Competition in wireless telecommunications

Miguel Angel Campo-Rembado and Arun Sundararajan[1]

Leonard N. Stern School of Business, New York University
44 West 4th Street, New York, NY 10012
(mcampo, asundara) @stern.nyu.edu
http://oz.stern.nyu.edu

May 2004

Abstract: We develop a model of competition in the wireless industry that explicitly characterizes the interdependence between network traffic, spectrum availability, infrastructure deployment, the generation of transmission technology, and service quality. We show that when spectrum and infrastructure are not abundant, service quality is endogenously affected by market share, and this externality leads to higher pricing power for providers. We then incorporate the effects of usage externalities and minimum infrastructure requirements in a two-stage vertical differentiation game, and establish that it has two distinct kinds of equilibria. Under the first, providers deploy the minimum possible network infrastructure, price symmetrically and split the market. We show that this equilibrium occurs at both very low and very high levels of average network traffic. For intermediate levels, the equilibria are asymmetric in both network deployment and pricing, though the extent of vertically differentiation is mediated by the presence of usage externalities. Our model explains why the average revenue per user (ARPU) that competing wireless firms get may flatten and decrease even as average per-user traffic and service value increase, and why a lower-quality provider may achieve higher equilibrium profits over a non-trivial range of situations. Our equilibrium strategies suggest relatively similar quality levels early in a market's evolution, when higher prices are maintained by an externality premium, followed by more aggressive quality differentiation as the market matures and traffic levels rise. We also identify a threshold level of demand at which equilibrium profits of both competing firms bilaterally approach zero, which we interpret as a natural industry migration point to the next generation of transmission technology. We conclude by discussing the policy implications of our results, and outlining current research.

JEL Codes:

1. Introduction

This paper develops a two-stage model of vertically differentiated duopoly in the wireless telecommunications industry. The objective is to present a model that captures distinguishing aspects of supply and demand in this industry, thereby explaining observed trends more accurately, and also providing guidelines for the timing of migration between successive generations of technology. The model may also generalize to other future telecommunications and networking industries that share similar characteristics. In the process, we also provide a generalization of the standard approach to vertically differentiated competition which incorporates usage externalities and minimum quality constraints.

Service quality in wireless telecommunications, which is related primarily to the network quality of competing providers, is an important basis for consumer choice. However, industry surveys suggest that there is not much variation in the quality levels actually chosen by different providers. Correspondingly, the prices chosen by different firms tend to be symmetric – each major U.S. provider has a menu of pricing plans which start with a minimum price of about thirty dollars, and measured price per minute-of-usage from 1997 to 2002 showed only small variations between the major U.S. providers. In contrast, standard vertical differentiation models (beginning with Shaked and Sutton, 1982) posit that since an increase in similarity in product quality increases the threat of price competition, equilibria always feature substantially different prices and quality levels. Additionally, equilibrium prices predicted by these models should increase when there is a uniform increase in product value across customers. This conflicts with observed behavior in the wireless industry, where despite an increase in the range of features adopted as well as the average usage of consumers, total prices for wireless service (as measured by ARPU: average revenue per user per month, a common industry measure) have not changed significantly[1].

Some distinguishing features of supply and demand in this industry may explain these deviations. First, wireless service quality depends on a set of inter-related technological choices, rather than being a simple, directly chosen strategic variable. Firms can influence quality by varying their deployed level of network infrastructure; this is a 'short-run' variable to some extent, and firms are continually investing in additions and upgrades to their networks. Additionally, given a fixed level of network infrastructure, quality is also influenced by a pair of related technological choices – the

---

[1] For example, according to CTIA's semi-annual cellular survey, average minutes of usage per user grew over 20% from 2002 to 2003, while average revenue per user was almost flat, increasing by less than 3%, which is roughly in pace with the CPI inflation rate.

amount of spectrum the firm owns, and the effectiveness of utilization of this spectrum by the type (generation) of technology used by the firm. These tend to be 'long-run' choices: a shift to a new generation of transmission technology is a multi-billion dollar undertaking, requiring an overhaul of the entire network infrastructure and simultaneous upgrades in consumer hardware; firms are also often restricted by regulatory constraints on spectrum availability and trading.

A second distinguishing feature is the usage externality displayed by wireless service. Unlike those customarily modeled (Katz and Shapiro, 1985), these are negative externalities. An increase in network traffic increases the fraction of traffic to an transmission tower that cannot be carried (and is therefore 'lost'), and service quality, measured by a function of this loss rate, is therefore endogenously affected by the equilibrium market shares of competing providers. For a fixed set of technology choices, the magnitude of these externalities are also influenced by the *average traffic* generated by each user. The level of average traffic has risen steadily over time, driven by increased voice calling, and this increase is projected to continue as features like web access, email, gaming and text messaging are adopted more widely.

Furthermore, each firm's minimum choice of quality is indirectly restricted by a number of factors. A minimum level of network infrastructure is needed to cover a region at any non-zero level of quality, since cell phone battery limitations, interference and certain kinds of distortion limit the maximum area that a single transmission tower can cover. In some cases, firms granted licenses in the 'beauty contests' for wireless spectrum are required by regulators to pre-commit to minimum levels of population coverage and network deployment. Additionally, most wireless service providers carry a substantial portfolio of debt, and the associated bond covenants may require them to own a minimum level of network assets. Therefore, while quality itself may not directly have a lower bound, there is likely to be a minimum level of network infrastructure required from any participating firm.

In our model, service quality of each firm is therefore influenced by three technological choices – network infrastructure, spectrum availability and effectiveness of spectrum utilization (the generation of transmission technology). Each competing firm is required to deploy a minimum level of network infrastructure. The service displays negative usage externalities stemming from increased loss rates as traffic increases. The firms' strategies are influenced by the average network traffic generated by each user in the market. We relate each of these factors to a single-dimensional measure of quality, based on a model of how wireless transmission actually works. For an exogenously specific generation of technology and a fixed amount of spectrum, we then analyze sequential network

infrastructure choices and pricing decisions for symmetric competing firms.

Our analysis begins by showing that the intensity of price competition is mediated to some extent by the presence of negative usage externalities. This is because for any fixed difference in quality levels, negative usage externalities bilaterally increase the slopes of the firms' profit functions. This leads to higher equilibrium prices, and admits the possibility of symmetric-quality equilibria that support positive profits for both firms. A similar effect has also been observed in exogenous-capacity models of competing service systems subject to congestion (Levhari and Luski, 1978), and in models of delivery-time-based competition (Lederer and Li, 1997), though their externalities stem from waiting times in a queue, rather than a loss of traffic. We decompose the equilibrium prices of our firms into a quality premium and an externality premium, and characterize some pricing and revenue trends in terms of this interpretation.

Next, we establish that our game has two kinds of equilibria. Under the first kind, both firms deploy identical levels of network infrastructure, leading to prices and quality levels that are symmetric. We establish that the unique symmetric equilibrium features bilateral deployment of the minimum level of infrastructure. As one might expect, this equilibrium occurs at very low levels of average traffic, and despite symmetric choices of prices and quality, equilibrium profits are strictly positive over a lower range of average traffic levels. Additionally, this is also the equilibrium at very *high* levels of average traffic. Intuitively, this is because when network traffic is sufficiently high, usage externalities compress the marginal quality impact of additional infrastructure (beyond the minimum level) to a point where it is too low to justify its cost.

In contrast, under the second kind of equilibrium, which occurs over a wide range of intermediate levels of average traffic, providers deploy asymmetric infrastructure levels. We establish that under this equilibrium, service quality is vertically differentiated across firms, and equilibrium prices are asymmetric. This is similar to the asymmetric equilibrium one would expect from a model of quality competition, though not identical, due to the presence of usage externalities. Under this equilibrium, prices are relatively flat as average traffic increases; this is consistent with the industry trend we highlighted earlier. For a non-trivial range of parameter values, the profits of the firm with higher service quality are lower than those of the lower-quality firm; this effect is driven by the reduction in marginal returns from additional capacity caused by usage externalities, which are more significant for the firm with the higher equilibrium demand (and quality). We show that market shares are closer, and equilibrium vertical differentiation is lower than in the absence of usage externalities, and characterize some other capital expenditure and technology usage trends.

Our model indicates three phases of equilibrium strategies in wireless market's evolution. Early on, when average traffic levels generated by each customer's usage are low, the quality levels of competing providers are quite close, and the externality premium provides each firm with a significant amount of pricing power. As traffic per user increases, however, the externality premium drops rapidly, and aggressive quality differentiation is necessary to sustain positive profit margins. Finally, at a threshold level of average traffic, the duopolists revert to the symmetric minimum-infrastructure equilibrium, profits approach zero for both firms at this threshold, and there is no equilibrium (symmetric or otherwise) with positive profits thereafter. Consequently, we interpret this threshold level of average traffic as a technology transition point, beyond which migrating to a new generation of technology is a necessity to remain in the industry. We do not explicitly model the transition, though our results form a basis for determining the timing of migration to a higher bandwidth technology, and we discuss this interpretation further in Section 4.

The literature on models of competition in the wireless industry is not extensive; most prior work in this area has focused on competition in wireline telecommunications (for instance, the articles in Spulber, 1995), auction design or spectrum bidding strategies (for instance, Ausubel et al., 1997), and interconnection pricing (Laffont, Rey and Tirole, 1998, Armstrong, 1998). In the literature we are aware of, our treatment of competition in wireless telecommunications as being vertically differentiated is not uncommon; for example, Reiffen et. al. (2000) provides an empirical study of the regulatory implications of vertical integration in the wireless industry, based on the standard model of quality-based competition. Our results contrast with those of Valetti (1999), who models capacity pricing in the wireless industry where coverage, rather than loss rates serve as a proxy for quality. This analysis is based on the assertion that all carriers have substantial excess capacity, which contrasts with our view that in equilibrium, network capacity and spectrum bandwidth are important constraints and drivers of quality. Besides, coverage levels for competing wireless providers tend to be comparable in most major markets. A related paper by Sweet et al. (2001) models constrained capacity for a monopolist who chooses between either buying additional spectrum or increasing their base station density; this enables them to characterize the value of spectrum in terms of the cost of capacity.

Our results highlight the importance of usage externalities in determining wireless service quality, and in inducing positive-profit equilibria even when quality and prices are identical across competitors. An alternate approach with a similar effect is presented by Heeb (2003), who introduces horizontal differentiation to relax price competition in a vertically differentiated market.

5

Earlier papers that examine the effect of negative externalities on the pricing strategy of duopolists with exogenously specified capacity include Luski (1976), Levhari and Luski (1978), Lederer and Li (1997), and Armony and Haviv (2003). These models do not consider endogenous infrastructure (capacity) investments (though Armony and Haviv endogenize heterogeneous customers); clearly, since their context is different from ours, they also do not explicitly model vertical differentiation or capture those aspects of supply and demand in the wireless market that we have chosen to highlight. A model more closely related is by Reitman (1991), who examines simultaneous choice of price and capacity for competing firms with congestion effects that are similar (to some extent) to our usage externalities. He does characterize the existence of vertically differentiated oligopoly equilibria; however, his focus is mainly on the case of perfect competition. Moreover, we show that the minimum network deployment constraint induces a second kind of (symmetric) duopoly equilibrium that does not arise in unconstrained models of quality differentiation that admit usage externalities.

The rest of this paper is organized as follows. Section 2 outlines the model and provides a preliminary analytical description of the direct and indirect effects of service quality on firm payoffs. Section 3 characterizes the subgame perfect equilibrium of the deployment-pricing game between the firms, and establishes some important properties of both symmetric and asymmetric equilibria. Section 4 discusses the variations in outcomes with changes in average traffic, spectrum bandwidth and cost structure, some revenue, capital expenditure and profit trends, and their implications for the timing of technology migration. Section 5 summarizes the paper's key results and concludes with an outline of ongoing research.

2. Model

2.1. Firms and customers

There are two firms A and B (henceforth called *providers*) who provide wireless communication services (henceforth called the *service*) over a pre-specified geographic region, whose area is normalized to 1. The customers in this area are homogeneous in their periodic level of the average traffic generated by each customer from their usage of the service. This level of *average traffic* is denoted E, and is measured in Erlangs (a standard measure of the 'amount' of demand in communications
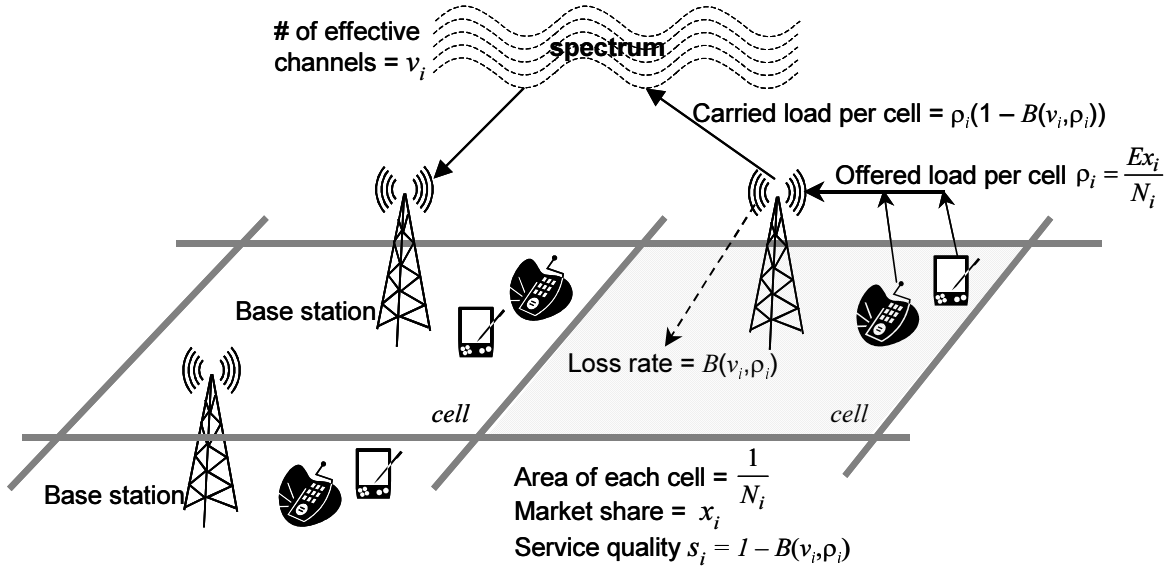
Figure 2.1: Illustration of some of the model's variables, for provider i

services[2]). The physical locations of the customers are assumed to be uniformly distributed over the area. The total number of customers per unit area is normalized to 1.

Each provider is assumed to have a combination of spectrum bandwidth and transmission technology that results in their having a total of v_i *effective channels*[3]. The provider makes two choices – their level of network infrastructure, which we model as the number of *base stations* N_i they deploy per unit area, and the total price p_i they charge each customer for their service. Each base station is assumed to have a single transmitter that can utilize all available v_i channels. The number of base stations per unit area directly determines the number of *cells* that carrier i divides each unit area into. Since customers are uniformly distributed in the region, we also assume that the deployment of the base stations is uniform across the area, and that all cells are of identical area[4]. With N_i base stations per unit area, the area of each cell is therefore $\frac{1}{N_i}$.

As a consequence, given N_i and E, if provider i has market share x_i, then the offered load of

---

[2]A population generates a traffic of z Erlangs when, on average, the population will demand the resources (in this case, effective channels) required to conduct an average of z calls per unit time.

[3]For instance, if a carrier has total spectrum of 3MHz (3000 kHz), and transmission technology is TDMA, which requires 30kHz of spectrum per channel, and allows three simultaneous send-recieve transmissions (using time-division multiplexing) on each 30kHz block, this results in a total of $\frac{3000 \times 3}{30} = 300$ effective channels. On the other hand, if GSM is the transmission technology, it requires the same 30kHz blocks, but allows upto 8 simultaneous send-receive pairs per block (due to superior time-division multiplexing), this results in $\frac{3000 \times 8}{30} = 800$ effective channels. For the CDMA protocol, the computation of effeanctive channels is based on the Shannon formula.

[4]For simplicity, we ignore corner issues, and any issues of efficient network topology. By assuming uniform base station deployment and therefore symmetric coverage, we are precluding the possibility of horizontal differentiation that might arise out of heterogeneity in the firms' relative coverage of different regions.

demand on the base station in each cell is

$$\rho_i = \frac{E x_i}{N_i}.$$

Service quality is assumed to be proportional to the fraction of demand that is successfully carried, and is determined by the *loss rate* at each base station[5]. Specifically, since the base station in each cell has access to v_i channels, and has a total offered load ρ_i, the service quality for provider i takes the form:

$$s_i = 1 - B(v_i, \rho_i),$$

where $B(\rho, v)$ is the Erlang loss function, defined as:

$$B(v, \rho) = \frac{\rho^v / v!}{\sum\limits_{i=1}^{v} (\rho^i / i!)}.$$

The Erlang loss function $B(\rho, v)$ is the loss rate from an $M/G/v/v$ queue with offered load ρ. It is widely used to model loss rates in telephony and data networks[6] (Jagerman, 1996).

Customers differ in their valuation of service quality s. Specifically, at a level of service quality s, and at a total price p, the preferences of a customer of type θ can be represented by the utility function $U(s, \theta, p)$:

$$U(s, \theta, p) = u(E)[w + \theta s] - p. \tag{2.1}$$

w represents the common (quality independent) value that every customer derives from the service (for instance, the value of having a cellular telephone in case of emergencies[7]). θ is assumed to have an absolutely continuous distribution $F(\theta)$ with support $[0, 1]$. The function $u(E)$ is strictly positive, non-decreasing and strictly concave in E, and captures increases in service value for customers as their (homogeneous) average level of demand E increases. We assume that w is high enough for the market to be fully covered in equilibrium, and all customers buy from one or the other provider.

---

[5] The primary driver of quality for voice communication is the fraction of dropped calls, which is directly related to the loss rate. Other aspects to quality include interference, noise, voice distortion and periodic silences, all of which are also related (though not entirely) to loss rates. In wireless data networks, the primary measure of quality is transmission speed, which is again directly related to the fraction of lost packets (since each lost packets must be resent, thereby reducing throughput and the effective network speed).

[6] An implicit assumption from choosing the Erlang loss function is that the region is sufficiently densely populated to approximate the demand process as a Poisson arrival process.

[7] Since the computation of E is customarily made at peak load, the term w may also capture the value that all customers get from a fraction of their off-peak usage, if this fraction is sufficiently low.
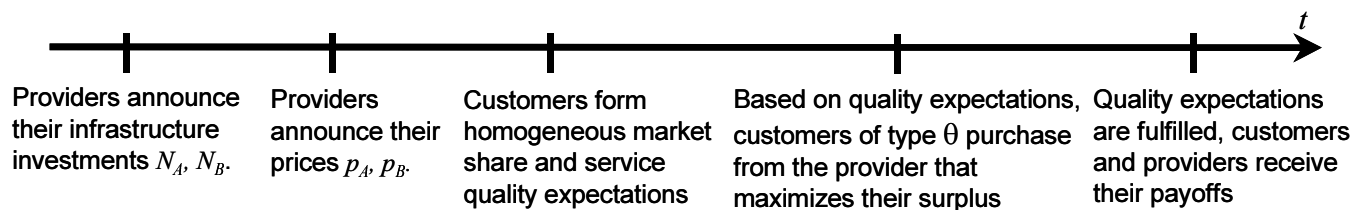
| Providers announce their infrastructure investments N_A, N_B. | Providers announce their prices p_A, p_B. | Customers form homogeneous market share and service quality expectations | Based on quality expectations, customers of type θ purchase from the provider that maximizes their surplus | Quality expectations are fulfilled, customers and providers receive their payoffs |

Figure 2.2: Timeline of events

The costs to the provider are assumed to be primarily infrastructure-related. Specifically, both providers have identical cost functions $c(N)$, where N is the number of base stations deployed. These costs per base station include networking equipment, civil work, permission acquisition, and fixed line costs to connect the base station to the land-based backbone network. We ignore interconnection pricing issues with the land-based network and the competing network. $c(N)$ is assumed to be strictly increasing and (weakly) convex in N.

We also assume that each firm is required to deploy at least at a minimum level $N_{\min}$ of network infrastructure. As discussed briefly, there are many reasons for this lower bound on network deployment. From a technological point of view, firms need to deploy a minimum number of base stations to cover any given area. Cell phone battery limitations, interference and Doppler distortion limit the maximum cell size, which in turn imposes a constraint on the minimum number of base stations required to offer service (of non-zero quality) to a given area. From a financial standpoint, wireless service providers are typically highly leveraged, and their bond covenants often require them to have deployed a minimum level of network assets. There are also minimum coverage and deployment requirements that the providers who are granted licenses for wireless spectrum through 'beauty contests' need to comply with.

2.2. Customer expectations and service quality

The timeline of events is summarized in Figure 2.2. In the analysis that follows, we assume that the two providers have identical spectrum bandwidth and transmission technology, and therefore, $v_A = v_B = v$. Suppose the providers have announced their capacity N_A and N_B, and their prices p_A and p_B. Without loss of generality, since the providers are identical, let $p_A \geq p_B$. Define $y \in [0, 1]$ as the expected indifferent customer type. The variable y is also referred to as the *market-share expectation* of the customers, since given y, the expected market share of provider A is $[1 - F(y)]$,

and the expected market share of provider B is $F(y)$. Therefore, the expected service quality of the providers is:

$$s_A \;=\; 1 - B(v, \frac{E(1 - F(y))}{N_A}) \tag{2.2}$$

$$s_B \;=\; 1 - B(v, \frac{EF(y)}{N_B}) \tag{2.3}$$

If $p_A > p_B$, the indifferent customer type must be one for whom $U(s_A, y, p_A) = U(s_B, y, p_B)$, which simplifies to the familiar condition:

$$y = \frac{p_A - p_B}{u(E)(s_A - s_B)} \tag{2.4}$$

Equations (2.2), (2.3) and (2.4) define the fulfilled-expectations equilibrium demand and quality. Any solution (y^*, s_A^*, s_B^*) to these three simultaneous equations is a fulfilled-expectations equilibrium consistent with the choices p_A, p_B, N_A and N_B, that specifies what equilibrium demand and service quality will be.

2.3. The effects of usage externalities on pricing power

Before characterizing the game and its equilibria in Section 3, we qualitatively analyze the effects of the consumption externality on equilibrium pricing behavior. Fix the capacity investments N_A and N_B, and denote the expected quality difference between the providers as

$$\Delta s = s_A - s_B = B(v, \frac{EF(y)}{N_B}) - B(v, \frac{E(1 - F(y))}{N_A}). \tag{2.5}$$

Equations (2.4) and (2.5) implicitly define a quality difference correspondence $\Delta s(p_A, p_B)$. Assume that this is single-valued. Based on (2.4), the second-period payoff functions for the two providers are therefore:

$$\pi^A(p_A, p_B) \;=\; p_A[1 - F(y)], \tag{2.6}$$

$$\pi^B(p_A, p_B) \;=\; p_B F(y), \tag{2.7}$$

where

$$y = \frac{p_A - p_B}{u(E)\Delta s(p_A, p_B)}. \tag{2.8}$$

(2.6), (2.7) and (2.8) yield the following equations[8]:

$$\pi_1^A(p_A, p_B) = [1 - F(y)] - \left[p_A \left(\frac{f(y)}{u(E)\Delta s(p_A, p_B)} \right) \right] + \left[p_A \left(\frac{\partial \Delta s}{\partial p_A} \right) \frac{y f(y)}{\Delta s(p_A, p_B)} \right]. \quad (2.9)$$

$$\pi_2^B(p_A, p_B) = [F(y)] - \left[p_B \left(\frac{f(y)}{u(E)\Delta s(p_A, p_B)} \right) \right] + \left[p_B \left(-\frac{\partial \Delta s}{\partial p_B} \right) \frac{y f(y)}{\Delta s(p_A, p_B)} \right]. \quad (2.10)$$

(2.9) and (2.10) provide an intuitive understanding of the different effects of a marginal increase in provider pricing. The first term in each equation is the direct positive effect of revenue increases from the price increase. The second term is the direct negative effect that arises out of the reduction in revenues from the direct reduction in demand that accompanies the price increase. The final term, which we term the *externality effect*, captures the mediating effect of the negative usage externality. When a provider's price increases, this shifts demand from their network to that of their competitor. Due to the dependence of quality on usage, this changes the service quality difference between the two providers, and the indirect quality effect captures the revenue impact of this change in quality difference. It is easily verified that this indirect quality effect is strictly positive in both equations.

Other things being equal, the presence of the indirect quality effect results in an equilibrium increase in prices, since it increases the slope of the second-period payoff functions of both providers. Therefore, at any fixed pair of infrastructure choices N_A and N_B, the usage externalities lead to higher prices in equilibrium.

3. Equilibrium

This section characterizes the equilibrium of the game introduced in Section 2. For ease of tractability, we assume that θ is uniformly distributed in $[0, 1]$, with corresponding distribution function $F(\theta) = \theta$ and density function $f(\theta) = 1$. We discuss the effects of relaxing this assumption in Section 5. The derivation of the equilibrium has three parts. First, we formalize the fixed-point equation that determines fulfilled-expectations outcomes after infrastructure and prices have been chosen. Next, we specify the Nash equilibrium of the second-period price competition game. Finally, we specify the first-period equilibrium subgame perfect choices of infrastructure.

---

[8] Throughout the remainder of the paper, numbered subscripts of *functions* represent partial derivatives with respect to the corresponding variable.

3.1. Fulfilled-expectations outcomes

As described in Section 2.2, given a set of prices (p_A, p_B) and capacities (N_A, N_B), customers form an expectation of the value of y, the indifferent customer type (also referred to as the *market-share expectation*). Given y, there is a unique expectation of quality levels, as originally specified in equations (2.2) and (2.3):

$$s_A(y) = 1 - B(v, \frac{E(1-y)}{N_A}). \tag{3.1}$$

$$s_B(y) = 1 - B(v, \frac{Ey}{N_B}). \tag{3.2}$$

Given an arbitrary y, N_A, N_B, define the *quality differential function* $Q(y, N_A, N_B)$ as the difference $[s_A(y) - s_B(y)]$. Using (3.1) and (3.2), this function is:

$$Q(y, N_A, N_B) = B(v, \frac{Ey}{N_B}) - B(v, \frac{E[1-y]}{N_A}). \tag{3.3}$$

The RHS of (3.3) has an expression identical to the function $\Delta s(p_A, p_B)$ used in Section 2.3. Given the expectations of quality, the *actual value* of the indifferent customer type, denoted $\Gamma(y)$, is specified below.

$$\text{When} \quad p_A > p_B : \Gamma(y) = \frac{p_A - p_B}{u(E)Q(y, N_A, N_B)}. \tag{3.4}$$

$$\text{When} \quad p_A = p_B : \begin{aligned} &\Gamma(y) = 0 \text{ if } Q(y, N_A, N_B) < 0; \\ &\Gamma(y) = 1 \text{ if } Q(y, N_A, N_B) > 0; \\ &\Gamma(y) = \frac{N_B}{N_A+N_B} \text{ if } Q(y, N_A, N_B) = 0. \end{aligned} \tag{3.5}$$

That is, if provider prices are different, then the indifferent customer type is as specified by (2.4). If prices are equal, but customer expectations are such that one provider has a higher quality, then that provider gets all the customers. On the other hand, if prices are equal, and if the market-share expectation y is such that service quality is equal, then the providers split demand based on network deployment levels. With equal prices and expected quality levels, any market share split is reasonable – (3.5) specifies the outcome which ensures that $\Gamma(y)$ is continuous[9].

---

[9] Often, the assumption made under this scenario is that firms split demand in any ratio. In fact, this may appear to be a critical assumption, since our equilibrium sometimes involves equal prices and quality. However, we will establish in Section 3.2 that with symmetric networks $N_A = N_B$, the only split in demand consistent with fulfilled expectations and second-period profit maximization is one in which providers split the market equally; moreover, with $N_A \neq N_B$, prices are never equal. Therefore, this assumption is without loss of generality.

Therefore, given a set of prices (p_A, p_B) and capacities (N_A, N_B), a market-share expectation y^* satisfies *fulfilled-expectations* if it is a solution to

$$y^* = \Gamma(y^*), \tag{3.6}$$

where $\Gamma(y)$ is as defined in (3.4) – (3.5).

3.2. Equilibrium prices

This section derives the symmetric second-period Nash equilibria which are consistent with a market-share expectation y that is fulfilled in equilibrium. The costs of N_A and N_B are already sunk, and therefore, any candidate pair of prices yield the following second-period payoffs to each provider:

$$\pi^A(p_A, p_B) = p_A[1 - y], \tag{3.7}$$

$$\pi^B(p_A, p_B) = p_B y, \tag{3.8}$$

where y satisfies fulfilled expectations. Using (3.4), (3.7) and (3.8), the conditions for y to satisfy fulfilled expectations can be restated as:

$$yu(E)Q(y, N_A, N_B) = p_A - p_B. \tag{3.9}$$

The following proposition establishes the conditions for a pair of prices (p_A, p_B) to be a Nash equilibrium of the second-period game. All proofs are in Appendix A.

Proposition 1. *Given N_A and N_B, any second-period Nash equilibrium price pair must take the following form:*

$$p_A = (1 - y)u(E)[Q(y, N_A, N_B) + y[1 - y]u(E)Q_1(y, N_A, N_B)]; \tag{3.10}$$

$$p_B = yu(E)Q(y, N_A, N_B) + y^2u(E)Q_1(y, N_A, N_B)], \tag{3.11}$$

where $y \in [0, 1]$ satisfies fulfilled expectations as specified by (3.9), given p_A, p_B and N_A, N_B

Proposition 1 specifies necessary (rather than sufficient) conditions for any Nash equilibrium price pair. The component of equilibrium price that depends on the direct difference in quality – that is, $(1 - y)u(E)Q(y, N_A, N_B)$ for firm A and $yu(E)Q(y, N_A, N_B)$ for firm B – is referred to as

the *quality premium*. These would be the equilibrium prices in the absence of usage externalities. The other component, which arises due to the indirect quality effect discussed in section 2.3 – $u(E)(1-y)yQ_1(y, N_A, N_B)$ for firm A, and $u(E)y^2Q(y, N_A, N_B)$ for firm B – is sometimes referred to as the *externality premium*.

Proposition 2 establishes some properties of a symmetric second-period price equilibrium.

Proposition 2. *(a) If $N_A = N_B = N$, then there exists a unique symmetric second-period Nash equilibrium (p^*, p^*), where:*

$$p^* = \frac{Eu(E)}{2N}B_2(v, \frac{E}{2N}).$$

(b) If $N_A \neq N_B$, then there is no symmetric second-period Nash equilibrium in pure strategies.

3.3. Equilibrium network infrastructure levels

This subsection characterizes equilibrium infrastructure deployment levels N_A^*, N_B^*, and shows that they are either at a symmetric minimum required level $N_{\min}$, or are asymmetric across the providers. Given any N_A, N_B, denote the corresponding second-period equilibrium prices as $p_A(N_A, N_B)$ and $p_B(N_A, N_B)$ respectively, and the corresponding equilibrium indifferent customer as $y(N_A, N_B)$. Since Proposition 1 is a necessary condition, it follows that:

$$p_A(N_A, N_B) = u(E)[1 - y(N_A, N_B)][Q(y(N_A, N_B), N_A, N_B) + y(N_A, N_B)Q_1(y(N_A, N_B), N_A, N_B)];$$
$$p_B(N_A, N_B) = u(E)y(N_A, N_B)[Q(y(N_A, N_B), N_A, N_B) + y(N_A, N_B)Q_1(y(N_A, N_B), N_A, N_B)],$$

and $y(N_A, N_B)$ is implicitly defined by:

$$u(E)Q(y(N_A, N_B), N_A, N_B)y(N_A, N_B) = p_A(N_A, N_B) - p_B(N_A, N_B). \tag{3.12}$$

The equilibrium infrastructure levels of the providers can now be characterized:

Proposition 3. *Any subgame perfect Nash equilibrium of the deployment-pricing game takes one of the following forms:*

(a) Both providers deploy the minimum infrastructure, and price symmetrically: $N_A^* = N_B^* = N_{\min}$ and $p_A^* = p_B^* = \frac{Eu(E)}{2N_{\min}}B_2(v, \frac{E}{2N_{\min}})$

(b) Providers deploy asymmetric network capacity, and charge different prices. That is, if either

$N_A^* > N_{\min}$ or $N_B^* > N_{\min}$, then $N_A^* \neq N_B^*$ and $p_A^* \neq p_B^*$.

Analysis of the first-period equilibrium reveals a set of distinct ranges of parameter values, each of which is associated with specific kinds of equilibrium pairs. Since the providers are identical (in terms of their cost functions and spectrum v), it is clear that if (N_A^*, N_B^*) is an equilibrium pair, then so is (N_B^*, N_A^*). We continue to denote provider A as the higher quality provider. We also fix the cost function $c(N)$ and number of channels v, and frame this discussion in terms of ranges of and increases in *average traffic* E, since this is the exogenous demand parameter that actually changes continuously over time. Figure 3.2 summarizes the equilibrium choices of deployment across the different ranges of E described below.

For convenience, the first-period profit functions are reproduced below:

$$\Pi^A(N_A, N_B) = u(E)[1 - y(N_A, N_B)]^2[Q(y, N_A, N_B) \tag{3.13}$$
$$+ y(N_A, N_B)Q_1(y(N_A, N_B), N_A, N_B)] - c(N_A)$$
$$\Pi^B(N_A, N_B) = u(E)[y(N_A, N_B)]^2[Q(y, N_A, N_B) \tag{3.14}$$
$$+ y(N_A, N_B)Q_1(y(N_A, N_B), N_A, N_B)] - c(N_B)$$

If one solves for the values of E that satisfy the first-order condition of the higher quality provider at the minimum deployment level $N_{\min}$:

$$\Pi_1^A(N_{\min}, N_{\min}) = 0 \tag{3.15}$$

this yields two solutions $E = \underline{E}$ and $E = \overline{E}$, with $\underline{E} < \overline{E}$. Outside this range of E, both providers choose equilibrium strategies according to part (a) of Proposition 3 – for $E \notin (\underline{E}, \overline{E})$, deployment is at the minimum level $N_{\min}$, and both providers price at the same level. The best response functions and the corresponding equilibrium is depicted in Figure 3.1(a).

As E crosses the threshold level $\underline{E}$, it is no longer optimal for provider A to choose $N_{\min}$, since $\Pi_1^A(N_{\min}, N_{\min}) > 0$ for $E > \underline{E}$. Denote the best response of provider A to a choice of $N_{\min}$ by provider B as $N_{\min}^A(E)$:

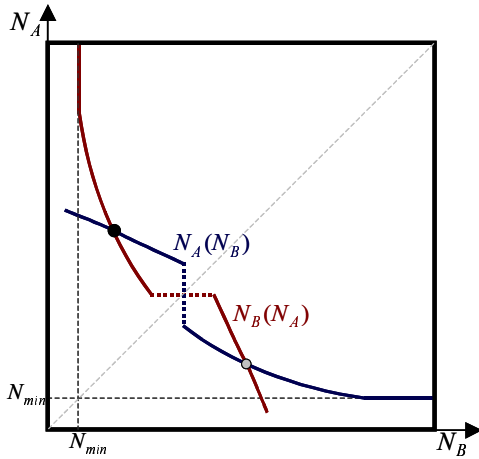$$N_{\min}^A(E) = \arg\max_N \Pi^A(N, N_{\min}). \tag{3.16}$$

For a range of values of $E > \overline{E}$, this is the equilibrium choice of provider A, and provider B continues to deploy at the level $N_{\min}$. This is depicted in Figure 3.1(b), and remains the equilibrium until a
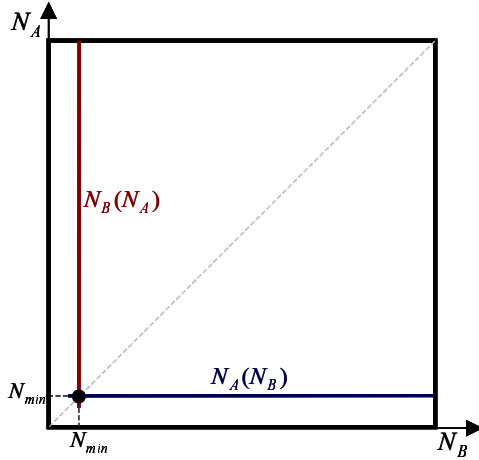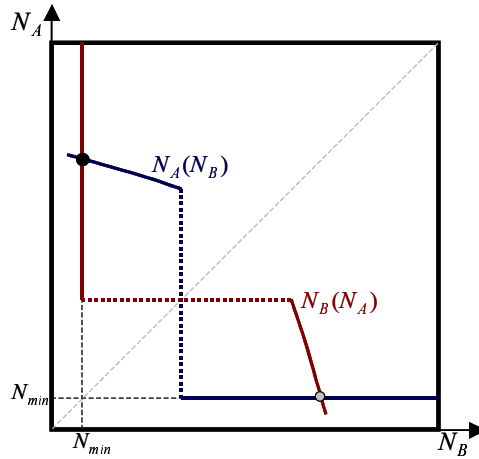
(a) Symmetric minimum infrastructure
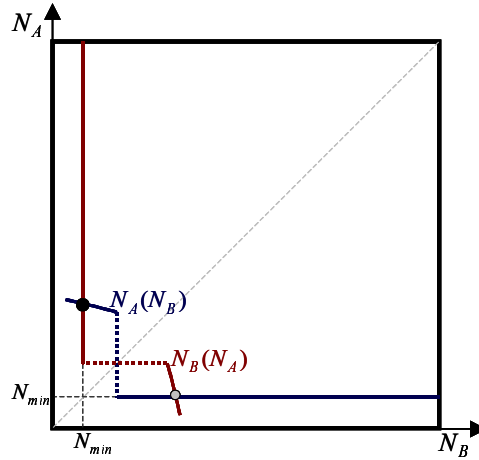
(b) Asymmetric, one provider deploys N_{min}.

(c) Interior solution with N_A, $N_B > N_{\text{min}}$

(d) Asymmetric, one provider deploys N_{min}.again

Figure 3.1: Depicts the best response functions for the two providers for different parameter values. The equilibrium points are circled. The solid back circle indicates the equilibrium point discussed in the text (where firm A is the higher quality provider), and the grey circle indicates the other equilibrium, which is identical except for a relabeling of providers. The discontinuity in the curves occurs when the firms 'switch positions' – that is, the point at which the deployment of provider j is high enough to make provider i's optimal response a choice which makes them the lower-quality, lower-capacity provider.
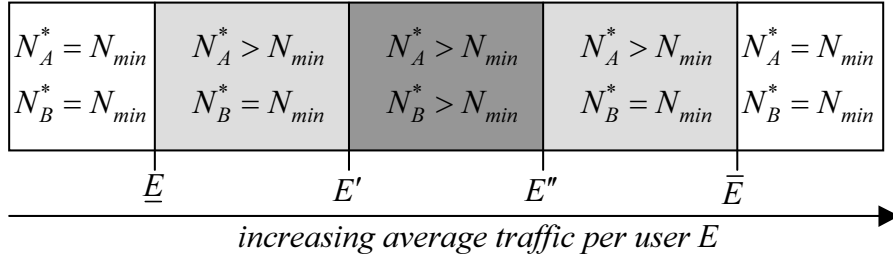
$$\begin{array}{|c|c|c|c|c|}
\hline
N_A^* = N_{min} & N_A^* > N_{min} & N_A^* > N_{min} & N_A^* > N_{min} & N_A^* = N_{min} \\
N_B^* = N_{min} & N_B^* = N_{min} & N_B^* > N_{min} & N_B^* = N_{min} & N_B^* = N_{min} \\
\hline
\end{array}$$

$$\underline{E} \qquad\qquad E' \qquad\qquad E'' \qquad\qquad \overline{E}$$

increasing average traffic per user E

Figure 3.2: Equilibrium deployment across different ranges of E.

second threshold value E', which is one of the two values of E that solve:

$$\Pi_1^B(N_{\min}^A(E), N_{\min}) = 0 \tag{3.17}$$

Beyond this point, $\Pi_1^B(N_{\min}^A(E), N_{\min}) > 0$, and both providers choose interior deployment levels $N_A^* > N_{\min}, N_B^* > N_{\min}$, as depicted in Figure 3.1(c).

As average traffic E increases further, provider B reduces its network infrastructure N_B after a point. At the next threshold value E'', which is the second value of E that solves (3.17), $N_B^* = N_{\min}$ again, and provider B remains at the minimum deployment level at all higher values of E. Provider A first continues to increase their deployment as E increases, yielding an equilibrium as depicted in Figure 3.1(d), then begins to reduce deployment until the point where E reaches the threshold $\overline{E}$ which is the second value of E that solves (3.15). At this point and beyond, provider A also chooses $N_{\min}$, since $\Pi_1^A(N_{\min}, N_{\min}) < 0$ for $E > \overline{E}$, and the equilibrium returns to the one depicted in Figure 3.1(a).

Figure 3.3 summarizes the evolution of the first-stage equilibrium deployment of infrastructure as average per-user traffic E varies. Interestingly, the outcomes are identical at very low or very high values of E. In the former case, when $E < \underline{E}$, there is insufficient traffic to justify infrastructure investment beyond the minimum level $N_{\min}$. In the latter case, when $E > \overline{E}$, the traffic is sufficiently high that the marginal impact of an additional base station on service quality is too low to justify its costs.

4. Discussion

This section discusses how revenues, profits, capital expenditure, service quality and market share vary with changes in average traffic E, customer value $u(E)$, spectrum v and infrastructure costs
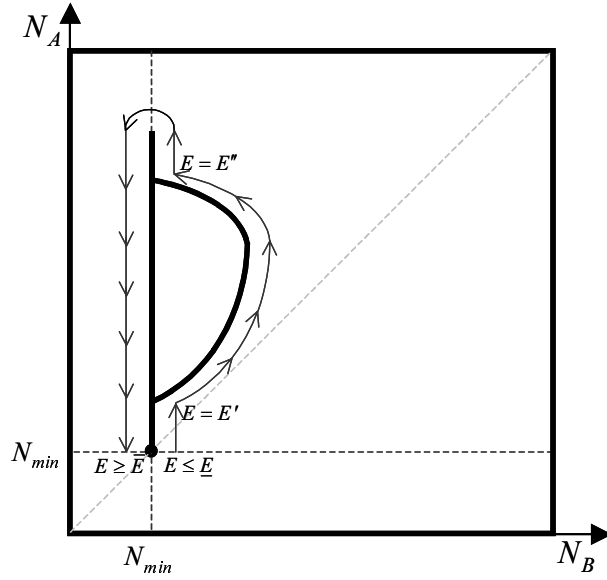
Figure 3.3: Evolution of the equilibrium infrastructure deployment point (first-stage choices of N) as average traffic per user E varies.

$c(N)$. Due to the complexity of the payoff functions derived in section 3, and the non-monotonicity of the higher derivatives of the Erlang loss function $B(v, \rho)$, comparative statics results are difficult to obtain analytically. We have therefore computed the equilibrium points of the game for a number of specific parameter values, and obtained the corresponding economic measures. We summarize those results from this exercise that appeared consistently across our numerical analysis. The specific results discussed below assume that infrastructure costs are linear (that is, $c(N) = cN$), and that $u(E) = E^\alpha$, where α measures the rate of increase in value as average per-user traffic increases – a higher α implies higher increases in value.

4.1. Revenue and capital expenditure trends

We start with a benchmark scenario in which $u(E)$ is constant in E (that is, $\alpha = 0$). This corresponds to a situation where there is no increase in the value perceived from the service when average traffic per user grows. For instance, when new bandwidth-intensive services are rolled out, sometimes they can cause substantial increases in traffic with minimal increases in the actual value customers feel they are getting from the service. This kind of increase in a technological measure of power or speed over time, with a minimal corresponding increase in willingness to pay often
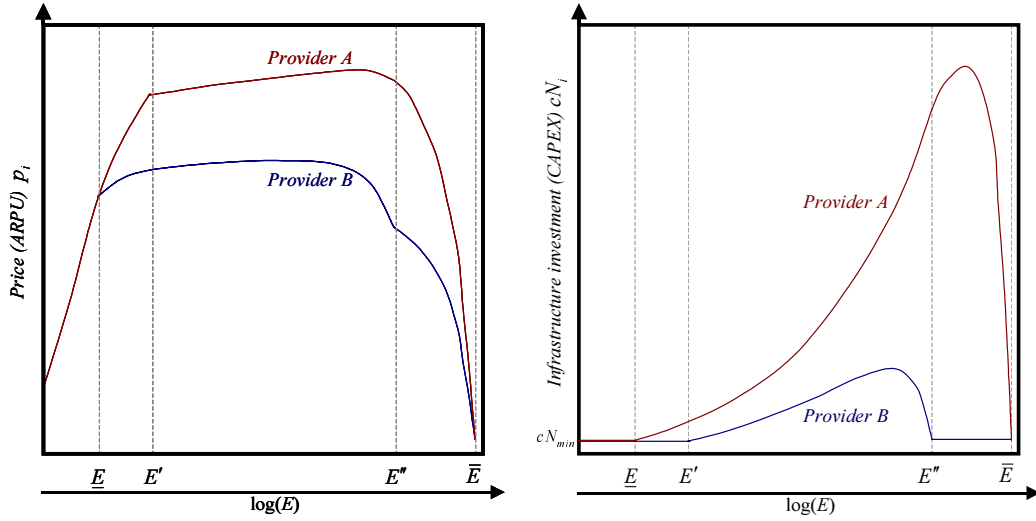
Figure 4.1: Variation in prices (or average revenue per user) and investements in network infrastructure (CAPEX) as average traffic E increases.

characterizes technology products[10] (personal computers being a common example). In this case, prices (or average revenue per user, or equivalently, revenues, since we have normalized the number of customers to 1) are strictly decreasing with increases in E. The driver of this trend is clearly a bilateral decrease in equilibrium provider pricing as traffic increases – the increase in E reduces service quality due to usage externalities, and there is no demand-side driven increase in pricing power.

When perceived value $u(E)$ is increasing and concave in average traffic E, the price (revenue) trends are sometimes altered directionally for lower values of E. Figure 4.1 illustrates the changes in prices and capital expenditure for a candidate set of values of $\alpha > 0$. Each of the Figures 4.1 through 4.3 uses exactly the same parameter values. If α is high enough, there is initially an increase in price as E increases.. However, at higher values of E, the quality deterioration from the negative usage externalities dominates the value increases, which results in a decrease in incremental infrastructure deployment and the corresponding drop-off in revenues. In equilibrium, the providers initially offset the reduction in quality partially by increasing their capital expenditure on infrastructure, as also illustrated in Figure 4.1. After a point, however, the marginal benefits of quality increases from infrastructure investments are outweighed by the costs, total infrastructure

---

[10] Of course, this is not a complete economic argument; it is merely an observed aspect of computing and communications products that some managers may find representative of the demand characteristics they face. Our model's validity and results do not rely on this assumption – this is a benchmark, and we subsequently discuss corresponding results when $u(E)$ is increasing in E.
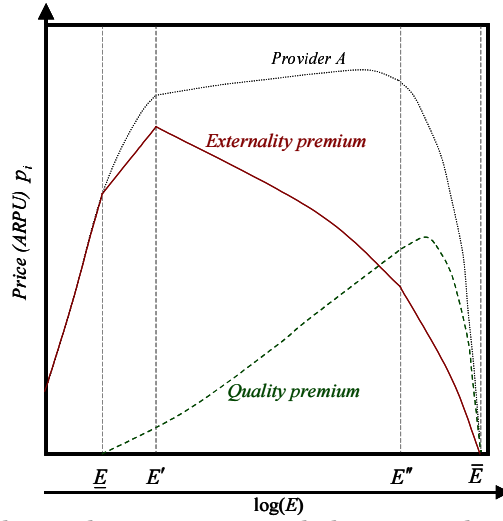
19

Figure 4.2: Contribution of the quality premium and the externality (congestion) premium to the price charged by Provider A as E varies.

deployment drops, and prices falls more sharply.

Figure 4.2 provides a clearer illustration of the relative contribution of the two components of provider pricing – the *quality* premium and the *externality* premium – derived in (3.10), for Provider A. The externality premium is the primary driver of pricing power for lower levels of per-user traffic E, because the equilibrium differential in quality is low. As E increases, Provider A invests proportionately more than Provider B, increasing the quality gap and the quality premium. However, the rate of change in the quality differential with changes in market share decreases as the number of base stations increases (this is due to the fact that the rate of change in this quality differential $Q_1(y, N_A, N_B)$ is linearly increasing in both E/N_A and E/N_B). As a consequence, the externality premium declines. At a point beyond E'', the higher quality provider begins to slow down their rate of infrastructure addition, and the quality differential peaks and begins to fall, more rapidly as Provider A reduces their infrastructure investment.

This analysis suggests two different strategies for firms at different stages of demand growth – uniform quality and congestion pricing with relatively low capacity investment for lower values of demand, followed by more aggressive vertical differentiation as E increases. We discuss this briefly in our concluding section.

An increase in channel capacity v does not change the qualitative features of the price or capital expenditure trends, but merely changes the range of values of E over which they occur. Specifically, an increase in v 'stretches' out the revenue and earnings trends, while a decrease tends to 'compress'
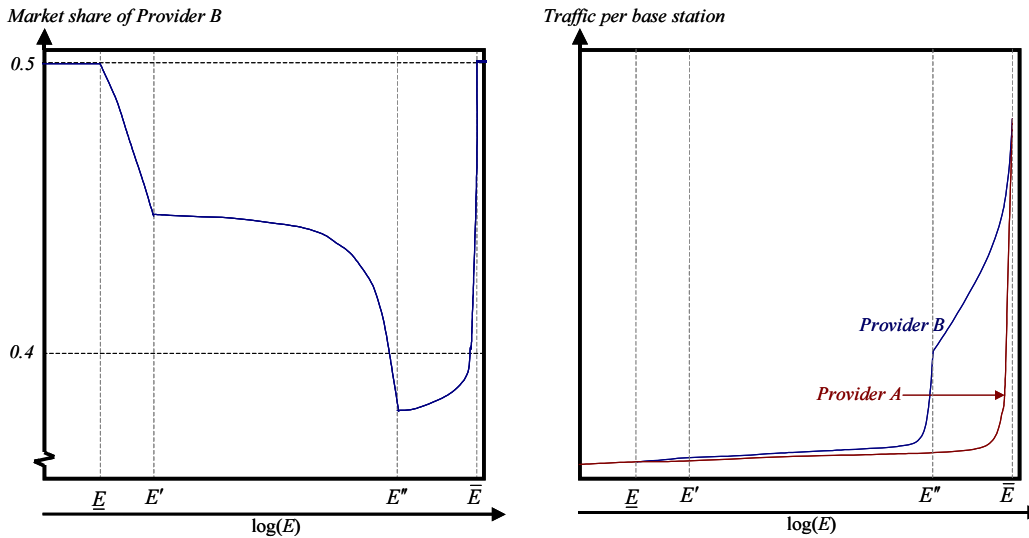
Figure 4.3: Variation in market share and average traffic per base station as average demand per user E increases.

them. Intuitively, the threshold values of E – that is, the values of $\underline{E}$, E', E'', and $\overline{E}$ defined in section 3.4 – are all increasing[11] in $v..$ The effects of an increase in variable infrastructure costs c are similar. In addition to the corresponding changes in the threshold values of E (which are caused by the fact that these values are also increasing in c), there is also a direct reduction in actual capital expenditure at each value of E, since the (decreasing) marginal revenue from infrastructure deployment is equated to a higher marginal cost.

4.2. Market share, profits and technology migration

The variation in relative market share across providers is illustrated in Figure 4.3. For $E \notin (\underline{E}, \overline{E})$, both providers split the market equally. As E increases from $\underline{E}$ to E'', the market share of provider A increases progressively, and peaks at $E = E''$. Subsequently, it decreases monotonically until $E = \overline{E}$, where the split is even again. This is in contrast to the corresponding model in the absence of usage externalities and with a direct choice of s_A and s_B, where it is straightforward to show that independent of c, firm B would have a market share of $\frac{1}{3}$, and firm A would have a market share of $\frac{2}{3}$. The presence of usage externalities therefore dampens a firm's need to (and ability to profitably) differentiate their product, and explains a more uniform equilibrium division of the market; the minimum deployment constraints suggest that it could be symmetric at extreme levels

---

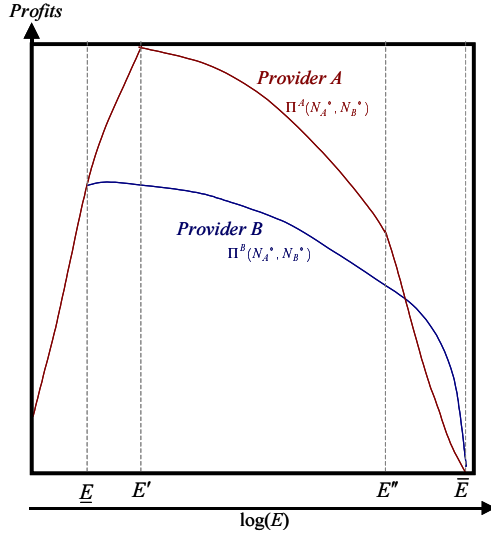[11]This can be proved by computing the total derivatives of (3.15) and (3.17).

Figure 4.4: Variation in equilibrium profits (earnings) $\Pi^A(N_A^*, N_B^*)$ and $\Pi^B(N_A^*, N_B^*)$ as average traffic per user E increases.

of average traffic.

The variation in quality across providers can be inferred from the traffic per base station depicted in Figure 4.3 – as depicted, the average traffic per base station increases gradually at first, and then exponentially as equilibrium infrastructure deployment starts to reduce.

The revenue and capital expenditure trends highlighted in Section 4.1 suggest that provider profits will be decreasing as average traffic increases, at least over the interval where revenues decrease while costs increase. This is indeed the case, as illustrated in Figure 4.4. Two features of the profit chart in Figure 4.4 are of particular interest. Firstly, over a portion of the higher range of average per-user traffic values, the earnings of provider A are lower than those of provider B, despite the fact that provider A has higher equilibrium quality and prices. This occurs over part of the range $E'' < E \le \overline{E}$, when the infrastructure deployment of provider B is constant at $N_{\min}$, while provider A continues to vary their deployment. This unusual outcome is a consequence of the reduction in marginal returns from additional capacity caused by usage externalities, which are more significant for Firm A who has higher equilibrium demand. This leads to disproportionately higher equilibrium infrastructure costs, and lower profits as a consequence.

Secondly, at $E = \overline{E}$, the earnings of both firms are consistently very close to zero, and this occurs across the entire range of parameter values we have studied. In other words, the point at which both providers stop investing in infrastructure (and choose equilibrium deployment $N_A^* = N_B^* = N_{\min}$)
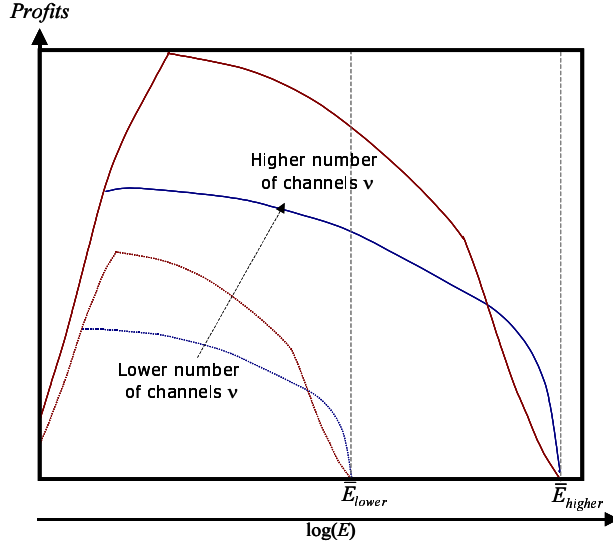
Figure 4.5: Impact of migrating to a new generation of technology. An increase in the number of effective channels v increases earnings bilaterally, and also increases the range of demand over which positive earnings are sustainable in equilibrium.

in response to increases in average per-user traffic coincides almost exactly with the point at which total provider profits are bilaterally zero.

The analytical basis of the latter result becomes more apparent when one recognizes that the equilibrium traffic per base station is always fairly high at $E = \overline{E}$. Now, suppose one uses the heavy-traffic approximation for the Erlang loss function:

$$B(v, \rho) \simeq 1 - \frac{v}{\rho}. \tag{4.1}$$

Under this approximation, the quality differential function is:

$$Q(y, N_A, N_B) = \frac{v}{E}\left(\frac{N_A}{1-y} - \frac{N_B}{y}\right), \tag{4.2}$$

and therefore,

$$yQ_1(y, N_A, N_B) = \frac{v}{E}\left(\frac{yN_A}{[1-y]^2} - \frac{N_B}{y}\right). \tag{4.3}$$

As a consequence, provider A's profit function takes the form:

$$\begin{aligned}
\pi^A(N_A, N_B) &= u(E)[1-y]^2[Q(y, N_A, N_B) + yQ_1(y, N_A, N_B)] \tag{4.4}\\
&= u(E)\frac{v}{E}N_A - cN_A. \tag{4.5}
\end{aligned}$$

23

Recall that $N_B^* = N_{\min}$ is unchanged in the neighborhood of $\overline{E}$. From equation (4.5), it is clear that $\Pi^A(N_A, N_{\min}) = 0$ and $\Pi_1^A(N_A, N_{\min}) = 0$ have the same approximate solution in N_A, which implies that the point at which provider A stops investing in infrastructure is the same as the point at which their profits are zero. The intuition here is that under heavy traffic, each part of the provider's network infrastructure is operating at close to full utilization (the approximation in (4.1) is analytically identical to the assumption of 100% utilization). The marginal revenue from each of the base stations (that is, each unit of infrastructure deployment) is therefore approximately equal, since each reduces the loss rate by about the same fraction $\frac{v}{E(1-y)}$. Therefore, provider A's profit function is approximately linear in N_A, and when marginal revenue from an additional base station reaches the marginal cost of the base station C, earnings tend to zero.

This result suggests a natural point of per-user traffic $\overline{E}$ at which providers should migrate to a new generation of wireless technology. As E increases beyond $\overline{E}$, equilibrium earnings become negative very soon, which makes technology migration (in conjunction with new debt financing) a necessity for continued business survival. As depicted by Figure 4.5, switching to a technology that admits a higher number of channels will bilaterally increase earnings, and this increase sustains even if variable costs of infrastructure for the new technology are higher.

5. Conclusions and ongoing work

We have developed a new model of duopoly in the wireless industry which admits negative usage externalities and minimum investment constraints into a two-stage vertical differentiation game. The specification of the model links network traffic, transmission technology, spectrum availability and infrastructure deployment with service quality, customer demand and equilibrium pricing, and we hope that it may form the basis for the analysis of related economic questions in this increasingly important industry sector.

We have shown that the presence of negative usage externality imposed by average per-user traffic on service quality actually reduces the competitive intensity of the industry. Oligopolists. in the wireless sector therefore have more pricing power than firms in traditional vertically differentiated industries, and equilibrium market shares are likely to be more uniform than those predicted by a standard model of vertical differentiation. This is a favorable observation from one point of view, since the sustainability of this industry depends on periodic capital investments in networks that are substantial. Note that even though market penetration in the U.S. is nearing saturation,

the major providers still invest in excess of \$100 billion annually on building and upgrading their networks. On the other hand, it suggests that measures of market power that rely on relative concentration (such as the Herfindahl-Hirschman index) may need to be adjusted when analyzing this industry, since they would tend to underestimate the pricing power of oligopolists. This is likely to be of particular interest if consolidation in the industry continues.

Our results also explain the seemingly puzzling trend of declining average revenues per user, despite increases in value that drive an increase in average traffic. This trend is a consequence of two primary drivers – the negative effect that usage externalities have on service quality, and the increase in competitive intensity as traffic per base station increases. This declining price trend eventually characterizes one phase of competition even when customer value from the service provided increases with average traffic. Moreover, as firms enter this phase of declining ARPU, the equilibrium earnings of the higher quality provider may often be lower than those of its lower-quality rival.

Our results also suggest three phases of equilibrium strategies for wireless providers. At an early stage in the market's evolution, when average traffic per user is low, our model suggests that firms should choose quality levels that are similar, and maintain relatively low infrastructure deployment. Significant quality differentiation is not optimal for either firm, since prices are kept high by the externality premiums, and any substantial increase in quality will increase the price elasticity of demand (informally, this is because the effect of a marginal transfer of market share on quality levels decreases as quality increases), and this is likely to intensify price competition. However, as the market matures and the average traffic per user increases significantly, firms should pursue more aggressive quality differentiation, as the externality premiums drop, and their interaction begins to resemble standard vertical differentiated competition. The profits of the higher quality firm are higher in this phase for the most part, though their profits are lower beyond a point. Absent any positive economies of scale, the lower quality firm will frequently earn a better return on their assets over a wide range of demand levels, on account of their lower infrastructure investments. These effects are caused by the reduction in marginal returns from additional capacity caused by usage externalities, which are more significant for the firm with the higher equilibrium demand (and quality).

Beyond a critical threshold of demand, providers are likely to be forced to upgrade their technology to one that provides them with a higher number of channels per base station. This will be a necessity for survival even if the (per-cell) variable infrastructure costs are higher under the

new technology[12]. More importantly, as new bandwidth-intensive services continue to drive up the average traffic per user, these newer networks will reach their corresponding critical $\overline{E}$ values, and this cycle of network upgrades is likely to persist. In this context, if calibrated and extended to a market with multiple operators, our model can help regulators price future spectrum, and can also predict equilibrium industry structure. This is one direction of research we continue to pursue actively.

Our model assumes that there are no switching costs. In the U.S., this seems like a more reasonable assumption following the FCC mandate of number portability. Our logical next step is to analyze the timing of technology migration more formally by explicitly modeling the technology migration game. Clearly, the relative infrastructure costs and channel capacities of each technology will play a critical role, as will the timing and structure of the game. From a social welfare perspective, it is not yet clear whether the timing of migration to next-generation technologies occurs at a point which is too early or too late, and whether the deterioration in service quality that precedes the migration can be avoided by appropriately timed government subsidies. A rigorous analysis may shed light on these issues, and we hope to answer some of these open questions in the near future.

References

1. Armony M. and M. Haviv, 2003, "Price and Delay Competition between Two Service Providers," *The European Journal of Operational Research*, 147 (1), 32-50.

2. Ausubel, L., P. Cramton, P. McAfee and J. McMillan, 1997, "Synergies in Wireless Telephony: Evidence from the Broadband PCS Auctions" *Journal of Economics and Management Strategy*, 6 (3), 497-527.

3. Heeb, R., 2003, "Innovation and Vertical Integration in Complementary Markets," *Journal of Economics and Management Strategy,* 12 (3), 387-417.

4. Katz, M. and C. Shapiro, 1985, "Network Externalities, Competition and Contracting," *American Economic Review*, 75, 424-440.

5. Laffont, J-J, P. Rey and J. Tirole, 1998, "Network Competition I. Overview and Nondiscriminatory Pricing," *Rand Journal of Economics*, 29 (1), 1-37

---

[12]In the short term, this is clearly a positive for equipment manufacturers for whom the impending migration of carriers to 3G technology represents a substantial revenue opportunity.

6. Lederer, P. J. and L. Li, 1997, "Pricing, Production, Scheduling, and Delivery-Time Competition," *Operations Research* 45 (3), 407-420.

7. Levhari, D. and I. Luski, 1978, "Duopoly Pricing and Waiting Lines," *European Economic Review*, 11, 17-35.

8. Luski, I., 1976, "On Partial Equilibrium in a Queuing System with Two Servers," *Review of Economic Studies*, 43, 519-525.

9. Reiffen D., L. Schumann and M. Ward, 2000, "Discriminatory Dealing with Downstream Competitors: Evidence from the Cellular Industry," *Journal of Industrial Economics* XLVIII, 253-286.

10. Reitman, D., 1991, "Endogenous Quality Differentiation in Congested Markets," *The Journal of Industrial Economics*, 39 (6), 621-647.

11. Shaked, A. and J. Sutton, 1982, "Relaxing Price Competition through Product Differentiation," *Review of Economics Studies*, XLIX, 3-13.

12. Spulber, D., 1995, "Introduction: Competition in Telecommunications," *Journal of Economics and Management Strategy,* 4 (2), 133-137.

13. Sweet, R., I. Viehoff, D. Linardatos, and N. Kalouptsidis, 2001, "Marginal Value-based Pricing of Additional Spectrum Assigned to Cellular Telephony Operators," *Information Economics and Policy* 14 (3), 371-384.

14. Valetti, T., 1999, "A Model of Competition in Mobile Communications," *Information Economics and Policy* 11, 61-72.

A. Proofs

Proof of Proposition 1: Any Nash equilibrium price pair must satisfy first-order conditions $\pi_1^A(p_A, p_B) = \pi_2^B(p_A, p_B) = 0$, given the market-share expectation y that satisfies fulfilled-expectations. Differentiating (3.7) with respect to p_A yields:

$$1 - y - p_A \frac{\partial y}{\partial p_A} = 0, \tag{A.1}$$

which rearranges to:

$$p_A = \frac{1-y}{\frac{\partial y}{\partial p_A}}.$$

(A.2)

Similarly, differentiating (3.8) with respect to p_B and rearranging yields:

$$p_B = \frac{-y}{\frac{\partial y}{\partial p_B}}.$$

(A.3)

Now, since y satisfies fulfilled-expectations, we can differentiate both sides of (3.9) with respect to p_A to get:

$$\frac{\partial y}{\partial p_A} u(E) \left(B(v, \frac{Ey}{N_B}) - B(v, \frac{E(1-y)}{N_A}) \right) + yEu(E)\frac{\partial y}{\partial p_A} \left(\frac{1}{N_B} B_2(v, \frac{Ey}{N_B}) + \frac{1}{N_A} B_2(v, \frac{E(1-y)}{N_A}) \right) = 1,$$

(A.4)

which implies that

$$\frac{\partial y}{\partial p_A} = \frac{1}{u(E) \left[B(v, \frac{Ey}{N_B}) - B(v, \frac{E[1-y]}{N_A}) + yE \left(\frac{1}{N_B} B_2(v, \frac{Ey}{N_B}) + \frac{1}{N_A} B_2(v, \frac{E(1-y)}{N_A}) \right) \right]}.$$

(A.5)

Similarly, differentiating both sides of (3.8) with respect to p_B yields:

$$\frac{\partial y}{\partial p_B} u(E) \left(B(v, \frac{Ey}{N_B}) - B(v, \frac{E(1-y)}{N_A}) \right) + yEu(E)\frac{\partial y}{\partial p_B} \left(\frac{1}{N_B} B_2(v, \frac{Ey}{N_B}) + \frac{1}{N_A} B_2(v, \frac{E(1-y)}{N_A}) \right) = -1,$$

(A.6)

which implies that

$$\frac{\partial y}{\partial p_B} = \frac{-1}{u(E) \left[B(v, \frac{Ey}{N_B}) - B(v, \frac{E[1-y]}{N_A}) + yE \left(\frac{1}{N_B} B_2(v, \frac{Ey}{N_B}) + \frac{1}{N_A} B_2(v, \frac{E(1-y)}{N_A}) \right) \right]}.$$

(A.7)

Now, differentiating $Q(y, N_A, N_B)$ with respect to y yields:

$$Q_1(y, N_A, N_B) = E \left(\frac{1}{N_A} B_2(v, \frac{E[1-y]}{N_A}) + \frac{1}{N_B} B_2(v, \frac{Ey}{N_B}) \right)$$

(A.8)

Comparing (A.5) and (A.7) with (A.8) and (3.3) yields:

$$\frac{\partial y}{\partial p_A} = \frac{1}{u(E)[Q(y, N_A, N_B) + yQ_1(y, N_A, N_B)]},$$

(A.9)

and

$$\frac{\partial y}{\partial p_B} = \frac{-1}{u(E)[Q(y, N_A, N_B) + yQ_1(y, N_A, N_B)]}.$$

(A.10)

28

The result follows by substituting (A.9) into (A.2) and (A.10) into (A.3).

Proof of Proposition 2: From (3.10) and (3.11) in Lemma 1, any equilibrium price pair (p_A, p_B) must satisfy

$$p_A = [1-y]u(E)[Q(y, N_A, N_B) + yQ_1(y, N_A, N_B)]. \tag{A.11}$$

$$p_B = yu(E)[Q(y, N_A, N_B) + yQ_1(y, N_A, N_B)]. \tag{A.12}$$

Therefore,

$$p_A - p_B = (1 - 2y)u(E)[Q(y, N_A, N_B) + yQ_1(y, N_A, N_B)]. \tag{A.13}$$

It is straightforward to establish that $(0, 0)$ is not an equilibrium, and therefore, $[Q(y, N_A, N_B) + yQ_1(y, N_A, N_B)] > 0$. As a consequence, since $u(E) > 0$, (A.13) implies that if $p_A = p_B$, then $y = \frac{1}{2}$.

(a) If $N_A = N_B = N$, then (3.5) reduces to

$$\Gamma(y) = \frac{N}{2N}, \text{ if } B(v, \frac{Ey}{N}) = B(v, \frac{E(1-y)}{N}). \tag{A.14}$$

From the properties of the Erlang loss function $B(v, \rho)$, we know that $B_2(v, \rho) > 0$, which in turn implies that the unique value of y for which $B(v, \frac{Ey}{N}) = B(v, \frac{E(1-y)}{N})$ is specified by $\frac{Ey}{N} = \frac{E(1-y)}{N}$, or $y = \frac{1}{2}$. This implies that when $N_A = N_B = N$, (3.5) is exactly the same as:

$$\Gamma(y) = \frac{1}{2}, \text{ if } y = \frac{1}{2}. \tag{A.15}$$

Therefore, (A.15) is consistent with a value of y such that $\Gamma(y) = y$. Substituting $N_A = N_B = N$ and $y = \frac{1}{2}$ into (3.10) and (3.11) yields the expression for p^*. It can easily be established that $\pi^A(x, p^*)$ and $\pi^B(p^*, x)$ are each strictly quasiconcave in x, and therefore, (p^*, p^*) is a Nash equilibrium.

(b) When $N_A \neq N_B$, there is no combination of N_A, N_B that can ensure that $\Gamma(\frac{1}{2}) = \frac{1}{2}$ if $p_A = p_B$. As a consequence, no symmetric price pair can be a Nash equilibrium.

Proof of Proposition 3: Recall that we have assumed $p_A \geq p_B$, and therefore, for an indifferent customer y, firm A has market share $(1-y)$ and firm B has market share y. Also recall that

the quality differential function, for any arbitrary y, N_A, N_B is:

$$Q(y, N_A, N_B) = B(v, \frac{Ey}{N_B}) - B(v, \frac{E[1-y]}{N_A}).$$
(A.16)

For reference, some partial derivatives of Q are listed below:

$$Q_1(y, N_A, N_B) = E \left(\frac{B_2(v, \frac{E[1-y]}{N_A})}{N_A} + \frac{B_2(v, \frac{Ey}{N_B})}{N_B} \right)$$

$$Q_2(y, N_A, N_B) = \frac{E[1-y]}{N_A^2} B_2(v, \frac{E[1-y]}{N_A})$$

$$Q_3(y, N_A, N_B) = \frac{-Ey}{N_B^2} B_2(v, \frac{Ey}{N_B})$$

$$Q_{11}(y, N_A, N_B) = E^2 \left(\frac{B_{22}(v, \frac{Ey}{N_B})}{N_B^2} - \frac{B_{22}(v, \frac{E[1-y]}{N_A})}{N_A^2} \right)$$

$$Q_{12}(y, N_A, N_B) = - \left(\frac{E}{N_A^2} B_2(v, \frac{E[1-y]}{N_A}) + \frac{E^2[1-y]}{N_A^3} B_{22}(v, \frac{E[1-y]}{N_A}) \right)$$

$$Q_{13}(y, N_A, N_B) = - \left(\frac{E}{N_B^2} B_2(v, \frac{Ey}{N_B}) + \frac{E^2 y}{N_B^3} B_{22}(v, \frac{Ey}{N_B}) \right)$$

The equilibrium second-period price pairs take the following form, as shown in Proposition 1:

$$p_A(N_A, N_B) = u(E)[1 - y(N_A, N_B)][Q(y(N_A, N_B), N_A, N_B)$$
(A.17)
$$+ y(N_A, N_B) Q_1(y(N_A, N_B), N_A, N_B)]$$

$$p_B(N_A, N_B) = u(E) y(N_A, N_B)[Q(y(N_A, N_B), N_A, N_B)$$
(A.18)
$$+ y(N_A, N_B) Q_1(y(N_A, N_B), N_A, N_B)]$$

where $y(N_A, N_B)$ satisfies $\Gamma(y) = y$, which implies that $y(N_A, N_B)$ satisfies:

$$y = [1 - 2y][\frac{Q(y, N_A, N_B) + y Q_1(y, N_A, N_B)}{Q(y, N_A, N_B)}],$$
(A.19)

or rearranging, implies that $y(N_A, N_B)$ satisfies the relationship:

$$[3y(N_A, N_B) - 1] Q(y(N_A, N_B), N_A, N_B) = [1 - 2y(N_A, N_B)] y(N_A, N_B) Q_1(y(N_A, N_B), N_A, N_B).$$
(A.20)

Differentiating both sides of (A.20) successively with respect to N_A and N_B and rearranging yields:

$$y_1(N_A, N_B) = \frac{[1-3y(N_A,N_B)]Q_1(y(N_A,N_B),N_A,N_B)+y(N_A,N_B)[1-2y(N_A,N_B)]Q_{12}(y(N_A,N_B),N_A,N_B)}{3Q(y(N_A,N_B),N_A,N_B)+[7y(N_A,N_B)-2]Q_1(y(N_A,N_B),N_A,N_B)-[1-2y(N_A,N_B)]y(N_A,N_B)Q_{11}(y(N_A,N_B),N_A,N_B)}$$

(A.21)

$$y_2(N_A, N_B) = \frac{[1-3y(N_A,N_B)]Q_1(y(N_A,N_B),N_A,N_B)+y(N_A,N_B)[1-2y(N_A,N_B)]Q_{13}(y(N_A,N_B),N_A,N_B)}{3Q(y(N_A,N_B),N_A,N_B)+[7y(N_A,N_B)-2]Q_1(y(N_A,N_B),N_A,N_B)-[1-2y(N_A,N_B)]y(N_A,N_B)Q_{11}(y(N_A,N_B),N_A,N_B)}$$

(A.22)

The payoff functions of the two firms are

$$\Pi^A(N_A, N_B) = u(E)[1 - y(N_A, N_B)]^2[Q(y, N_A, N_B) \tag{A.23}$$
$$+y(N_A, N_B)Q_1(y(N_A, N_B), N_A, N_B)] - c(N_A)$$
$$\Pi^B(N_A, N_B) = u(E)[y(N_A, N_B)]^2[Q(y, N_A, N_B) \tag{A.24}$$
$$+y(N_A, N_B)Q_1(y(N_A, N_B), N_A, N_B)] - c(N_B)$$

The first-order conditions are $\Pi_1^A(N_A, N_B) = \Pi_2^B(N_A, N_B) = 0$. *(For the remainder of this proof, we drop the variables in the expressions for brevity. For instance, we refer to $y_1(N_A, N_B)$ as y_1, $Q_{11}(y(N_A, N_B), N_A, N_B)$ as Q_{11} and so forth).*

Differentiating (A.23) with respect to N_A and (A.24) with respect to N_B yields:

$$\Pi_1^A = u(E)\{(1 - y)^2(Q_2 + yQ_{12}) + y_1(1 - y)^2(2Q_1 + yQ_{11}) \tag{A.25}$$
$$-2y_1(1 - y)(Q + yQ_1)\} - c_1; \tag{A.26}$$
$$\Pi_2^B = u(E)\{y^2(Q_3 + yQ_{13}) + y^2(y_2)(2Q_1 + yQ_{11}) + 2y(y_2)(Q + yQ_1)\} - c_1, \tag{A.27}$$

which when rearranged, yield the first-order conditions:

$$(1 - y)^2(Q_2 + yQ_{12}) + y_1\left[y(1 - y)^2Q_{11} + 2y(1 - y)(1 - 2y)Q_1 - 2(1 - y)Q)\right] = \frac{c_1}{u(E)} \tag{A.28}$$

$$y^2(Q_3 + yQ_{13}) + y_2\left[y^3Q_{11} + 4y^2Q_1 + 2yQ)\right] = \frac{c_1}{u(E)} \tag{A.29}$$

Recall from (A.21) and (A.22) that

$$y_1 = \frac{(1-3y)Q_1 + y(1-2y)Q_{12}}{3Q + (7y-2)Q_1 - y(1-2y)Q_{11}} \tag{A.30}$$

$$y_2 = \frac{(1-3y)Q_1 + y(1-2y)Q_{13}}{3Q + (7y-2)Q_1 - y(1-2y)Q_{11}} \tag{A.31}$$

and therefore we have the first order conditions purely in terms of our primitive functions:

$$(1-y)^2(Q_2 + yQ_{12}) + \frac{[(1-3y)Q_1 + y(1-2y)Q_{12}][y(1-y)^2 Q_{11} + 2y(1-y)(1-2y)Q_1 - 2(1-y)Q)]}{3Q + (7y-2)Q_1 - y(1-2y)Q_{11}} = \frac{c_1}{u(E)} \tag{A.32}$$

$$y^2(Q_3 + yQ_{13}) + \frac{[(1-3y)Q_1 + y(1-2y)Q_{13}][y^3 Q_{11} + 4y^2 Q_1 + 2yQ)]}{3Q + (7y-2)Q_1 - y(1-2y)Q_{11}} = \frac{c_1}{u(E)} \tag{A.33}$$

If $N_A = N_B = N$, this implies that $p_A = p_B = p$ (from Proposition 1) and that $y = \frac{1}{2}$. It therefore follows that:

$$Q_1 = \frac{2E}{N} B_2(v, \frac{E}{2N}) \tag{A.34}$$

$$Q_2 = -Q_3 = \frac{E}{2N^2} B_2(v, \frac{E}{2N}) \tag{A.35}$$

$$Q_{11} = 0 \tag{A.36}$$

$$Q_{12} = Q_{13} = -\left(\frac{E}{N^2} B_2(v, \frac{E}{2N}) + \frac{E^2}{2N^3} B_{22}(v, \frac{E}{2N}) \right) \tag{A.37}$$

and therefore:

$$\pi_1^A(N, N) = \frac{1}{4}(Q_2 + \frac{Q_{12}}{2}) - c_1 \tag{A.38}$$

$$\pi_2^B(N, N) = \frac{1}{4}(Q_3 + \frac{Q_{13}}{2}) - \frac{Q_1}{3} - c_1, \tag{A.39}$$

which upon substituting (A.35) and (A.37) yields:

$$\Pi_1^A(N, N) = -\frac{E^2}{16N^3} B_{22}(v, \frac{E}{2N}) - c_1(N) \tag{A.40}$$

$$\Pi_2^B(N, N) = -\frac{E^2}{16N^3} B_{22}(v, \frac{E}{2N}) - \left(\frac{E}{4N^2} + \frac{2E}{3N} \right) B_2(v, \frac{E}{2N}) - c_1(N). \tag{A.41}$$

Since $B_2(v, \rho) > 0$ for all v, ρ, the two first-order conditions cannot be satisfied simultaneously, and the result follows.