure. The optimal access charge may, when non-price measures are not possible, be lower than marginal cost of providing access.

JEL Classifications: D82, L13, L22, L51

Keywords: regulation, vertical relations, duopoly, foreclosure

*I would like to thank Øystein Foros, Kare P. Hagen, Derek Clark, Nils-Henrik von der Fehr, Petter Osmundsen, Patrick Rey and Jean Tirole for very helpful comments and suggestions. Financial support from Telenor, NFR, and Norges Banks Fond til Økonomisk Forskning is gratefully acknowledged. A previous version of this paper was presented at the Econometric Society Summer Meeting 1999 in Santiago de Compostela, Spain, and EARIE 2000 in Lausanne, Switzerland.

1 Introduction

In vertically related markets, the production of final products makes use of (essential) inputs produced in a complementary market. The producers of these essential inputs usually have opportunities to earn positive economic profits. The extent to which this is possible depends, among other things, both on regulatory and competition policies. For instance, firms offering Internet access to end-users must purchase access to consumers from local access providers.[1] The pricing of local access is more often than not subject to regulation. It is reasonable to expect that the regulation of access charges in general results in prices which are different from the prices an unregulated firm would choose. If the access charge is set below the monopoly level, regulation effectively restricts the potential for monopoly rents in access provision. The main reason for introducing regulation of access charges is to stimulate competition, by ensuring that rival firms can obtain access to end-users at reasonable terms. This means setting a low access charge. However, as we shall see, restricting the integrated firm's earning potential upstream results in an increased incentive to foreclose its rivals. Furthermore, in the case of U.S. legislation, restrictions are to some extent imposed on which firms are allowed into the long-distance markets. By restricting local access providers' opportunity to serve long-distance markets (i.e., refusing vertical integration), competition authorities may restrict the potential for earning monopoly rent upstream even without regulating access charges. This is a result well known in the literature on vertical relations.[2]

Allowing vertical integration opens up the possibility for foreclosure activities by the vertically integrated firm. A combined policy of allowing vertical integration and regulation of monopoly rent upstream (through access charge regulation) may seriously affect downstream competition. When we allow for the possibility for non-price discrimination by the vertically integrated firm, an interesting question in this setting is whether abandoning access charge regulation all together (or, equivalently, set a non-binding access charge) results in a qualitatively different outcome than does

[1] The local access providers are often the incumbent telecommunications companies.

[2] This result is a consequence of the upstream firm's inability to credibly commit to charging monopoly prices (Rey and Tirole, 1997)

the situation with binding access charge regulation.

The contribution of this model to the theory of regulation is to allow for the possibility that a vertically integrated firm (the network owner) may foreclose its rivals (partially or complete).[3] The foreclosure may be thought of as either increasing the cost of purchasing network inputs (in addition to the exogenously given access price), or equivalently as degrading the quality of such inputs. The increased costs of the rivals could, for instance, be due to legal expenses incurred when attempting to obtain access on equal terms with the network owner's downstream subsidiary, or more direct costs due to lower quality of access.

Throughout the paper I assume that the regulator cannot regulate the downstream sector due to reasons external to the model presented below.[4] A general model of optimal regulation should be able to explain the deregulation of the downstream industry following as the optimal outcome of a complete regulatory setup. In the present model, following a partial approach, I assume that the downstream industry is deregulated. This is the case, e.g., in the telecommunications industry. Often, the regulation of access charges is easier to implement due to a vast array of different products on the market. It also seems to be the case that industry regulators focus more of their attention on the regulation of access terms.

The issue of optimal (regulated) access charges has been examined by many others (e.g., Armstrong, Doyle and Vickers (1996), Laffont and Tirole (1990, 1996)). In contrast to their work, I focus on the relation between the regulated access charge and the vertically integrated firm's incentive to foreclose its rival along other dimensions. This has some resemblance to the multi-task principal-agent literature (see Holmström and Milgrom, 1991). The vertically integrated firm has in reality two main methods of foreclosing rival firms. Either by means of setting high access charges, or by non-price behaviour. In the model presented here, however, the access charge is regulated and the only method of foreclosure that remains open is through

[3]Partial foreclosure means that the vertically integrated firm supplies access to the downstream competitor, but at less favourable conditions than its own downstream subsidiary. Complete foreclosure is a situation where the competitors are in effect denied access.

[4]This is also the starting point of Vickers (1995).

non-price behaviour. The incentives for non-price behaviour is then obviously affected by the regulation of the access charge.

The access terms offered to rival firms consist of two main elements - the price paid for access and the quality of access. I assume that the access charge is subject to regulation. However, the network provider has ample opportunities to degrade the quality of access offered to its competitors. The decision of the network provider (the vertically integrated firm) on access *quality* is *not* regulated in the model, and may seriously affect the competition downstream. This relates the model to the literature on raising rivals' costs and foreclosure (see, for instance, Rey and Tirole, 1997, and Economides, 1998a, 1998b). The model presented in this paper can be seen as a combination of Vickers (1995) and Economides (1998a, 1998b), with a primary focus on the incentives to reduce the quality of the inputs offered to downstream rivals, which is equivalent to raise rivals' costs. The issue of raising rivals' costs is the focus of Economides (1998a, 1998b), but his model does not incorporate the optimal regulation of access charges. Vickers (1995) presents a paper on regulation in vertically related markets, but his primary concern is to investigate the importance of the industry structure (vertically separation or integration) on optimally regulated access prices. The primary focus in my paper is rather on how regulating the access charge affects the incentive to foreclose rival firms, and how access charges in this context should be set to achieve the social optimum.

The model fits into the essential facilities, or bottleneck doctrine: One dominant firm owns the supply of a non-standard input, to which there are no, or few, substitutes. A firm is then said to be foreclosed if the dominant firm denies proper access to the essential facility, with the intent of maintaining its market power (Rey and Tirole, 1997). If the upstream firm is not allowed to vertically integrate with one of the downstream firms, the upstream firm may be in a situation similar to that of a durable-good monopolist; it will be difficult to credibly commit to maintaining monopoly pricing, since once the upstream monopolist has sold to one of the downstream firms at monopoly price it will have incentives to reduce the price to other downstream firms to effectively exploit the residual demand. Thus, the upstream monopolist may not be able to exploit its monopoly power. Vertical integration, on the other hand, eliminates the commitment problem. Thus, allowing vertical integration enables the upstream firm to earn monopoly profits. In the model presented below, this negative effect (from a static efficiency point of view) is restricted since access charges are regulated.

The remaining part of the paper is organised as follows. In section 2 the basic model is presented and analysed, in section 3 I investigate the optimal regulation of access charges both when the vertically integrated firm can and cannot use non-price discrimination. In section 4 the results are summarised and policy implications are discussed.

2 A model of vertically related markets

There are two firms. Firm i is a producer of a final product. Firm v is a vertically integrated firm, that supplies an essential input for the production of the final product. In the final product market there is Cournot competition. The vertically integrated firm can be perceived as a firm providing both long-distance services and local calls, whereas the other firm provides only long-distance services. Alternatively, the vertically integrated firm provides both content and network services, whereas the other firm provides content only.

The regulator offers a menu of contracts contingent on the report made by the regulated firm, specifying an access charge and awarding a transfer to the regulated firm. The network provider makes a report of its own (upstream) efficiency level, which is subject to private information. We allow for differences in the efficiency levels upstream and downstream. When setting the access charge, the regulator must take into account that the regulated firm may take some unverifiable actions which affect the costs of its downstream rival (and that the incentive to do so may depend on the level of the access charge). The regulator can thus indirectly affect both prices and consumers' surplus through the determination of the access charge.

For purposes of comparison, I consider both the cases where the vertically integrated firm can and cannot take some unverifiable action (non-price discrimination). The absence of non-price discrimination can be taken to mean that regulators are able to write complete contracts related to the access terms offered to other firms. This can be considered as a benchmark case to investigate the effects of foreclosure on the level of optimal access charges.

2.1 The demand side

In the downstream market firms are facing the inverse linear demand function, $P(Q) = a - Q$, where $Q = q^v + q^i$ is total production downstream.

Net consumers' surplus is in this case given by:

$$CS(Q) = \frac{1}{2}Q^2 \tag{1}$$

2.2 Firms and costs

The vertically integrated firm earns profit in two different markets - the upstream and downstream market. This implies that foreclosing rival firms downstream entails an opportunity cost - the reduction in profits in the upstream activity resulting from a lower overall production level downstream. In addition, foreclosure entails a monetary cost. Activities designed to foreclose rival firms are normally not consistent with competition laws. Consequently, firms that undertake such activities must conceal their actions, and it is realistic to assume that this is costly. Foreclosure is socially wasteful both since total downstream production is reduced and since it involves a monetary cost of the unproductive activity.[5]

The profit function of the competing firm downstream is given by:

$$\Pi^i = \left(P(Q) - r - \beta^i - w\right) q^i \tag{2}$$

where β^i is the efficiency level, and q^i is the production level of firm i. The profit function of the vertically integrated firm is given by:

[5] An alternative justification for $C(r)$ could be that a more pronounced level of foreclosure makes it more likely that the competition authorities reveals the unwanted practice, which may lead to a fine being imposed on the firm practicing foreclosure.

$$\Pi^v = \left(P(Q) - \beta^d - w\right) q^v + (w - \beta^u) Q - C(r) + t - F \tag{3}$$

where β^d and β^u are the downstream and upstream efficiency levels, respectively. q^v is the downstream production of the vertically integrated firm, Q is the total production downstream, t is the transfer from the regulator, and F is a fixed cost related to upstream production. There are no capacity constraints upstream, and upstream inputs are available to any firm which is willing to pay the prevailing price.[6] Let r denote the degree of foreclosure in the access terms for the competitors. It is chosen by the firm and is unverifiable by the regulator. A high level of r is interpreted as a low level of access quality offered to the rival firm, which reduces the willingness to pay for the product sold to end-users.[7] The cost of foreclosure, $C(r)$, is increasing and strictly convex ($C_r > 0$, $C_{rr} > 0$), and is for the sake of simplicity assumed to have the following quadratic form: $C(r) = \varphi r^2/2$. In the case where the integrated firm cannot foreclose its rival, the parameter r is exogenously given and normalised to zero, with $C(0) = 0$.

The regulated variable w is the price all downstream firms pay per unit of the inputs purchased from the upstream firm. All downstream firms obtain access at the same price, including the downstream subsidiary of the upstream firm. For the vertically integrated firm the access charge is, however, simply an internal transfer.

[6] With no capacity constraints upstream, one may argue that price compatition is as likely as quantity competition. However, as an ad hoc justification of the use of quantity competition, it is assumed that each of the downstream firms must decide on a level of capacity prior to entering into the downstream market (e.g., firms must lease lines from the local access provider which have a fixed capacity prior to producing the services sold to end-users). This implies that at the last stage of the game, each firm has a limited level of capacity. In such a situation, and under certain conditions about the capacity levels and the rationing rule, Kreps and Scheinkman (1983) shows that the unique outcome of a price competition game with capacity constraints is the Cournot outcome.

[7] The effect of modelling foreclosure as a quality degradation rather than as a cost raising strategy is that the price elasticity of demand is affected by the level of foreclosure. Increasing the level of foreclosure makes the rivals' demand more elastic.

2.3 Welfare and the regulator

The regulator is assumed to maximise a utilitarian welfare function, where transfers awarded to the regulated firm are socially costly due to distortions imposed on other sectors of the economy to raise the revenue. The welfare function is given by:

$$W = CS + \Pi^{v} + \Pi^{i} - (1+\lambda)\, t \tag{4}$$

where $(1+\lambda)$ is the social cost of transfers to the regulated firm, with $\lambda > 0$. The welfare function is assumed to be concave in w. The regulatory agency has imperfect knowledge of the costs upstream, and it only knows the distribution, $G(\beta^{u})$ with the strictly positive density function $g(\beta^{u}) > 0$, and the support of upstream costs, $\beta^{u} \in [\underline{\beta}, \overline{\beta}]$, with $\underline{\beta} \geq 0$. The upstream and downstream costs of the vertically integrated firm are assumed to be independently distributed, which implies that observation of the downstream costs yields no information about upstream efficiency to the regulator.

2.4 Timing of the game

The stages of the regulatory game are as follows:

- **Stage 1:** The regulator offers a contract of the form $M = \left\{ w\left(\widehat{\beta}^{u}\right), t\left(\widehat{\beta}^{u}\right) \right\}$ to the regulated firm.
- **Stage 2:** The regulated firm reports a type $\widehat{\beta}^{u}$ to the regulator, and the contract is executed. The regulator assigns the firm a transfer, t, and an access charge, w.
- **Stage 3:** The regulated firm decides on the (unregulated) quality of access terms to downstream rivals (the level of foreclosure), r.
- **Stage 4:** Firms compete á la Cournot in the downstream market.

2.5 Solving the model

In order to solve this model, we start at the last stage of the game, and solve by backward induction. The regulator must take into account the effects that the regulatory mechanism he proposes has on all subsequent stages. We first solve for the outcome of the post-regulation stages. Using his knowledge of the outcome in the last two stages, the regulator designs a second-best optimal regulatory mechanism. For now, I will assume that there is a fixed number of firms in the downstream industry.

2.5.1 Cournot competition

There is quantity competition between the firms downstream. Firm v chooses the level of quantity q^{*v} which maximises Π^v, where Π^v is defined by eqn. (3). Firm i solve a similar maximisation problem, maximising (2) with respect to q^i.

The vertically integrated firm's optimal quantity choice is determined by:

$$P_Q q^{*v} + \left(P\left(Q^*\right) - \beta^d - \beta^u\right) = 0 \tag{5}$$

For firm i, the optimal quantity is determined by:

$$P_Q q^{*i} + \left(P\left(Q^*\right) - \beta^i - w - r\right) = 0 \tag{6}$$

Solving for equilibrium quantities, we obtain the following:

$$q^{*v} = \frac{a - 2\left(\beta^d + \beta^u\right) + \beta^i + w + r}{3} \tag{7}$$

and

$$q^{*i} = \frac{a - 2\left(\beta^i + w + r\right) + \beta^d + \beta^u}{3} \tag{8}$$

Total downstream quantity is

$$Q^* = q^{*v} + q^{*i} = \frac{2a - \beta^i - \beta^d - \beta^u - w - r}{3}$$

Using the equilibrium quantities, we can express the stage 4 equilibrium profit for firm v as:

$$\Pi^{*v} = \left(q^{*v}\right)^2 + \left(w - \beta^u\right) q^{*i} - \frac{\varphi}{2} r^2 + t - F \tag{9}$$

and for firm i as:

$$\Pi^{*i} = \left(q^{*i}\right)^2 \tag{10}$$

2.5.2 Choosing the level of foreclosure

The regulated firm can in some circumstances foreclose rival firms, by degrading the quality of inputs sold to its rivals. When firm v can undertake some unverifiable action, it chooses a level of r to maximise Π^v. The maximisation is subject to $q^v = q^{*v}$ and $q^i = q^{*i}$, where q^{*v} and q^{*i} are the equilibrium quantities in the quantity competition stage. Consequently, $r^* = r\left(\beta^i, \beta^u, w, \beta^d\right)$.

The profit maximising level of quality degradation r^*, is governed by the following relationship:

$$\frac{d\Pi^v}{dr} = \underbrace{\frac{\partial \Pi^v}{\partial q^v}\frac{\partial q^{*v}}{\partial r}}_{=0} + \frac{\partial \Pi^v}{\partial q^i}\frac{\partial q^{*i}}{\partial r} + \frac{\partial \Pi^v}{\partial r}$$

which can be rewritten as:

$$\left[(w - \beta^u) + P_Q q^{*v}\right]\frac{\partial q^{*i}}{\partial r} - C_r = 0 \tag{11}$$

Observing eq. (11) we note that if $\left[(w - \beta^u) + P_Q q^{*v}\right] > 0$, or equivalently, if net marginal profit of an increase in the rival's quantity is positive, then firm v chooses the level of foreclosure as low as possible; $r^* = \underline{r}$.[8] In line with the reasoning of Economides (1998b), we observe that $\left[(w - \beta^u) + P_Q q^{*v}\right] > 0$ can only be positive *and* consistent with the existence of profit-maximising downstream rivals if the downstream subsidiary of the vertically integrated firm is sufficiently inefficient.[9]

[8]Thus, if we allowe for negative levels of r, it is possible that a vertically integrated firm may find it in its own interest to upgrade, rather than downgrade, the quality of inputs sold to its rival, implying that downstream market shares are shifted in favour of the rival firm. However, in this paper we maintain the assumption that $r \geq 0$.

[9]Assume $\left[(w - \beta^u) + P_Q q^{*v}\right] > 0$. The first-order condition for the profit-maximising quantity choice for the vertically integrated firm is given by: (*) $\left[P_Q q^{*v} + (w - \beta^u)\right] + P - \beta^d - w = 0$. Then, the condition $\left[P_Q q^{*v} + (w - \beta^u)\right] > 0$ implies $P - w < \beta^d$. The first-order condition for the rival firm is given by: (**) $P_Q q^{*i} + P - \beta^i - w - r = 0$. We observe that if $\beta^i = \beta^d$, condition (**) cannot be satisfied if $r \geq 0$. However, condition (**) can be satisfied if $\beta^d > P - w > \beta^i + r$; i.e., if β^d is sufficiently greater than β^i.

Consequently, the vertically integrated firm may wish to shift downstream market shares in favour of its own rival if its own downstream subsidiary is sufficiently inefficient. This is achieved by providing higher quality access to the rival firm than is available to its own subsidiary. The intuition behind this result is that when the rival firm produces the final product much more efficiently, it is better for the vertically integrated firm to concentrate on the provision of access and leave as much as possible of the production of the final product to the rival. An interpretation of this could be that the vertically integrated firm in effect outsources the production of the final product to its rival. In Economides (1998b), in contrast, $(w - \beta^u) + P_Q q^{*v} > 0$ can never be positive and consistent with the existence of profit-maximising rival firms, as all downstream firms have the same marginal costs.

Lemma 1 *If* $\beta^d > P - w > \beta^i + r$, *then* $[(w - \beta^u) + P_Q q^{*v}] > 0$.

Eqn. (11) can, by using the linear demand function and quadratic cost function, be rewritten as:[10]

$$r^* = \frac{2}{9\varphi - 2}\left(a + \beta^i - 2\beta^d - 2w + \beta^u\right) \tag{12}$$

The profit maximising level of foreclosure thus has the following properties:[11]

$$\begin{aligned} dr^*/d\beta^d &= -4/(9\varphi - 2) < 0 \\ dr^*/d\beta^u &= 2/(9\varphi - 2) > 0 \\ dr^*/d\beta^i &= 2/(9\varphi - 2) > 0 \\ dr^*/dw &= -4/(9\varphi - 2) < 0 \end{aligned}$$

The degree of foreclosure is reduced when the downstream inefficiency of the vertically integrated firm increases; $dr^*/d\beta^d < 0$. The intuition is that higher downstream costs makes it relatively more profitable to offer access downstream (which

[10]The second-order condition for r^* to be a maximum is $\varphi > 2/9$, which is assumed to be satisfied. If, on the other hand, the second-order condition is not satisfied, the vertically integrated firm will choose a level of r which completely forecloses its rival.

[11]These properties are also valid for more general specifications of the inverse demand function.

corresponds to a lower r), rather than to compete in services downstream. Consequently, the vertically integrated firm sets a low level of foreclosure to increase the profit earned in the downstream segment. The intuition behind $dr/d\beta^u > 0$ is as follows: Higher upstream costs makes it less profitable to offer access, and the vertically integrated firm chooses to compete downstream (i.e., the vertically integrated firm wants to set a higher r) rather than to offer access. The fact that $dr^*/dw < 0$ seems quite reasonable, as r and w are basically substitute foreclosing activities. A firm selling inputs that are essential to competing downstream firms, can improve its own downstream subsidiary's position by either increasing the access charge, or by reducing the quality of access.

Foreclosure benefits the regulated firm through two effects: first, foreclosing the downstream rival reduces its equilibrium output, and thus increases the vertically integrated firm's equilibrium output (since quantities are strategic substitutes; $dq^{*v}/dq^i < 0$). Secondly, foreclosure increases the price on the inframarginal units of firm v's output downstream. These benefits are traded off against the opportunity cost of foreclosure, which is due to the contractive effect on upstream profit resulting from lower quantity sold to its rivals.

2.5.3 Report to the regulator

The upstream firm makes its report, and the game is played according to the terms of the contract (if accepted). Given the level of efficiency β^u, firm v chooses the report

$$\begin{aligned} \widetilde{\beta}^u &= \arg\max_{\widehat{\beta}^u} \Pi^v \qquad (13) \\ &= \arg\max_{\widehat{\beta}^u} [\left(P\left(Q^*\left(\widehat{\beta}^u,\beta^u\right)\right) - \beta^d - w\left(\widehat{\beta}^u\right)\right) q^{*v}\left(\widehat{\beta}^u,\beta^u\right) \\ &\quad + \left(w\left(\widehat{\beta}^u\right) - \beta^u\right) Q^*\left(\widehat{\beta}^u,\beta^u\right) - C\left(r^*\left(\widehat{\beta}^u,\beta^u\right)\right) + t\left(\widehat{\beta}^u\right)] \end{aligned}$$

3 Optimal regulation

The regulator maximises the expected value of the welfare function, with expectations taken over the upstream type of the regulated firm, β^u, subject to participation and incentive constraints. As for the participation constraint, it is required that the regulated firm earns non-negative profits in aggregate (that is, the sum of upstream and downstream profits), $\Pi^v \geq 0$. This implies that we need not be concerned with conditions to ensure that the downstream subsidiary of the vertically integrated firm earns positive profits. It should, however, be noted that when the participation constraint is applied to aggregate profits this opens up the possibility for cross-subsidisation from the regulated activity (upstream production) to the competitive segment (the downstream industry). A participation constraint applied to aggregate profit also simplifies somewhat the analysis, as the incentive constraint is related to aggregate profits (see below).

3.1 Unregulated access charge

In an unregulated environment, the vertically integrated firm can choose to foreclose rival firms using either the access price or through non-price discrimination. Since non-price discrimination is costly for the integrated firm, it is reasonable to assume that a preferable (i.e., less costly) method of foreclosure is to set a high access charge. The unregulated firm solves the following maximisation problem:

$$\max_{w} \Pi^v = \left(q^{*v}\right)^2 + \left(w - \beta^u\right) q^{*i} - F$$

$$\text{subject to } q^{*v} \geq 0 \text{ and } q^{*i} \geq 0$$

The solution to the problem, w^{ur}, is to set:

$$w^{ur} = \frac{5a - 4\beta^i - \beta^d + 5\beta^u}{10}$$

or, if $\beta^i = \beta^d = c$, we have $w^{ur} = \left(a - c + \beta^u\right)/2$. This implies that $q^{*i} = 0$; the vertically integrated firm is a monopolist in both markets. If firms are not symmetric, the rival firm will be active in the downstream market provided that it is more efficient than the downstream subsidiary of the vertically integrated firm.[12]

3.2 First-order incentive constraint

The regulator realises that if the network owner is also present downstream, truthful revelation must be based on the joint profit function for the vertically integrated firm, i.e., equation (3). The reason is that the report of efficiency made to the regulator internalises any effects that the report (and resulting infrastructure quality) has on the downstream profits.

Lemma 3 *Local incentive compatibility requires that:*

$$\frac{d\Pi^v}{d\beta^u} = -Q^* + \left[(w - \beta^u) + P_Q q^{*v}\right] \frac{dq^{*i}}{d\beta^u}$$

Proof: Apply the envelope theorem to (3), given that (q^v, q^i), and r are chosen optimally (in stages 4 and 3, respectively). More generally, we can write the first-order incentive constraint as: $\frac{d\Pi^v}{d\beta^u} = \frac{\partial\Pi^v}{\partial\beta^u} + \frac{\partial\Pi^v}{\partial q^i}\frac{dq^{*i}}{d\beta^u} + \frac{\partial\Pi^v}{\partial r}\frac{dr^*}{d\beta^u} + \frac{\partial\Pi^v}{\partial q^v}\frac{dq^{*v}}{d\beta^u}$. However, using the first-order conditions determining r^* and q^{*v}, this simplifies to $\frac{d\Pi^v}{d\beta^u} = \frac{\partial\Pi^v}{\partial\beta^u} + \frac{\partial\Pi^v}{\partial q^i}\frac{dq^{*i}}{d\beta^u}$. The expression $dq^{*i}/d\beta^u$ will be different in the two subcases we discuss (i.e., foreclosure or no foreclosure). QED.

From lemma 3 we observe that the regulated firm may face countervailing incentives,[13] which come from the process of internalising the effects on downstream profit. A necessary, but not sufficient condition for countervailing incentives to be present is that the expression in the squared brackets is positive, or equivalently if $\beta^d > P - w > \beta^i + r$ (see lemma 1). The first component of the incentive constraint, $-Q^*$, is also found in, e.g., the Baron and Myerson (1982) model of regulation of a monopolist with unknown costs. Reducing the quantity produced for less efficient types makes it less desirable for efficient types to imitate less efficient types. In the

[12]Inserting for the (unregulated) access charge in the rival's equilibrium quantity yields $Q^{*i} = 2\left(\beta^d - \beta^i\right)/5$.

[13]See Lewis and Sappington (1989).

model presented here, there is in addition an effect on the downstream equilibrium of changing β^u. By the envelope theorem, the change of firm v's downstream quantity from changing β^u does not affect the incentive constraint. However, the change in equilibrium quantity for the competitor(s) affects the price of the final product, and consequently, the incentives change. This must be taken into account when formulating the incentive constraint.

If the sign of the incentive constraint is unambiguously negative for all values of β^u in the support, the firm has only incentives to overstate its upstream costs. The positive element implies that the firm has incentives to understate its costs upstream for some realisations of the efficiency parameter. The countervailing incentive stems from the fact that a lower level of efficiency (i.e., a higher β^u) effectively increases the equilibrium quantity of firm i. This has two effects, working in opposite directions: 1) For a given quantity for firm v, the price on inframarginal units falls. To retain incentive compatibility, the information rent must increase. 2) Higher upstream (marginal) costs implies higher quantity for firm i, which has a positive impact on profits since it increases the revenue that the vertically integrated firm earns on its upstream operations as the rival requires more access capacity.[14] This effect tends towards a lower level of information rents.

Thus, regulators may want to impose countervailing incentives in order to reduce information rents. In certain cases (e.g., if lump-sum transfers can be used, or if there is no break-even constraint on upstream profit), we would expect that optimal access charges should be set below marginal cost of providing access. The reason for this is that there is unregulated Cournot competition downstream, and lowering the marginal costs (of the rival firm) corrects for the inefficiency downstream by raising equilibrium quantities. Higher quantities result in increased consumer surplus. However, there is a social cost associated with using lump-sum transfers. Whether the optimal access charge is higher or lower than the marginal cost of providing access depends on which distortion is the more costly at the margin - the cost of

[14]Note that the negative effect on upstream profits from the negative impact that increasing β^u has on q^{*v} disappears by use of the envelope theorem. The only effect to consider is the one related to the change in the rival's quantity.

public funds, or the deadweight-loss downstream.

3.2.1 No foreclosure

If the vertically integrated firm cannot discriminate against the rival using non-price behaviour, the incentive constraint in lemma 3 can be rewritten as (using the linear demand function):

$$\begin{aligned}\frac{d\Pi^v}{d\beta^u} &= -Q^* + \left[w - \beta^u - q^{*v}\right]\frac{\partial q^{*i}}{\partial \beta^u} \\ &= -\left[\frac{7a - 2\beta^i - 5\beta^d - 5w - 2\beta^u}{9}\right]\end{aligned} \tag{14}$$

We observe that the information rent will be decreasing in w, as $d^2\Pi^v/d\beta^u dw = 5/9$. This implies that by increasing the access charge, the information rent is reduced. By increasing the access charge w the first-order incentive compatibility constraint becomes less negative, and consequently, the gain from imitating less efficient types becomes smaller. Note that since $\partial q^{*i}/\partial\beta^u > 0$, a necessary condition for creating countervailing incentives is $w > \beta^u$.

3.2.2 Foreclosure

Assume that the vertically integrated firm *can* discriminate the rival using non-price behaviour. In this case there are two effects of changing β^u on q^{*i}. The *direct effect*, $\partial q^{*i}/\partial\beta^u$, which is positive, is the effect when the level of foreclosure is fixed. Upstream (marginal) costs are part of the total marginal costs, and increasing either upstream or downstream marginal costs adversely affects the downstream quantity produced by the vertically integrated firm. The *indirect effect*, which is negative, takes into consideration that the level of foreclosure is also affected when β^u changes. Increasing upstream inefficiency increases the incentive to foreclose the rival firm, and increasing foreclosure leads to a lower level of equilibrium quantity for the rival as its perceived costs are increased. Which effect dominates depends on the characteristics of the cost function $C(r)$. If the cost function is sufficiently convex in r, or $\varphi > 2/3$, the direct effect dominates and $dq^{*i}/d\beta^u > 0$.

Using the fact that $dq^{*i}/d\beta^u = \partial q^{*i}/\partial\beta^u + (\partial q^{*i}/\partial r)(dr^*/d\beta^u)$, the incentive constraint in lemma 3 can be rewritten as (using the linear demand and quadratic cost functions):

$$\frac{d\Pi^v}{d\beta^u} = -\left[q^{*i} + \frac{12\varphi - 4}{9\varphi - 2}q^{*v}\right] + (w - \beta^u)\frac{3\varphi - 2}{9\varphi - 2} \tag{15}$$

The asymmetry of information may cause the optimally regulated access charge to switch from being less than marginal cost to being higher than marginal cost, since setting $w > \beta^u$ may reduce the informational costs of access provision. Provided that $w > \beta^u$ (which turns out to be the case in equilibrium), the condition $\varphi > 2/3$ also ensures that there is a countervailing incentive present.

As it turns out, the parameter φ also determines the sign of the marginal information cost. The single-crossing condition is given by:

$$\frac{d^2\Pi^v}{d\beta^u dw} = \frac{(3\varphi - 2)(15\varphi - 2)}{(9\varphi - 2)^2} \tag{16}$$

From eqn. (16) we observe that the marginal information cost, $d^2\Pi^v/d\beta^u dw$ is positive if $\varphi > 2/3$, equal to 0 if $\varphi = 2/3$, and negative if $2/15 < \varphi < 2/3$. If $d^2\Pi^v/d\beta^u dw > 0$ it implies that the information rent awarded to the regulated firm decreases in w. Note that if the cost function is not convex enough (that is, $\varphi < 2/3$), then the information rent is in fact increasing in w. For $\varphi = 2/3$, the asymmetry of information does not affect the optimal solution - the full information solution coincides with the solution under asymmetric information.

3.3 Second order incentive constraints

The necessary condition for the mechanism to be implementable, with the control variable w, is given in the following lemma (adapted from Fudenberg and Tirole, 1991):

Lemma 5 *The mechanism $M = \left\{w\left(\widehat{\beta}^u\right), t\left(\widehat{\beta}^u\right)\right\}$ is implementable if and only if:*

$$\frac{\partial}{\partial\beta^u}\left(\frac{\partial\Pi^{*v}}{\partial w}\right)\frac{dw}{d\beta^u} \geq 0$$

The proof of lemma 5 is in an appendix.

It can be shown that $\frac{\partial^2 \Pi^v}{\partial w \partial \beta^u}$ is indeed positive in the absence of foreclosure. When foreclosure is possible $\frac{\partial^2 \Pi^v}{\partial w \partial \beta^u}$ is positive provided that $\varphi > 2/3$.[15] Then, the more efficient the upstream firm is, the lower is the access charge. This is summarised in a proposition:

Proposition 1 *When* $\frac{\partial^2 \Pi^v}{\partial w \partial \beta^u} > 0$, *the following condition is sufficient for implementation of the mechanism* M*:* $dw/d\beta^u \geq 0$.

A higher level of β^u results in a higher level of foreclosure, r, which implicitly increases the costs of the rivals. As for the access price, w, we observe that an increase in the access price reduces the incentive to foreclose rival firms. Foreclosure and pricing of access are substitute activities to improve the competitive situation downstream for the vertically integrated firm. A sufficient condition on w for implementation of the mechanism M is that w increases as the inefficiency level of the regulated firm increases. If this is not satisfied, the regulated firm will use its possibilities to improve its own situation through the unregulated activity; i.e., by foreclosing the rivals. The profit-maximising level of r is increasing in β^u, and w is increasing in β^u (second-order sufficient condition for implementation). The instruments r and w are substitutes for the vertically integrated firm when it comes to imposing higher costs on its competitors, through which it can improve its own situation in the competitive downstream market. This can also be related to the multi-task literature:[16] If the regulated firm is not given enough margin along the regulated dimension, $w - \beta^u$, this may have a negative effect on the unregulated (unobservable) dimension, r. A small margin in the regulated dimension effectively means that the regulated firm is given strong incentives to foreclose.

When foreclosure is possible and the cost function is not sufficiently convex (i.e., $\varphi < 2/3$), we obtain a qualitatively different requirement for implementation:

Proposition 2 *When* $\frac{\partial^2 \Pi^v}{\partial w \partial \beta^u} < 0$, *the following condition is sufficient for implementation of the mechanism* M*:* $dw/d\beta^u \leq 0$.

[15] $\frac{\partial^2 \Pi^v}{\partial w \partial \beta^u} \geq 0$ is also known as the single-crossing condition.

[16] See Holmström and Milgrom (1991).

As is mentioned above, the information rent is increasing in the access charge when $\varphi < 2/3$; i.e., when the convexity of the foreclosure costs is less pronounced. In addition, there are no countervailing incentives present.

3.4 Full information benchmark

The absence of asymmetric information is used as a benchmark case to investigate the consequences that the regulator's imperfect information has on optimal regulatory policy. This is not the first-best solution, as the number of firms in the unregulated downstream market is exogenously given (i.e., no free entry). The regulator maximises the following problem (substituting for transfer):

$$\begin{aligned}\max_{w} W &= \frac{1}{2}Q^2 + (1+\lambda)\left[(q^v)^2 + (w-\beta^u)\,q^i - \frac{\varphi}{2}r^2\right] \\ &\quad + \left(q^i\right)^2 - (1+\lambda)\,F - \lambda\Pi^v\end{aligned} \tag{17}$$

subject to

$$\Pi^v\left(\beta^u\right) \geq 0 \tag{PC}$$

$$r^* = \arg\max_{r} \Pi^v\left(q^{*v}, q^{*i}\right) \tag{S.3}$$

$$q^v = q^{*v} \geq 0,\ q^i = q^{*i} \geq 0 \tag{S.4}$$

That is, r^* is the profit-maximising choice of foreclosure from stage 3 (constraint S.3), and constraint S.4 represents the Nash-equilibrium in stage 4 of the game. This is to say that when deciding on the optimal policy, the regulator must take into account the effects that a change in the quality of the infrastructure has on both firm v's decision on the level of foreclosure activity (r), and the effect in the quantity competition game. Since rents to the regulated firm are costly, the regulator will determine transfers such that $\Pi^v\left(\beta^u\right) = 0$. The solution to the maximisation problem of the regulator is determined by:

$$
\begin{aligned}
0 &= Q^* \frac{dQ^*}{dw} + (1+\lambda)\left(2q^{*v}\frac{dq^{*v}}{dw} + q^{*i} + (w-\beta^u)\frac{dq^{*i}}{dw}\right. \\
&\quad \left. -\varphi r^* \frac{dr^*}{dw}\right) + 2q^{*i}\frac{dq^{*i}}{dw}
\end{aligned}
$$

When foreclosure is not an option, we assume that $r = 0$. Consequently, we only need to consider the *direct* effect of changes in w on q^{*v}, q^{*i} and Q^*.

3.4.1 Access charge regulation (no foreclosure)

The full information optimal level of access charges when the vertically integrated firm cannot use non-price discrimination is denoted w_{nr}^f, and can be written as:

$$
w_{nr}^f = \frac{a(5\lambda-1) - \beta^d(4+\lambda) + \beta^i(5-4\lambda) + \beta^u(2+5\lambda)}{1+10\lambda}
$$

Optimal access charge when there is no foreclosure has the following properties (the magnitude of the social cost of public funds, λ, plays an important part in determining the signs):[17]

$$
\begin{aligned}
\partial w_{nr}^f/\partial a &= (5\lambda-1)/(1+10\lambda) > 0 \text{ if } \lambda > 1/5 \\
\partial w_{nr}^f/\partial \beta^d &= -(4+\lambda)/(1+10\lambda) < 0 \\
\partial w_{nr}^f/\partial \beta^i &= (5-4\lambda)/(1+10\lambda) > 0 \text{ if } \lambda < 5/4 \\
\partial w_{nr}^f/\partial \beta^u &= (2+5\lambda)/(1+10\lambda) > 0
\end{aligned}
$$

We see that the optimal access charge is used to ensure that market shares are distorted in favour of the more efficient firm (in the end-user production), with the access charge being (unambiguously) decreasing in β^d and increasing in β^i. A similar result is obtained by Lewis and Sappington (1999). However, the latter property is

[17] If both firms have identical downstream marginal costs ($\beta^d = \beta^i = c$), then w_{nr}^f is decreasing in c provided that $\lambda > 1/5$. For a high enough λ, the optimal access charge is determined to yield a positive access margin. Furthermore, the regulated firm earns positive profits in the downstream market. This implies that the transfer is a negative number. In this case, the primary concern of the regulator is to ensure that the profit of the regulated firm is high, which it is when total quantity is high, as this yields a large transfer *to* the regulator.

only valid when the social cost of public funds is not too large. When $\lambda > 5/4$, we observe that the optimal access charge is smaller the higher is the marginal cost of the rival firm. In such a situation, efficiency in production becomes less important whereas the cost of transfers is the dominating concern.[18]

Proposition 3 *When the vertically integrated firm cannot discriminate against its rival (neither by setting a high access charge nor by non-price means), the access charge is (normally) distorted away from the marginal cost of providing access in order to favour the more efficient downstream firm.*

In general, we cannot determine the magnitude of w_{nr}^f relative to the upstream marginal cost, β^u. What we see in the present model is that the optimal access charge is in fact increasing in the upstream marginal cost.[19] Using numerical examples, it is possible to show the following:[20]

Proposition 4 *If $\lambda = 0$ and the rival firm (if any) has a cost advantage, then $w_{nr}^f < \beta^u$. The higher the downstream marginal cost of the integrated firm is relative to the rival firm the further below upstream marginal cost is the access charge. If the integrated firm enjoys a cost advantage downstream ($\beta^i > 0, \beta^d = 0$), $w_{nr}^f > \beta^u$ for some parameter values.*

With no social cost of public funds ($\lambda = 0$), the optimal access charge policy requires that access is subsidised to correct the distortion caused by imperfect competition in the downstream market. This causes an access deficit for the vertically integrated firm, but the regulatory agency can use costless lump-sum transfers to finance this deficit.

[18] When the shadow cost is large, we know that the vertically integrated firm has a positive access margin. In the downstream industry there is at least a non-zero profit. Since there is full information, transfers are set such that the regulated firm's profit is exactly 0, and, consequently, transfers are negative if we assume $F = 0$ - i.e., the regulated firm pays the regulator. So when $\lambda > 5/4$, the gain to the regulator from receiving money from the regulated firm outweighs the cost of the distorted production downstream.

[19] If the optimal access charge is *not* increasing in the upstream marginal cost, this implies that market shares are distorted in favour of the downstream rival.

[20] In the calculations, I assume that $\beta^d, \beta^i, \beta^u \in [0, 1]$.

If the social cost of public funds is large, it becomes very costly to use lump-sum transfers to subsidise the access deficit. Consequently, optimal policy calls for an access surplus, except for the case where the downstream subsidiary is highly inefficient. When the downstream subsidiary is highly inefficient it is better to subsidise the rival to distort market shares to the integrated firm's disadvantage.

Proposition 5 *For $\beta^d = \beta^i = 0$, $w^f_{nr} > \beta^u$ when $\lambda \gtrapprox 0.2$. For shadow cost of public funds below this level, the optimal access charge is always less than β^u.*

It is reasonable that the degree to which there is an access deficit ($w^f_{nr} < \beta^u$) is smaller the higher is the social cost of public funds. When λ becomes large, the cost of transferring a lump-sum payment to the regulated firm becomes a heavier burden. Consequently, optimal policy on access charges shifts from subsidising the rival firm and using compensating lump-sum transfers, to allowing the regulated firm to earn a positive margin on access provision.

We assume that the number of firms competing downstream is exogenously given (and equal to 2), which implies that there is an inefficiency in this market. Subsidising the rival's cost (by subsidising access) is one way to correct for this inefficiency. However, there is also a second distortion from the fact that transfers to the regulated firm must be financed through taxing other sectors of the economy. Which of these distortions is the more pronounced will determine the appropriate policy on access pricing. When the latter effect is more important, access cannot be subsidised.

3.4.2 Access charge regulation (foreclosure)

Under full information, and allowing for non-price discrimination, the constrained optimal access charge, w^f_r, is given by:

$$\begin{aligned} w^f_r &= \frac{1}{N}\left[a\left(\lambda\left(45\varphi^2 - 28\varphi + 4\right) - 9\varphi^2 + 38\varphi - 16\right)\right. \\ &\quad + \beta^d\left(16 + 8\varphi - 36\varphi^2 + \lambda\left(20\varphi - 9\varphi^2 - 4\right)\right) + \beta^i\left(45\varphi^2 - 46\varphi + \lambda\left(8\varphi - 36\varphi^2\right)\right) \\ &\quad \left. + \beta^u\left(\lambda\left(45\varphi^2 - 28\varphi + 4\right) + 18\varphi^2 + 14\varphi - 12\right)\right] \end{aligned}$$

where $N \equiv (9\varphi^2 + 52\varphi - 28 + 2\lambda\left(45\varphi^2 - 28\varphi + 4\right))$. In this case, the optimal access charge will depend not only on relative efficiencies and the social cost of public funds, but also on the convexity of the cost function $C(r)$ (or, equivalently, the magnitude of the parameter φ). When we open up for the possibility of foreclosure activities by the vertically integrated firm, a more general statement can be given with regard to the access charge relative to the upstream marginal cost.

Proposition 6 *When the integrated firm can use non-price discrimination, we find that $w_r^f > \beta^u$ for all realisations of parameter values. Furthermore, the optimal access charge is higher when the cost function is less convex in the foreclosure variable r (i.e., the smaller is φ).*

We have seen above that the degree to which the vertically integrated firm chooses to foreclose its rival depends (in part) on the level of the access charge. The higher is the access charge, the lower is the level of foreclosure. From a social point of view, foreclosure is costly and should be kept as low as possible. Consequently, the optimal access charge accounts for the opportunity to foreclose rival firms by setting the access charge higher than is the case when foreclosure is not an option. The fact that the optimal access charge is higher when the cost related to foreclosure is less convex in r reflects the fact that a less convex cost function implies that it is less costly to foreclose rival firms. Thus, when foreclosure is costly the level of foreclosure is less responsive to changes in the access charge (i.e., $\partial r^*/\partial w$ is less negative when φ is large) and, consequently, the optimal access charge is lower compared to the case when foreclosure is less costly (i.e., a lower φ).

Proposition 7 *When non-price discrimination is an option for the vertically integrated firm, the access charge is used to distort downstream market shares in favour of the more efficient firm only when the costs of foreclosure is sufficiently high. For $2/3 < \varphi < 1$ and $\lambda \lesssim 5/4$, the access charge is set in order to attempt to cancel out any differences in the marginal costs of the downstream firms.*

This can be seen from the comparative static results for the optimal access charge when foreclosure is possible, since we observe that $\partial w_r^f/\partial\beta^u > 0$, $\partial w_r^f/\partial\beta^d \gtrless 0$, $\partial w_r^f/\partial\beta^i \gtrless 0$, and $\partial w_r^f/\partial a > 0$.[21]

[21] When the downstream firms have identical marginal costs (i.e., $\beta^d = \beta^i = c$), the optimal access charge is unambiguously decreasing in c.

If $2/3 < \varphi < 59/75$, then $\partial w_r^f / \partial \beta^d > 0$.[22] We know from above that the incentive to foreclosure is negatively related to the downstream cost β^d (and more so the less costly foreclosure is), and that increasing the access charge reduces the incentives to foreclose rivals. Furthermore, we know that when the cost of foreclosure is relatively low, the vertically integrated firm will choose a high level for r, which is detrimental to welfare. An increase in β^d will subsequently lower the level of foreclosure both directly as determined by eqn. (12), and also indirectly through raising the access charge.

When φ is large, the level of foreclosure is less responsive to changes in β^d (or any other of the exogenous parameters), so when β^d (or the access charge w) increases the level of foreclosure is reduced less than when foreclosure entails a lower cost. Consequently, with a high cost of foreclosure the cost of raising the access charge to counter the foreclosure (measured by the detrimental effect on welfare through reduced total quantity) exceeds the benefits of a level playing field.

Furthermore, $\partial w_r^f / \partial \beta^i$ is negative for all levels of the shadow cost of public funds provided that $\varphi \leq 1$. If $\varphi > 1$, $\partial w_r^f / \partial \beta^i > 0$ when the shadow cost of public funds is sufficiently low.[23] Thus, contrary to the case where non-price discrimination is not possible, the regulator does not (necessarily) use the access charge to distort the market share in favour of the more efficient firm.

3.5 Asymmetric information

In order to maximise the welfare function $W(w)$, the regulator offers the incentive compatible contract

$$\left(w\left(\widehat{\beta}^u\right), t\left(\widehat{\beta}^u\right)\right)$$

[22] If we restrict our attention to the case where $\varphi > 2/3$ (to avoid problems of discontinuities) and assume that $\lambda \in (0, 5)$, I find that for low enough φ (that is, $\varphi < 59/75$), $\partial w_r^f / \partial \beta^d > 0$, for all levels of the shadow cost of public funds. The higher is φ above this level, the higher must the shadow cost of public fund be for $\partial w_r^f / \partial \beta^d$ to be positive. When $\varphi > 13/10$, $\partial w_r^f / \partial \beta^d < 0$ for all permissible values of λ.

[23] The more convex the cost of foreclosure is, the higher can λ be. However, even when φ becomes very large, the optimal access charge is decreasing in β^i if $\lambda \gtrsim 5/4$.

to the upstream firm, using his knowledge of the distribution and support of the unknown parameter, and of the way the game is played in the subsequent stages. The contract specifies the access charge (w), and transfer to the regulated firm (t).

If the efficiency level of the network provider is not known to the regulator, he must offer a regulatory contract which induces the regulated firm to reveal its true efficiency level. The regulator will then maximise welfare, subject to incentive and participation constraints.

The regulator's optimisation problem is given by:

$$\max_{w} \int_{\underline{\beta}}^{\overline{\beta}} \Big[CS\left(w, \beta^{u}\right) + (1+\lambda) \left[\left(q^{v}\right)^{2} + \left(w - \beta^{u}\right) q^{i} - \frac{\varphi}{2} r^{2} \right] + \left(q^{i}\right)^{2} \quad (18)$$
$$- (1+\lambda) F - \lambda \Pi^{v} \Big] dG\left(\beta^{u}\right)$$

subject to

$$\frac{d\Pi^{v}}{d\beta^{u}} = -Q^{*} + \left[\left(w - \beta^{u}\right) + P_{Q} q^{*v}\right] \frac{dq^{*i}}{d\beta^{u}} \quad \text{(IC)}$$

$$\Pi^{v}\left(\beta^{u}\right) \geq 0 \quad \text{(PC)}$$

$$\begin{aligned} dw/d\beta^{u} &\geq 0 \quad \text{if } d^{2}\Pi^{v}/dwd\beta^{u} > 0 \\ dw/d\beta^{u} &\leq 0 \quad \text{if } d^{2}\Pi^{v}/dwd\beta^{u} < 0 \end{aligned} \quad \text{(SIC)}$$

$$r^{*} = \arg\max_{r} \Pi^{v}\left(q^{*v}, q^{*i}\right) \quad \text{(S.3)}$$

$$q^{v} = q^{*v} \geq 0,\; q^{i} = q^{*i} \geq 0 \quad \text{(S.4)}$$

The constraint S.3 is the profit-maximising level of foreclosure, and S.4 is the Cournot-equilibrium. The constraint SIC, is the second-order incentive constraint, which we will ignore for now but check *ex post*. The participation constraints must be satisfied to induce voluntary participation. Since there are countervailing incentives, the standard procedure of assuming that (PC) binds for the least efficient type is no longer valid in the general case. Instead, the participation constraint is introduced explicitly into the optimisation problem.[24] However, I will assume that the sign of the incentive constraint does not change over the relevant interval for β^u, and use the standard methodology to solve this problem. This means that the participation constraint will be binding only for the least efficient firm (i.e., for $\beta^u = \overline{\beta}$).[25] Integrating the incentive constraint by parts, using the fact that [PC] binds for the least efficient firm $\overline{\beta}$, we get an expression for the regulator's virtual surplus (i.e., welfare adjusted to take account of the informational costs of inducing truthful revelation):[26]

$$\begin{aligned} VS \;=\; & \int_{\underline{\beta}}^{\overline{\beta}} \Big\{ CS\left(w,\beta^u\right) + (1+\lambda)\left[\left(q^v\right)^2 + \left(w-\beta^u\right)q^i - \frac{\varphi}{2}r^2\right] + \left(q^i\right)^2 \qquad (19) \\ & - (1+\lambda)F + \lambda\frac{G\left(\beta^u\right)}{g\left(\beta^u\right)}\left[\left(\left(w-\beta^u\right) - q^{*v}\right)\frac{dq^{*i}}{d\beta^u} - Q^*\right]\Big\}\, dG\left(\beta^u\right) \end{aligned}$$

Maximising expression (19) pointwise with respect to w, defines the optimal access charge under asymmetric information, w^a_{nr}:

$$\frac{\partial VS}{\partial w} = \frac{\partial W}{\partial w} + \lambda\frac{G\left(\beta^u\right)}{g\left(\beta^u\right)}\frac{\partial^2 \Pi^v}{\partial \beta^u \partial w} = 0$$

where W refers to the welfare function under full information. We see that the optimal access charge under asymmetric information, w^a_{nr}, is equal to the optimal

[24] See Maggi and Rodriguez-Clare (1995) for a general treatment of the problem of countervailing incentives.

[25] For the case without forclosure, a sufficient condition for ensuring that $d\Pi^v/d\beta^u < 0$, and consequently that the condition $\Pi^v\left(\overline{\beta}\right) \geq 0$ is sufficient to guarantee participation from all types, is $a > w$.

[26] See appendix 1.

access charge under full information plus an incentive correction term. The magnitude and sign of the incentive correction term will depend on whether the regulated firm can or cannot foreclose its rival. We will consider these two subcases below.

3.5.1 Access charge regulation (no foreclosure)

When the regulated firm cannot discriminate against the rival firm using non-price means, the incentive correction form is given by:

$$\lambda \frac{G(\beta^u)}{g(\beta^u)} \frac{5}{9} \geq 0 \tag{20}$$

Assuming $\lambda > 0$, eqn. (20) is zero only when $\beta^u = \underline{\beta}$. Thus, we have the standard "no distortion at the top" - i.e., the allocation for a regulated firm of the most efficient type is not distorted relative to the full information benchmark.[27]

FIGURE 1 HERE

Figure 1: No foreclosure - optimal access charges under full information (dashed line) and asymmetric information (solid line). Uniform distribution with parameter values: $(\beta^i = \beta^d = 0, a = 2, \lambda = 0.25)$

Proposition 8 *For all types of the regulated firm except $\beta^u = \underline{\beta}$, the optimal access charge is higher under asymmetric information relative to the full information case.*

The intuition is clear. By setting a high access charge the regulator effectively restricts total output downstream since $\partial Q^* / \partial w < 0$, which reduces the information rent necessary to induce truthful revelation. Furthermore, for a given β^u we see that $(w^a_{nr} - \beta^u) > (w^f_{nr} - \beta^u)$, such that the countervailing incentive is more pronounced. As is discussed above, creating countervailing incentives can be beneficial from a social welfare point of view as this reduces the information rents. The comparative static results for w^a_{nr} are the same as in the full information case, w^f_{nr}.

[27] A sufficient condition for the second-order incentive constraint to be satisfied is if the inverse hasard rate is increasing in β; i.e., $\frac{d}{d\beta^u} \frac{G(\beta^u)}{g(\beta^u)} \geq 0$. This assumption is satisfied for a number of distributions.

3.5.2 Access charge regulation (foreclosure)

If the regulated firm can discriminate against rival firms, the incentive correction term is given by the following expression:

$$\lambda\frac{G\left(\beta^{u}\right)}{g\left(\beta^{u}\right)}\left(\frac{3\varphi-2}{9\varphi-2}\right)\left(\frac{6-33\varphi}{9\varphi-2}\right)$$

which sign and magnitude depends on both the convexity of the cost function and the efficiency of the regulated firm. Observe that the solution is not affected by the asymmetry of information provided that $\beta^{u}=\underline{\beta}$, and/or $\varphi=2/3$.

Proposition 9 *Provided that $\lambda>0$, optimal access charge is higher under asymmetric information relative to the full information solution for $\beta^{u}>\underline{\beta}$ and $\varphi>2/3$; i.e., $w_{r}^{f}<w_{r}^{a}$. For $\varphi=2/3$ and/or $\beta^{u}=\underline{\beta}$, there is no distortion relative to the full information solution. For $\varphi<2/3$, $w_{r}^{f}>w_{r}^{a}$.*

The intuition is similar to the no-foreclosure case. By increasing the access charge, total downstream output is reduced and the countervailing incentives becomes more pronounced. Part of the reason that the incentive correction term disappears when $\varphi=2/3$ is that $dq^{*i}/d\beta^{u}=0$; i.e., the (positive) direct effect is exactly offset by the (negative) indirect effect. This leaves us with a first-order incentive constraint equal to $d\Pi^{v}/d\beta^{u}=-Q^{*}$. For $\varphi=2/3$, we observe from eqn. (12) that $dr^{*}/dw=-1$. Thus, any increases in the access charge is matched exactly by a reduction in the level of foreclosure. This implies that $dQ^{*}/dw=\partial Q^{*}/\partial w+\left(\partial Q^{*}/\partial r\right)\partial r^{*}/\partial w=0$, since $\partial Q^{*}/\partial w=\partial Q^{*}/\partial r$. Consequently, any changes in the access charge has no effect on the incentive constraint at the margin for $\varphi=2/3$.

FIGURE 2 HERE

Figure 2: Foreclosure - optimal access charge under asymmetric information (solid line) and full information (dashed line). Uniform distribution with parameter

values: $\left(\beta^i = \beta^d = 0, a = 2, \lambda = 0.25, \varphi = 1\right)$

When $\varphi < 2/3$, we know from above that the second-order incentive constraint requires that $dw/d\beta^u \leq 0$. When solving the regulator's maximisation problem without taking the second-order constraint into account, the solution candidate for optimal access charge is *increasing* in β^u. This cannot be the optimal solution. This will imply that the optimal solution is characterised by "bunching"; i.e., that different types are given the same contract.[28]

The comparative static results for w_r^a are as follows: $\partial w_r^a/\partial\beta^u > 0$, $\partial w_r^a/\partial a > 0$, $\partial w_r^a/\partial\beta^d \gtrless 0$, and $\partial w_r^a/\partial\beta^i \gtrless 0$. To avoid problems of discontinuities I will assume that $\varphi > 2/3$. When $\varphi > 4/3$, the optimal access charge is decreasing in the marginal cost of the downstream subsidiary (i.e., $\partial w_r^a/\partial\beta^d < 0$).[29] The optimal access charge is decreasing in β^i whenever $\varphi < 1$, and may be increasing in β^i if $\varphi > 1$ provided that λ is not too large. This is essentially the same as is the case for optimal access charge under full information with foreclosure, w_r^f, and the intuition is explained in section 3.4.2.

3.6 Welfare comparisons

In this model, the welfare considerations are quite straightforward. As it turns out, the level of welfare is always higher whenever non-price discrimination is *not* possible, both under full and asymmetric information. The level of welfare is highest under full information and no foreclosure, and is lowest when there is asymmetric information and foreclosure. The reason for this is simply that foreclosure entails two types of costs (and virtually no benefits in this model); the monetary cost $C(r)$, but also the cost in terms of reduced consumer surplus of a reduction in the total downstream production.

When there is no access charge regulation, the rival firm is foreclosed completely if both firms' downstream costs are identical (that is, if $\beta^d = \beta^i$).[30] In this situation, the vertically integrated firm acts as a monopolist both upstream and downstream.

[28] The optimal policy in this case will not be expressed formally.

[29] For lower φ, $\partial w_r^a/\partial\beta^d < 0$ if the shadow cost of public funds is low enough.

[30] See fig. 3 for an illustration of the welfare levels under full information. The parameter values

Under full information, welfare is always highest under regulation when foreclosure is not possible. However, welfare may be higher in the unregulated case than in the regulated case if foreclosure is possible, provided that the cost of foreclosure is sufficiently high (with $\varphi > 1$). Consequently, the regulator cannot do better than the unregulated outcome when it cannot control the quality of access provided by the vertically integrated firm, provided that the cost of foreclosure is sufficiently high.[31] The reason for this is that when the integrated firm is unregulated, it can choose to either foreclose its rival using non-price means or by setting a high access charge. Since non-price foreclosure is costly, the firm will set a high access charge and eliminate competition downstream. This will lead to an elimination of the monetary costs associated with foreclosure. When the firm is regulated, it will attempt to foreclose its rival by (costly) non-price means. Consequently, when the costs of such behaviour is sufficiently high, these costs outweigh the negative effect on welfare of the lower monopoly quantity in the unregulated case.

Contrary to the full information case, the unregulated welfare may be higher than what is the case under regulation and no foreclosure when there is asymmetric information (assuming that $\beta^d = \beta^i$). See fig. 4 and 5 for illustrations (fig. 4 with $\varphi = 1$, and fig. 5 with $\varphi = 2$).[32] If the marginal cost of upstream production becomes sufficiently high, welfare is highest in the unregulated case. The reason is a combination of several things. First of all, the regulator does not need to award information rents to the vertically integrated firm in the unregulated case. Furthermore, in the unregulated case where the rival is completely foreclosed, the problem of double marginalisation is avoided which obviously has a positive impact on welfare. An finally, the total downstream quantity falls faster when upstream marginal cost increases in the regulated case than in the unregulated case (as a consequence of

are as follows: $a = 5, \beta^d = \beta^i = 0, F = 0, \lambda = 1/4$ and $\varphi = 2$. The solid line represents welfare with no foreclosure, the short dashed line represents unregulated welfare, and the long dashed line represents welfare with foreclosure.

[31]The regulator will actually never do worse if it does not interfere even if φ is smaller than 1.

[32]The parameter values are as follows: $a = 5, \beta^d = \beta^i = 0, F = 0, \lambda = 1/4$. The solid line represents welfare with no foreclosure, the short dashed line represents unregulated welfare, and the long dashed line represents welfare with foreclosure.

the way the optimal access charge is set). When upstream marginal cost increases, the access margin (i.e., $w - \beta^u$) increases if there is asymmetric information. This is contrary to the case with full information. Since total downstream quantity falls when β^u increases, the distortion in the downstream market becomes larger. To mitigate this distortion, the regulator decides to increase the access charge by less than the increase in β^u which leads the access margin to fall. However, in the presence of asymmetric information, we know that the magnitude of the access margin is important for the level of the information rent necessary for truthful revelation. More specifically, we know that the higher the access margin is, the lower is the information rent. Consequently, the regulator may choose to accept a larger distortion in the downstream market when β^u increases in order to reduce the costly information rents payable to the regulated firm.

4 Summary

Regulatory authorities in a number of countries, including the US and EU, recommend using a cost based policy with respect to the pricing of access to an essential facility - the local loop ("the last ten miles" of the telecommunications network). In the US 1996 Telecommunications Act, prices are recommended to be non-discriminatory and cost based, and may include a reasonable profit to cover non-traffic sensitive costs. The main reason for pursuing such a policy seems to be that regulators feel that this is the best method of ensuring rival firms access to the essential facility at reasonable terms. If the vertically integrated firm - the owner of the network - is allowed to determine access charges without regulatory intervention, it will do so to foreclose rival firms completely. Consequently, there is obviously some scope for the regulation of access charges. However, by focusing too heavily on the costs of providing access in the determination of optimal access charges, the regulators may run the risk of ignoring the possibility that the network owner may be able to discriminate against potential rivals by means of non-price behaviour (such as degrading the quality of access offered to rival firms, which then translates into a lower quality of the rival's final product).

The analysis above suggests that if regulators cannot write complete contracts with respect to the access terms offered to rivals (i.e., the regulators can only control the price a network owner charges, but not the quality of the access), the optimal access charge should be distorted away from marginal costs of providing access to yield a positive profit margin on access. Even if we assume that the regulator can control all aspects of the access terms offered to rival firms (i.e., a complete contract situation), the optimal access charge will in general be different from the marginal cost of providing access in situations where there is imperfect competition downstream. In the absence of any foreclosure activities, the socially optimal access charge may yield either a positive profit margin on access or an access deficit, depending on how socially costly transfers to the regulated firm are. When the integrated firm can practice foreclosure, the regulated firm will always earn a positive access margin, which is not necessarily the case when foreclosure is not an option. Furthermore, the access margin under asymmetric information is always higher than under full information. Consequently, opting for a cost-based regulation of access charge, where cost-based usually is interpreted as the cost of providing access, is not necessarily socially optimal.

Furthermore, as seems quite intuitive, I find that optimal access charges should be determined not only by looking at the costs of providing access, but also examine the costs of providing end user services (for both the vertically integrated firm and its downstream rival) - the relative efficiencies of the service providers. The optimal access charge should not only attempt to price access solely based on the cost of providing access, but it should also pay attention to the distribution of the production downstream between the two firms. Furthermore, my results suggest that the regulation of the access charge is not always used to award the more efficient downstream firm a larger market share.

5 References

Armstrong, M., S. Cowan and J. Vickers (1994), *Regulatory reform: Economic analysis and the British experience*, Cambridge, MA: MIT-Press

Armstrong, M., C. Doyle and J. Vickers (1996), "The access pricing problem: A synthesis", *Journal of Industrial Economics* **44**, 131-150

Baron, D.P. and R. Myerson (1982), "Regulating a monopolist with unknown cost", *American Economic Review* **50**, 911-930

Economides, N. (1998a), "Raising rivals' costs in Complementary Goods Markets: LECs entering into long distance and Microsoft bundling Internet Explorer", mimeo, Stern School of Business, NY

Economides, N. (1998b), "The incentive for non-price discrimination by an input monopolist", *International Journal of Industrial Organization* **16**, 271-284

Fudenberg, D. and J. Tirole (1991), *Game theory*, Cambridge: MA, MIT-Press

Holmström, B. and P. Milgrom (1991), "Multi-task principal-agent analysis: Incentive contracts, asset ownership and job design", *Journal of Law, Economics and Organization* **7**, 26-53

Kreps, D. and Scheinkman (1983), "Quantity precommitment and Bertrand competition yield Cournot outcomes", *Bell Journal of Economics* **14**, 326-337

Laffont, J.J. and J. Tirole (1990), "The regulation of multiproduct firms: I and II", *Journal of Public Economics* **43**, 1-66

Laffont, J.J. and J. Tirole (1996), "Competition through interconnection: Theory and practice", *Journal of Regulatory Economics* **10**, 227-256

Lewis, T., and D. Sappington (1989), "Countervailing incentives in agency problems", *Journal of Economic Theory* **49**, 294-313

Lewis, T., and D. Sappington (1999), "Access pricing with unregulated downstream competition", *Information Economics and Policy* **11**, 73-100

Maggi, G., and A. Rodriguez-Clare (1995), "On countervailing incentives", *Journal of Economic Theory* **66**, 238-263

Rey, P., and J. Tirole (1997), "A primer on foreclosure", forthcoming in *Handbook of Industrial Organization*, vol. 3, Armstrong, M., and R.H. Porter (eds.), North-Holland

Tirole, J. (1988), *The theory of industrial organization*, Cambridge, MA: MIT-Press

Vickers, J. (1995), "Competition and regulation in vertically related markets", *Review of Economic Studies* **62**, 1-17

Weisman, D. (1995), "Regulation and the Vertically Integrated Firm: The Case of RBOC Entry into Interlata Long Distance", *Journal of Regulatory Economics* **8**, 249-266

White, Lucy (1998), "Foreclosure with incomplete information", mimeo, Nuffield College, Oxford

6 Appendicies

Appendix 1

Integrating the local incentive constraint by parts

Our problem is to get the objective function of the regulator to a form which allows us to maximise the integrand pointwise. Inserting for the information rent into the objective function results (in part) in the following expression:

$$\int_{\underline{\beta}}^{\overline{\beta}} \Pi^v(\beta^u)\, g(\beta^u)\, d\beta^u = -\int_{\underline{\beta}}^{\overline{\beta}} \left[\int_{\beta^u}^{\overline{\beta}} \frac{d\Pi^v}{d\beta^u} d\beta^u\right] g(\beta^u)\, d\beta^u$$

where $\frac{d\Pi^v}{d\beta^u} = -Q^* + [(w-\beta^u) + P_Q q^{*v}] \frac{dq^{*i}}{d\beta^u}$ is the local incentive constraint. We will integrate this expression by parts to eliminate the integral over the interval $[\beta^u, \overline{\beta}]$. Define $\left[\int_{\beta^u}^{\overline{\beta}} \frac{d\Pi^v}{d\beta^u} d\beta^u\right] = u(\beta^u)$, and $v(\beta^u) = G(\beta^u)$, which imply that $v'(\beta^u) = dv(\beta^u)/d\beta^u = g(\beta^u)$, and $u'(\beta^u) = du(\beta^u)/d\beta^u = \frac{d\Pi^v}{d\beta^u}$. The formula for integrating by parts is given by: $\int_{\underline{\beta}}^{\overline{\beta}} u'v dx = uv \,|_{\underline{\beta^u}}^{\overline{\beta^u}} - \int_{\underline{\beta}}^{\overline{\beta}} uv' dx$. Using this formula, and the definitions above, we obtain:

$$\begin{aligned} -\int_{\underline{\beta}}^{\overline{\beta}} \left[\int_{\beta}^{\overline{\beta}} \frac{d\Pi^v}{d\beta^u} d\beta^u\right] g(\beta^u)\, d\beta^u &= \left[\int_{\beta^u}^{\overline{\beta}} \frac{d\Pi^v}{d\beta^u} d\beta^u G(\beta^u)\right] \Big|_{\underline{\beta^u}}^{\overline{\beta^u}} \\ &\quad - \int_{\underline{\beta}}^{\overline{\beta}} \frac{d\Pi^v}{d\beta^u} G(\beta^u)\, d\beta^u \end{aligned}$$

Using the fact that $G(\underline{\beta}) = 0$, $G(\overline{\beta}) = 1$, and $\int_{\overline{\beta}}^{\overline{\beta}} \frac{d\Pi^v}{d\beta^u} d\beta^u = 0$, we can rewrite the above as:

$$\int_{\underline{\beta}}^{\overline{\beta}} \Pi^v(\beta^u)\, g(\beta^u)\, d\beta^u = \int_{\underline{\beta}}^{\overline{\beta}} \frac{d\Pi^v}{d\beta^u} G(\beta^u)\, d\beta^u$$

Substituting the right-hand side of this expression for Π^v into the objective function, we obtain the "virtual surplus" function.

Appendix 2

The second-order incentive constraint

In the first stage the regulator proposes an incentive compatible mechanism, $M = \left\{w\left(\widehat{\beta}^u\right), t\left(\widehat{\beta}^u\right)\right\}$. As described in the model chapter, the firm makes a report on its level of efficiency to the regulator (stage 2). When the firm makes this report, it maximises its profits with respect to $\widehat{\beta}^u$ (the reported parameter). This solves the first-order condition: $\frac{\partial \Pi^v}{\partial \widehat{\beta}^u} = 0$. In order for this to be a maximum, the second-order condition $\frac{\partial^2 \Pi^v}{\partial \left(\widehat{\beta}^u\right)^2} \leq 0$ must also be satisfied. At this optimum, changing the value of its report has a marginal net benefit of zero, so taking the total derivative wrt. to $\widehat{\beta}^u$ evaluated at $\widehat{\beta}^u = \beta^u$ yields: $\frac{\partial^2 \Pi^v}{\partial \left(\widehat{\beta}^u\right)^2}\frac{d\widehat{\beta}^u}{d\beta^u} + \frac{\partial^2 \Pi^v}{\partial \widehat{\beta}^u \partial \beta^u} = 0$. The first element is non-positive (from the second-order condition), and the second element $\frac{\partial^2 \Pi^v}{\partial \widehat{\beta}^u \partial \beta^u} = \frac{\partial^2 \Pi^v}{\partial w \partial \beta^u}\frac{dw}{d\beta^u}$ must be non-negative for the equality to hold (lemma 5).

This is the (standard) implementation requirement that the product of the single-crossing condition and the monotonicity condition is non-negative. If for instance the inverse demand is linear in Q and costs, $C(r)$, are quadratic in r, then the information rent is decreasing in w (equivalent to $\frac{\partial^2 \Pi^v}{\partial w \partial \beta^u} > 0$). Sufficient conditions for the information rent being decreasing in w are 1) concave inverse demand with respect to quantity ($P_{QQ} \leq 0$), 2) the regulated firm having a sufficiently large market share of the total downstream market relative to the rival firm, and 3) costs being sufficiently convex in the foreclosure variable, r. We shall investigate this in more detail below. Sufficient conditions for second-order incentive compatibility is then: $\frac{dw}{d\beta^u} \geq 0$. Thus, the second-order incentive condition requires that the access price increases in the upstream inefficiency. QED.

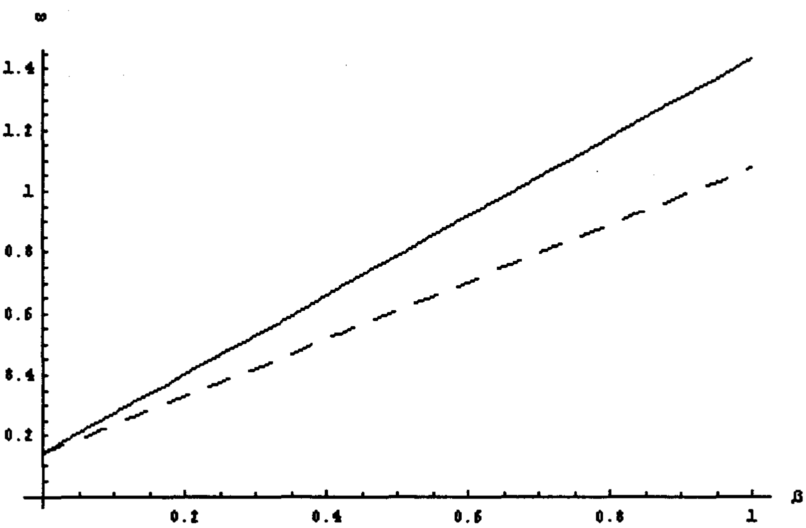


Figure 1:

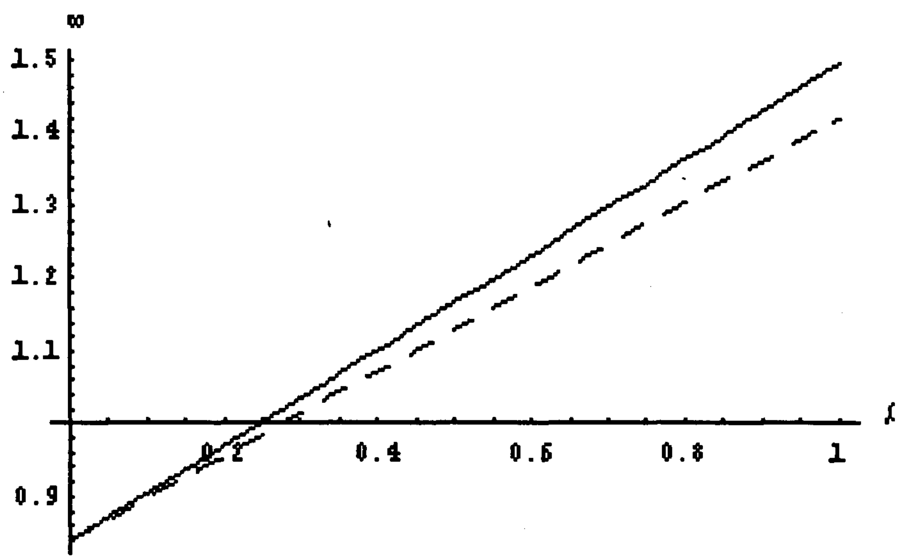


Figure 2:

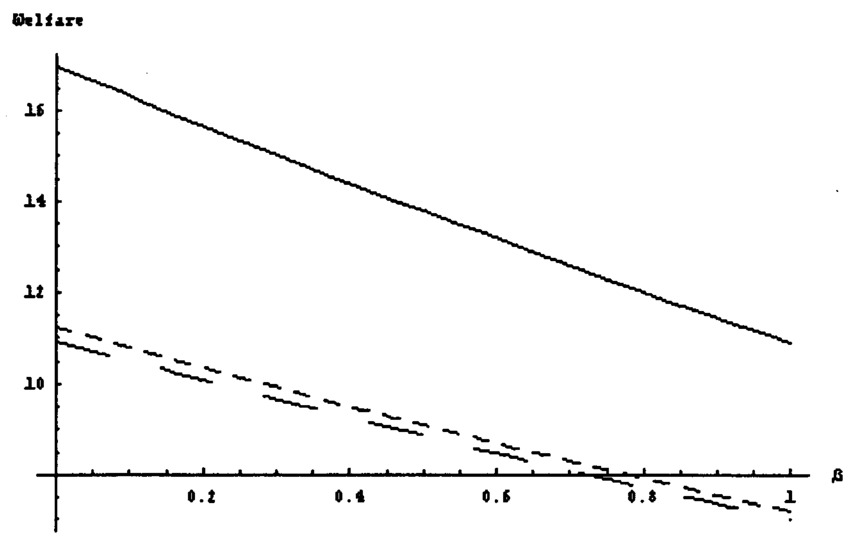


Figure 3:

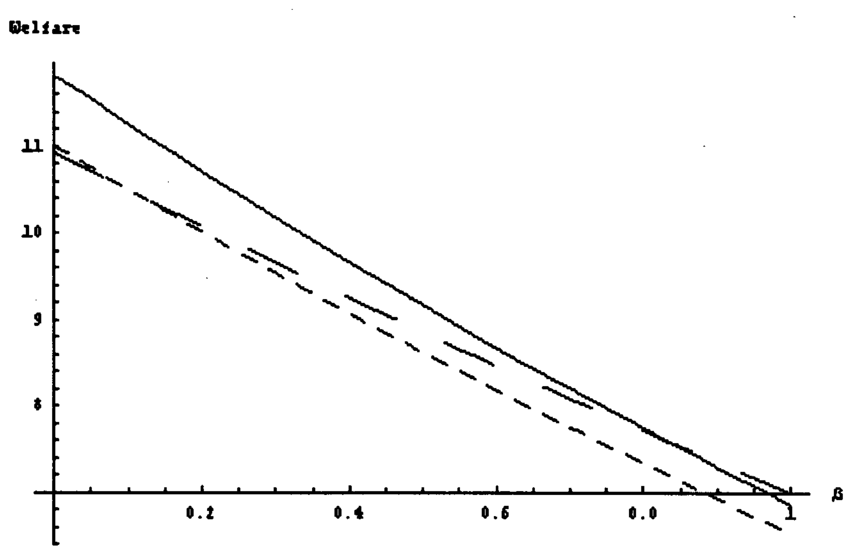


Figure 4:

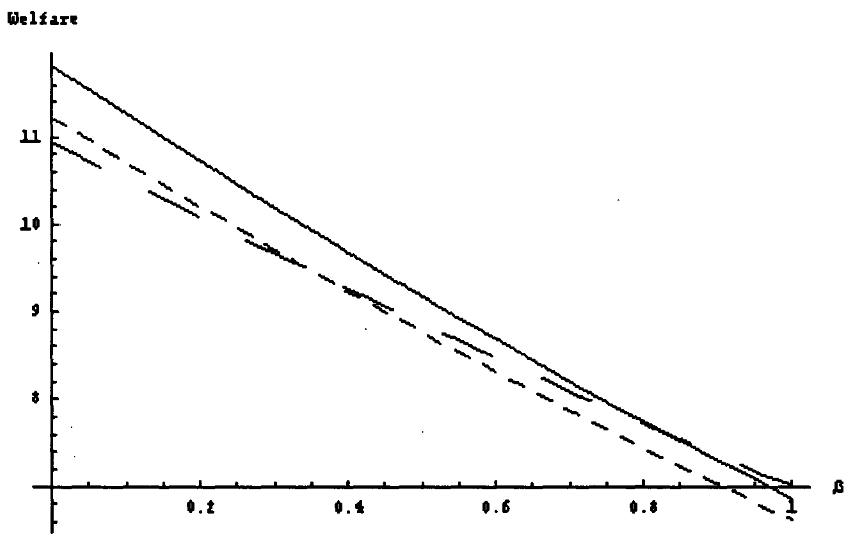


Figure 5:

ESSAY 4

Market structure and regulation

Jan Y. Sand*

University of Tromso,

NFH, Department of Economics, N-9037 Tromso, Norway

Jan.Sand@nfh.uit.no

October 2001

Abstract

This paper considers regulation of market structure and access charges in vertically related markets. The main trade-off in determining the appropriate access charge is static efficiency versus fixed cost duplication. The subgame perfect regulation is characterised, and I find that regulating access charges combined with no structure regulation is always the (weakly) best option. For an interval of the downstream fixed cost, no regulation of the access charge yields the same level of welfare as the regulated case. Furthermore, I find that the excess entry result of Mankiw and Whinston (1986) does not generally carry over to vertically related markets.

JEL Classification: L13, , L22, L51, L96

Keywords: Market structure, regulation, excess entry

*I am grateful for very valuable comments from Kåre P. Hagen, Nils-Henrik von der Fehr, Derek Clark, and Gianni De Fraja. Part of this work was undertaken while visiting Department of Economics and Related Studies at the University of York, whose hospitality is greatly appreciated. This work has received financial support from Ingegerd og Arne Skaugs Forskningsfond and Norges Banks Fond til Økonomisk Forskning.

1 Introduction

Allocative efficiency and entry conditions are two important issues that occupy many regulators, as ensuring that scarce resources are used efficiently is one on the main goals of regulation. Thus, regulators aim at determining the correct (regulated) prices from a static allocative efficiency point of view and prices that ensure that the degree of entry into the industry is socially optimal. This involves, among other things, removing inefficient entry barriers so that the correct number of firms may enter an industry. In the absence of fixed costs, we know that more competition yields a better result. However, if there are fixed costs the welfare loss due to fixed cost duplication must be measured against the benefits of increased competition.

The main intentions of this paper are twofold: First, to determine the socially optimal policy (or combination of policies) in terms of access charge determination and choice of market structure, where the access charge is determined either by the upstream monopolist, or by a regulator. Second, to examine how the number of entrants under free entry compares to the socially optimal level. With imperfect competition and free entry there will, under certain conditions, be a tendency towards excessive entry (Mankiw and Whinston, 1986). The main trade-off in the determination of the access charge is then whether it should be set to create a level playing field or to limit the potentially excessive entry. The literature on excessive entry with imperfect competition typically only examines one market. One of the goals of this paper is to examine whether the excessive entry result carry over to a setting with vertically related markets, where the input market may be subject to price regulation due to monopolisation and where the input monopolist may not be allowed to enter the downstream market. Mankiw and Whinston (1986) show, under the assumptions made in the present model (specifically ignoring the integer constraint), that more firms will enter in equilibrium than is socially optimal. In their model, however, there is only one market to consider. In the present model, firms operate in vertically related markets and the firm supplying the essential input may also serve the final product market. Furthermore, the price of the essential input may or may not, be subject to regulation. I show that the excessive entry result obtained by Mankiw and Whinston (1986) arises as a special case in the present model.

It is assumed that there is free entry into the downstream market for all independent firms, whereas the upstream monopolist may face restrictions on his entry opportunities. Any obstacles to entry is only posed by the fixed cost incurred when entering and the level of the access charge chosen by either the network monopolist or the regulator. With imperfect competition, it may be necessary to subsidised entry to foster competition, but, on the other hand, subsidies must normally be financed by distortionary taxation which creates efficiency losses elsewhere in the economy. Realistically, such subsidies will be difficult to advocate to politicians and bureaucrats in most situations. In particular, it will be difficult in global markets where international trade agreements and cross-border competition policy agreements restrict the opportunity for subsidising production.

The role of the access charge from a social point of view is twofold, and in setting the appropriate access charge there are some potentially opposing effects. First, it can be used to correct for the potential allocative loss downstream as a result of imperfect competition. Second, the access charge can be used to limit socially costly duplication of fixed costs. As long as there is imperfect competition downstream, there will be an allocative loss. If we examine the upstream market in isolation, static allocative efficiency calls for pricing the input at marginal cost.[1] However, from the theory of second-best we know that when there are distortions in the economy, first-best pricing is generally not welfare maximising.[2] When there is Cournot-competition in a vertically related market (the downstream market), the optimal access price implies setting price below marginal cost in a regime where the regulator has full information about all relevant aspects.[3] Such a policy results

[1] If there are fixed costs upstream, the access charge must be in excess of the marginal cost of providing access. Furthermore, dynamic efficiency aspects may call for access charges in excess of marginal costs in order to provide the appropriate investment incentives for the network owner. Such dynamic aspects are not considered in the current paper.

[2] The classical reference on second-best theory is Lipsey and Lancaster (1956).

[3] Access charges below marginal costs require transfers to the regulated firm to compensate for the access deficit. Access subsidies are only optimal if the cost of transfers to the regulated firm is

in increased downstream quantity, which consequently leads to consumers' welfare increasing. On the other hand, setting too low an access charge may encourage excess entry and too much duplication of fixed costs. These two elements suggest that marginal cost pricing of access, as proposed by e.g. the EU, is in general not the best policy.[4] If the regulator can subsidise access deficits and provided that the shadow cost of public funds is sufficiently low, the optimal access charge is set below the marginal cost of providing access (see Sand, 2000, 2001).[5] The main assumption in this paper is, however, that the access charge must be sufficient to cover costs, so the details of first- and second-best regulation will not be discussed (see Sand, 2001).

The broader question we may ask is whether firms enjoying a monopoly position in one market should be allowed to enter vertically related markets. This is an issue that has been debated in great detail in relation to many network industries. In the early 1980's, the conclusion for the telecommunications industry in the US was that local and long-distance telephony should be provided by separate entities, and lead to the break-up of AT&T. Regulators and competition policy authorities are again becoming less reluctant to allowing integration between firms providing local and long-distance telephony.

When there is only one firm producing an essential input for downstream production, there is a danger that the monopolist may exploit his monopoly powers. However, contrary to the majority of the literature on vertical relations (see Rey and Tirole, 1997, for a summary of this literature), the regulator may choose to restrict the upstream firm's potential for earning monopoly rents by regulating the price charged for the intermediate product (the intermediate product is access to end-users). A completely unregulated vertically integrated firm will, if firms are symmetric, have incentives to foreclose rivals in the competitive segment and there-

sufficiently low (see Sand, 2000, for more details).

[4]If the network provider has a dominant position, EU suggests that access charges should be based on actual cost plus a reasonable rate of return on investments. See, e.g., Article 7 of Directive 97/33/EC of the European Parliament and of the Council of 30 June 1997.

[5]When the shadow cost of public funds is too large, the social cost of the distortion in the downstream industry must be weighed against the welfare loss of transfers.

fore directly affect the degree of entry. In my model, the regulator may implicitly determine whether firms find it profitable to enter the downstream industry through the determination of the access charges. Consequently, the actions of the regulator (i.e., determining access charges) have similar effects to foreclosure activities undertaken by vertically integrated firms, or raising rivals' costs (see Salop and Scheffman, 1983). Foreclosure is frequently defined as "any dominant firm's practice that denies proper access to an essential input it produces to some users of this input, with the intent of extending monopoly power from one segment of the market (the bottleneck segment) to the other (the potentially competitive segment)" (Rey and Tirole, 1997; p.1). If the upstream firm is restricted to charge only a linear price for access, the only way to extract the surplus in the downstream segment in an efficient way is to vertically integrate into this segment. However, since there is imperfect competition downstream, the vertically integrated firm can only extract the entire surplus downstream if it completely forecloses all rivals. Alternatively, the upstream monopolist could use an appropriately designed two-part tariff (franchise fees) to extract the total industry surplus without having to vertically integrate. Another alternative method of preserving the monopoly profit for the upstream firm is the use of resale price maintenance (RPM).[6]

In the present paper, I will be concerned with the analysis of vertically related markets. The vertical structure of the type of model considered here is compatible with a great number of industries. All network industries typically consist of vertically related markets. The market for Internet access is one example. An essential input for the Internet access providers (IAPs) is access to the local loop and end-users, a service which usually is provided by telecommunications firms. Many of these firms have substantial market powers in the local access market, and the pricing of local access is usually subject to regulation by national regulators. The final product market can be thought of as a potentially new market, e.g., a market for broadband communication services, where firms may or may not enter. I have in mind a situation where the upstream firm provides transportation network ser-

[6]See Rey and Tirole (1997) and Tirole (1988) for more detailed discussions about both RPM and franchise fees.

vices that may be used as an essential input to produce services to consumers, but the analysis applies more generally than this. The model presented below is very stylised, but it is still general enough to be appropriate for the analysis of other cases where there is a distribution network (essential facility) and imperfect competition in a downstream market (electricity transmission and production, railroad, airline market and landing slots, etc.).

In the present model it is assumed that the regulator has two main regulatory instruments at his disposal. First, the regulator may impose entry restrictions on the upstream monopolist. Entry into downstream markets by the upstream monopolist may be restricted if the regulator fears that a firm producing an essential input (an upstream monopolist) may foreclose rival firms in the downstream industry.[7] Second, the regulator may decide on the appropriate access charge, which still is a disputed matter for instance in the widely deregulated telecommunications industry. It is assumed throughout the paper that the regulator cannot subsidise access by setting an access charge lower than the upstream monopolist's marginal cost of providing the access.

The primary reason for regulating at all should be that such a regulation yields a higher level of welfare than no regulation does, and it therefore seems appropriate to use the unregulated scenario as the benchmark case. As it turns out, regulating access charges is not generally the best policy. In this paper, I examine how well regulation (both structure regulation and access charge regulation) does compared to the unregulated case. In Vickers (1995), the focus is on how the vertical structure affects the choice of the level of the (regulated) access charge, and consequently, the unregulated access charge case is not considered. Furthermore, contrary to Vickers (1995) the regulator cannot give lump-sum transfers to the regulated firm and it is assumed that the regulator has full information about all relevant factors. This implies that the financial loss access subsidies impose on the upstream firm cannot be financed by transfers from the regulator.

The rest of the paper is organised as follows: Section 2 presents the model, and

[7]An alternative to control the number of downstream firms is to restrict entry by the use of licenses. This is, however, not considered in the present paper.

section 3 provides the reader with some basic intuition as well as the formal analysis. Section 4 provides the analysis of welfare under various policy combinations and evaluates how free entry compares to the social optimal degree of entry. In section 5, some concluding remarks are made.

2 The framework

In the model considered here, I assume that there is only one firm upstream. The number of downstream firms is determined endogenously, and potential downstream firms enter if they find it profitable to do so and if they are allowed to do so by the regulator.

The analysis is conducted in the setting of a multi-stage game. In the *first stage* either the upstream firm or the regulator decides on an access charge. In the *second stage*, firms choose simultaneously to enter or not. In the *final stage*, firms compete simultaneously in quantities. The final stage of the game is unregulated. The choice of whether the upstream monopolist should be subject to entry restrictions is taken by the regulator prior to the access charge is being determined.

The downstream firms are assumed to compete in quantities. The justification of Cournot competition in the final product market is that firms, prior to the final stage competition, need to choose the capacity of the transport network. This capacity choice amounts to building up a transport network, or leasing transport capacity from another firm (see Hansen, 1999). Thus, the quantity choices these firms make in the final stage of the game is in reality a choice of capacity.[8]

If the upstream monopolist is allowed to enter the downstream market, its profits consist of both the profit it earns from selling access and the profit from selling the

[8] It is often argued that a more realistic assumption is that firms compete in prices, not quantity. However, the Cournot outcome can, provided that certain assumptions are met, be seen as the outcome of a two-stage game where capacity choice precedes price competition (Kreps and Scheinkman, 1983).

final product to consumers. Firm v's upstream profit only is given by:

$$\pi^u = (w - \beta) \sum_{i=0}^{n} q_i \tag{1}$$

where $q_0 \geq 0$ is the output of the downstream subsidiary of the upstream monopolist and $q_i \geq 0$, $i = 1...n$, are the downstream output levels of the independent firms. The parameters w and β are the access charge and upstream marginal cost, respectively. I will assume that the upstream firm must be financially viable as a separate entity, and that the regulator cannot (or will not) use transfers to compensate the upstream monopolist for any operating loss it may incur if the regulator chooses to use access subsidies (i.e., determining an access charge below marginal cost).[9] I will therefore make the following assumption:[10]

Assumption 1

The regulated access charge must ensure coverage of costs; i.e., $w \geq \beta$.

Fixed costs upstream is an important characteristic of most local access technologies, and is a major reason for the lack of competition in the local loop ("the last mile" of telecommunication networks). In many networks the investments in infrastructure are already sunk, and play no role in the determination of the access charge. Thus, an ad hoc justification of leaving fixed costs in network provision out of the analysis is that the game which is played in the current paper takes place after investments are sunk. Furthermore, the level of these fixed costs is assumed to be high enough to deter entry into the upstream segment.[11]

[9] The upstream firm will, of course, never want to determine an access charge below marginal costs of providing access.

[10] Note that assumptions 1-3 all put restrictions on the magnitude of the endogenous variable w. However, it turns out that both the unregulated and regulated access charges all satisfy the restrictions imposed by assumptions 1-3 if the level of the fixed costs is sufficiently low; i.e., if $a - c - \beta - 2\sqrt{F} \geq 0$.

[11] A consequence of ignoring fixed costs upstream is that the assumption which requires the access charge to cover the marginal cost of providing access is sufficient for ensuring that the uptream firm earns non-negative profits. If there are fixed costs present, then the access charge must be strictly above marginal cost for Ramsey reasons.

The profit obtained downstream for firms $i = 0, 1, ..., n$, is given by:

$$\pi_i = (P - w - c)\, q_i - F \tag{2}$$

where $P = a - \sum_{i=0}^{n} q_i$, is the inverse demand function. Let us define Q as the total quantity produced downstream, which is given by: $Q = \sum_{i=0}^{n} q_i$. Downstream marginal cost is the same for all firms, and is denoted c. It is assumed that all downstream firms pay the same access charge w, but for a vertically integrated firm the access charge is simply a transfer price. We assume that all firms must pay the same fixed costs, F, for establishing downstream operations. This could be costs associated with setting up a distribution and sales network, marketing expenses, and it is reasonable that both firms face the same fixed costs. These fixed costs could also be attributed to a USO-fee (USO - Universal Service Obligation), which all firms operating downstream must pay.[12] The presence of the fixed cost implies that firms may not want to enter into the industry even if equilibrium quantities are positive. A sufficient condition for the upstream firm to enter the downstream industry is that the downstream operation, in isolation, earns positive profits.[13]

The regulator's welfare is given by:

$$W = \begin{cases} CS^{nsr} - (c + \beta) \sum_{i=0}^{n} q_i - (n+1) F \\ CS^{sr} - (c + \beta) \sum_{i=1}^{n} q_i - nF \end{cases} \tag{3}$$

where $CS^k = aQ_k - (Q_k)^2/2$, for $k = nsr, sr$, is the gross consumers' surplus under the cases of no structure regulation (nsr) and structure regulation (sr), respectively. In the latter case, the upstream monopolist is not allowed to enter the downstream market.

[12] An extension to the present model could be to say that the regulator determines the level of the fixed costs associated with entry into the downstream market. This could, for instance, be the case where firms who enter are required to contribute towards the costs of maintaining networks in unprofitable geographical areas (e.g., part of the entry costs contribute towards the costs of Universal Service Obligations, USO).

[13] In general, we would expect that the upstream firm evaluates the profit earned as a monopolist access provider with no downstream operation against the profit earned as a vertically integrated firm (i.e., compare $(P - c - w)\, q_0 - F + (w - \beta) \sum_{i=0}^{n} q_i$ against $(w - \beta) \sum_{i=1}^{n} q_i$.

2.1 The game

We are looking for the subgame perfect equilibrium in the game, and start at the last stage to reveal the equilibrium path. We first solve for the Cournot-equilibrium. Then we solve for the optimal access charge; i.e., the price paid for access which maximises either the upstream firm's profit (the unregulated case) or the regulator's welfare function (the regulated case).

The access charge can also be used as an instrument to limit the degree of costly duplication of fixed costs. The regulator may choose to set the access charge high enough to foreclose the rival firms if the fixed costs are sufficiently high.[14] An alternative policy to the regulation of access charge to avoid duplication of fixed costs is either to restrict the upstream monopolist's opportunity to enter into vertically related markets, or to limit entry by independent firms.

One might argue that regulating the final product prices is another alternative. However, as is argued by, e.g., Laffont and Tirole (2000), the more severe monopoly problem is in the upstream market with (potentially) significant economies of scale. The downstream market is more competitive. This leads us to the conclusion that the more appropriate regulatory policy would be to direct the attention to the bottleneck segment (network services) which is the real problem, and regulate access charges rather than regulating the prices of the final products.

3 The regulated and unregulated cases

In this section I will determine both the unregulated firm's choice of access charge and the socially optimal access policy, when the number of firms downstream is determined endogenously.

Let us assume that firm v is active in both markets. We now allow for potential rival firms to enter freely in the downstream market. In this case, the inverse demand function is given by: $P = a - \sum_{i=0}^{n} q^i$, where firms $i = 1, ..n$ are the independent

[14] If we allow for differences in marginal cost downstream, the regulator may decide to foreclose (some of) the rival firm(s) if the rival firm is very inefficient relative to the downstream subsidiary of the access provider.

rival firms and firm 0 is the downstream subsidiary of the upstream monopolist. When there is free entry, the total number of rivals will be determined by the zero profit condition for all independent firms i:

$$\left(a - \sum_{i=0}^{n} q^i - c - w\right) q^i \geq F \tag{4}$$

Ignoring problems of indivisibilities, we can assume that the inequality is satisfied as an equality in equilibrium.[15] Since we assume that all potential entrants have identical marginal costs, all active downstream rivals produce the same quantity; $q^i = q$, $\forall i = 1, ..n$. I also assume that the downstream subsidiary of the upstream monopolist face the same downstream marginal cost as the independent firms. Any cost differences between the downstream competitors will therefore be due to the fact that the access charge is different from the marginal cost of providing access. Thus, the vertically integrated firm may therefore have a different level of production. Consequently, the total downstream quantity is given by $Q = nq + q_0$.

Cournot equilibrium In the final stage of the game firms compete in quantities, taking the access charge as given. It can be shown that when there are n entrants and the upstream monopolist is allowed to enter the competitive segment, the (unique) equilibrium quantities for a given level of the access charge will be:

$$q^* = \frac{a - c - 2w + \beta}{n + 2} \tag{5}$$

$$q_0^* = \frac{a - c - (n+1)\beta + nw}{n + 2} \tag{6}$$

where q^* and q_0^* represent the equilibrium quantities for each rival and the vertically integrated firm, respectively. Note that in some of the cases considered below, the vertically integrated firm may enjoy a monopoly situation downstream, with the resulting monopoly output $Q_0^m = (a - c - \beta)/2$.

[15] In reality, there are of course indivisibility and the zero profit condition is satisfied as an equality only by coincidence. The true number of firms which will enter is the largest positive integer which satisfies the zero profit condition, and this will in general imply that firms earn positive profits downstream. However, to simplify the analysis I abstract from this problem.

The total downstream quantity is in equilibrium given as:

$$Q^* = q_0^* + nq^* = \frac{(n+1)(a-c) - nw - \beta}{n+2} \tag{7}$$

Observe that in the competitive limit when $n \to \infty$, $\lim_{n\to\infty} q_0^* = w - \beta$, which is (strictly) positive if firm v earns a (strictly) positive profit margin upstream. Furthermore, $\lim_{n\to\infty} q^* = 0$ and $\lim_{n\to\infty} Q^* = (a - c - w) > 0$, where the latter limit is equal to the competitive output $Q^{comp} = (a - c - \beta)$ if $w = \beta$.

To ensure that the competitive output is positive, I assume the following:

Assumption 2

$w < a - c$.

Assumption 2 ensures both that the competitive limit is positive when there are no vertical restrictions, and, as we will see below, that the output of the independent downstream firms is positive when the upstream monopolist is not allowed to enter into the competitive segment. If assumption 2 is not satisfied it indicates that the market size is small relative to the marginal costs of the independent firms, and this implies that there will never be any entry by independent firms in the competitive segment (not even in the absence of fixed costs).

If the network monopolist is not allowed to enter the downstream market, then it can be shown that the symmetric (and unique) equilibrium quantity in the Cournot game is given by:

$$\widetilde{q}^* = (a - c - w)/(n+1) \tag{8}$$

In this case, we see that assumption 2 ensures that there will be production in equilibrium (provided that the fixed costs are sufficiently low). If, on the other hand, $w > a - c$, then we are in the uninteresting case of a market with no activity.

Entry decisions The entry decision by each independent firm is taken at the intermediate stage of the game, in which the equilibrium quantities from the final stage of the game are used to determine the profitability of entry. The zero-profit condition, eqn. (4), dictates the total number of rivals the industry can support for a given set of parameter values. By inserting for eqn. (5) into (4), we obtain the following expression which determines the number of firms, $n \geq 0$, that enters the downstream industry:

$$n \leq \widehat{n} \equiv \frac{a - c - 2w + \beta - 2\sqrt{F}}{\sqrt{F}} \tag{9}$$

where $\widehat{n}$ is the value of n which ensures that (4) is satisfied as an equality. The actual number of firms, n, that enters will be the largest positive integer satisfying (9). We observe that the number of firms which enter is a decreasing function of the level of fixed costs and the access charge. After we have calculated the optimal access charge, we can then work out how many firms that actually enter using (9).

When the upstream monopolist cannot enter the downstream market, the number of firms entering is determined by:

$$n \leq \widetilde{n} \equiv \frac{a - c - w - \sqrt{F}}{\sqrt{F}} \tag{10}$$

If there are no fixed costs associated with entry into the downstream market and if the access charge is set equal to marginal cost of providing access ($w = \beta$), then $q^* = \widetilde{q}^* = (a - c - w) / (n + 2)$, and q^* is non-negative since $a > c + w$ by assumption 2.[16] In this case, the number of firms entering the industry tends to infinity ($\Pi^i \rightarrow 0$ as $n \rightarrow \infty$). This is the same result as in Mankiw and Whinston (1986), who find that if the fixed cost of entry approaches zero the bias towards excessive entry tends to infinity. However, they prove that the welfare loss caused by having too many firms approaches zero in this case.

The entry dynamics implies that the regulator can determine the degree of entry into the industry if the access charge is regulated. We know from Mankiw and Whinston (1986) that there is a tendency for excess entry in markets with imperfect competition when the business stealing effect is significant. In this case, the profit of new firms entering the industry comes at the expense of incumbent firms' profits. This implies that the gain to society of a new firm entering is less than the gain to

[16] If assumption 2 is not satisfied, there will be no production in the downstream industry (and consequently, no production upstream), as firms are symmetric and $q^* = q_0^*$ under the assumptions in the current paragraph.

the entering firm. Consequently, the regulator can by means of the access charge restrict entry such that only the socially optimal number of firms enters into the industry, if he so desires.

I will assume that the following condition is satisfied, unless stated otherwise:

Assumption 3

To ensure that downstream output for the independent firms is positive, we must have the following: $w \leq (a - c + \beta)/2$.

A necessary, but not sufficient, requirement for entry by independent firms is that the equilibrium output is positive, which is ensured by assumption 3. It also ensures that the price-cost margin of the downstream firms is non-negative. Assumption 3 also ensures that the total equilibrium downstream output, Q^*, increases when more firms enter; i.e., $\partial Q^*/\partial n \geq 0$. Thus, under assumption 3 we know that the introduction of another firm in the downstream market implies that there is a *market expansion effect.*

Furthermore, assumption 3 ensures that each downstream firm's equilibrium quantity decreases in the number of firms: i.e., $\partial q^*/\partial n \leq 0$ and $\partial q_0^*/\partial n \leq 0$. This implication of assumption 3 captures the *business stealing effect* of new entry. For each new entrant in the market, that particular firm brings an added social gain due to the market expansion effect. However, part of the profit of the potential new entrant comes from stealing some of the existing firms' market shares and profits. Thus, from a social point of view the profit of a given new entrant, which is the basis for the entry decision of that firm, is higher than the value to society of that new entrant.[17]

Consequently, there are opposing effects on welfare due to entry. First, consumers are better off due to the fact that quantity increases in the number of firms entering downstream. On the other hand, there are real economic costs due to entry, due to higher output and more duplication of fixed costs. The optimal entry implies balancing these costs and benefits.

[17] Assumption 3 satisfies all the conditions laid down by assumption 1-3 in Mankiw and Whinston (1986), in which case free entry in imperfectly competitive markets tends towards excessive entry.

3.1 The basic trade-off: Outsourcing of production

Before I analyse the various cases in detail, it will be beneficial to give a short overview of the main trade-offs the firm and the regulator, respectively, face when the level of the access charge is to be determined. An understanding of the basic economic forces at work will help understand the results presented below.

Both in the unregulated case (the upstream firm determines the level of the access charge) and in the regulated case (a regulator determines the access charge), the process of determining the level of the access charge may be seen as a problem of *outsourcing* a production activity. In the unregulated case, the determining factor in choosing whether to price access to allow independent downstream firms compete is whether such a policy increases the network monopolist's profit. The idea is to price the access to achieve a production mix which maximises profits. By pricing access low enough to make entry by independent firms profitable, the network monopolist effectively outsources part of (or the total) production downstream to these independent firms. Whether this is in the network monopolist's own interest depends on several factors. The most commonly used explanation of why a firm chooses to outsource (part of) its production relates to cost aspects. Certain production activities, either final products or inputs to the production process, may be better undertaken by independent firms if these firms are more cost efficient than the in-house supplier. With a sufficiently low access charge, independent firms will enter and earn a positive level of profits. However, the network monopolist will be able to capture at least some of the profit from the downstream market by choosing the access charge appropriately. In the present model, this effect is not present since all downstream firms are equally efficient.

Part of the profit earned by outsourcing production to independent firms may, as mentioned, be captured by the network monopolist through setting an access charge in excess of marginal cost, but some profit will have to be retained by the independent firms in order to induce their entry. Consequently, there will be a profit-shifting effect from the network monopolist to the independent suppliers. This effect reduces the network monopolist's incentives to outsource its (final good) production, and this effect will push for determining a higher access charge to limit the degree of outsourcing (or potentially eliminate outsourcing altogether). Furthermore, by setting the access charge low enough to induce entry this implies that the total production downstream increases due to lower market price of the final product. This leads to more profit from selling access, but less profit in the final product market (due to lower output by the network monopolist's own downstream subsidiary and lower price on all units sold). However, this expansion in output cannot improve profits since a vertically integrated monopolist (with monopoly in both segments) will follow a price strategy to maximise its overall level of profit. The vertically integrated monopolist can expand the total production level by increasing its own production level, in which case it will be able to capture the entire profit from the increase (a complete internalisation). However, if the increased production is undertaken by an independent firm, the network monopolist cannot capture the entire gain from increasing production through the linear access price. This will also tend towards foreclosure, or a higher level of the access charge.

From society's point of view, the problem of determining the level of the access charge is also a problem of outsourcing. Equivalently, the regulator faces the problem of setting the access charge to ensure that the resources are utilised to maximise some measurement of welfare (in the present model, the sum of profit and consumers' surplus). The trade-offs facing the regulator are to some extent identical to the unregulated firm's problem. Production should take place with the most efficient firm, and the determination of the socially optimal access charge reflects this (see, e.g., Sand, 2000 and 2001, or Lewis and Sappington, 1999).[18] The more efficient an independent firm is relative to the network monopolist's own subsidiary, the lower is the access charge set to distort market shares in favour of the more efficient firm. However, in the present model all firms are equally efficient, so a distortion in market shares due to differences in efficiency is not an issue. Furthermore, contrary to the

[18]In the regulatory debate, a more pronounced view is that the access charges should be set to level the playing field, which is often interpreted to imply that the access charge should be set close to marginal cost. However, this does not necessarily imply that the total production is undertaken most efficiently.

unregulated case, the profit-shifting from the network monopolist to independent rival firms is of no importance since it is simply a transfer (all firms' profits have equal weight in the welfare function).

In addition, and most importantly in the present model, the regulator faces a trade-off between increasing the downstream production through entry and the socially costly duplication of fixed costs. There are fixed costs associated with production in the downstream market, and since entry implies that more firms produce a lower output each, there is less utilisation of the economies of scale. However, entry implies a higher level of total production which results in an increase in consumers' surplus. Thus, the main trade-off facing a benevolent regulator is to balance the increase in consumers' surplus as a result of entry (which tends towards a lower access charge) against the social cost of utilising the economies of scale to a lesser extent (which tends towards a higher access charge). Essentially, the access charge has from society's point of view two main roles to play in the case with free entry. It is an instrument for correcting the allocative inefficiency in the downstream market, which (under certain conditions) calls for an access charge below marginal cost of providing access. The access charge can also used to ensure that there is no inefficient entry, which often calls for an access charge in excess of marginal cost to mitigate the excess entry problem. Thus, the optimal access charge is determined by the optimal trade-off between the regulator's concern to achieve allocative efficiency (to which the process of free entry yields insufficient entry) and the business stealing effect (which tends towards excess entry).

3.2 Unregulated access charge

In this section I investigate how an unregulated, profit maximising firm will set the access charge in the first stage of the game. I will focus on two cases: 1) The upstream firm is allowed to enter the downstream market, and 2) there are restrictions on the upstream monopolist entry into vertically related markets (i.e., the upstream firm cannot enter the downstream market). Rival firms in the downstream market enter as long as their profit is non-negative. I will look for subgame perfect equilibria in the access charge.

3.2.1 No structure regulation

When there is free entry, the equilibrium profit of the vertically integrated firm is modified as follows:

$$\Pi_v(n,w) = \left(q_0^*(n,w)\right)^2 + (w-\beta)\,nq^*(n,w) - F \qquad (11)$$

where q^* is defined by eqn. (5) and q_0^* defined by eqn. (6).

If the upstream monopolist is allowed to enter into the downstream market and is free to determine an access charge to maximise its own profit, then the upstream firm will set w to maximise eqn. (11). Since the access charge is determined prior to the entry decision by potential entrants, the number of firms that actually enters is a function of the access charge. Consequently, the unregulated firm must take the interaction between the degree of entry and the access charge into account in his maximisation problem. We will assume that, as an approximation, eqn. (9) holds as an equality, and therefore determines the exact number of firms that enters. The profit maximising access charge will either be given by the first-order condition when maximising Π_v with respect to w, which we denote w^{ur}. However, this does not necessarily yield the global maximum for firm v's profit. Thus, alternatively, firm v will determine an access charge $\widehat{w}$, such that $\pi_i < 0$, for all $i = 1, .., n$, in equilibrium. Such a policy will imply complete foreclosure of all rival firms.

The result is summarised in the following proposition:

Proposition 1 *The vertically integrated firm will never accommodate rival firms' entry when all downstream firms are symmetric (including the subsidiary of the upstream firm). The subgame perfect equilibrium implies setting an access charge high enough to deter all entry.*

Proof. The upstream firm will choose the access charge, $\widehat{w}$ or w^{ur}, which yields a global maximum for profits. If $\Pi^v(w = \widehat{w}) \geq \Pi^v(w = w^{ur})$, the subgame perfect equilibrium implies deterring entry of all rival firms. First, it can be shown that, after inserting for the equilibrium in quantities, $\partial \Pi_v^e / \partial n < 0$. Thus, if there is entry, firm

v will decide an access charge to limit entry as much as possible, which means $n = 1$. Then, let us compare firm v's profit when $n = 1$ to the monopoly profit when $n = 0$, for any given level of the access charge. This entails comparing eqn. (11), given that $n = 1$, against firm v's monopoly profit: $\Pi_v \mid_{n=1} - \Pi_v^m = -\frac{5}{36}(a - c - 2w + \beta)^2$, which is negative per assumption. Consequently, the upstream firm has no interest in accommodating entry. ■

The intuition can be related to the discussion above where the fundamental trade-offs facing the firm (and regulator) are explained. When access charge is unregulated and all downstream firms are equally efficient in producing the final product, there are two effects and both give incentives to foreclose rival firms. By allowing independent firms to enter through setting a low access charge, there is a profit-shifting effect from the network monopolist's own downstream subsidiary to the independent firms which has a negative impact on the network monopolist's total profit. Furthermore, the increase in profit from the access segment due to entry will not increase overall profit for the network monopolist (see discussion above). Consequently, the vertically integrated firm has no incentives to outsource the production of the final product to its rivals.

3.2.2 Structure regulation

When the upstream monopolist is not allowed to enter the downstream industry, we need to maximise eqn. (1) with respect to the access charge, taking into account that the number of firms entering is endogenously determined by eqn. (4). Again, we will ignore the integer constraint with respect to entry to simplify the exposition.

The solution to the upstream monopolist's maximisation problem is given by:

$$\widehat{n}\widetilde{q}^* + (w - \beta)\frac{\partial \widehat{n}}{\partial w}\widetilde{q}^* + (w - \beta)\,\widehat{n}\frac{\partial \widetilde{q}^*}{\partial w} = 0 \tag{12}$$

where $\widetilde{q}^*$ is defined by eqn. (8), and is the equilibrium quantity of each of the downstream firms, and $\widetilde{n}$ is defined by eqn. (10). The closed-form solution, w_r^{ur}, to this problem is given by:

$$w_r^{ur} = \frac{1}{2}\left(a - c + \beta - \sqrt{F}\right) \tag{13}$$

The comparative static results for the unregulated access charge shows that the unregulated access charge is: 1) Increasing in upstream marginal cost, 2) decreasing in downstream marginal cost, 3) decreasing in the level of fixed costs, and 4) increasing in the market size.

The first effect is straightforward. The second effect can be explained as follows: We observe that total downstream quantity is decreasing in c, which implies a lower level of sales of local access for the upstream monopolist. Thus, to compensate for this reduction in the profit from the upstream operation, the upstream monopolist lowers the access charge the independent firms must pay. The third effect, $\partial w_r^{ur}/\partial F < 0$, is due to a straightforward trade-off between reducing the access charge to maintain high level of sales of access to downstream firms and lower level of sales at a larger profit margin. The effect on the access charge of changes in the market size is straightforward.

In the case where firm v can enter the downstream market, we find that in equilibrium no firms enter as a result of the non-accommodating access pricing strategy that firm v chooses. Because of this, the regulator may choose not to allow firm v to enter the downstream market. When entry by the upstream monopolist is restricted, we find that the degree of entry is negatively correlated to both the upstream and downstream marginal costs (i.e., $\partial\widetilde{n}/\partial\beta < 0$ and $\partial\widetilde{n}/\partial c < 0$, respectively), provided that the access charge is determined by eqn. (13).[19]

In contrast to the case with no structure regulation, the firm that determines the access charge is now *not* active in the downstream market. This implies that changes in the upstream marginal cost only enter through the effect on the access charge, and we know that $\partial w_r^{ur}/\partial\beta > 0$. Thus, increasing the marginal cost upstream raises the access charge which reduces the profitability of each of the downstream firms. Thus, there will be fewer firms entering when β goes up. We also know that an increase in c (downstream marginal cost) leads to a reduction in the access charge, the *indirect* effect, which by itself would tend towards a higher entry. However, the *direct* effect of increasing c is to reduce the profitability downstream, and it is this

[19]We see this by examining the comparative statics of eqn. (10), when we have inserted for $w = w_r^{ur}$. This yields a level of entry given by: $\widetilde{n}^{ur} = \left(a - c - \beta - \sqrt{F}\right)/2\sqrt{F}$.

latter (direct) effect which is dominant. The upstream monopolist must balance the loss of profitability from reducing the access charge against the loss of profitability from a lower level of traffic.

The direct effect of increased fixed costs is to reduce entry. However, an indirect effect due to an increased level of fixed costs is that the unregulated access charge is reduced in this case, which, *ceteris paribus*, increases the degree of entry into the market, but this effect is dominated by the direct effect. When the size of the market increases, we have seen that firm v raises the access charge which in isolation tends towards lower equilibrium profit downstream and less entry. However, an increase in the size of the market works in the opposite direction and more than cancels out the negative effect on entry of the increase in the access charge.

3.3 Optimal access charge

In this section, it is assumed that the regulator determines the socially optimal level of the access charge, and the number of rival firms that actually enter is determined by the zero-profit condition, eqn. (4).[20]

3.3.1 No structure regulation

Simplifying the expressions, social welfare is given as (when inserting for the final stage equilibrium quantities):

$$W = CS\,(n,w) - (c+\beta)\,Q^*\,(n,w) - (n+1)F \tag{14}$$

where $CS = aQ^*\,(n,w) - [Q^*\,(n,w)]^2\,/2$ is the gross consumers' surplus.[21] Let $Q^*\,(n,w) = q_0^*\,(n,w) + nq^*\,(n,w)$ be the total equilibrium downstream quantity when the upstream monopolist is allowed to enter into the downstream market,

[20]The results with respect to the regulated access charge (in both the following subsections) are the same as in Vickers (1995) if we ignore the asymmetry of information about costs, or, alternatively, if consumers' surplus and firms' profits are weighted equally in the welfare function.

[21]The gross consumers' surplus is in general given by: $\int_0^Q P\,(s)\,ds$. We can rewrite this as net consumers' surplus + total expenditure, which in equilibrium turns out to be (with linear inverse demand): $CS = aQ^* - \frac{1}{2}\,(Q^*)^2$.

and $P^*(Q^*)$ be the resulting price. The social cost of production is given by $(c+\beta)\,Q^*(n,w)+(n+1)F$.

In this section, the upstream monopolist is allowed to enter the downstream market. The number of firms that enters the industry is $\widehat{n}$, and is determined by eqn. (9). Thus, the regulator's problem is then to maximise (14) with respect to w, subject to the constraint that entry is endogenous and determined by eqn. (9) and zero-profit restrictions. The solution to this maximisation problem, w^*, is determined by:[22]

$$\frac{dW\left(\widehat{n}\left(w\right),w\right)}{dw}=\frac{dCS\left(\widehat{n}\left(w\right),w\right)}{dw}-(c+\beta)\frac{dQ^*\left(\widehat{n}\left(w\right),w\right)}{dw}-\frac{\partial\widehat{n}}{\partial w}F=0 \tag{15}$$

For all permissible parameter values we find that $dQ^*/dw<0$, which implies that $dCS/dw<0$. Thus, increasing the access charge reduces the equilibrium quantity and the gross consumers' surplus.

A socially optimal access charge when the upstream monopolist is allowed to enter into the downstream market is determined by eqn. (15) (Vickers, 1995):

Proposition 2 *When the upstream monopolist is allowed to enter the downstream market, the regulated access charge is determined by:* $w^*=\beta+\sqrt{F}$.

In the absence of fixed costs, we know that the number of firms entering the downstream market approaches infinity. Furthermore, when $n\rightarrow\infty$, we know that $Q^*\rightarrow(a-c-w)$ which equals the competitive output if $w=\beta$. Thus, when $F=0$ and $n\rightarrow\infty$, it can be shown that $\partial W/\partial w<0$. Consequently, the regulator will choose to determine an access charge as low as possible without violating the constraint $w\geq\beta$, which implies that $w^*=\beta$. When there are no fixed costs, there is obviously no costly duplication of fixed costs. Then, the only concern for the regulator is to implement a level playing field for all firms operating

[22] I assume that the second-order condition with respect to w is satisfied: $\frac{d^2W}{dw^2}=\frac{d^2CS}{dw^2}-(c+\beta)\frac{d^2Q^*}{dw^2}<0$. For all permissible parameter values, the second-order condition is indeed satisfied. The total derivatives of CS and Q are given by $dCS/dw=\partial CS/\partial w+(\partial CS/\partial n)(\partial\widehat{n}/\partial w)$ and $dQ^*/dw=\partial Q^*/\partial w+(\partial Q^*/\partial n)(\partial\widehat{n}/\partial w)$.

in the downstream market. However, if there are fixed costs associated with entry downstream, the regulator must incorporate the social cost of duplicating these when determining the socially optimal access charge. Thus, the optimal access charge may be in excess of the upstream marginal cost to avoid too much costly duplication.

A benevolent regulator will want to determine an access charge to obtain the socially optimal mix between maximising consumers' surplus and utilising the economies of scale that are present in the downstream industry. If maximising consumers' surplus is the regulator's only concern, then this is an argument for marginal cost pricing of access which implies a higher total quantity in the downstream market. However, the regulator will also be concerned with the duplication of fixed costs associated with entry, which results in utilising the economies of scale for the network monopolist's own less. This latter concern implies that the regulator will want to set an access charge in excess of marginal cost to limit the degree of entry.

In a social optimum, the price of access should then reflect the true social cost of expanding output, which will consist of two elements: the marginal cost of access and the social cost of utilising the economies of scale of the network monopolist's own subsidiary less.

3.3.2 Structure regulation

When no vertical integration is allowed by regulatory authorities, all downstream firms are on equal terms. Consequently, the total equilibrium quantity produced and sold downstream is then $\widetilde{Q}^*(n,w) = n\widetilde{q}^*(n,w)$, where $\widetilde{q}^*$ is given by eqn. (8). The welfare is in this case given by:

$$W = \widetilde{CS}(n,w) - (c+\beta)\widetilde{Q}^*(n,w) - nF \tag{16}$$

where $\widetilde{CS}$ is the gross consumers' surplus when firm v is not allowed to enter.

In this section, the firm providing the essential input, firm v, is not allowed to enter the downstream market. The number of firms that enter the downstream market is $\widetilde{n}$, which is determined by eqn. (10). The regulator's problem is then to maximise (16) with respect to w, subject to endogenous entry and zero-profit conditions. The solution, w_r^*, is then determined by:[23]

$$\frac{dW(\widetilde{n}(w),w)}{dw}=\frac{d\widetilde{CS}(\widetilde{n}(w),w)}{dw}-(c+\beta)\frac{d\widetilde{Q}^*(\widetilde{n}(w),w)}{dw}-\frac{\partial\widetilde{n}}{\partial w}F=0 \qquad (17)$$

When there are fixed costs and the upstream monopolist is not allowed to enter into the downstream market, we find, when inserting for the inverse demand function and eqn. (10), that the socially optimal access charge is given by $w_r^*=\beta$. In this case, the welfare function is decreasing in the access charge for all values of w. When there are no vertical restrictions (see the previous subsection), the welfare function is strictly decreasing in the access charge only when there are no fixed costs. When there are vertical restrictions imposed on the upstream monopolist, this is the case even if fixed costs are positive. Thus, we have the following result (Vickers, 1995):

Proposition 3 *When the upstream monopolist cannot enter the downstream market, the socially optimal access charge implies marginal cost pricing of access: i.e.,* $w_r^*=\beta$.

The main difference to the case with no vertical restrictions is that the socially optimal access charge under vertical restrictions is simply marginal cost pricing of access, without any mark-up to counter the duplication of fixed costs. Transferring production between firms entails reducing the production of one independent firm and increasing it for another, and there are no losses in the economies of scale of the network monopolist's own downstream activities (of which there are none in this scenario).

4 Welfare comparisons and entry

In this section I will compare the level of welfare under various combinations of policies, and the main objective is to find the subgame perfect regulation. There

[23]The total derivatives of consumers' surplus and quantity are given by: $d\widetilde{CS}/dw=\partial\widetilde{CS}/\partial w+\left(\partial\widetilde{CS}/\partial n\right)(\partial\widetilde{n}/\partial w)$ and $d\widetilde{Q}^*/dw=\partial\widetilde{Q}^*/\partial w+\left(\partial\widetilde{Q}^*/\partial n\right)(\partial\widetilde{n}/\partial w)$. Observe that $d\widetilde{Q}^*/dw<0$ if assumption 2 is satisfied, which implies that $d\widetilde{CS}/dw<0$.

are four different cases (i.e., policy combinations) to consider: 1) No structure regulation and access charge regulation, 2) no structure regulation and unregulated access charge, 3) structure regulation and access charge regulation, and 4) structure regulation and unregulated access charge. The outcomes in terms of access charge and the number of firms downstream, are summarised in the following table (*AR* denotes the case of access charge regulation, whereas *UR* denotes the unregulated access charge case):

	No structure regulation	*Structure regulation*
AR	(1) $w^* = \beta + \sqrt{F}$ $\widehat{n}^* + 1 = \frac{a-c-\beta-3\sqrt{F}}{\sqrt{F}}$	(3) $w_r^* = \beta$ $\widetilde{n}^* = \frac{a-c-\beta-\sqrt{F}}{\sqrt{F}}$
UR	(2) $w^{ur} \to$ *complete foreclosure* $\widehat{n}^{ur} + 1 = 1$	(4) $w_r^{ur} = \frac{1}{2}\left(a - c + \beta - \sqrt{F}\right)$ $\widetilde{n}^{ur} = \frac{a-c-\beta-\sqrt{F}}{2\sqrt{F}}$

It should be noted that the welfare levels reported in this section are, of course, only valid provided that at least one downstream firm is active. If the level of fixed costs is sufficiently high, there may not be any firms that find it profitable to enter, in which case the welfare level is zero. The welfare levels for the four cases are, respectively:

$$\begin{aligned}
1)\ W_{nsr}^* &= (a-c-\beta)\left(a-c-\beta-2\sqrt{F}\right)/2 + F \\
2)\ W_v^m &= \frac{1}{8}\left(3(a-c-\beta)^2 - 8F\right) \\
3)\ W_{sr}^* &= (a-c-\beta)\left(a-c-\beta-2\sqrt{F}\right)/2 + F/2 \\
4)\ W_r^{ur} &= \frac{3}{8}\left(a-c-\beta-\sqrt{F}\right)^2
\end{aligned}$$

4.1 No regulation versus structure regulation

When there are no restrictions put on firm v, it enters the downstream market and chooses an access charge to deter all other firms from entering. Consequently, firm v enjoys a monopoly situation both upstream and downstream, and the welfare level is given by W_v^m which is determined by eqn. (3) when inserting for the monopoly output. The total downstream quantity is $Q_0^m = (a-c-\beta)/2$, and fixed costs are only incurred once.

If, on the other hand, firm v cannot enter the downstream market (the case of structure regulation), the chosen access charge is w_r^{ur} which is determined by eqn. (13). Provided that $w = w_r^{ur}$, the total downstream quantity is in equilibrium $Q_r^{ur} = \left(a - c - \beta - \sqrt{F}\right)/2$, and the level of entry is given by $\widetilde{n}^{ur} = \left(a - c - \beta - \sqrt{F}\right)/2\sqrt{F}$. Note that at least one firm enters the downstream market in the current case (i.e., $\widetilde{n}^{ur} \geq 1$) only if $F \leq \left[(a - c - \beta)/3\right]^2$.

Comparing the welfare levels in the two cases with unregulated access charge, with or without structure regulation, yields the following result:

Proposition 4 *When the access charge is unregulated, the level of welfare is always at least as large without structure regulation compared to the case with structure regulation.*

Proof. Define $\Delta_1 \equiv W_v^m - W_r^{ur}$. By inserting for $w = w_r^{ur}$ defined by eqn. (13) in W_r^{ur}, it can be shown that $\Delta_1 = \frac{6}{8}\sqrt{F}\left(a - c - \beta - \frac{11}{6}\sqrt{F}\right)$. If at least one independent firm enters (with structure regulation), then $a - c - \beta \geq 3\sqrt{F}$, which implies that $\Delta_1 > 0$. If $a - c - \beta < 3\sqrt{F}$, then $W_r^{ur} = 0$ and $\Delta_1 \equiv W_v^m > 0$. ■

Proposition 4 tells us that a vertically integrated monopoly is preferred in terms of the welfare level compared to a situation with vertical separation with independent firms downstream in an unregulated access market. By applying structure regulation, the network monopolist can only recoup his monopoly profit in the access market and since the access charge is not subject to regulation he will be able to earn a positive level of profit. To obtain a strictly positive level of profit, the network monopolist must set the access charge strictly in excess of marginal cost of providing access, which will imply that the independent firms will face higher costs of production than a subsidiary of the network monopolist. Thus, to a vertically integrated firm's ability to avoid the problem of double marginalisation, the output in the final product market is lower with structure regulation than without. Without structure regulation, the vertically integrated firm will foreclose all rival firms and produce the monopoly quantity downstream.

4.2 Access charge regulation

As already noted, the regulator may also subject access charges to regulation. In this section I will compare the welfare levels in the access charge regulation scenario, with and without structure regulation.

Proposition 5 *When the access charge is subject to regulation, a policy of no structure regulation is preferred by the regulator (i.e., $W^*_{nsr} > W^*_{sr}$).*

Proof. If there is no structure regulation, we know that the regulated access charge is given by $w^* = \beta + \sqrt{F}$. Inserting for w^* into $\widehat{n}$, where $\widehat{n}$ is determined by eqn. (9), yields actual entry given by $\widehat{n}^* = \left(a - c - \beta - 4\sqrt{F}\right)/\sqrt{F}$. Welfare is given by eqn. (14), which is denoted by W^*_{nsr} after we have inserted for $w = w^*$ and $n = \widehat{n}^*$. Note that $\widehat{n}^* \geq 1$, i.e., that at least one rival firm enters, only if $a - c - \beta \geq 5\sqrt{F}$. If there is structure regulation, the regulated access charge is given by $w^*_r = \beta$, and actual entry is given by $\widetilde{n}^* = \left(a - c - \beta - \sqrt{F}\right)/\sqrt{F}$. In this case, welfare is given by eqn. (16), and denoted W^*_{sr} after we have inserted for $w = w^*_r$ and $n = \widetilde{n}^*$. Define $\Delta_2 \equiv W^*_{nsr} - W^*_{sr}$. Inserting for access charge and entry in the two policy regimes, we find that $\Delta_2 = F/2 > 0$ for $F > 0$. ■

In the present model with endogenous entry, actual entry is larger under structure regulation, since $\widehat{n}^* < \widetilde{n}^*$. The reason being that with structure regulation, the regulated access charge rule is to simply use marginal cost pricing of access, whereas without structure regulation the regulator determines an access charge strictly in excess of marginal costs to limit the degree of costly duplication of fixed costs. The number of firms entering when there is no structure regulation does not include the network monopolist's own subsidiary. Thus, the total number of firms is then $\widehat{n}^* + 1$. However, it is easily shown that $\widetilde{n}^* > \widehat{n}^* + 1$, and consequently, there is more competition with structure regulation. With more competition, the (gross) consumers' surplus is higher (under assumption 3). However, since $\widehat{n}^* < \widetilde{n}^*$ we also know that the duplication costs are larger, and Proposition 5 tells us that the increase in consumers' surplus due to more competition and higher total output in the final product market is not enough to offset the cost of allowing more firms to enter (i.e., the increase in fixed costs).

4.3 Access regulation, structure regulation or no regulation?

As we have seen above, the welfare levels in both the regulated and unregulated scenarios are higher if the regulator does not impose structure regulation. In this section I will examine which policy or combination of policies, that yields the best outcome in terms of welfare for different levels of the fixed costs. The first question we may ask is whether there are conditions under which regulation of access charges is always better in terms of welfare than not regulating access charges. If we can prove that the level of welfare in the scenario with a regulated access charge and with structure regulation is larger than the welfare without structure regulation and without access charge regulation; i.e., $W^*_{sr} > W^m_v$, then regulation of access charges will always prove welfare optimal in the present model. However, this turns out to be the case only when the size of the market is large relative to the level of fixed costs. Thus, it may in certain situations be better in terms of welfare not to regulate the access charge.

For us to be able to make the welfare comparisons we need to know when there is entry by independent firms. This information is summarised in Lemma 1:

Lemma 1 *We know the following about independent firms' entry:*
1) $\widehat{n}^{ur} = 0$,
2) $\widehat{n}^* \geq 1$ *if* $a - c - \beta - 5\sqrt{F} \geq 0$,
3) $\widetilde{n}^{ur} \geq 1$ *if* $a - c - \beta - 3\sqrt{F} \geq 0$, *and*
4) $\widetilde{n}^* \geq 1$ *if* $a - c - \beta - 2\sqrt{F} \geq 0$.

We know from Propositions 4 and 5 that $W^m_v > W^{ur}_r$ and $W^*_{nsr} > W^*_{sr}$. Provided that at least one independent firm enters under structure regulation (with and without access charge regulation), it can easily be proven that $W^*_{nsr} > W^*_{sr} > W^{ur}_r$ (with the first inequality being provided by Proposition 5). Thus, welfare when regulating access charges (with or without structure regulation) is always strictly higher compared to not regulating access charges and subjecting the upstream monopolist to structure regulation.

Furthermore, it can be shown that W^*_{nsr} always yields the highest level of welfare in the free entry case, and strictly higher welfare than the (best) alternative if at least one rival firm enters the downstream industry. The requirement for entry by rival firms is that $\widehat{n}^* \geq 1$, or equivalently, $a - c - \beta - 5\sqrt{F} \geq 0$. This implies that rival firms enter only if the size of the market is sufficiently large relative to, among other things, the level of fixed costs. If $\widehat{n}^* \geq 1$, we can show that $W^*_{nsr} > W^m_v$ and if, in the regulated access charge scenario, no firms enter this implies that firm v will maintain a monopoly position both upstream and downstream and the welfare level is the same as in the unregulated access charge scenario W^m_v.

We now know that of the four different policy combinations the case with unregulated access charge and with structure regulation yields the lowest level of welfare. Furthermore, we know that welfare with access charge regulation and without structure regulation can never do worse than the best alternative policy combination. What can also be shown is that $W^*_{sr} \geq W^m_v$ only if $a - c - \beta - 6\sqrt{F} \geq 0$, with $W^*_{sr} > W^m_v$ if $a - c - \beta - 6\sqrt{F} > 0$.

The results of the welfare comparisons are gathered in the following proposition:

Proposition 6 *Welfare comparisons:*

1. If $a - c - \beta \geq 6\sqrt{F}$, in which case at least one independent firm enters (independent of regulatory structure), then $W^*_{nsr} > W^*_{sr} \geq W^m_v > W^{ur}_r$.

2. If $6\sqrt{F} > a - c - \beta \geq 5\sqrt{F}$ in which case at least one independent firm enters (independent of regulatory structure), then $W^*_{nsr} > W^m_v > W^*_{sr} > W^{ur}_r$.

3. If $5\sqrt{F} > a - c - \beta \geq 4\sqrt{F}$ in which case no independent firms enter without structure regulation, then $W^*_{nsr} = W^m_v > W^*_{sr} > W^{ur}_r$.

Consequently, we find that when the level of fixed costs is relatively low compared to the size of the market, regulating access charges will always yield at least the same level of welfare as the next best policy combination. For sufficiently low fixed costs, regulation of the access charge is always strictly better than the best alternative. If, however, fixed costs are sufficiently high, not regulating access charges in combination with the absence of structure regulation may yield higher welfare than regulating access charges and imposing structure regulations. Furthermore, for an interval of fixed costs, leaving access charges unregulated may be as good as regulating access charges (assuming that there is no structure regulation). This is due to the fact that no independent firms find it profitable to enter if fixed costs are sufficiently high when $w^* = \beta + \sqrt{F}$.

Thus, we see that the lower the level of fixed costs associated with entry into the downstream market, the better does the regulation of access charges do. Intuitively, it may be reasonable to assume that low levels of fixed costs, or equivalently low barriers to entry, necessitates less regulatory intervention. This is, however, not the case. Low fixed costs implies a higher degree of entry, *ceteris paribus*, provided that the access charge is regulated. If there is no regulation of the access charge we have seen that the subgame perfect equilibrium entails a complete foreclosure of all rival firms (cf. Proposition 4). This suggests that at least some form of regulatory intervention may be desirable. Regulating the access charge only will, provided that the fixed costs are low enough, result in entry which is beneficial for consumers, and will never be worse than a double monopoly. For sufficiently low fixed costs, the expansion in output due to entry - entry aided by access charge regulation - outweighs the social cost of duplication. Another alternative is to only impose structure regulation, which may be superior in terms of welfare compared to the complete foreclosure case, as it entails at least some degree of entry and output expansion. If the fixed costs are not too high, the gain to consumers' surplus outweighs the cost of duplication.

4.4 Excessive entry?

If there is free entry into a market and imperfect competition, we know from Mankiw and Whinston (1986) that free entry under certain conditions results in a socially excessive number of firms in the industry. The socially optimal number of firms is, of course, not necessarily equal to the number of firms that ensures zero profit. In this section we will look at how entry varies (under the zero profit assumption) with different policy combinations. I will focus on how the established result that free entry together with the business stealing effect leads to excessive entry when there is imperfect competition may change if (1) an upstream firm may serve a vertically related market, and (2) when the access charge may be subject to regulation.[24] The Mankiw and Whinston (1986) model is essentially extended to cover further cases which are of particular interest in vertically related markets, with their excessive entry result as a special case. In all of the analysis below the inverse demand is assumed to be linear.

The socially optimal level of entry will normally be defined by the first-order condition with respect to n on the appropriate welfare function (depending on whether there are restrictions on the network monopolist's opportunity for entry into the final product market). The socially optimal level of entry will then be compared to the level of entry that takes place in the four cases of policy combinations. The socially optimal level of entry will also be compared to the sub-game perfect policy combination.

When the regulator allows the network monopolist to enter the downstream market, the number of firms entering is determined by eqn. (9), whereas when he is not allowed to enter the number of firms is determined by eqn. (10). When the regulator chooses to determine the access charge, he implicitly affects the degree of entry. Similarly, if the regulator chooses *not* to regulate the access charge, he also makes an implicit choice about the level of entry. In this section, the analysis is not an attempt to shed light on subgame perfect policies, but rather to examine whether the excess entry result of Mankiw and Whinston (1986) carries over to each of the four scenarios analysed in the present paper. This implies that I only examine whether free entry results in excessive, socially optimal, or insufficient entry for each of the cases (i.e., for the given level of access charge in each case).

[24]The regulator could (in addition to, or instead of access charge regulation) potentially use direct entry restrictions as another regulatory instrument.

4.4.1 Structure regulation

Let us first consider the case where the network monopolist only serves the access market (i.e., the situation with structure regulation). In this case, the only major difference to Mankiw and Whinston (1986) is that the marginal cost of the downstream firms is endogenously determined. If the access charge is regulated we observe that entry is twice as large as in the unregulated case ($\widetilde{n}^* = 2\widetilde{n}^{ur}$), which implies that a benevolent regulator will choose a level of access charges such that more firms find it profitable to enter. We know that an unregulated network monopolist can only capture his monopoly profit through the access charge, and will, since he is restricted to a linear tariff scheme, choose an access charge in excess of marginal cost to capture some of this profit (cf. eqn. 13), which results in a lower level of entry. A welfare maximising regulator, however, chooses to use marginal cost pricing of access (cf. eqn. 17). This result is summarised in the following remark:

Remark *If the market for the intermediate good (access) is unregulated and the monopolist provider of access cannot operate in the final product market, then free entry and imperfect competition results in a lower level of entry than if the access charge is regulated.*

When the regulator restricts entry into the vertically related market, the socially optimal number of firms is in this case found by maximising eqn. (16) with respect to n, which is denoted n_r^{**}, and is implicitly given by:

$$\frac{\partial \widetilde{CS}(n,w)}{\partial n} - (c+\beta)\frac{\partial \widetilde{Q}^*}{\partial n} - F = 0 \tag{18}$$

In this case, assumption 2 is sufficient to ensure that total equilibrium output increases in the number of firms downstream, and that the output of each individual firm is decreasing in the number of firms (the latter effect is the business stealing effect). Without loss of generality, I will assume that the marginal cost of providing the final product is normalised to zero; i.e., $c = 0$.

We can rewrite eqn. (18) as:

$$(n_r^{**}+1)^3 = \frac{(a-w)^2}{F} + \frac{(a-w)(w-\beta)(n_r^{**}+1)}{F} \tag{19}$$

where n_r^{**} is the socially optimal level of entry when the network monopolist cannot enter the downstream market. Rewriting the free entry condition when the network monopolist is not allowed to enter the downstream market yields (n_r^e denotes the level of entry under free entry with structure regulation):

$$(n^e + 1)^2 = (a - w)^2 / F \tag{20}$$

The comparison between the actual entry and the socially optimal level of entry is summarised in the following proposition:

Proposition 7 *If the network monopolist is not allowed to enter the downstream market, then regulation of the access charge implies that the excess entry result of Mankiw and Whinston (1986) is retained. However, if the network monopolist determines the access charge, then there is neither excess nor insufficient entry.*

Proof. By comparing eqns. (19) and (20) we observe that if $w = \beta$ (i.e., in the regulated access charge scenario), the outcome is identical to the linear demand example in Mankiw and Whinston (1986): If the socially optimal number of firms is equal to 3 (i.e., $n_r^{**} = 3$), then free entry implies that 7 independent firms enter. If $n_r^{**} = 5$, then $n_r^e = 13$, and if $n_r^{**} = 8$, then $n_r^e = 26$.[25] When $w > \beta$ (i.e., the unregulated case), the comparison between n^* and n_r^e can be done by writing n_r^e as a function of n_r^{**}. Without loss of generality, we can normalise β to 0. Then we have the following:

$$(n_r^e + 1)^2 = (n_r^{**} + 2)^3 - (a - w)(n_r^{**} + 1) w/F \tag{21}$$

When the upstream monopolist determines an unregulated level of access charge in excess of the marginal cost of providing access (i.e., $w_r^{ur} = \left(a - \sqrt{F}\right)/2 > 0$), the latter element of eqn. (21) implies that the actual level of entry is lower than

[25] As is pointed out by Mankiw and Whinston (1986), the difference between the socially optimal number of firms and the actual entry is not the correct measure of welfare loss. They show that the welfare loss due to excess entry decreases as n^* increases. Obviously, as the free entry number of firms get larger, the equilibrium quantity becomes closer to the competitive limit which is beneficial to welfare, which partly offsets the cost of duplicating fixed costs.

when $w = 0$, and the degree of excess entry is lower. For $w = w_r^{ur}$, then $n_r^{**} = n_r^e$ if $n_r^{**} = \left(a - \sqrt{F}\right)/2\sqrt{F}$. By inserting for $n_r^{**} = \left(a - \sqrt{F}\right)/2\sqrt{F}$ in eqn. (19), the equality is satisfied and it is verified that $n_r^e = n_r^{**} = \left(a - \sqrt{F}\right)/2\sqrt{F}$. Note that for a given market size a, an increase in the socially optimal degree of entry must be accompanied by a reduction in the fixed costs (e.g., Let $a = 10$; if $n^* = 3$, then $F \approx 2.04$ to satisfy eqn. (19), if $n^* = 5$, then $F \approx 0.83$, etc.). ■

Consequently, when the network monopolist is not allowed to enter the downstream market, the degree of excess entry is lower without access charge regulation. Furthermore, as the result immediately above suggests, the unregulated monopolist will determine an access charge which induces the socially optimal number of firms to enter. This result is contrary to what is obtained by Mankiw and Whinston (1986). Thus, a welfare maximising regulator will choose an access charge which induces more entry than what is socially optimal. Levelling the playing field by using marginal cost pricing of access does not incorporate the effect of business stealing by entering firms, but takes only into account the direct cost of entry (the cost of production). The unregulated network monopolist is, contrary to the regulated firm, interested in capturing some of the (total) downstream profit by determining an access charge in excess of marginal cost, and he realises that the some of the profit a new entrant earns is a result of business stealing. Consequently, the network monopolist finds it less profitable to let a new firm enter and internalises the external effect posed by business stealing and prices access to induce optimal entry. When determining the socially optimal access charge, the welfare maximising regulator is not interested in the distribution of profits between the downstream firms and the upstream monopolist, but is simply concerned with ensuring the highest possible level of welfare which implies excessive entry.[26]

[26]However, when undertaking welfare comparisons between the various policy combinations, we find that welfare in the (structure regulation, no access charge regulation)-case yields the lowest level of welfare of the four combinations we consider.

4.4.2 No structure regulation

If we examine the situation where there is no structure regulation where the network monopolist can serve the final product market, then a similar result is obtained. When the access charge is determined by the network monopolist, the access charge is set high enough to deter all entry by independent firms (cf. Proposition 1), which implies that there is only one active firm in the downstream market. If, however, the regulator determines the level of the access charge he cannot do worse in terms of entry than in the unregulated case. Provided that the level of fixed costs is not too high, regulating access charges will lead to a higher degree of entry than in the unregulated case (specifically, if $a - c - \beta \geq 4\sqrt{F}$, then $\widehat{n}^* \geq \widehat{n}^{ur}$). This can be summarised as follows:

Remark *Assume there is free entry and imperfect competition downstream. Setting socially optimal access charges when the network provider is allowed to serve the downstream market will always result in at least the degree of entry that prevails if the access charge is set by the network provider.*

For any given level of the access charge, the socially optimal degree of entry is the solution to maximising eqn. (14) with respect to n is defined as n^{**}, and is implicitly given by the following expression:[27]

$$\frac{\partial CS^g(n,w)}{\partial n} - (c+\beta)\frac{\partial Q^*}{\partial n} - F = 0 \tag{22}$$

We know from assumption 3 that entry increases the total downstream output. We can rewrite eqn. (22) in the following manner:

$$(n^{**}+2)^3 = \frac{(a-2w+\beta)^2}{F} + \frac{(a-2w+\beta)(w-\beta)(n^{**}+2)}{F} \tag{23}$$

Rewriting the free entry condition when the network monopolist is allowed to enter, eqn. (4), we find that:

$$(n^e+2)^2 = (a-2w+\beta)^2/F \tag{24}$$

[27] A closed-form solution for the number of firms entering is difficult to obtain.

where n^e is the number of firms entering under free entry (without structure regulation).

Proposition 8 *Assume that the network monopolist is allowed to enter the downstream market. Then there will be no entry if the monopolist determines the access charge. If the access charge is subject to regulation, then the degree of excess entry is less pronounced without structure regulation than in the case with structure regulation.*

Proof. In Proposition 1, the first part this result is proven. The comparison between the socially optimal level of entry and the actual entry is then determined by:

$$(n^e+2)^2 = (n^{**}+2)^3 - (a-2w+\beta)(n^{**}+2)(w-\beta)/F \qquad (25)$$

If $w = \beta$, then the latter element of (25) vanishes, and $n^e = (n^*+2)^2 - 2$. This implies that the degree of excess entry will be substantial, e.g., with $n^e = 23$ if $n^{**} = 3$; $n^e = 47$ if $n^{**} = 5$; $n^e = 98$ if $n^{**} = 8$. By setting $w > \beta$, which is always the case when there is no structure regulation, the tendency for excess entry is partially mitigated. Without loss of generality, we can normalise β to 0. For $w = w^* = \beta + \sqrt{F}$, we know that if $a = 10$, then the following combinations of optimal entry and fixed costs satisfy eqn. (23): $(n^{**} = 3, F \approx 0.83)$, $(n^{**} = 5, F \approx 0.17)$, $(n^{**} = 8, F \approx 0.07)$. Assuming $a = 10$ and $c = \beta \equiv 0$, then if $n^{**} = 3$, then $n^e = 6$; if $n^{**} = 5$, then $n^e = 11$, if $n^{**} = 8$, then $n^e = 23$. ■

Consequently, leaving the pricing of access services unregulated results in a degree of entry which is lower than the level induced by a welfare maximising regulator, due to complete foreclosure (cf. Proposition 1). Again, contrary to the result obtained by Mankiw and Whinston (1986), there is no excessive entry in the present model of vertically related markets provided that the price of the intermediate product is unregulated (even if there is both imperfect competition and a business stealing effect of new entry). In this particular case, there may be socially insufficient entry. However, if the access charge is subject to regulation then there is excess entry, but to a lesser extent than in the case with both access charge regulation and structure regulation. Note that with marginal cost pricing of access, there is a substantial degree of excess entry when the network monopolist is allowed to enter the downstream market.

5 Concluding remarks

This article has studied socially optimal regulatory policies in vertically related markets, where the regulatory instruments available are access charge and structure regulation. Furthermore, it is examined whether the excess entry result obtained by Mankiw and Whinston (1986) carries over to a situation with vertically related markets. It is shown that free entry may or may not induce excessive entry in an imperfectly competitive downstream market depending on the regulatory policy chosen.

The analysis above is undertaken in a very stylised model, which does not capture all aspects of vertically related industries, and network industries in particular. First of all, the model assumes that there are no economies of scope between network provision and service provision. This may be a simplification in, for instance, the telecommunications industry where network providers often argue that there are substantial synergy effects between these two production elements. This implies that there are additional benefits to society from vertical integration that are not taken into account in the present analysis. The presence of economies of scope will only strengthen the result that the subgame perfect regulatory policy involves regulation of access charges and vertical integration. Allowing for economies of scope may, however, change the ranking of some of the socially suboptimal regulatory policies. In other related settings, it is not so obvious that there are economies of scope - for instance between content provision and network services in the Internet industry. The data flow from content provision provided over the Internet is often transported through the traditional telecommunications network, and it is not necessarily the case that the traditional telecom firms are better at providing content than independent firms. In some cases, for instance news and certain types of information, it seems reasonable that some independent firms are better equipped for producing content.

The present paper does not either examine network externalities, which is essential in the industries I have had in mind. How the introduction of such effects will influence the outcome depends on the way the network externalities work. Let us, for instance, assume that the level of the network effect is determined, in part, by the total quantity (of, e.g., network subscriptions) in the downstream market. If all networks are perfectly interconnected, all firms enjoy the same level of network effects for all configurations of the industry. This will increase the size of the total market, but not the distribution of market shares. If the network externalities are increasing in the downstream output, then the regulatory policy that results in the highest level of output will also generate the highest level of externalities. This will be an additional social benefit not taken into account in the present analysis. In the present model, the output is largest with structure regulation combined with regulation of the access charge.

6 References

Directive 97/33/EC of the European Parliament and of the Council, "On interconnection in Telecommunications with regard to ensuring universal service and interoperability through application of the principles of Open Network Provision (ONP)". Official Journal L 199, 26/07/1997 p. 0032 - 0052

Hansen, Bjørn (1999), "Competition and Broadband Upgrades in a Vertically Separated Telecommunications Market", Telenor Research & Development, R 22/99

Kreps, D. and J.A. Scheinkman (1983), "Quantity Precommitment and Bertrand Competition yield Cournot Outcomes", *The Bell Journal of Economics* **14**, 326-337

Laffont, J.J. and J. Tirole (2000), *Competition in telecommunications*, Cambridge, MA: MIT-Press

Lewis, T. and D. Sappington (1999), "Access pricing with unregulated downstream competition", *Information Economics and Policy* **11**, 73-100

Lipsey, R. G. and K. Lancaster (1956), "The General Theory of Second Best", *Review of Economic Studies* **24**, 11-32

Mankiw, N. G. and M. D. Whinston (1986), "Free entry and social inefficiency", *Rand Journal of Economics* **17**, 48-58

Rey, P. and J. Tirole (1997), "A primer on foreclosure", forthcoming in Armstrong, M. and R.H. Porter (eds.) *Handbook of Industrial Organization*, vol. 3, North-Holland

Salop, S. and D. Scheffman (1983), "Raising rivals' costs", *American Economic Review* **73**, 267-271

Sand, J. Y. (2000), "Regulation and foreclosure", mimeo, University of Tromso, Norway

Sand, J.Y. (2001), "A note on first- and second-best access charge regulation", mimeo, University of Tromso, Norway

Tirole, J. (1988), *The theory of industrial organization*, Cambridge, MA: MIT-Press

Vickers, J. (1995), "Competition and regulation in vertically related markets", *Review of Economic Studies* **62**, 1-17

ESSAY 5

A note on first- and second-best access charge regulation

Jan Y. Sand*

University of Tromso

Department of Economics

NFH

N-9037 Tromso, Norway

July 2001

Abstract

This paper considers access charge determination in vertically related markets with imperfect competition between downstream firms, and when the number of firms is exogenously determined. The properties of both the unregulated and regulated access charges are discussed, both with and without transfers between the regulator and the regulated firm.

JEL Classification: L13, L22, L51

Keywords: duopoly, access charge regulation, vertical relations

1 Introduction

In industries which are characterised by vertically related markets, the issue of pricing the access to essential facilities becomes important from a social point of view. In many industries, the providers of essential inputs needed to produce a product that

*I am grateful for valuable comments from Kåre P. Hagen and Nils-Henrik von der Fehr.

can be sold to end-users possess market power and may have incentives to choose a price which is higher than the socially optimal price. An additional problem is that the providers of the essential input may be active in the end-user market, competing with rival firms, which implies that there may be incentives to foreclose rival firms. This is frequently the case in the telecommunications industry (the providers of local access also supplies long-distance services), and can also be observed in the market for computer software (the operating system provider also produces application software).

In the present paper, I examine under which circumstances a monopolist provider of an essential input will choose not to completely foreclose rivals firms producing products for end-users. The pricing of access can be seen as a problem of outsourcing all or part of the production of end-user products. When the access charge is not subject to regulation, the provider of the essential input will determine a level of the access charge such that the rival firm producing for end-users is only allowed to produce if it is sufficiently more efficient than the monopolist's own subsidiary. The regulator's problem when choosing the access charge is also related to choosing which firm to produce the end-user product, but the focus is then on the socially optimal mix of production between the two firms. Furthermore, the first- and second-best access charges are examined. The first-best access charge is obtained if the regulator can use compensating transfers at no social cost, whereas no transfers can be used in the case of the second-best access charge.

In light of the potential for foreclosure of rival firms in the end-user market, the question of whether a monopolist provider of an essential input should be allowed to enter a vertically related market is also examined in a stylised setting. Foreclosure in the setting presented below takes the form of setting a high access charge. Alternatively, foreclosure can be exercised by the use of non-price means, e.g., by degrading the quality of the input supplied to rivals (see e.g., Economides, 1998 and Sand, 2000).

The remainder of the paper is organised as follows: In section 2, the model is described and the solution to the final stage of the game is presented. In section 3, both the unregulated and socially optimal access charges are analysed. In section 4, I compare various combinations of regulatory policies to find the subgame perfect regulatory regime. In section 5, some concluding remarks are presented.

2 The framework

I assume that there is one firm upstream (firm v) and only two potential firms downstream (firm s and firm v). The two downstream firms must decide whether they wish to be active in the final product market (or a regulator decides whether they are allowed to enter this market), where one of the potential downstream firms is a subsidiary of the upstream monopolist. The final product market can be thought of as a new market, e.g., a market for broadband communication services, where firms may or may not enter. The upstream firm provides transportation network services which is an essential input for broadband services.

The analysis is conducted in the setting of a multi-stage game. In the *first stage* either the upstream firm or the regulator decides on an access charge. In the *second stage*, firms choose simultaneously to enter or not. In the *final stage*, there is either a monopoly or a duopoly (depending on the strategies chosen by the regulator and firms v and s in the preceeding stages). The final stage of the game is unregulated and firms compete in quantities.

The entry game can be thought of as follows, where firms v and s choose either to enter (E) or not enter (NE):

		Firm s	
		E	NE
Firm v	E	Duopoly	Monopoly to firm v
	NE	Monopoly to firm s	No activity downstream

There are potentially four different cases to be considered in the current model, but I assume that the case where no firms enter is not an option. The access charge and level of fixed costs will play an important role in respect to which firm(s) enters, and will be discussed in detail below.

Firm v's profit earned in the upstream market only is given by (if firm v is vertically integrated, it will in addition earn profit downstream):

$$\pi_v^u = (w - \beta)(q_v + q_s) \tag{1}$$

where $q_i \geq 0$, for $i = v, s$, are the downstream output levels. Fixed costs upstream is an important characteristic of most local access technologies, and is a major reason for the lack of competition in the local loop ("the last mile" of telecommunication networks). In many networks the investments in infrastructure are already sunk, and plays no role in the determination of the access charge. Thus, an ad hoc justification of leaving fixed costs out of the analysis is that the game which is played in the current paper takes place after investments are sunk. Furthermore, the level of these fixed costs is assumed to be high enough to deter entry into the upstream segment.[1] The parameters w and β are the access charge and upstream marginal cost, respectively.

Firm i's downstream profit, for $i = v, s$, is given by:

$$\pi_i = (P - w - c_i)\, q_i - F \tag{2}$$

where $P = a - q_i - q_j$, for $i \neq j$, and $i, j = v, s$, is the inverse demand function. Downstream marginal cost for firm i is denoted c_i. We assume that both firms must pay the same fixed costs, F, for establishing downstream operations. This could be costs associated with setting up a distribution and sales network, marketing expenses, and it is reasonable that both firms face the same fixed costs. These fixed costs could also be attributed to a USO-fee (USO - Universal Service Obligation), which all firms operating downstream must pay. The presence of the fixed cost implies that firms may not want to enter into the industry even if equilibrium quantities are positive. A sufficient condition for the upstream firm to enter the downstream industry is that the downstream operation, in isolation, earns positive profits.[2]

[1] A consequence of ignoring fixed costs upstream is that the assumption which requires the access charge to cover the marginal cost of providing access is sufficient for ensuring that the uptream firm earns non-negative profits. If there are fixed costs present, then the access charge must be strictly above marginal cost for Ramsey reasons.

[2] In general, we would expect that the upstream firm evaluates the profit earned as a monopolist

The regulator's welfare is given by:

$$W = \begin{cases} CS_v^m + \left(P_v^m - c_v - \beta\right) q_v^m - F & (a) \\ CS^d + \left(P^d - c_v - \beta\right) q_v^d + \left(P^d - c_s - \beta\right) q_s^d - 2F & (b) \\ CS_s^m + \left(P_s^m - c_s - \beta\right) q_s^m - F & (c) \end{cases} \tag{3}$$

where $CS_i^m = \int_0^{q_i^m} P(s)\, ds - P_i^m q_i^m$, for $i = v, s$, and $CS^d = \int_0^{q^d} P(s)\, ds - P^d \left(q_v^d + q_s^d\right)$ are the net consumers' surplus under monopoly and duopoly, respectively.[3] The welfare if only firm v is active downstream is given by (a), if both firms are active (b), and if only firm s is active by (c). Observe that a duopoly outcome implies a duplication of the fixed costs, which is detrimental to welfare.

2.1 Solving the game

We are looking for the subgame perfect equilibrium in the game, and starts at the last stage to reveal the equilibrium path. We first solve for the Cournot-equilibrium. Then we solve for the optimal access charge; i.e., the price paid for access which maximises either the upstream firm's profit (the unregulated case) or the regulator's welfare function (the regulated case).

2.1.1 Final stage

The Duopoly case The vertically integrated firm v maximises the sum of its upstream and downstream profit (i.e., the sum of equations 1 and 2). Firm v's and firm s's equilibrium profit is given by, respectively:

$$\Pi_v^d = \left(Q_v^d\right)^2 + (w - \beta)\, Q_s^d - F \tag{4}$$

$$\Pi_s^d = \left(Q_s^d\right)^2 - F \tag{5}$$

access provider with no downstream operation against the profit earned as a vertically integrated firm (i.e., compare $(P - c_v - w)\, q_v^d - F + (w - \beta)\left(q_v^d + q_s^d\right)$ against $(w - \beta)\, q_s^m$, where superscript m and d represent monopoly and duopoly, respectively).

[3]Superscript m and d represent monopoly and duopoly, respectively, and subscript v and s represent the two firms.

where Q_v and Q_s are the Cournot-equilibrium quantities, where

$Q_v = (a - 2(c_v + \beta) + c_s + w)/3$ and $Q_s = (a - 2(c_s + w) + c_v + \beta)/3$. In order to have a set of non-negative profit levels we must have: $Q_v^d \geq \sqrt{F - (w - \beta) Q_s^d}$, and $Q_s^d \geq \sqrt{F}$.

Monopoly case If firm v enjoys a monopoly position downstream, the access charge will simply be an internal transfer and, consequently, we are in the standard case of a profit-maximising monopolist. It is obvious that the role of the access charge disappears when firm v has monopoly both upstream and downstream, as it is simply an internal transfer. In this case, there is no need for an examination of the optimal access charge.

Total industry profit is given by:

$$\Pi_v^m = (Q_v^m)^2 - F \tag{6}$$

If firm s is the monopolist downstream, total industry profit includes firm v's profit from upstream operations:

$$\Pi_s^m = (Q_s^m)^2 - F \tag{7}$$

$$\Pi_v^u = (w - \beta) Q_s^m \tag{8}$$

The monopoly profit maximising quantities are $Q_v^m = (a - c_v - \beta)/2$ and $Q_s^m = (a - c_s - w)/2$ in the two different cases. To ensure that either firm v or s is active downstream when there is a monopoly, I assume the following:

Assumption 1

i) $a - (c_v + \beta) > 2\sqrt{F}$

ii) $w < a - c_s - 2\sqrt{F}$

Part *(i)* of assumption 1 ensures that the total industry profit in equilibrium is positive when firm v is the only active firm (upstream *and* downstream), whereas part *(ii)* ensures that firm s earns a non-negative level of profit in equilibrium as a monopolist. Thus, assumption 1 tells us that a firm which enjoys a monopoly situation will have a positive level of production, and, furthermore, that a firm's monopoly profit downstream is non-negative (i.e., $Q_i^m \geq \sqrt{F}$ for $i = v, s$). To ensure this, there must be a ceiling on the access charge (assumption 1, part ii). If it is the downstream subsidiary of the upstream monopolist who is the sole downstream provider, then the level of the access charge obviously plays no role.

To ensure that the upstream firm earns a non-negative level of profit, we assume the following:

Assumption 2

The regulated access charge must ensure coverage of costs; $w \geq \beta$.

If it is the rival downstream firm that enjoys a monopoly situation downstream, then the upstream monopolist will always prefer to provide access to firm s at a price which implies that firm s makes a positive level of profit (provided, of course, that $w \geq \beta$). An access charge such that $w > \beta$, and such that $\Pi_s^m \geq 0$, yields a strictly positive profit level for firm v. If the upstream monopolist sets an access high enough to deter entry by firm s, and provided that firm v cannot enter downstream itself, yields zero profit for firm v. Assumption 2 ensures that production upstream earns at least zero profit.

2.1.2 Entry decisions

Firms v and s will choose to enter into the downstream industry if they find it profitable. Furthermore, for firm v to enter, such entry must in addition be allowed by the regulator.

Provided that firm v enters the downstream industry, firm s chooses to remain in the industry if $\Pi_s^d \geq 0$. If firm v does not enter and firm s enters, firm s gets a profit of $\Pi_s^m > 0$ and 0 otherwise (i.e., no entry). Consequently, firm s will always choose to enter into the industry if firm v chooses *not* to enter, or if firm v is not allowed to enter. Obviously, the profit obtained in either situation will depend on the access charge set by the government in the regulated case, and firm v in the unregulated case.

One of the roles that the access charge play from a social point of view is to correct for the distortion that exists in the downstream market.[4] Since there are only two firms and these firms compete in quantities, there will be an allocative loss downstream. If we examine the upstream market in isolation, the price of access should be set equal to marginal cost.[5] However, from the theory of second-best we know that when there is a distortion in one market, it may be socially desirable to introduce distortions in other markets to correct for the existing distortion.[6] If the regulator can subsidise access deficits and provided that the shadow cost of public funds is sufficiently low, the optimal access charge is set below the marginal cost of providing access (see Sand, 2000).[7] However, the main assumption in this paper is that the access charge must be sufficient to cover costs, and transfers will only be introduced to show the characteristics of the first-best solution.

The access charge can also be used as an instrument to limit the degree of costly duplication of fixed costs. Provided that firm v finds it in its interest to enter, the regulator may choose to set the access charge high enough to foreclose the rival firm if the fixed costs are sufficiently high, or if the rival firm is very inefficient relative to the downstream subsidiary of the access provider. An alternative policy to the regulation of access charge to avoid duplication of fixed costs is to restrict the upstream monopolist's opportunity to enter into vertically related markets.

The main reason that regulation of final product prices is not considered here is that the (potentially) most competitive market is the downstream market, and it is more likely that this market is in less need of regulatory corrections. Arguments along this line is found in, e.g., Laffont and Tirole (2000), who argues that the more severe monopoly problem is in the upstream market with (potentially) significant

[4]If we consider the case of free entry into the downstream market, the role of the access charge from a social point of view is twofold (see Sand, 2001). First, it can be used to correct for the potential allocative loss downstream as a result of imperfect competition. Second, the access charge can be used to limit socially costly duplication of fixed costs.

[5]If there are fixed costs upstream, the access charge must be in excess of the marginal cost of providing access.

[6]The classical reference on second-best theory is Lipsey and Lancaster (1956).

[7]When the shadow cost of public funds is too large, the social cost of the distortion in the downstream industry must be weighed against the welfare loss of transfers.

economies of scale. Consequently, provided that a complete deregulation is the ultimate goal of the emerging telecommuncations legislations the more appropriate regulatory policy seems to be to direct the attention to the bottleneck segment (i.e., network services), and regulate access charges rather than regulating the prices of the final products.

3 First- and second-best access charges in a duopoly

With an exogenous number of potential firms downstream (two potential firms), the main question is to investigate in which situations it is socially optimal to have a monopoly or duopoly downstream. The regulator determines the access charge prior to entry decisions and Cournot competition. We will assume that the regulator has complete information about all relevant aspects to the game.

The regulator calculates the level of welfare in equilibrium for all possible outcomes, and determines a policy with respect to the access charge and entry to maximise welfare. If the access charge is unregulated, the local access provider is free to determine the level of the access charge that maximises his profits.

3.1 Unregulated access charge

Let us assume that firm v enters the downstream industry. It will then choose an access charge which maximises its profits, given by eqn. (4), subject to non-negativity constraints on equilibrium quantities. By setting a high access charge it ensures that its own downstream subsidiary obtains a dominant position in the final product market, which is positive for firm v's profit. However, by setting a high access charge, firm v sells less to its rival of access services, which implies has a negative impact on firm v's profit. This trade-off results in the following proposition.[8]

Proposition 1

[8] This is similar to the result in Lewis and Sappington (1999).

The vertically integrated firm will only foreclose its downstream rival completely using the access charge if its own subsidiary is equally, or more efficient that the rival firm.

The proof of this result is explained here. The solution to the vertically integrated firm's maximisation problem, w^{ur}, is given by:

$$w^{ur} = \frac{5a - 4c_s - c_v + 5\beta}{10} \tag{9}$$

We observe that if firms have identical downstream marginal costs, i.e., $c_s = c_v$, then $w^{ur} = (a - c + \beta)/2$ and $Q_s^d = 0$, which yields a negative profit for firm s; $\Pi_s^d = -F < 0$, and firm s does not enter. However, if the rival firm has a cost advantage over firm v's own downstream subsidiary, then $Q_s^d > 0$. We then need to examine whether this output level is sufficiently large to cover the fixed costs. If the rival firm enjoys a cost advantage over firm v, then the vertically integrated firm faces a problem which is essentially equivalent to the problem of choosing whether to outsource part of its own production to an independent firm. A firm seeking to maximise its own profit will choose to outsource (all, or part of) its production if the cost advantage of the rival firm is sufficiently large.

When firm s has a cost advantage, the equilibrium quantity for firm s when inserting for the profit maximising choice for the access charge is given by: $Q_s^d = 2(c_v - c_s)/5$. We know that in order for firm s to enter we must have $(c_v - c_s) \geq \left(5\sqrt{F}\right)/2$, which implies that the rival firm will enter only if the vertically integrated firm's cost disadvantage is sufficiently large. The higher the entry costs, the larger must firm v's cost disadvantage be for the vertically integrated firm to not force firm s out of the market. The reason for not exercising full foreclosure in this case is that firm v finds it more profitable to outsource some of the downstream production to firm s. It is better for the vertically integrated firm to concentrate on providing the monopoly service (i.e., local access), rather than to compete head to head downstream where it has a disadvantage. Furthermore, eqn. (9) is the subgame perfect equilibrium access charge only if firm v's profit when determining the access charge according to this formula is at least as large as firm v's monopoly profit. It can be shown that firm v's profit when choosing an accommodating strategy is indeed larger than firm v's entry deterring profit. The reason for this is that firm s's entry is only accommodated provided that it is sufficiently more efficient than firm v, in which case firm v finds it relatively more profitable to outsource some of the downstream production rather than to be an inefficient sole producer of the final product.

When firm v does not enter the downstream market (e.g., if the regulator restricts entry directly), there is a monopoly both upstream and downstream. Firm v then maximises its profit given by eqn. (1), subject to firm s choosing its monopoly quantity Q_s^m. The access charge, w', that solves this maximisation problem is then given by:

$$w' = \frac{a - c_s + \beta}{2} \tag{10}$$

which yields a positive quantity for firm s downstream. Note that we will make the assumption that $w' \geq \beta$ to ensure that the upstream monopolist earns a non-negative level of profit. This is summarised in assumption 3:

Assumption 3

The following is necessary for the unregulated upstream monopolist to earn positive profits when the rival firm is a monopolist downstream: $a - c_s - \beta \geq 0$.

Thus, assumption 3 ensures that an upstream monopolist selling access to a downstream monopolist will make non-negative profits when it cannot enter into the downstream market, since 1) $w' \geq \beta$ is ensured and 2) $Q_s^m > 0$.

3.2 Optimal access charge with an exogenous number of firms

The regulator maximises a utilitarian welfare function, which is the sum of consumers' surplus and the firms' profits, defined by (3), subject to non-negativity constraints on equilibrium quantities and profits. We will for now restrict our attention to the case where the regulated access charge is at least as large as the marginal cost of providing access (assumption 2).[9]

This implies that in the presence of imperfections in the downstream market, the access charge which maximises welfare in our case is not necessarily the unrestricted optimal access charge. As discussed above (and in more detail in Sand, 2000), a feature of optimal access charges when there is a Cournot duopoly downstream is that the regulator finds it optimal to induce an access deficit when transfers are allowed (and if the shadow cost of the transfers is sufficiently low). Contrary to Sand (2000), it is assumed that transfers are *not* allowed, unless stated otherwise. This implies that the regulated access charge must cover the costs of providing access (assumption 2).[10]

We can rewrite the welfare function as follows:

$$\begin{aligned} W_s^m &= CS_{gross}^s - (c_s + \beta) Q_s^m - F \\ W^d &= CS_{gross}^d - (c_v + \beta) Q^v - (c_s + \beta) Q^s - 2F \\ W_v^m &= CS_{gross}^v - (c_v + \beta) Q_v^m - F \end{aligned} \tag{11}$$

where we have inserted for final stage equilibrium quantities. CS_{gross}^i, represents the gross consumers' surplus, which is equal to: $CS_{gross}^i = \left(aQ_i^* - (Q_i^*)^2/2\right)$, where $i = d, s, v$ denotes duopoly, monopoly to firm s and monopoly to firm v, respectively. Q_i^* is the total equilibrium downstream quantity in the three cases.

Ignoring first the constraint that the access charge must be at least as large as the marginal cost of providing access, and let the regulator have access to transfers at zero social cost. We then obtain the *unrestricted* optimality conditions (i.e., the first-best regulation) for the two cases of downstream monopoly to firm s and the

[9]Transferring public funds to the firm is socially costly, and this will limit the extent to which the regulator whishes to subsidise the rival's cost of purchasing access. In this paper, I assume that the regulator cannot use lum-sum transfers. However, the case of costless transfers is used as a first-best benchmark.

[10]In the full information scenario *with* transfers, the regulated firm's profit is set equal to the value of the outside options for the regulated firm (usually normalised to zero). Profits to the firm are socially costly if there are efficiency losses associated with public funds, and are therefore taxed away. In the present model, any positive profit may be retained by the regulated firm.

case of duopoly:

$$w_s^* = 2\beta + c_s - a$$

$$w_d^* = 2\beta + 5c_s - 4c_v - a$$

The optimal access charge when firm v is the sole provider both upstream and downstream is indeterminate, as it is simply an internal transfer within the firm and plays no allocative role. It is furthermore worth noting that when firm s is a monopolist downstream, the unrestricted access charge is below marginal cost of providing access, since $w_s^* - \beta = -(a - c_s - \beta)$ which must be negative if firm s is active downstream. Since the welfare function is concave in the access charge, any access charge in excess of the optimal access charge will reduce welfare.[11] Consequently, if the regulator cannot use transfers he will determine an access charge as low as possible without violating the constraint $w \geq \beta$. This implies setting an access charge equal to marginal cost of providing access; $w_s^{*r} = \beta$.

However, let us for a moment contiune to examine what would happen if the regulator could costlessly subsidise an access deficit and sets an access charge equal to $w_s^* = 2\beta + c_s - a$. Inserting the optimal access charge into the monopoly quantity of firm s yields an equilibrium output $Q_s^* = a - c_s - \beta > 0$, which corresponds to the competitive output level when the (social) marginal cost of producing is $c_s + \beta$.[12] Consequently, by setting the access charge below marginal cost of providing access, the regulator is able to design a perfect correction for the distortion in the downstream market through the regulation of the upstream market. When we restrict our attention to the no transfer case, the distortion in the downstream market cannot fully be corrected for, and the total downstream quantity is 1/2 of the competitive output level, provided that $w_s^* = \beta$.

[11] Since $\partial^2 W_s^m / \partial w^2 < 0$, then for any $w > w_s^*$, where w_s^* solves the regulator's unconstrained maximisation problem, welfare is reduced.

[12] To see this, simply equate the inverse demand function $P = a - Q$ with the marginal cost of producing as perceived by society which is $c_s + \beta$ (the sum of firm s's marginal cost and the marginal cost of providing access). This yields $Q^{comp} = a - c_s - \beta$.

The optimal unrestricted access charge when there is a duopoly, w_d^*, is above marginal cost of providing access only if firm s's marginal downstream cost is sufficiently higher than firm v's marginal cost, in which case firm s actually chooses not to enter. Inserting for $w_d^* = 2\beta + 5c_s - 4c_v - a$ into firm v and s's equilibrium quantities in the final product market yields the following quantities: $Q_v^d = 2(c_s - c_v)$, and $Q_s^d = (a - \beta) + 3c_v - 4c_s$. Thus, firm v is only active downstream if its own downstream subsidiary is more efficient than the rival. Rearranging the expressions, we see that for the access charge w_d^* to be larger than marginal cost of access and firm s being active in the downstream market, we need parameter values such that $(c_s - c_v) \geq Q_s^d \geq \sqrt{F}$. The first inequality implies that for the access charge w_d^* to be larger than β, the efficiency parameters of firms s and v must be such that the equilibrium quantity of firm v is at least twice as large as firm s. The latter inequality simply ensures that firm s finds it profitable to enter.

We observe that if firms have identical downstream marginal costs, i.e., $c_v = c_s = c$, the optimal *unrestricted* access charge of the form w_d^* is not consistent with both firms producing positive quantities in equilibrium. When firms have identical downstream costs and when the access charge is determined by w_d^*, firm v does not produce in equilibrium. In this case, the access margin $w_d^* - \beta = -(a - c - \beta)$ must be negative if any production is to take place in equilibrium. Consequently, since $w_d^* < \beta$, firm s's perceived marginal cost $(c + w_d^*)$ of producing is lower than firm v's perceived cost of producing $(c + \beta)$. Another observation that we can make, is that in the identical cost case the unrestricted regulation results again in a total quantity downstream which corresponds to the competitive level of output. If we reintroduce differences in downstream marginal costs, we find that $Q_{unrestr}^* = a - \beta + c_v - 2c_s$.

Let us consider the case where the upstream firm is allowed to enter the downstream market. In many situations, the unrestricted access charge given by w_d^* will not satisfy the constraint $w \geq \beta$. Since the welfare function W^d is concave on w, the regulator will set an access charge as low as possible, while satisfying the constraints. This implies that the restricted optimal access charge (when the conditions above are not met) is given by: $w_d^{*r} = \beta$. In this case, both firms produce positive quantities in equilibrium, and the only reason for producing different level of output is that $c_v \neq c_s$.[13]

When inserting for the restricted access charge, we find that total downstream quantity in equilibrium is given by: $Q^*_{restr} = (2a - c_v - c_s - 2\beta)/3 > 0$. The difference in the output levels in the unrestricted and restricted case can be written as: $\Delta Q^* \equiv Q^*_{unrestr} - Q^*_{restr} = ((a - \beta - c_s) + 4(c_v - c_s))/3$ or $(a - \beta + 4c_v - 5c_s)/3$, which is positive only if $w^*_d - \beta < 0$. Thus, the equilibrium output is larger in the first-best regulation case only if the regulator uses access subsidies. In this case the entrant is effectively subsidised by the choice of regulatory policy, and the playing field is tilted in favour of the rival firm to correct for the distortion in the downstream market.

The results on the optimal access charge is summarised in proposition 2:[14]

Proposition 2

1. If the upstream monopolist does not enter the downstream market, the regulated access charge is $w^*_s = 2\beta + c_s - a$ if the resulting access deficit can be costlessly subsidised, and $w^{*r}_s = \beta$ otherwise.

2. When both firms are active in the downstream industry, the regulated access charge is equal to $w^*_d = 2\beta + 5c_s - 4c_v - a$ if $-(a - c_s - \beta + 4(c_v - c_s)) \geq 0$ or if the access deficit can be costlessly subsidised, and $w^{*r}_d = \beta$ otherwise.

If we simplify the analysis by only considering firms with identical marginal costs ($c_v = c_s = c$) and retain assumption 2, then the regulated access charge will always be set equal to the marginal cost of providing access. Thus, in this simplified environment we have that $w^{*r}_s = w^{*r}_d = \beta$ which is assumed to be lower than the unregulated access charge. This implies that both firms will be active in the market as a consequence of the regulation and will be producing the same quantity. If an access deficit can be costlessly subsidised, then the first-best optimal regulation results in only firm s being active downstream. In the unregulated case, we know that firm s is completely foreclosed and does not produce in equilibrium.

[13] Assumption 1 ensures that both firms produce a positive level of output.

[14] The requirement in part 2 of Proposition 2 that $-(a - c_s - \beta + 4(c_v - c_s)) > 0$ is equivalent to requiring that $w^*_d \geq \beta$.

The regulator can leave it to the market (read, the vertically integrated firm) to determine the access charge, the regulator (or competition authority) can restrict firm v's opportunity to enter into vertically related markets, and/or the regulator can determine the access charge.[15] When the regulator decides which policy should be chosen (i.e., the subgame perfect policy), he must take into account the welfare loss in the duopoly case due to duplication of fixed costs. This is the topic for next section.

4 Welfare comparisons

In this section, we will examine the level of welfare under different sets of policies with respect to the access charge and entry. We will consider only the restricted case, where no transfers can be used and where there cannot be an access deficit. In addition to regulation of the access charge, the regulator may also restrict the upstream firm's opportunity to enter in the downstream industry (structure regulation). To simplify the analysis, I will only consider the case of symmetric downstream firms where $c_v = c_s = c$.

4.1 No regulation versus structure regulation

We have learnt from the analysis above that when firms face the same downstream marginal costs, the *unregulated* situation without any structure regulation implies that the only active firm is the vertically integrated firm (which enjoys a monopoly situation both upstream and downstream); i.e., the vertically integrated firm will

[15] If there are fixed costs, $K > 0$, in the upstream market, then a zero-profit condition on access provision implies the following: $(w - \beta) \geq K/Q$, where Q is the downstream quantity (which is either a duopoly or a monopoly). Thus, the access charge must be set strictly in excess of the marginal cost of providing access. Let us assume that $c_v = c_s = c$, which implies that firm v always will produce a higher quantity than firm s ($Q_v > Q_s$ in the duopoly case, and $Q_v^m > Q_s^m$ in the two monopoly cases). The consequence of introducing fixed costs upstream is that $w_s^* > w_v^* > w_d^* > \beta$ (i.e., the access charge is highest when firm s is a monopolist and lowest under a duopoly, but always strictly in excess of the marginal cost of providing access).

foreclose its rival completely. Consequently, there is no duplication of fixed costs and no problems of double marginalisation.[16] This situation yields the welfare level W_v^m as defined in expression (11), with the access charge w^{ur} given by eqn. (9).

If the regulator chooses to only restrict firm v's opportunity to enter into the downstream market (*structure regulation*), firm v is only active upstream and firm s is the only firm producing downstream. The welfare will be given by W_s^m also defined in expression (11), with the access charge w' defined by eqn. (10). In both these two cases, the upstream monopolist decides on the level of the access charge.

Proposition 3

Assume that firms are symmetric (i.e., $c_v = c_s$). If the access charge is unregulated, then the overall welfare level is higher without structure regulation.

Proof. Define $\Delta W^m \equiv W_s^m - W_v^m$ and $\Delta Q^m = Q_s^m - Q_v^m$, where $Q_s^m\,(w = w') = (a - c - \beta)/4$ and $Q_v^m = (a - c - \beta)/2$. The access charge w' is defined by eqn. (10). Simplifying the expressions, $\Delta Q^m = -(a - c - \beta)/4$, which is negative provided that assumption 1 is satisfied.

Then $\Delta W^m = (a - c - \beta)\,\Delta Q^m - \left[(Q_s^m)^2/2 - (Q_v^m)^2/2\right]$. Under assumption 1 we observe that $(a - c - \beta)\,\Delta Q^m < 0$. The expression $\left[(Q_s^m)^2/2 - (Q_v^m)^2/2\right]$ can, after simplifications, be written as $-3\,(a - c - \beta)^2/32$, which is negative under assumption 1. Consequently, we can write ΔW^m as: $\Delta W^m = -5\,(a - c - \beta)^2/32 < 0$. ■

If only firm v produces the final product, it will avoid the problem of double marginalisation and will consequently choose a higher level of output than is the case when the upstream monopolist sells access to the monopolist firm s. By observing the form of the welfare function, we can see that welfare is increasing and concave in quantity since it is assumed that a monopolist downstream has a positive profit margin (which requires $a - c - \beta > 0$). Thus, a welfare maximising regulator, when choosing between no regulation and structure regulation, will do better by allowing

[16]Double marginalisation is a problem if firm s produces downstream and the access charge is larger than the marginal cost of providing access. The classical reference is Spengler (1950).

firm v to enter the downstream industry and effectively foreclose the rival firm (which then chooses not to enter the industry) as the total output level increases. Duplication of fixed costs, however, is not an issue as the downstream industry will be monopolised in both cases.

If firms have different marginal costs upstream (i.e., $c_v \neq c_s$) and if $c_v < c_s$, the result will be the same as in Proposition 3. In this case, firm v has an even greater cost advantage in downstream production. However, if $c_v > c_s$ the analysis becomes more complicated. Then, if access is still unregulated, the absence of structure regulation is only welfare optimal if the downstream subsidiary of firm v is not too inefficient relative to firm s.

4.2 Access charge regulation

The regulator may in addition to, or instead of the structure regulation, choose to regulate access charges. If there is regulation of access charges, the regulator faces two options. He may regulate only access charges, or he may use a combination of access charge regulation and structure policy. We then need to consider two different cases. One case in which access charge regulation is used in combination with structure regulation, which effectively awards firm s a monopoly in the downstream industry. The other case consists of access charge regulation only. In both these cases, it is assumed that the regulator cannot use transfers to subsidise any access deficits, if this were to occur in a fully optimal regulatory policy.

If the regulator restricts the entry of firm v into the downstream industry, it will choose as the optimal access charge $w_s^{*r} = \beta$.[17] The corresponding output by the downstream monopolist is then: $Q_s^m = (a - c - \beta)/2$, which is the same output level as firm v would choose if it is allowed monopoly power downstream since marginal cost pricing of access and identical downstream marginal costs imply that the two firms face the same total costs. The welfare in this case is denoted W_{sr}, which can be rewritten as $W_{sr} = (a - c - \beta) Q_{sr}^m - (Q_{sr}^m)^2/2 - F$.[18]

[17] We know from above that the first-best regulation would result in an access charge, w_s^*, below the marginal cost of providing access.

[18] Subscript sr denotes structure regulation.

If the regulator abstains from using structure regulation in this case we know from above that both firms will be active in the downstream industry, and we also know that the regulator will determine a policy of marginal cost access pricing; i.e., $w_d^{*r} = \beta$. The equilibrium levels of output for the two firms are, since firms are symmetric: $Q_s^d = Q_v^d = (a - c - \beta)/3$. The level of welfare in this situation is given by: $W^d = (a - c - \beta)(Q_v^d + Q_s^d) - (Q_v^d + Q_s^d)^2/2 - 2F$.

We now want to compare W^d to W_{sr} to examine which of these two policies yield the highest level of welfare. Note that when there is no structure regulation, there is duplication of fixed costs.

Proposition 4

Assume that firms are symmetric (i.e., $c_v = c_s$). If access charge is subject to regulation and in the absence of transfers, a welfare maximising regulator will do better by not subjecting firms to structure regulation when fixed costs are not too large.

Proof. When using transfers is not an option, we have seen that the regulator will determine access charges $w_d^{*r} = w_s^{*r} = \beta$. Define $\Delta W^d \equiv W^d - W_{sr}$ and $\Delta Q^d \equiv Q^d - Q_{sr}^m$, where $Q^d = Q_v^d + Q_s^d$ and $Q_v^d = (a - c - \beta)/3 = Q_s^d$ when we have inserted for the regulated access charge. We can rewrite ΔW^d as follows: $\Delta W^d = (a - c - \beta)\Delta Q^d - \left[(Q^d)^2/2 - (Q_{sr}^m)^2/2\right] - F$. Since the regulator chooses a plicy of marginal cost pricing of access, the two firms face identical marginal costs of producing the final product. We then know that $\Delta Q^d = (a - c - \beta)/6 > 0$. Furthermore, $\left[(Q^d)^2/2 - (Q_{sr}^m)^2/2\right] = 7(a - c - \beta)^2/72$, which implies that $\Delta W^d = (a - c - \beta)^2/6 - 7(a - c - \beta)^2/72 - F$, or $\Delta W^d = 5(a - c - \beta)^2/72 - F$. Consequently, $\Delta W^d > 0$ if $(a - c - \beta) > \sqrt{72F/5}$. ■

The intuition behind this result is, of course, that a duopoly produces a higher level of output than does a monopoly. In our case, both the downstream subsidiary of firm v and firm s have identical costs, so there is no issue related to choosing the more efficient producer of the two. The only aspects related to the welfare considerations of the different policies are whether the total output differs in the two cases, and the importance of the duplication of the fixed costs. When the detrimental welfare effect of fixed costs duplication does not dominate the positive welfare effect of increased production, then the regulator should choose to only regulate access charge and allow the upstream monopolist into the downstream industry.

4.3 Access regulation, structure regulation or no regulation?

What then remains to be examined is which of these policies, or combination of policies, that yields the highest level of overall welfare. We know from Proposition 3 that if the regulator chooses not to regulate access charges, the level of welfare is highest if firm v is free to enter the downstream industry (i.e., when there is no structure regulation). Consequently, we need to compare the welfare level W_v^m (which corresponds to the highest possible welfare level when the access charge is unregulated) to W^d and W_{sr} (which is the welfare level without and with structure regulation, respectively, when access charge is regulated). We know that the level of the fixed cost of entry plays an important role in determining which policy is best when the access charge is regulated. We will need to consider both the case of F large and F small.[19]

Large fixed costs In this case, we know from Proposition 4 that $W^d < W_{sr}$. This implies that if the regulator regulates the access charge, then the best the regulator can do in terms of overall welfare is to combine this regulation with structure regulation. Thus, we need to compare W_{sr} to W_v^m to determine which policy combination that yields the highest level of welfare.

Proposition 5

Assume that firms are symmetric (i.e., $c_v = c_s$). When fixed costs are sufficiently large, a welfare maximising regulator is indifferent between structure regulation combined with marginal cost pricing of access, and no regulation. In addition,

[19] I.e., we must consider the case where $\sqrt{75F/5} > a - c - \beta$ (F large), and the case when the inequality is reversed (F small).

these mixes of policies are strictly preferred to only regulating access charges (i.e., $W_v^m = W_{sr} > W^d$). Furthermore, among the regulatory mechanisms we consider, only the first-best regulation will yield a higher level of welfare.

The proof of the proposition is straightforward and is omitted. The result rests on the fact that marginal cost pricing of access together with $c_v = c_s$ result in the same monopoly output downstream independent of whether the production is undertaken by firm v or s. Since the total output is identical and fixed costs are only paid once in the two situations, the level of welfare will remain unchanged. Consequently, we can conclude that when fixed costs are sufficiently large and the regulator cannot use transfers, the regulator can do no better than to leave the market unregulated.

If, however, the regulator may use subsidies with a shadow cost of public funds equal to zero, then the regulator can achieve the first-best outcome. The first-best situation amounts to producing the competitive output level, and this will be achieved with only firm s producing downstream. This is discussed in detail in a previous section.

Small fixed costs When fixed costs are sufficiently small we know from Proposition 4 that $W^d > W_{sr}$. Thus, if the regulator chooses to regulate access charges, the best he can do is *not* to impose structure regulation. Consequently, we need to compare W^d to W_v^m. The question here is whether access regulation will result in higher welfare compared to the completely unregulated case (where the latter situation will lead to firm v being a monopolist in both markets).

Proposition 6

Assume that firms are symmetric (i.e., $c_v = c_s$). When fixed costs are sufficiently small, a welfare maximising regulator is indifferent between structure regulation combined with marginal cost pricing of access, and no regulation. In addition, these mixes of policies do strictly worse than only regulating access charges (i.e., $W_v^m = W_{sr} < W^d$). Furthermore, among the regulatory mechanisms we consider, only the first-best regulation will yield a higher level of welfare.

Since firms are identical and access is priced at marginal cost, the trade-off in welfare is the same as when comparing access charge regulation to a combined access charge regulation and structure regulation. Thus, in the current situation with low levels of the fixed costs, regulation performs better than no regulation. Furthermore, the regulator should not use policies directed at the structure of the markets.

5 Concluding remarks

The main aim of this note has been to find the socially optimal regulatory policy when the regulator has two main regulatory instruments at its disposal, structure regulation and access charge regulation, in a stylised model.

The analysis suggests that the level of the fixed costs is of great importance in choosing the combination of policies to achieve the highest possible level of welfare, given the imperfections that are assumed to be present (imperfect competition downstream, and in feasibility of giving lump-sum transfers). In this particular model we observe that access charges should be subject to regulation only when fixed costs of establishing firms in the competitive market downstream are sufficiently low; i.e., when entry barriers are (sufficiently) low. It may seem counter-intuitive that the regulator should regulate access charge to ensure a level playing field only when it is easy to enter an industry, whereas when entry is difficult it should make entry even more difficult by not subjecting access charges to regulation. However, the explanation is quite straightforward. Regulation of access charges imply that there will duplication of fixed costs, provided that there are no structure regulation. However, having two firms downstream implies that the total quantity produced is higher, which implies that consumers' surplus is higher. For sufficiently low fixed costs, the social cost of duplication is less than the gain to consumers' surplus. No regulation at all, on the other hand, implies in the case of symmetric firms, complete foreclosure of the rival firm and no duplication of fixed costs. When fixed costs are sufficiently large, the cost of duplication is dominating the gain to consumers' surplus of having more firms in the downstream market.

An issue that may be of some importance when analysing different vertical market structures, but which is not considered in the model presented above, is the issue of economies of scope. It may in certain markets and situations be reasonable to assume that a vertically integrated firm may enjoy an additional cost advantage over a vertically separated firm which is not due to the elimination of double marginalisation. If economies of scope between the upstream and downstream operations are present, it is no longer the case that a regulator is indifferent between not regulating the market at all (in which case the vertically integrated firm will monopolise the downstream market) and a policy which combines marginal cost pricing of access with structure regulation. In such a case, choosing not to regulate the market at all (i.e., neither access charge regulation nor structure regulation) will be preferred by the regulator. In the case of large fixed costs, economies of scope may also modify the ranking between welfare levels with only access charge regulation, and welfare with both access charge regulation and structure regulation. The reason for this is that the cost advantage of the vertically integrated firm may, if sufficiently pronounced, outweigh the cost of duplicating fixed costs. However, it is likely that the overall conclusion that no regulation is best is still valid, since the no regulation scenario (in which the network provider, firm v, will be a vertically integrated monopoly) also will benefit from the cost advantage due to economies of scope. In the case if small fixed costs, economies of scope seems only to strengthen the conclusion that regulating access charges is the welfare optimal solution.

6 References

Economides, N. (1998), ""The incentive for non-price discrimination by an input monopolist", *International Journal of Industrial Organization* **16**, 271-284

Hansen, Bjørn (1999), "Competition and Broadband Upgrades in a Vertically Separated Telecommunications Market", Telenor Research & Development, R 22/99

Kreps, D. and J.A. Scheinkman (1983), "Quantity Precommitment and Bertrand Competition yield Cournot Outcomes", *The Bell Journal of Economics* **14**, 326-337

Laffont, J.J. and J. Tirole (2000), *Competition in telecommunications*, Cambridge, MA: MIT-Press

Lewis, T. and D. Sappington (1999), "Access pricing with unregulated downstream competition", *Information Economics and Policy* **11**, 73-100

Lipsey, R. G. and K. Lancaster (1956), "The General Theory of Second Best", *Review of Economic Studies* **24**, 11-32

Rey, P. and J. Tirole (1997), "A primer on foreclosure", forthcoming in Armstrong, M. and R.H. Porter (eds.) *Handbook of Industrial Organization*, vol. 3, North-Holland

Salop, S. and D. Scheffman (1983), "Raising rivals' costs", *American Economic Review* **73**, 267-271

Sand, J. Y. (2001a), "Regulation and foreclosure", mimeo, University of Tromso, Norway

Sand, J.Y. (2001b), "Market structure and regulation", mimeo, University of Tromso, Norway

Spengler, J. (1950), "Vertical integration and anti-trust policy", *Journal of Political Economy* **58**, 347-352

Tirole, J. (1988), *The theory of industrial organization*, Cambridge, MA: MIT-Press

Vickers, J. (1995), "Competition and regulation in vertically related markets", *Review of Economic Studies* **62**, 1-17

Part 3:
Investment incentives and regulation

ESSAY 6

Demand-side Spillovers and Semi-collusion in the Mobile Communications Market[#]

Øystein Foros[*], Bjørn Hansen[**] and Jan Yngve Sand[***]

Abstract: We analyze roaming policy in the market for mobile telecommunications. Firms undertake investments in network infrastructure to increase geographical coverage, capacity in a given area, or functionality. Prior to investments, roaming policy is determined. We show that under collusion at the investment stage, firms' and a benevolent welfare maximizing regulator's interests coincide, and no regulatory intervention is needed. When investments are undertaken non-cooperatively, firms' and the regulator's interests do not coincide. Contrary to what seems to be the regulator's concern, firms would decide on a higher roaming quality than the regulator. The effects of allowing a virtual operator to enter are also examined. Furthermore, we discuss some implications for competition policy with regard to network infrastructure investment.

Keywords: Mobile communications, roaming, competition, virtual operators

JEL classification: L13, L51, L96

This work has been carried out as part of a research project at Telenor R&D - DN 0401. The authors are grateful to the editor and an anonymous referee for comments. Furthermore we thank Kåre P. Hagen, Lars Sørgard, Lars Mathiesen and Derek Clark, participants at the 19th Annual Conference on the "Competitive challenge in network industries", May 2000, Lake George, New York, EUNIP 2000, December 2000, Tilburg, Netherlands, Helsinki Conference on Anti-trust Issues in Network Industries 2001, and EARIE 2001, Dublin, Ireland, for comments and discussions.

* Norwegian School of Economics and Business Administration and Telenor R&D, Norway. Corresponding address: Norwegian School of Economics and Business Administration, Department of Economics, Hellev. 30, N-5045 Bergen, Norway. E-mail: oystein.foros@nhh.no.

** Telenor R&D, Norway.

*** University of Tromso and Telenor R&D, Norway.

1 Introduction

We analyze competing mobile telephony providers' incentives to invest in, and share infrastructure. Furthermore, we analyze whether the regulator should intervene into the firms sharing agreements, and whether the regulator should allow the firms to coordinate their investments. The infrastructure investment we have in mind is upgrading of the mobile networks from second generation (2G) to third generation (3G) systems. Agreements on sharing infrastructure are called roaming in mobile markets.

The main improvement of 3G networks (e.g. UMTS) compared to the current 2G mobile networks (e.g. GSM) is to increase the speed of communication in the access network and thereby give access to new services and new functionality for existing services. Investments will consequently increase consumers' willingness to pay for mobile access. The basic mechanism driving our results is that investments carried out by one firm increase the value of the service provided by other firms when there is a roaming agreement between the firms.

An analysis of consequences of coordinating investments and infrastructure sharing seems to be more relevant in 3G as compared to 2G networks. First, while the providers of 2G networks (GSM) made their investment non-cooperatively, we now see that several providers of UMTS are coordinating their investments in infrastructure (e.g. in Sweden and Germany). There has been a heated debate whether the regulator should allow the firms to cooperate at the investment stage. Second, the benefit from sharing agreements through roaming seems to be higher in 3G than 2G networks. In current 2G networks the consumers have access to a given capacity (9.6 kbit/s), while in 3G the available capacity for data transmission may be allocated in a more dynamic way. If there are free resources in the network, a consumer may be given a capacity of up to 2 megabit/s. However, as the number of users in a given area at a given time increases, each user will have less capacity available. This may increase the value of infrastructure sharing agreements. There are in general potential gains from sharing network capacity when the load in two networks is not perfectly correlated. Let us illustrate this by a simple example. At a given time, operator A has no free capacity in its network whereas operator B has idle capacity. Suppose now that a subscriber of A tries to download a huge amount of content and needs 2 megabit/s. If there is no capacity sharing agreement between A and B, the customer will not be able to access such services at that time. However, if a sharing agreement is established between the operators, the service will be available to the consumer. In this situation, it is obvious that an investment in capacity by B will increase the willingness to pay for subscriptions from firm A.

We analyze a stylized multi-stage model where the firms first agree on roaming quality, second choose their investment non-cooperatively or cooperatively, and finally compete à la Cournot. We investigate whether it is welfare improving to let the firms semi-collude by choosing their investment cooperatively before they compete in the downstream market. Alternatively we may have semi-collusion where the firms compete at the investment stage and collude in the retail market such as in Brod and Shiwakumar (1999), Fershtman and Gandal (1994), and Steen and Sørgard (1999). More generally the latter form of semi-collusion seems realistic since the firms typically will collude on the most observable variable. Usually this will be the price in the retail market. However, our motivation is that we observe collusion at the investment stage in the market we consider, and the effects of allowing such collusion are what we want to investigate. Therefore, we do not consider collusion in the final product market.[1]

In our basic model we assume that there are two symmetric facility-based firms. If the investments are set non-cooperatively we show that voluntary roaming leads the firms to agree on a too high roaming quality compared to the social optimum. Moreover, the investments are strategic complements, and firms will then invest less with voluntary than mandatory roaming. In contrast, if the investments are set cooperatively, the firms' choices on the roaming quality coincide with the regulator's interests. We show that the firms should be allowed to semi-collude in the way described above, since this yields the highest welfare.

In an extension of the basic model we assume that there is a non-facility-based firm, or a virtual operator, in addition to the two facility-based firms. Whether the virtual operator should be allowed to enter the market and to which extent the presence of

[1] See Busse (2000) and Parker and Röller (1997) for analysis of tacit collusion in the mobile market.

such an operator will affect the incentives to invest in infrastructure has been a hot topic in the industry and amongst regulators. This debate started when the Scandinavian virtual mobile operator Sense Communication attempted to get access to the facility-based mobile operators' networks. The facility-based firms were reluctant to grant Sense Communication access.[2] A much more friendly reception was given to Virgin Mobile in the UK.

The roaming quality between the entrant and the two-facility-based firms in our model is assumed to be weakly lower than the roaming quality between the two facility-based firms. We show that when the investments are set non-cooperatively between the facility-based firms an increase in the roaming quality of the incumbents may now increase the investments. This is in contrast to the basic model.

We analyze a type of semi-collusion where the firms may collude at the investment stage and compete in the retail market. Our model is an extension of the multiple stage models of d'Aspremont and Jacquemin (1988) and Kamien *et al.* (1992) considering R&D investment. Through the roaming agreements the investments in infrastructure give rise to spillover effects similar to those considered in models of strategic R&D investments. In the majority of these models, the externality is exogenous. In our model, we focus on the situation where the level of the externality is endogenously determined.[3] Furthermore, in contrast to the majority of the R&D literature we introduce asymmetry between the firms. That is, there are firms that invest in infrastructure and firms that do not invest.[4]

We also make some other key assumptions in our model. First, for the sake of simplicity we ignore the issue of interconnection (agreements that give access to

---

[2] Sense wanted to issue their own SIM cards, but the Scandinavian facility-based operators refused this. Sense filed a complaint to all national regulators, but only the Norwegian regulator supported it. Telenor's appeal to the Norwegian regulator was still pending when Sense filed for bankruptcy in March 1999 (Matthews, 2000). Now, Sense Communication, along with several other virtual operators, has an agreement with Telenor to resell airtime.

[3] Katsoulacos and Ulph (1998a,b) introduce endogenous levels of spillovers between firms. Contrary to their models we assume that the investments undertaken by firms result in product innovation with probability one. Furthermore, our focus is on a context where firms (or a regulator), in the terminology of Katsoulacos and Ulph, choose the degree of information sharing and not research design. We are accordingly examining investments with firms operating in the same industry, but pursue complementary research.

[4] See e.g. De Bondt (1997) for a survey of the R&D literature of strategic investments.

rivals' customer bases) and focus on roaming only. We give roaming a wider interpretation than pure geographical coverage. Roaming agreements extend availability, such that (i) subscribers can make and receive calls via the infrastructure coverage of a rival operator, (ii) when there is congestion a customer may take advantage of the infrastructure of the rival, and (iii) give access to new functionality/services in the rivals' network.

Second, to simplify we make the assumption that there is no side payments between firms engaged in roaming. If firms have the ability to write complete contracts in all dimensions of the infrastructure sharing (roaming), all external effects from the investment can in principle be internalized through the price mechanism. Then, the problem of spillovers through roaming analyzed in this paper may vanish. Since sharing agreements in the next generation systems should ensure a dynamic capacity allocation, it is however rarely possible to write complete contracts in all dimensions. Thus, even if a price mechanism for roaming is implemented, it will not be able to internalize all external effects. This is similar to what we see in the Internet, where infrastructure sharing of backbones is common (Crémer, Rey and Tirole, 2000). Note that regulation may also constrain the firms' ability to internalize external effects through pricing. In particular, this will be important in the interaction between the facility-based firms and the non-facility based firm (virtual operators).

Third, we make the simplifying assumption that consumers only pay for subscription, not for usage. On the one hand this is evidently a restricting assumption since mobile providers typically employ various types of nonlinear pricing. On the other hand, it is far from evident what the alternative is, and in particular, whether mobile providers will choose to price discriminate between calls originated off-net and calls originated on-net.[5] The focus in our model is however on how availability in various dimensions (capacity, speed etc.) affects the choice of supplier. Consequently, our focus is on

---

[5] In addition to the pricing issues, by introducing a call volume dimension we will have to model the cost structure for calls originating and terminating on the same net by a subscriber of that network. Under such a generalization of our model it would also be natural to relax assumption 2 and introduce a volume price on roaming (as well as a volume price on interconnection). A proper modeling of all these pricing and cost components will lead to a very complex model. By disregarding both the revenue and the cost side of traffic we also avoid the so-called "bill and keep fallacy".

demand for subscriptions, not usage and then it is sufficient to consider pricing as a fixed per period fee.

Fourth, we assume Cournot competition in the retail market. We interpret the quantity firms dump in the retail market as the number of subscriptions they sell. A justification for assuming Cournot competition is that, due to the fact that the amount of radio spectrum available is scarce, there are both physical and technological limits to capacity. Furthermore, the firms must choose a capacity level which is built (or rented) in both the backbone network and the access network (number of base stations) *prior* to the competition in the retail market.[6] This will be more important with 3G systems where the capacity needed increases. However, as shown by Kreps and Scheinkman (1983), strong assumptions are required to ensure that a capacity constrained price game result in identical results as a Cournot game. Nevertheless, this seems more appropriate than assuming a Bertrand game without capacity limits.

The rest of the paper is organized as follows: In section 2, the model with only facility-based firms is presented and analyzed. In section 3, we provide an extension to the basic model by introducing a virtual operator. In section 4, some concluding remarks are made.

2 The model

In our basic model we will look into a duopoly case where we assume the following three-stage game:

Stage 1: Either the firms or the regulator determines roaming quality

Stage 2: The firms determine infrastructure investments either non-cooperatively or cooperatively

Stage 3: Cournot competition

[6] The basic structure in a mobile network is that coverage in a given area is achieved through a number of base stations covering given areas (cells). Hence, a mobile network consists of a net of such cells. The spectrum band allocated for mobile use limits the total bandwith a cell can handle at a given point of time. Thus total capacity measured is limited and one bandwith hungry user occupying 2 Mbit/s is crowding out approximately 200 ordinary voice calls. In situations with capacity problems

There are four different variants of the game depending on the stage 1 and stage 2 strategies:

		Stage 2: Investments	
		Competitive	Collusive
Stage 1: Roaming	Voluntary	**Game 1**	**Game 3**
	Mandatory	**Game 2**	**Game 4**

Figure 1 The four variants of the game

The choice of whether firms cooperate when determining their investment levels will depend on whether such cooperation is approved by the competition authorities. Regarding stage 2 and 3 the structure in our model is fairly similar to Kamien *et al.* (1992) and d'Aspremont and Jacquemin (1988). The generic feature of the investment is that it leads to product innovation increasing the quality of the service.

One interpretation of the timing in our model is that roaming policy may be part of the licenses to the operators in the case where the degree of roaming is mandatory, and therefore chosen *prior* to investments taking place. When roaming is voluntary, we assume that firms can commit to a policy on roaming *prior* to undertaking the investments. Indeed, the timing of the roaming quality decision relative to infrastructure investments can obviously be different. To be more specific, the infrastructure investment may be decided *prior* to a decision on roaming quality. Such timing may involve problems with investment hold-ups, but this will not be our main focus. Thus, in our choice of timing we implicitly assume that the firms/the regulator can credibly commit to a given policy on roaming. As far as the regulator is concerned, the issued licenses may serve as a commitment device, whereas the commitment problem under voluntary roaming is solved, e.g. if a given roaming policy is embedded in the network design (e.g. type of interfaces).[7]

---

it is possible to invest in higher capacity through what is called cell splitting. Cell splitting implies that a given area is served with a higher number of smaller cells.

[7] Poyago-Theotoky (1999) considers a model of R&D where the degree of spillover is endogenous. In her model, the timing of the game is different from ours, in that the R&D investment decision (which is equivalent to our infrastructure investment decision) is made *prior* to the decision on how much information to share with competitors. In addition, she allows firms to choose different levels of

2.1 The demand side

When firm i invests in its infrastructure it impacts on the quality of the services its own customers are offered, but there may also be an impact on the quality of the services offered by the rival firm j, and vice versa. Given a roaming policy β and investment decisions x_i, we can now write the total quality offered to consumers by firm i:

$$a_i = a + x_i + \beta x_j \tag{1}$$

where x_i is network investment undertaken by firm i, and x_j indicates the investment by the rival. We assume that $\beta \in [0,1]$ is a parameter indicating the degree of roaming. This parameter measures the externality effect from sharing infrastructure. If $\beta = 1$, there is an agreement on full roaming, while $\beta = 0$ implies minimum roaming quality.

The inverse demand function faced by firm i:

$$p_i = a_i - q_i - q_j$$

The price, p_i, is the subscription fee for availability (i.e., a monthly fee). The externality introduced above is such that when firm i invests in infrastructure, the marginal willingness to pay for the final products produced by both firms is increasing.

2.2 The supply side

We assume a linear cost function in the final stage for firm i, given by $C_i = cq_i$. The cost c is the direct cost associated with access connection of one user. We assume that firms face quadratic (network infrastructure) investment costs, given by $TC_i(x_i) = \varphi x_i^2 / 2$, where

$\varphi > 4/3$. We will later demonstrate that the restriction on φ ensures a unique and stable equilibrium. Overall profit for firm i is then:

$$\pi_i = (p_i - c)q_i - \varphi x_i^2 / 2 \tag{2}$$

---

spillover. In our model, the degree of spillover (interpreted as roaming quality) is reciprocal, in that the degree of spillover is identical in both directions. When firms choose R&D cooperatively they choose to fully disclose their findings, whereas when there is competition in R&D firms choose minimal disclosure. The latter result is very different from what we find in our model.

2.3 Welfare

We assume that the regulator maximizes welfare given by the sum of producer and consumer surplus:

$$W = CS + \pi_1 + \pi_2 \tag{3}$$

Since firms are symmetric and the inverse demand functions are linear with identical slopes, we can write consumer surplus as $CS = 2q^2$, where q is the symmetric production level of each firm.

2.4 Cournot-competition (stage 3)

At stage 3, firm *i* maximizes the profit function: $\pi_i = (p_i - c_i)q_i$. Combining the first order conditions for the two firms we obtain equilibrium quantities:

$$q^*_i = \frac{a - c + x_i(2-\beta) + x_j(2\beta - 1)}{3}$$

Note that in a symmetric equilibrium ($a_i = a_j$), the equilibrium quantity is given by $q^* = (a_i - c)/3$.[8] This quantity is monotonously increasing in $a_i = a_j$, and this implies that consumer surplus is monotonously increasing in a_i. Firm *i* obtains stage 3 equilibrium profits given by $\pi_i = (q^*_i)^2$.

2.5 Infrastructure investment (stage 2)

When firms invest in infrastructure at stage 2 of the game, they take into account the effect such investments has on the stage 3 equilibrium.

2.5.1 Non-cooperative solution

At stage 2, the firms maximize the profit function (2), subject to Cournot equilibrium quantities at stage 3, which implies that the (symmetric) equilibrium investment is given by:

[8] As it turns out, the unique equilibrium in investment is indeed the symmetric equilibrium.

$$x^*_{nc} = \frac{(4-2\beta)(a-c)}{9\varphi - 2(2-\beta)(1+\beta)} \tag{4}$$

In equilibrium, firms' profits are non-negative for all permissible parameter values, and consequently firms will participate in the game. Furthermore, the symmetric equilibrium is the unique equilibrium.[9] Our result is analogous to d'Aspremont and Jacquemin (1988) and Kamien *et al.* (1992), and is summarized in the following lemma:

Lemma 1 *If* $\beta > 1/2$*, then* x_i *and* x_j *are strategic complements; that is,* $(\partial x_i/\partial x_j)^{nc} > 0$. *Reversing the inequality makes* x_i *and* x_j *strategic substitutes.*

When one firm invests in its infrastructure, the equilibrium quantity of the other firm may or may not increase as a result of the spillover through the roaming agreement. As pointed out by Kamien *et al.* (1992), the spillover externality (or investment externality) is positive if and only if $\beta > 1/2$. In other words, the spillover externality is positive only when x_i and x_j are strategic complements.

2.5.2 Cooperative solution

In the cooperative case the two firms coordinate their infrastructure investments at stage 2, and compete à la Cournot at stage 3 in the same way as above. When determining the profit-maximizing choice of investments at stage 2, firm *i* maximizes the joint profit of the two firms (i.e. industry profits):

$$\max_{x_i} \Pi = (q_i)^2 + (q_j)^2 - \frac{\varphi}{2}(x_i^2 + x_j^2) \quad \text{for } i \neq j$$

The following first-order condition yields the equilibrium investment for a given firm under collusion:[10]

$$x^*_c = \frac{2(a-c)(1+\beta)}{9\varphi - 2(1+\beta)^2} \tag{5}$$

---

[9] The second-order condition requires that $2(2-\beta)^2 - 9\varphi < 0$. In order to have a stable equilibrium the slope of the reaction function has to have an absolute value below unity. It is straightforward to demonstrate that this condition is fulfilled for $\varphi > 4/3$. Hence, the second order condition is also fulfilled when we have assumed that $\varphi > 4/3$. Finally, it can be shown that the symmetric equilibrium is indeed unique.

[10] In the same way as in game 1, the second-order condition, $2(1+\beta)^2 - 9\varphi < 0$, is satisfied with our parameter restrictions.

The cooperative solution yields lower infrastructure investment than the non-cooperative, $x_c^* < x_{nc}^*$, if and only if $\beta < 1/2$. This is equivalent to the results obtained by d'Aspremont and Jacquemin (1988) and Kamien *et al.* (1992).

Observe furthermore that for $\beta < 1/2$, equilibrium quantity in the final stage of the game is lower under collusion relative to the non-cooperative case. Consequently, consumer surplus is, on the one hand, lower under collusion relative to non-cooperation if the roaming quality is sufficiently low. If, on the other hand, $\beta > 1/2$, consumers' surplus is higher under collusion at the investment stage. In addition, firms' profits under collusion are always at least as large as under non-cooperation. For $\beta = 1/2$ firms are indifferent between collusion and non-cooperation at stage 2. Consequently, a welfare-maximizing regulator would, provided that the roaming quality is sufficiently high (i.e. $\beta > 1/2$), choose to allow collusion at stage 2 and such collusion would be in the firms' interests. In the next section we examine which level of roaming quality a welfare-maximizing regulator would choose.

2.6 Roaming quality (stage 1)

In this section, we extend the model by considering the two cases where: 1) Firms decide roaming quality (*voluntary roaming*) and 2) the regulator decides roaming quality (*mandatory roaming*). These two cases combined with the two ways of determining investments at stage 2 yields four different games (see Figure 1). The analysis in this section can be seen as an extension to the basic model with exogenous R&D spillovers to examine endogenous spillovers.

2.6.1 Game 1, voluntary roaming when investments are determined non-cooperatively

Recall that equilibrium infrastructure investment (under non-cooperation) is given by equation (4). Direct differentiation of (4) yields the following result:

Lemma 2 *When firms determine the investments non-cooperatively, the infrastructure investment decreases as the roaming quality increases since* $\partial x_{nc}^* / \partial \beta < 0$.

The intuition behind Lemma 2 is as follows: We are considering the equilibrium with *reciprocal* spillover levels, implying that both firms' final products increase in value to consumers by the same proportion (contrary to what is the case in section 3). Hence, there is no product differentiation gain from investments for any of the firms. When firms do not cooperate at the investment stage, they do not internalize the effect of the investment on the other firm's profit. An increase in the degree of roaming quality is thus affecting the marginal revenue from investing in infrastructure adversely, due to the fact that the investing firm is unable to capture the effect on the rival's profit. When examining the stage 3 equilibrium we observe that equilibrium quantities increase in both the degree of roaming quality and the investment level. Since quantities are strategic substitutes, the rival firm will increase its production if you reduce yours. Each firm will then have to be more cautious and ration its production of the final product more than is the case if they collude in the investment stage. To achieve a substantial enough reduction, any given firm will have to limit its production even more. Consequently, since q_i increases when β increases and $(q_i + q_j)$ increases when x_i increases, the investing firms can ration the final product market by reducing the level of investments when the degree of roaming quality increases.

The competitive equilibrium infrastructure investment is given by equation (4); $x_{nc}^{*} = x_{nc}(\beta)$. From Lemma 2 we know that $x_{nc}'(\beta) < 0$. Using the symmetry of the problem the equilibrium profit is accordingly:

$$\pi = \left(\frac{a + (1+\beta)x_{nc}(\beta) - c}{3} \right)^2 - \varphi \frac{(x_{nc}(\beta))^2}{2} \tag{7}$$

By differentiating equilibrium profit with respect to β, we find the roaming quality preferred by the firms.

$$\frac{2}{9}\left(a - c + (1+\beta)x_{nc}(\beta)\right)\left(x_{nc}(\beta) + (1+\beta)x_{nc}'(\beta)\right) - \varphi x_{nc}'(\beta) = 0 \tag{8}$$

It can be shown that the optimal β for the firms is independent of $(a\text{-}c)$. The solution to equation (8) will be a function of the convexity of the investment cost function, i.e. of the value for φ. The profit maximizing choice with respect to β is increasing and concave in φ and strictly larger than 9/10. As an example; when $\varphi = 3/2$ the expression given by equation (8) is concave over the interval where it is defined and the first order condition is satisfied for $\beta = 0.941$. If the firms determine roaming quality, the firms will accordingly agree on $\beta = 0.941$ as the preferred level of roaming.

However, there may be a commitment problem for the firms. After firms have chosen their level of expenditure on network infrastructure (at stage 2), both firms have an incentive to renegotiate the roaming quality between stages 2 and 3. The reason for this is that for a given level of x_i, the (stage 3) Cournot-equilibrium profit of both firms is strictly increasing in β. If firms cannot commit to the roaming quality chosen at stage 1 of the game, we may thus experience hold-up problems in network infrastructure investments.[11] As indicated earlier, this commitment problem is solved if roaming policy is embedded in the network design, such that the firms cannot change roaming policy after the investments have been made. Furthermore, in section 3, we demonstrate briefly that the introduction of a virtual operator can eliminate the potential commitment problem for the firms.

2.6.2 Game 2, mandatory roaming when investments are determined non-cooperatively

In this section we investigate the roaming quality a welfare-maximizing regulator would choose, which can either be considered as a benchmark case or as the chosen roaming quality under *mandatory roaming*. We assume that the regulator maximizes the objective function given by equation (3). Producer surplus is calculated above (equation 7). Consider now the roaming quality preferred by consumers. We know that the equilibrium is symmetric in the sense that the two firms invest in the same level of infrastructure and that they offer the same quantity at stage 3 of the game. Let x^* and q^* denote the profit-maximizing choices of investment and quantity, respectively. In the equilibrium, we have the following equilibrium price:

$$p = a + (1+\beta)x^* - 2q^*$$

Inserting for the non-cooperative equilibrium investment given by equation (4), and $q^* = \left((a-c)+(1+\beta)x_{nc}^*\right)/3$ the consumer surplus becomes:

---

[11] Since the firms' profit is increasing in roaming quality for a given level of investment, firms have incentives to increase β to its maximum after the investments are sunk. The hold-up problem arises because we know that when investments are made non-cooperatively, the marginal revenue of the investments is reduced when β increases. Thus, if firms know that, ultimately, roaming quality is chosen to give perfect roaming, they will hold back on investments.

$$CS_{nc} = \frac{2\left[(a-c)+(1+\beta)x_{nc}(\beta)\right]^2}{9} \tag{9}$$

By differentiating consumer surplus given by equation (9) with respect to β, we find that the first order condition is satisfied for roaming quality $\beta = 0.5$.[12] Consequently, a roaming quality determined at $\beta = 0.5$ is the consumer surplus maximizing quality level.[13]

The roaming level maximizing consumer surplus can be compared to the roaming quality maximizing profits for the firms, (equation 8). In section 2.6.1 we demonstrated that the roaming level maximizing producer surplus is strictly larger than 9/10. It is accordingly evident that the firms prefer a roaming level exceeding the roaming level preferred by consumers. Since welfare is the sum of consumer and producer surplus, the welfare maximizing roaming level (game 2) is below the level preferred by the firms (game 1). This result is summarized in Proposition 1:

Proposition 1 *Assume that investments are undertaken non-cooperatively. Voluntary roaming induces firms to choose a higher level of roaming quality relative to the social optimum. Consequently, in a voluntary roaming regime firms invest less in network infrastructure than in a mandatory regime.*

The intuition behind this result is as follows: For a given quantity of the final product, the consumers benefit from high levels of infrastructure investments. Firms also benefit from high levels of investments, *ceteris paribus*. However, the investing firms cannot capture all of the benefits from the investments (due to the spillover externality through the roaming agreement) and, in addition, firms face convex costs of investing. Consequently, firms will choose a lower level of investment compared to the level that maximizes consumers' surplus. Furthermore, from Lemma 2 we know that incentives to invest in infrastructure are worse the higher the degree of roaming quality. In a situation with high roaming quality, each of the investing firms will attempt to be a free rider on the other firm's investments. Consequently, the equilibrium investment level will be lower. On the other hand, a welfare-maximizing regulator will be able to internalize all the externalities of roaming. Thus, when roaming is voluntary firms choose a high level of roaming quality to reduce

---

[12] The solution that $\beta = 0.5$ maximizes consumers' surplus is also independent of the convexity of the investment cost function, i.e. of the value for φ.

[13] The second-order condition for CS is satisfied for all permissible values of β.

investment incentives. When roaming is mandatory, a lower level of roaming quality is chosen to induce higher levels of investment in network infrastructure.

If firms are not allowed to collude at the investment stage, regulatory intervention may be needed. In this case, firms would set a higher roaming quality than the socially optimal level. Consequently, firms would agree on more compatibility than what would be beneficial to consumers and society as a whole. For consumers, an increase in roaming quality has two potentially opposing effects. Increased roaming quality implies (under non-cooperation) that investment is reduced. This implies that the size of the market is reduced (the inverse demand function is shifted inwards). For a given quantity, this results in a reduction in the price charged to consumers. This has a positive impact on consumers' welfare. However, a reduction in infrastructure results in a reduction in quantities sold in the last stage of the game. This has a negative impact on consumers' welfare. The overall effect of increasing quality is positive for low levels of roaming quality, and negative for high levels. For a sufficiently high roaming quality, an increase in roaming quality reduces consumers' welfare. Firms, on the other hand, can by setting a high roaming quality restrict the size of the infrastructure investments. As a result, the output of the final product in stage 3 of the game is also restricted, with an increase in equilibrium price.

2.6.3 Game 3 and 4, voluntary and mandatory roaming when investments are determined cooperatively

Recall that the equilibrium investment as a function of roaming quality is given by equation (5). Direct differentiation yields the following result:

Lemma 3 *When firms collude at the investment stage, the infrastructure investment increases as the roaming quality increases since* $\partial x_c^* / \partial \beta > 0$.

Contrary to the case of non-cooperative investment, the incentives to invest are in fact higher when the roaming quality is high. In this case, firms achieve a better coordination due to the fact that the effect on the other firm's profits is internalized. Thus, all benefits from the investments are credited to the investing firm, which changes investment incentives qualitatively. There is no longer the problem that each of the investing firms will have incentives to free ride on the other firm's investments. If firms determine a high (low) level of roaming quality, each firm's investment presents a large (minor) positive external effect the firms jointly can internalize. By inserting for the collusive equilibrium investment, x_c^* (given by equation 5), and third stage equilibrium quantities, we obtain consumer surplus (under collusion), *CS_c*:

$$CS_c = 2\left(\frac{a-c}{3}+\frac{2(a-c)(1+\beta)^2}{3\left(9\varphi-2(1+\beta)^2\right)}\right)^2 \tag{10}$$

When firms decide on infrastructure investments collusively, consumer surplus given by equation (10) is increasing and convex in roaming quality over the relevant interval for β. Consequently, the optimal β for consumers is equal to unity (or maximal roaming quality). Quantity is increasing in roaming quality both directly and through the effect of roaming quality on investment. Furthermore, consumer surplus is increasing in quantity. Thus, a higher roaming quality implies a higher level of consumers' surplus.

Firms maximize their equilibrium profit under collusion with respect to β, which results in:

$$\frac{\partial\pi}{\partial\beta} = 2q_c^*\left(\frac{x_c^*}{3}+\left(\frac{1+\beta}{3}\right)\frac{\partial x_c^*}{\partial\beta}\right)-\varphi x_c^*\frac{\partial x_c^*}{\partial\beta} \tag{11}$$

where q_c^* is the stage 3 equilibrium quantity for each firm if firms collude at the investment stage. In the collusion case, the profit function is increasing and convex in β over the interval [0,1]. This implies that the optimal roaming quality for firms corresponds to the maximal roaming quality. We summarize our findings in Proposition 2:

Proposition 2 *Assume that firms collude at stage 2. The (unregulated) profit maximizing choice of roaming quality is identical to the choice of a regulator maximizing welfare. Consequently, the level of investments in infrastructure is identical in both the voluntary and mandatory roaming regimes.*

Since firms' and consumers' interests coincide there is no reason for governmental intervention and there is no need for considering game 3 and 4 separately. As in the non-cooperative case, consumers benefit from high investments in infrastructure for a given production level. However, to achieve high levels of investments, consumers now choose a *high* level of roaming quality. The main reason for the difference in our result is that firms, when they are allowed to coordinate their investments, are able to capture all benefits from the investments. Because of this, firms' investment incentives are changed and they now seek high roaming quality to induce high levels of investments, whereas in the non-cooperative case they seek high roaming quality to induce low levels of investments.

2.7 Collusive and competitive investments compared

At stage 2 of the game the level of investment in the network infrastructure is determined. As already stated, firms can either compete (games 1 and 2) or they can collude when determining the investment level (games 3 and 4). We may think of the decision to allow firms to collude or not as a decision taken by the competition authorities *prior* to commencement of the 3-stage game we analyze above.

Assuming that roaming is voluntary we can compare equilibrium under collusive and competitive investments, respectively (i.e. we compare games 1 and 3), to determine whether the investing firms should be allowed to collude or not at the investment stage. Under voluntary roaming the firms set the β so as to maximize profits. Firms are evidently better off under collusion as compared to competitive investments.[14] Furthermore, it follows from the calculations above that equilibrium consumer surplus is higher under collusion (game 3 and 4) as compared to equilibrium consumer surplus under competitive investments and mandatory roaming (game 1).[15] The intuition behind this result is as follows: Consumers' surplus increases, both in roaming quality for given investments and in investment for a given roaming quality. Since we have demonstrated that roaming is at the highest possible level under collusive investments and that investments are higher under collusion provided that $\beta > 0.5$, consumer surplus will indeed be higher under collusion.

---

[14] When firms collude they can always mimic the outcome under competition, if they choose to deviate from this outcome it is because they are better off.

[15] This result is derived by first observing that consumer surplus (CS) is higher in game 2 as compared to game 1. Then a sufficient condition for demonstrating that CS is higher in game 3 and 4 as compared to game 2 is to assume that the regulator in game 2 determine the roaming at the level maximizing CS ($\beta = 0.5$). Then we can compare CS under collusion, equation (10), for $\beta = 1$ and compare it CS under competitive investments, equation (9), for $\beta = 0.5$. Then we find that $CS_{game\ 3\ and\ 4} \geq CS_{game\ 2} \geq CS_{game\ 1}$.

In the figure below we illustrate the welfare effects of chosen roaming policy for a regulator under the two investment regimes for the parameter values $a = 3$, $c = 1$, and $\varphi = 3/2 > 4/3$:

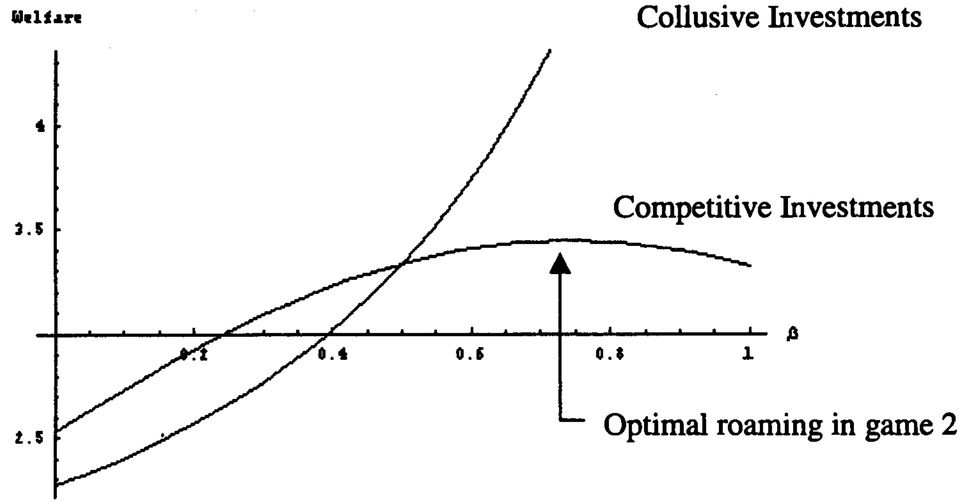


Figure 2 Welfare comparison of collusion versus non-cooperation

It is evident from the figure above that social welfare is maximized when the regulator allows firms to collude at stage 2, provided that the quality of roaming is sufficiently high (i.e. for β higher than 0.5). We know from the analysis above that a welfare maximizing regulator indeed will set the quality of roaming equal to unity when firms collude at the investment stage. The figure is also illustrating the, perhaps, counter intuitive results that it is detrimental to welfare to determine mandatory roaming at too high a level under competitive investments.

Consequently, the competition authorities (or a regulator) can never do worse than allowing collusion at the investment stage, provided that the firms/regulator can commit to a sufficiently high roaming quality. The reason is that allowing collusion allows the positive external effects to be internalized.

3 Entry of a non-facility-based firm

In this section we analyze a situation where there are two different types of firms. One type is *facility-based* and invests in its own infrastructure. We assume that there are two incumbents which both are facility-based as in the previous section. The other type is an entrant that is a *virtual operator*. The timing of the game is the same as in the above analysis, but it is of course only the facility-based firms which undertake investments at stage 2 and we assume that the facility-based firms have all the bargaining power in determining the quality of roaming for the virtual operator.

The model is amended to incorporate the fact that there are now two types of firms. In the case where the entrant is a virtual operator we have:

$$a_i = a + x_i + \beta^f x_j$$
$$a_v = a + \beta^v (x_i + x_j),$$

where subscript $i, j = 1,2,\ i \neq j$ represents the two incumbents, and subscript v represents the virtual operator.[16] The parameter β^f represents the degree of the roaming quality between the facility-based firms, while β^v is the roaming quality from a facility-based operator to a virtual operator. We restrict the analysis to the case where $\beta^f \geq \beta^v$.[17] The parameters β^v and β^f can also be interpreted as the virtual operator's and facility-based operator's capabilities of transforming the inputs into a final product of high quality, respectively.

The profit functions of the facilities firms are equal to equation (2), while the profit of the virtual operator is $\pi_v = (p_v - c)q_v$. By combining the stage 3 first order conditions for the three firms we obtain the following equilibrium quantities:

$$q_i^* = [(a-c) + x_i(3 - \beta^f - \beta^v) + x_j(3\beta^f - \beta^v - 1)]/4$$
$$q_v^* = [(a-c) + (x_i + x_j)(3\beta^v - \beta^f - 1)]/4$$

---

[16] Ceccagnoli (1999) gives a similar formulation with process innovation. In contrast to us he only focuses on the case where the R&D-investment is set non-cooperatively.

[17] This seems to be an appropriate assumption if the virtual operator is simply a reseller of airtime (e.g. Sense Communication in Scandinavia). However, if the virtual operator has a well known brand name and possesses detailed knowledge about certain segments of the market (e.g. Virgin in the UK), it may be reasonable to assume that investments benefit the virtual operator more than they benefit the incumbents.

3.1 Incumbents set the investments non-cooperatively

When the two facility-based firms set their investment non-cooperatively, we use the symmetry between the two firms, and we find the equilibrium investment for each facility-based firm: [18]

$$x = \frac{(a-c)(3-\beta^f-\beta^v)}{8\varphi-2(3-\beta^f-\beta^v)(1+\beta^f-\beta^v)} \tag{12}$$

We now examine how the equilibrium investment level x changes when β^f and β^v change:

Proposition 3 *When introducing a virtual operator and the investments are made non-cooperatively by the facility-based firms, increased roaming quality between the investing firms may result in a higher equilibrium investment level; i.e.* $\partial x / \partial\beta^f > 0$ *if and only if* $(\beta^f+\beta^v) < 2\sqrt{\varphi}-3$. *Furthermore, increased roaming quality to the virtual operator reduces investment incentives,* $\partial x / \partial\beta^v < 0$.

Proof:

$$\frac{\partial x}{\partial\beta^f} = \frac{-(a-c)[8\varphi-2(3-\beta^f-\beta^v)^2]}{[8\varphi-2(3-\beta^f-\beta^v)(1+\beta^f-\beta^v)]^2}$$

$$\frac{\partial x}{\partial\beta^v} = \frac{-(a-c)[8\varphi+2(3-\beta^f-\beta^v)^2]}{[8\varphi-2(3-\beta^f-\beta^v)(1+\beta^f-\beta^v)]^2} < 0$$

The first expression is positive if and only if $[8\varphi-2(3-\beta^f-\beta^v)^2] < 0$. This implies that $\beta^f+\beta^v < -3 \pm 2\sqrt{\varphi}$. Since $\beta^f+\beta^v \in [0,2]$ the only root possibly satisfying the inequality is $(\beta^f+\beta^v) < 2\sqrt{\varphi}-3$ .QED.

Recall the basic model with only two facility-based firms. Then, if investments were undertaken non-cooperatively, the incentives to invest were lower the higher the degree of roaming quality was. For sufficiently high values of the roaming qualities we obtain a similar result in the presence of a virtual operator. We observe from Proposition 3 that, contrary to the basic model, the incentives to invest may in fact be improved the higher the roaming quality is between the facility-based firms. This will

[18] The second order condition is fulfilled if $\varphi > 9/8$, which also ensures stability.

be the case given a combination of a small sum of $(\beta^f + \beta^v)$ and sufficiently large φ.

In stage 1 of the game, either the investing firms or the regulator choose the level of roaming quality. Our findings suggest that if roaming is *voluntary* and investments are made non-cooperatively, the investing firms will choose to set the roaming quality between the investing firms as high as possible ($\beta^f = 1$). The roaming quality between an investing firm and the virtual operator is set as low as possible ($\beta^v = 0$).

It may however be the case that profit for the virtual operator is negative in this solution. A sufficient condition for ensuring non-negative profits for the virtual operator is that the convexity of the investment cost function is sufficiently large (i.e. for φ sufficiently large). It can be shown that $\varphi > 2$ ensures non-negative profits for the virtual operator under competitive investments. The intuition behind this result is that when $(3\beta^v - \beta^f - 1) < 0$ any investments in infrastructure undertaken by the facility-based firms will reduce the equilibrium output of the virtual operator, and consequently, infrastructure investments may be used to deter entry. If, however, the cost of investing is sufficiently convex, then the equilibrium level of investment is low enough not to deter entry by the virtual operator.

We only consider the case where the facility-based firms choose accommodation of entry by the virtual operator. In order to ensure this we assume that the costs are sufficiently convex (high value of φ). As long as we have accommodation of entry proposition 3 holds. However, for lower values of φ, it may be optimal for the facility-based firms to invest such that the virtual operator is foreclosed from the market. In this case, the higher the β^v compared to β^f, the more the facility-based firms have to invest in order to deter entry.[19]

When roaming is *mandatory*, a welfare-maximizing regulator decides the appropriate levels. The regulator will choose a roaming policy that corresponds to the voluntary roaming case. This seems to correspond well to intuition. By keeping β^v low and β^f

---

[19] In a similar context, the case where a facility-based firm over-invests to deter entry is analyzed by Foros (2000).

high, firms have better incentives to invest in infrastructure. Furthermore, a higher level of investment leads to more output being produced in the final product market, which is a direct benefit to consumers. Consequently, there is little scope for regulatory intervention in this case. The main result in this section then suggests that the provision of the right investment incentives is of greater importance than allowing the virtual operator entry at equal terms to the incumbents.

3.2 Incumbents set the investments cooperatively

When the facility-based firms set investments cooperatively they maximize the joint profit, and by inserting for optimal stage 3 quantity we obtain the following equilibrium investment level: [20]

$$x = \frac{(a-c)(1+\beta^f-\beta^v)}{4\varphi-2(1+\beta^f-\beta^v)^2} \tag{13}$$

Hence we have the following results:

Proposition 4 *When facility-based firms invest in infrastructure cooperatively, the investment level is unambiguously increasing in the degree of roaming quality between the cooperating firms,* $\partial x / \partial \beta^f > 0$, *and decreasing in the roaming quality between facility-based firm and the virtual operator,* $\partial x / \partial \beta^v < 0$.

The proof is straightforward and hence omitted.

Similar to our findings when investment is undertaken non-cooperatively, we observe that the equilibrium investment level is reduced whenever the roaming quality between the investing firms and the virtual operator increases. The intuition is the same as in the non-cooperative case. One important reason for investing in infrastructure is to differentiate its product from that of the competitor. If the roaming quality between the investing firms and the virtual operator is sufficiently high, the importance of the differentiation effect diminishes. As mentioned above, the degree of roaming quality may be interpreted either as the quality of the access provided to the virtual operator or as the virtual operator's capability of transforming the product innovation resulting from the investment made by the facility-based firms into high

[20] Note that for the equilibrium investment level to be positive, the investment cost needs to be sufficiently convex; this requires $\varphi > 2$.

quality final products. Consequently, our findings suggest that any increase in the quality of input or in the virtual operator's capabilities reduces the facility-based firms' incentives to invest in infrastructure.

Our results for both voluntary and mandatory roaming when investments are undertaken cooperatively suggest that there is little scope for regulation. This is also the case when investments are undertaken non-cooperatively. Both the investing firms' interests and the interests of a welfare-maximizing regulator coincide. When the facility-based firms cooperate at the investment stage, they maximize joint profit with respect to (β^f, β^v), whereas a welfare maximizing regulator chooses (β^f, β^v) to maximize the sum of profits and consumers' surplus. The solution in both cases yields maximum roaming quality between the facility-based firms and minimal roaming quality between facility-based firms and the virtual operator. In order to ensure nonnegative profits for the virtual operator when $\beta^v = 0$ we must however make further restrictions on the convexity of the investment cost function. A sufficient condition is that $\varphi > 4$.

For a given investment level, an increase in the degree of roaming quality to the virtual operator is a social benefit and adds to consumers' surplus. However, increasing the roaming quality to the virtual operator adversely affects the facility-based firms' incentives to invest in infrastructure, and higher levels of investments is also a benefit to consumers. When a virtual operator is allowed to enter and investments are undertaken collusively the trade-off is the same as when investments are undertaken non-cooperatively, and the investment incentives dominate.

3.3 Some remarks on entry of a virtual operator

One remaining question is whether the competition authorities (or a regulator) should allow a virtual operator to enter or not. It is reasonable to assume that such entry should be encouraged if the entry implies that welfare is higher than is the case in the absence of a virtual operator. We assume that the decisions to allow entry by a virtual operator and whether to allow collusion are taken *prior* to commencement of the three-stage game analyzed above. Without going into details, it can be shown that welfare is higher if collusion at the investment stage is allowed (as is also the case without a virtual operator), and the reason is again that the positive externality can be internalized more easily under collusion.

Furthermore, it can also be shown that welfare is indeed higher when entry of a virtual operator is allowed when collusion is allowed. Since we know that welfare under collusion is higher than in the non-cooperative case, we need to show that welfare under collusive investments *with* a virtual operator is higher than without the virtual operator. Then, we have essentially proven that the subgame-perfect equilibrium policy for the government is to allow entry by a virtual operator and allow facility-based firms to cooperate at the investment stage.[21] Furthermore, we have seen that the optimal roaming policy implies that the virtual operator is only given minimal roaming quality, which means that the entry for the virtual operator is not on particularly generous terms.

A final point to be made is that the introduction of a virtual operator can eliminate the commitment problem for firms. Note that firms choose the roaming quality *prior* to undertaking investments, and in the absence of a virtual operator firms may have incentives to change the quality of roaming after investments have been sunk. When a virtual operator is allowed to enter, this commitment problem is eliminated, and firms have no incentives to change the quality of roaming after investments are sunk.

4 Concluding remarks

We have discussed roaming policy (both voluntary and mandatory), and we have also briefly discussed some competition policy aspects related to sharing infrastructure in the mobile communications market. In particular, we have focused on the interaction between roaming policy and investment incentives in the third generation mobile networks (e.g. UMTS). We have shown that all involved are better off under collusion provided that roaming quality is set sufficiently high. Furthermore, in our model, the chosen level of roaming quality is indeed sufficiently high in all cases. This implies that when a regulator or competition authority chooses whether collusion at the investment stage should be allowed, they know that in whatever policy they choose with respect to roaming, they can never do worse than allowing collusion. In some of the Nordic countries, major players in the mobile communications market have

[21] The proof of this result involves messy, but straightforward algebraic manipulations.

decided to cooperate in the process of setting up the next generation mobile networks, which corresponds to the collusion case in our model. In danger of stretching our model a little too far, we have shown that such collusion is actually beneficial in terms of welfare. Consequently, competition authorities should not interfere with such cooperation.

When we introduce a virtual operator into the game relying on the facility-based firms for infrastructure access, we find that the relationship between roaming quality and investment incentives is qualitatively different. Furthermore, our findings suggest that there is little scope for regulation of roaming quality when there is a virtual operator present, both under cooperative and non-cooperative investments. This is also different from the case without the virtual operator present, where the social optimum does not correspond to the unregulated outcome if investments are undertaken non-cooperatively.

References

Brod, A. and Shivakumar, R., "Advantages of Semi-Collusion", The Journal of Industrial Economics, vol. 47, pp. 221-230. 1999.

Busse, M.,"Multimarket contact and price coordination in the cellular telephony Industry", Journal of Economics & Management Strategy, vol. 9, pp. 287-320. 2000.

Cave, M. and Valletti, T., "Competition in the UK mobile communications", Telecommunication Policy, vol. 22, pp. 109-131.1998.

Ceccagnoli, M. "The Effects of Spillovers and Rivalry among Firms with Heterogeneous Technological Capabilities on R&D Efforts", Mimeo, 1999.

Crémer, J., Rey; P. and Tirole; J., "Connectivity in the commercial internet", The Journal of Industrial Economics, vol. 48, pp. 433-472. 2000.

d'Aspremont, C. and Jacquemin, A., "Cooperative and Noncooperative R&D in Duopoly with Spillovers", American Economic Review, vol. 78, pp. 1133-1137.1988.

De Bondt, R., "Spillovers and innovative activities", International Journal of Industrial Organization, vol. 15, pp.1-28. 1997.

Fershtman, C. and Gandal, N., "Disadvantageous Semicollusion", International Journal of Industrial Organization, vol. 12(2), pp. 141-154. 1994.

Foros, Ø. "Strategic Investments with Spillovers, Vertical Integration and Foreclosure in the Broadband Access Market". Manuscript.

Kamien, M., Muller, E. and Zang, I., "Research Joint Ventures and R&D Cartels", American Economic Review, vol. 82, pp. 1293-1306. 1992.

Katsoulacos, Y. and Ulph, D. "Endogenous spillovers and the performance of research joint ventures", The Journal of Industrial Economics, vol. 46, pp. 333-357.1998a.

Katsoulacos, Y. and Ulph, D., "Innovation spillovers and technology policy", Annales d'Economie et de Statistique, pp. 589-607. 1998b.

Katz, M. "An analysis of cooperative research and development", Rand Journal of Economics, vol. 17, pp. 527-543. 1986.

Kreps, D.M. and Scheinkman, J.A., "Quantity Precommitment and Bertrand

Yield Cournot Outcomes", Bell Journal of Economics, vol. 14(2), pp. 326-337. 1983.

Leahy, D. and Neary. J.P., "Public policy towards R&D in oligopolistic Industries", American Economic Review, vol. 87, pp. 642-662. 1997.

Matthews, J. "Virtual Mobile Services: Strategies for fixed and mobile Operators". London, Ovum (Ovum reports). 2000.

Parker P.M. and Röller, L.H., "Collusive conduct in duopolies: mulitmarket contact and cross-ownership in the mobile telephone industry", Rand Journal of Economics, vol. 28, pp. 304-322. 1997.

Poyago-Theotoky, J. "A note on endogenous spillovers in a non-tournament R&D duopoly", Review of Industrial Organization, vol. 15, pp. 253-262. 1999.

Salant, S. and Schaffer, G., "Optimal asymmetric strategies in research joint-Ventures", International Journal of Industrial Organization, vol. 16, pp.195-208.1998.

Steen, F. and Sørgard, L., "Semicollusion in the Norwegian Cement Market", European Economic Review, vol. 43(9), pp. 1775-1796. 1999.