View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

UNIVERSITY OF SURREY

Discussion Papers in Economics

OPTIMAL ADMINISTERED INCENTIVE PRICING OF SPECTRUM

By

Paul Levine
(University of Surrey)
&
Neil Rickman
(University of Surrey)

DP 10/07

Department of Economics
University of Surrey
Guildford
Surrey GU2 7XH, UK
Telephone +44 (0)1483 689380
Facsimile +44 (0)1483 689548
Web www.econ.surrey.ac.uk
ISSN: 1749-5075

Optimal Administered Incentive Pricing of Spectrum*

Paul Levine
University of Surrey

Neil Rickman
University of Surrey & CEPR

October 13, 2007

Abstract

Administered Incentive Pricing (AIP) of radio spectrum as advocated by Smith/NERA (1996) and recently assessed by Indepen (2003) envisages an incremental path towards efficient pricing, with revealed and stated preference methods being used to reveal opportunity costs. We build on the latter to develop and optimal pricing scheme that allows for consumer surplus, interference constraints and their implications for productive efficiency, revenue implications and market structure. We demonstrate the subtle relationship between the interference constraints and the pricing and channel use decisions of network operators. We proceed to show that the optimal AIP is higher in sectors where spectrum can be shared and that it acts as Ramsey tax across sectors of the economy, i.e., is inversely related to the elasticity of demand. As a special case of our model we examine optimal pricing where the regulator is constrained to ignore the revenue implications. Then optimal spectrum prices are lower and the relationship between prices and the ability to share spectrum is reversed.

JEL Classification: L10, L50, L96

Keywords: radio spectrum, spectrum pricing, administered incentive pricing.

*We are grateful for financial support from the Radiocommunications Agency and Ofcom for the research reported in this paper. Any views expressed are those of the authors and do reflect those of Ofcom; nor did they, at any time, reflect those of the Radiocommunications Agency.

Contents

1 Introduction

Wireless communications services require radio spectrum as a factor of production. The range of frequencies available is finite so this input, as for other inputs, is a scarce resource. However, the nature of interference constraints that require different frequencies for some communications, but the possibility of the same shared frequency for others, gives this input into production a unique characteristic which has important implications in several areas of radio services. In this paper we focus on one of these – administered spectrum pricing.[1]

Until recently spectrum was priced to recover the costs of administering licence application and licencees. A drive towards more efficient use and valuation of spectrum led to Smith/NERA's "administered incentive pricing" (AIP) scheme in 1996. As a recent Indepen (2003) report makes clear, this attempts to price spectrum on the basis of opportunity cost across its alternative uses—the original idea here dates back to Levin (1970). The approach envisages an incremental path towards efficient pricing, with revealed and stated preference methods being used to reveal opportunity costs. Indepen (2004) provides a helpful summary of the assumptions underlying this approach. In particular, markets are perfect (market structure issues are not discussed) and spectrum allocation (or channel assignment) has already taken place. These, of course, are potential weaknesses of an optimal pricing regime that should aim to maximise welfare subject to various constraints (including those relating to channel interference). In addition, the approach provides little help for any regulator charged (as we understand Ofcom to be) with the task of encouraging competition—oligopolistic equilibria must be addressed here.

In this paper we show how to combine the above generalisations in order to derive an optimal price for spectrum. We incorporate interference constraints (using basic graph theory) in the allocation of spectrum and show how this can affect downstream

[1]Other areas include the benefits of cooperation between users in mesh networks that may avoid the tragedy of the commons in the licence-exempt sector, and the effects of spectrum trading (see Jones *et al.* (2004)).

retail prices. We also allow for oligopoly. Thus, the regulator in effect chooses both a spectrum licence price and the number of firms in the industry. In this setting, 'optimal AIP' must take account of consumer surplus, the reduction in tax distortions resulting from licence revenue, productive efficiency and the effect on market structure. We show that in such a setting, the optimal AIP is is higher in sectors where spectrum can be shared and that it acts as Ramsey tax across sectors of the economy, i.e., is inversely related to the elasticity of demand. As a special case of our model we examine optimal pricing where the regulator is constrained to ignore the revenue implications. Then optimal spectrum prices are lower and the relationship between prices and the ability to share spectrum is reversed. Moreover the Ramsey tax effect no longer applies. We show that the Smith/NERA AIP provides a incremental approach to changing spectrum prices that takes into account productive efficiency, but ignores the other effects that feature in our optimal pricing, namely the consumers' willingness to pay, revenue and imperfect competition.

The rest of the paper is organized as follows. Sections 2, 3 and 4 together provide a general framework for spectrum pricing that combines the productive and allocative efficiency aspects of spectrum allocation and therefore places the market valuation of (non-auctioned)[2] spectrum on a theoretically sound footing.[3] We ask how a regulator can allocate spectrum when account is taken of the possible interference between channels and sites. We show that the mathematical channel assignment problem (between an exogenous set of demands) can be nested within a wider economic channel allocation problem that endogenises these demands. The regulator is assumed to allocate spectrum on the basis of prices charged to users (i.e. firms). In what follows a *licence fee* for a given bandwidth of spectrum is a per-

[2]Auctions provide a mechanism for pricing spectrum when the regulator cannot tell how valuable spectrum is to operators. In this paper we go back a stage and study the complete information problem where there is no asymmetric information between the regulator and the firms. This provides an essential benchmark against which to assess mechanisms for addressing asymmetric information and it allows us to focus on hitherto neglected spatial aspects of spectrum allocation

[3]These sections draw upon, and develop, work in Leese *et al.* (2000).

period rent paid to use the spectrum in a particular geographical location or over several locations.[4] We show how these prices are determined by interactions between the economy (consumer demands, firms' production decisions and the regulator's preferences) and the technical requirements of channel assignment.

Section 5 then applies this framework to network licence pricing for the analytically tractable case of a linear technology for which spectrum is a pure complement and cannot be substitutes by capital or labour. We demonstrate the subtle relationship between the interference constraints and the pricing and channel use decisions of network operators and we derive results on optimal pricing referred to above. Section 6 relaxes this assumption and provides an outline of how optimal prices could be computed given a general production technology. Section 7 provides a comparison between our scheme and the AIP of Smith-NERA. Finally section 8 provides conclusions and an indication of possible future research.

2 The Channel Assignment Problem

Productive efficency implies the uses of inputs, including radio spectrum, to minimize the cost of particular goods and services. This aspect involves both the allocation of spectrum to particular broad use categories (e.g., broadcasting, fixed-wireless applications) and, within each use, the assignment of spectrum rights to particular users within a particular frequency band. The latter *channel assignment problem* involves the mathematical problem of how to assign channels to a competing, but pre-determined, set of demands while satisfying co-site and inter-site constraints imposed by the need to avoid excessive interference so as to either minimize the span required, or to use the maximum span and minimize some interference cost function (Leese, 1998).

[4]Note that we assume *linear* spectrum prices. A more complicated access structure would be for the regulator set a non-linear sliding scale of access charges.

We consider licences that apply to both a single market in a particular locality or to a network of markets over many localities.

Thus the channel assignment problem deals with the competing wishes to provide sufficient radio coverage while at the same time avoiding unacceptable interference between groups of transmitters (see Leese (1998)). The problem specification must therefore include information about the requirements for spectrum across the system, and also a set of constraints, designed to limit the interference levels, that a channel assignment should respect. In the version of the problem used most widely in practice, the spectrum requirements are given by specifying the number of distinct channels that each transmitter site requires. So, for instance, if there are n transmitter sites, called $T_1, T_2, \ldots, T_n$, then we have a corresponding set of demands $m_1, m_2, \ldots, m_n$, where site T_i requires the assignment of m_i distinct channels.

There are several ways of specifying these constraints. The commonest usual route, which reflects the use of protection ratios in the radio community, is to have a set of constraints each relating to either a single transmitter site (called co-site constraints), or a pair of transmitter sites (called inter-site constraints). To be explicit, suppose the channels are labelled by integer values, corresponding to their positions in the frequency spectrum. Then the co-site constraints require that if $f_1^{(i)}$ and $f_2^{(i)}$ are different channels both assigned to some transmitter site T_i then

$$|f_1^{(i)} - f_2^{(i)}| \geq \kappa_i$$

for some specified minimum channel separation κ_i. Likewise, the inter-site constraints require that if $f^{(i)}$ and $f^{(j)}$ are channels assigned to two different transmitter sites T_i and T_j, respectively, then

$$|f^{(i)} - f^{(j)}| \geq \kappa_{ij}$$

for some specified minimum channel separation κ_{ij} (equal to κ_{ji}). The constraints are therefore completely specified by the numbers κ_i and κ_{ij}, which are usually written in the form of a matrix, called the constraint matrix. The κ_i make up the diagonal entries and the κ_{ij} the off-diagonal entries.

The final part of the problem specification is the objective, for which there are two natural choices. The first and most widely studied to date is the *minimum* *span problem*, in which the aim is to find an assignment satisfying all spectrum requirements and all interference constraints, for which the span, defined as the difference between the highest and lowest channels used, is as small as possible. This would tend to be the concern of spectrum regulators and system designers. The second possibility, which we adopt in this paper, is the *fixed spectrum problem*, in which a maximum span is given (corresponding to the amount of spectrum available) and the aim is to assign channels to as many spectrum requirements as possible, within the given span and without violating any constraints. This would tend to be the concern of system operators, as they attempt to manage existing services. A variant on the fixed spectrum problem would assign channels to all spectrum requirements and try to minimize the number of violated constraints.

The above specification assumes that the transmitter locations and powers are fixed (they are effectively taken account of by the constraint matrix). More general formulations could have locations and powers as extra variables, to be optimized along with the channels, but there has been very little theoretical work on such problems to date.

The channel assignment problem has exercised many researchers over many years. The standard formulation includes, as a special case, the celebrated *graph-colouring problem.* A graph in this context is a collection of abstract 'nodes', some pairs of which are joined by 'edges'. The colouring problem is to attach a colour to each node in such a way that a pair of joined nodes should receive different colours and the total number of colours used should be as small as possible. The smallest number of colours needed is called the *chromatic number* of the graph. If we think of the nodes as transmitter sites and the colours as channels then we have precisely a minimum span channel assignment problem, in which all the m_i are 1, and the κ_{ij} are 1 if the nodes T_i and T_j are joined and 0 otherwise. (Since only one channel is required at each site, the values given to the co-site constraints κ_i are immaterial.) In physical terms, we are modelling only co-channel interference, with the edges in the graph indicating the rough location of potential coverage blackspots.

We now relate this very general formulation of the channel assignment and graph colouring problems to an economic model that explains demands for channels in terms of cost and market conditions. Each 'transmitter site' or node in the graph incorporates a local market consisting of an oligopoly of firms producing a service which we assume to be homogeneous. A spatial interpretation of nodes or sites is to regard them as equal cells (e.g. squares) comprising the region under consideration. A 'transmitter site' then consists of all the transmitters used by the firm, and it is possible that firms can share transmitters, perhaps charging an access price. We propose this as our 'core' economic model.

Each member of the oligopoly requires radio channels to provide the service. In each cell a local oligopoly provides a local service. Each firm within this market purchases a licence from the regulator to use a number of channels which depends on the volume of service. The proximity of firms implies that no channel re-use is possible within a cell and we assume that there is only co-channel interference (i.e. $\kappa_i = 1; \kappa_{ij} = 0$ or 1). The demands for radio channels in each market is the sum of individual demands of the firms in that market. We model these demands in the next section.

Each cell can now be given a colour and a shared colour indicates the possibility of 'channel re-use' or 'sharing' between these regions. Figures 1 and 2 illustrate this description of the channel assignment problem using a 4-node graph. The chromatic number of the graph in Figure 1. Figure 2 is the coloured map using two colours. Numbers inside circles are market demands and in our example demand increases as we proceed from West to East. Because diagonal squares can share channels the total demand of 60 channels can be serviced by a minimum of 40 distinct channels: say, 1-10 in market A, 11-30 in market B of which 11-20 can be re-used in market C. In market D 1-10 can be re-used from market A leaving a further distinct group of channels, say 31-40, to complete the radio channel requirements of these four markets, given the constraint matrix.

3 The Core Economic Model

We now turn to details of the core economic model. We consider a single local market with N competing firms providing a homogeneous service at a market price P.[5] Firm k produces output q_k, $k = 1, 2, \cdots, N$ and output $Q = \sum_{k=1}^{N} q_k$. The demand curve is given by $Q = D(P)$; $D'(P) < 0$ and we assume that $\lim_{P\to\infty} PD(P) = 0$. In what follows we write the inverse demand curve as $P = D^{-1}(Q) = P(Q)$ for short. Units of output are customer-minutes of some service requiring radio channels as an input per unit of time (say, the financial year).

Dropping the firm subscript for now, on the supply side labour (L), capital (K) and radio spectrum (Z) combine as inputs to produce output given a production. Let us first consider the following very general CES production function which we later specialize for reasons of tractability:

$$q = T\left[\gamma_1 L^{\eta} + (1-\gamma_1)[\gamma_2 Z^{\xi} + (1-\gamma_2)K^{\xi}]^{\eta/\xi}\right]^{\frac{1}{\eta}} \tag{1}$$

where T is a total factor productivity, a measure of technical efficiency. In (1) we have grouped capital and spectrum together with an elasticity of substitution equal to $\frac{1}{1-\xi}$. The elasticity of substitution between labour and the grouped inputs Z and K is $\frac{1}{1-\eta}$. Then if spectrum and capital are substitutes, but labour is a complement to the other inputs we would choose $\xi \in (0, 1)$ and $\eta < 0$.

Alternatively we could model spectrum as a complement to the other two substitutable inputs by grouping inputs as follows:

$$q = T\left[\gamma_1 Z^{\eta} + (1-\gamma_1)[\gamma_2 L^{\xi} + (1-\gamma_2)K^{\xi}]^{\eta/\xi}\right]^{\frac{1}{\eta}} \tag{2}$$

In the limit as η and ξ tend to 0, both (1) and (2) tend to the Cobb-Douglas form

$$q = TL^{\theta_1} Z^{\theta_2} K^{\theta_3}\,;\ \ \sum_{i=1}^{3} \theta_i = 1$$

Given a production function in one of these forms and given factor prices (w, r, a) per unit of labour, capital and spectrum respectively, we can formulate a *minimum*

[5] Later we introduce sectors and in each sector we allow firms to provide the service across a number of local markets.

cost function per unit of output $c(w, r, a)$ in the standard way. Associated factor demands per unit of output are $L(w, r, a)$, $K(w, r, a)$ and $Z(w, r, a)$. Standard analysis gives $\frac{\partial L}{\partial w}, \frac{\partial K}{\partial r}, \frac{\partial Z}{\partial a} < 0$. We assume that each firm is a price taker in factor markets and in the market for licences which incorporates all local markets such as the one modelled in this section. We assume that the price elasticity of demand in the market, $\epsilon(Q) = -\frac{PdQ}{QdP}$, is constant with respect to total output Q. We assume that $\epsilon > 1$ for reasons which will become apparent.

The sequence of actions is as follows:

1. The regulator sets the licence price for spectrum

2. Firms compete in the market given the licence prices and other factor prices.

The appropriate equilibrium concept is a subgame perfect equilibrium found by backward induction. We assume a Cournot-Nash equilibrium at stage 2 of the game. Thus in setting the licence price at stage 1 the regulator acts as a Stackelberg Leader. The next two sub-sections solve for stage 2 of the game, considering in turn the cases of an exogenous number of firms and then an endogenous number determined by a free entry zero-profit condition.

3.1 A Symmetric Equilibrium with an Exogenous Number of Firms

. Given the core model, profits for the *kth* firm are

$$\Pi_k = \Pi_k(q_k, w, r, a) = [P - c_k(w, r, a)]q_k - F \tag{3}$$

The firm's problem is to choose q_k to maximize Π_k.

Write total output $Q = q_k + \tilde{q}_k$. In a market-clearing equilibrium this is equated with total demand $D(P)$. Then in a Cournot-Nash equilibrium firm k takes the output of all other firms, $\tilde{q}_k$, as given along with the inverse demand curve $P = D^{-1}(Q) = D^{-1}(q_k + \tilde{q}_k)$ and the access price. Note that firms then act strategically with respect to other firms' choices of output. However at stage 2 of the game, as followers in a leader-follower game they are price-takers with respect to the access price a. This rules out strategic bidding for licences which the auction literature considers.

Firm k maximizes profits given by (3) with respect to q_k given $\tilde{q}_k$, a and the market-clearing condition $P = D^{-1}(Q) = D^{-1}(q_k + \tilde{q}_k)$ where we recall that $P = D^{-1}(Q)$ is the inverse demand curve. The first-order condition for profit-maximization is then

$$P'q_k + P - c_k(w, r, a) = 0 \tag{4}$$

and the second-order condition is

$$2P' + q_k P'' < 0 \tag{5}$$

Rearrangement of (4) gives the familiar mark-up pricing result for an oligopolist

$$P = \frac{c_k(w, r, a)}{1 - \frac{q_k}{\epsilon Q}} \tag{6}$$

where $\epsilon = \epsilon(Q) = -\frac{PdQ}{QdP}$ is the elasticity of demand. As the number of firms increases the price tends to the marginal cost (including the marginal cost of the channel), $c_k(w, r, a)$. Since we assume that firms are identical, $Q = Nq_k = Nq$, $c_k(w, r, a) = c(w, r, a)$ say. Then (6) becomes

$$P = \frac{c(w, r, a)}{1 - \frac{1}{\epsilon N}} = P(a, N) \tag{7}$$

say. Note that since factor prices w, r are exogenous in the model (determined by the general equilibrium in which the market model is embedded) we omit them as arguments in the price function (7). $\epsilon N > \epsilon > 1$ ensures the price is always positive and is also a second-order condition for profit-maximization. To see this write a constant elasticity demand curve as

$$Q = AP^{-\epsilon} \tag{8}$$

Differentiating twice with respect to Q we have

$$P'' = \frac{(\epsilon + 1)(P')^2}{P} \tag{9}$$

Substitute for P'' in (5) and also put $q_k = q = Q/N$. Then using $P' < 0$ and $\epsilon = -\frac{P}{Q}\frac{dQ}{dP}$, a little algebra gives the second order condition as

$$\epsilon > \frac{1}{2N-1} \tag{10}$$

Since $2N - 1 \leq 1$ for $N \geq 1$, clearly $\epsilon > 1$ is sufficient for (10) to be satisfied. The intuition behind this condition is that if the elasticity of demand is too low, then firms can allow the market price and profits to increase indefinitely by reducing output.

In an N-identical firm Cournot-Nash equilibrium output of each firm is given by $q = Q/N = D(P(a, N))/N$ with corresponding profits:

$$\Pi = \Pi(a, N) = [P(a, n) - c(w, r, a)]D(P(a, N))/N - F \tag{11}$$

This leads to our first proposition:

Proposition 1

Assume $\epsilon > 1$. Then profits $\Pi(a, N)$ decreases with respect to a and N.

Proof: See Appendix.

The intuition behind the proposition is as follows. An increase in the licence price increases cost and with that the retail price. Demand falls and if the elasticity of demand is greater than unity, revenue falls resulting in an decrease in profits. An increase in the number of firms reduces the mark-up and has two opposing effects on profits: the price falls increasing total revenue; but this revenue is now shared between more firms. With a constant elasticity of demand $\epsilon > 1$ the latter dominates and profits per firm fall.

3.2 A Free Entry Symmetric Equilibrium

Up to now we have taken the number of firms N as exogenous. There are two ways of making N endogenous. The first, is to make N a policy variable, chosen by the regulator when she issues the channel licences. The second, considered in this subsection, is to assume that there are no barriers to entry except a participation constraint that profits cannot be negative. Firms will then enter up to the point at which profits become negative.

The number of firms in equilibrium is given by N^* which, given the licence price, satisfies[6]

$$\Pi(a, N^*) = 0 \tag{12}$$

Since from Proposition 1, $\Pi(a, N)$ is decreasing in N and becomes negative for large N, if we assume that a monopolist would enjoy positive profits (ie $\Pi(a, 1) > 0$), then there exists a *unique* $N^*(a)$ satisfying (12). Furthermore differentiating (12) with respect to a and using proposition 1 we have

$$\frac{dN^*}{da} = -\frac{\frac{\partial \Pi}{\partial a}}{\frac{\partial \Pi}{\partial N}} < 0 \tag{13}$$

From Proposition 1, $\Pi(0, 1) > \Pi(a, 1) > 0$. Also from $\lim_{P\to\infty} PD(P) = 0$, $\Pi(a, 1)$ becomes negative for sufficiently large a. Hence even for a monopolist there exists an access price that would force profits below zero and drive the firm out of business.

Associated with $N^*(a)$ the total demand for channels $m(a)$ in the market is given by

$$m^*(a) = N^*(a)Z(a) \tag{14}$$

Differentiating (14) with respect to a we have:

$$\frac{dm(a)}{da} = \frac{dN^*}{da}Z(a) + N^*(a)\frac{dZ}{da} \tag{15}$$

Since $\frac{dN^*}{da} < 0$ and $\frac{dZ}{da} < 0$ we arrive at $\frac{dm(a)}{da} < 0$. We can now gather our results together as:

Proposition 2

(1) Given the access price a, and the total demand $D(P)$ there exists an unique number of firms $N^*(a)$ providing this service.

(2) $N^*(a)$ decreases with a.

[6] Note that free entry will lead to suboptimal duplication of fixed costs. See Perry (1984).

(3) There exists a sufficiently high access price that drives all firms out of business.

(4) Total demand for channels $m^*(a)$ decreases with a.

The final result here is the crucial one. Demand falls as access prices increase through two effects. First for a given number of firms, the price increases as firms raise prices with rising costs. Second, higher access prices force some firms out of business leaving the remaining ones with more market power. They then use this greater market power to raise their mark-up over the higher costs.

4 Optimal Spectrum Pricing

Now consider p sectors of the economy using radio spectrum and providing a homogeneous service. We assume services across sectors are neither substitutes nor complements. In each sector $i = 1, 2, \cdots, p$ there are ℓ_i local markets such as the one considered previously in the core model. Assume that there is no substitution by consumers between products provided in each sector. The assignment of spectrum is subject to interference constraints discussed in Section 3.1. Each local market requires a number of transmitters which are too close to share radio channels. We can treat these transmitters as one node in our previous discussion so a node represents both a market and a cluster of transmitters which can be considered as one transmitter on one site. The spectrum allocation problem can now be embedded in the following wider economic allocation problem:

1. Calculate the total demand for radio channels in each local market $j = 1, \cdots, \ell_i$ for each sector $i = 1, 2, \cdots, p$. Because there is no substitutability nor complementarity across sectors, assuming the same across markets demand in local market j of sector i will depend only the price P_{ij} evaluated according to the previous section.

2. We choose a standard measure of *social welfare function* W of the form:

$$W = \sum_{i=1}^{p}\sum_{j=1}^{\ell_i} S_{ij} + (1+\Lambda)R + \sum_{i=1}^{p}\sum_{j=1}^{\ell_i} \Pi_{ij} \tag{16}$$

   where S_{ij} is the *consumer surplus* in market j of sector i defined by:

$$S_{ij} = \int_{P_{ij}}^{\infty} D_{ij}(p)dp, \tag{17}$$

   R is the revenue from access prices given by

$$R = \sum_{i=1}^{p}\sum_{j=1}^{q_i} m_{ij}(a_{ij})a_{ij}, \tag{18}$$

   In (16), $1+\Lambda$ is the cost of public funds where $\Lambda > 0$ captures the distortionary effects of taxes that would otherwise be required in the absence of this revenue. Another interpretation of (16) is as a regulator's objective function that incorporates obligations imposed by law. Under UK and EU law pricing of spectrum must be limited to spectrum management considerations and must not be used as an instrument for raising taxes. This suggests that regulators must ignore the revenue term in (16) and therefore choose $\Lambda = -1$. In what follows we retain a general social\objective function of the form (16) and substitute $\Lambda > 0$ or $\Lambda = -1$ after the optimization exercise.

3. The social planner can now maximize (16) with respect to access prices $a_{ij}, i = 1, \ldots, p,\ j = 1, \ldots, \ell_i$ and the numbers of firms providing each service in each market, $N_j,\ j = 1, p$, subject to the engineering constraints discussed in Section 3.1 and the spectrum resource constraint.

$$\sum_{i=1}^{p}\sum_{j=1}^{\ell_i} m_{ij}(a_{ij}) \leq Z^{max} \tag{19}$$

   Licences are provided for a limited number of firms. In general profits will therefore not be driven down to zero by the process of free entry, and total profits must be included in the social welfare of a utilitarian regulator.

4. If the spectrum regulator is not responsible for market structure she would maximize (16) with respect to spectrum prices only subject to a free-entry equilibrium condition

$$\Pi_{ij}(a_{ij}, N_{ij}) = 0\,;\;\; i = 1, \cdots, p\,;\, j = 1, \cdots, \ell_i \tag{20}$$

and the engineering constraints and total spectrum constraint.

5. In principle the welfare maximizing regulator should choose different spectrum prices in each market to reflect different demand and cost conditions and the nature of the inter-market interference constraints. However practical concerns may dictate that the same spectrum price is administered in a sector (i.e., $a_{ij} = a_i$). Indeed, if channel trading between markets is allowed, arbitrage would constrain the regulator in this way.

6. In this very general set-up the number of firms varies across markets in each sector. We can think of these firms as *local operators*, each firm providing a service in a single market. An alternative is to consider each firm as a *network operators* providing a homogeneous service across all markets in the sector. Then in sector i and market j we have that $N_{ij} = N_i$. This assumption is adopted in the example that follows.

5 Optimal Spectrum Pricing with Linear Technology

To illustrate this framework we examine the optimal *network* licence price across sectors consisting of N_i network operators providing a service across ℓ_i local markets. Firms and markets are identical except for the interference constraints. Firms have access to a band of radio channels which can be used in all ℓ_i markets, interference constraints permitting. For each channel they pay a fixed licence price a_i payable to the regulator per unit of time. No channel sharing is possible within a local market.

First consider the case of $\ell_i = 3$ in a particular sector i. There are now four interference graphs to consider. The two most straightforward are two *homogeneous* cases: a complete graph which has edges between every pair of nodes and the other extreme of a graph with no edges. Less straightforward and more interesting are the *inhomogeneous* cases of a single edge graph and graph with two edges shown in Figure 3.

We first focus on the pricing and output decisions of network operators given the licence price in a particular sector. Until we come to examine sector-specific prices in subsection 5.4 below we drop the sector subscript i. To simplify the analysis in the following subsections, we modify the core economic model by specializing the production function to:

$$q = [\gamma_1 L^\eta + (1 - \gamma_1) K^\eta]^{\frac{1}{\eta}} \; ; \; z \geq q \tag{21}$$

Thus we consider a *linear technology* for spectrum for which spectrum is a pure complement and cannot be substituted by capital or labour; output q is produced using a Leontief technology. Without loss of generality we can choose units such that one unit of 'output' requires one radio channel and output capacity equals the total number of channels available. Thus for firm k to produce output q_{kj} per period in a particular local market $j = 1, 2, 3$, it requires $Z_k \geq z_{kj} \geq q_{kj}$ radio channel licences where Z_k is the total spectrum held and z_{kj} are the channels available to firm k in market j. These will depend on the nature of the interference graph. We assume that the licence fee a is independent of the firm and its location. Total costs include a set-up cost F so total costs for firm k are given by

$$C_k(\{q_{kj}\}, Z_k, a) = F + aZ_k + c \sum_{j=1}^{3} q_{kj} \tag{22}$$

where $c = c(w, r)$ is the cost function associated with the CES production function of labour and capital in (21).

As set out above, the sequencing of events is that the regulator first sets the licence price for spectrum in a particular sector i and second, firms compete in the market given the licence price a and other factor prices. Proceeding by backward induction, given factor prices (w, r) firm k chooses labour and capital to minimize the cost $c(w, r)q_{kj}$ of producing output q_{kj} in market $j = 1, 2, 3$. By choice of units and Leontief production this requires q_{kj} radio channels in market j. Prices in each market are P_j to reflect the radio environment. The firm purchases a licence for Z_k radio channels at a price a. Taking the relevant interference graph into consideration, this permits the firm to use $z_{kj} \leq Z_k$ channels in market j. In a Cournot-Nash equilibrium in each market, firm k then chooses outputs q_{kj} and radio channels Z_k to maximize profits

$$\Pi_k = \sum_{j=1}^{3} (P_j - c)q_{kj} - aZ_k - F \tag{23}$$

given the channel availability constraint

$$q_{kj} \leq z_{kj} \leq Z_k \tag{24}$$

and given the output of other firms $\tilde{q}_{kj}$ in markets $j = 1, 2, 3$.

5.1 Homogeneous Graphs

These are the most straightforward cases. For the graph with no edges, all channels are available in each market and therefore $z_{kj} = Z_k$. Let $\lambda_j \geq 0$ be the shadow price associated with the constraint (24). Then firm k maximizes a Lagrangian

$$\mathcal{L}_k = \Pi_k - \sum_{j=1}^{3} \lambda_j (q_{kj} - Z_k) \tag{25}$$

with respect to q_{kj}, Z_k, $\{\lambda_k\}$ given the corresponding decisions of other firms. The first order condition are for $j = 1, 2, 3$

$$q_{kj} \ : \ (P_j - c) + q_j P_j' - \lambda_j = 0 \tag{26}$$

$$Z_k \ : \ -a + \sum_{j}^{3} \lambda_j = 0 \tag{27}$$

$$\text{CS} \ : \ \lambda_j (q_{kj} - Z_k) = 0 \tag{28}$$

Equation (26) equates the marginal return from providing output in each market with the shadow price of spectrum in that market. Equation (27) equates the price of spectrum with the network shadow price. (42) are the Kuhn-Tucker complementary slackness conditions. If a constraint does not bind the corresponding shadow price takes a zero value.

The solution for this homogeneous case is very simple. By symmetry $\lambda_j = \lambda = \frac{a}{3} > 0$ and the constraints bind. Proceeding as before in a symmetric Nash equilibrium, the Lerner price in all markets is given by

$$P_j = P = \frac{c + \frac{a}{3}}{1 - \frac{1}{\epsilon N}} \tag{29}$$

Output is given by $Q = AP^{-\epsilon}$ in each market, profits $\Pi(a, N) = (P - c - \frac{a}{3})\frac{Q}{N} - F$ per market. Then either a free-entry condition $\Pi(a, N) = 0$ or a regulator's choice of the number of firms to be allowed licences determines N.

The other homogeneous case where all nodes are joined is very similar. Now given Z_k, $z_{kj} = \frac{Z_k}{3}$ channels are available per market. By a similar analysis we then arrive at the price in each market given by

$$P_j = P = \frac{c + a}{1 - \frac{1}{\epsilon N}} \tag{30}$$

Effectively firms are now paying more for their spectrum because they cannot share the costs across markets, so from Proposition 1 profits in this case are lower and less firms will enter this market in free entry equilibrium.

5.2 A Graph with Two Edges

In Figure 6 let markets $j = 1, 2, 3$ be at nodes A, B and C. Then for firm k if all Z_k channels are available in market 1 and $q_{k1} \leq Z_k$ are used, then $z_{k2} = z_{k3} = Z_k - q_{k1}$ are available in markets 2 and 3. Then firm k maximizes a Lagrangian

$$\mathcal{L}_k = \Pi_k - \lambda_1(q_{k1} - Z_k) - \lambda_2(q_{k2} - Z_k + q_{k1}) - \lambda_3(q_{k3} - Z_k + q_{k1}) \tag{31}$$

with respect to q_{kj}, Z_k, λ_k given the corresponding decisions of other firms. The first order condition are

$$\begin{aligned} q_{k1} &: (P_1 - c) + q_1 P_1' - \sum_j^3 \lambda_j = 0 && (32) \\ q_{k2} &: (P_2 - c) + q_2 P_2' - \lambda_2 = 0 && (33) \\ q_{k3} &: (P_3 - c) + q_3 P_3' - \lambda_3 = 0 && (34) \\ Z_k &: -a + \sum_j^3 \lambda_j = 0 && (35) \\ \text{CS} &: \lambda_{kj}(q_{kj} - z_{kj}) = 0 && (36) \end{aligned}$$

The solution to these conditions sees market 1 releasing spectrum for the other two markets k so $\lambda_1 = 0$. Spectrum is fully utilized in markets 2 and 3 so $\lambda_2, \lambda_3 > 0$. In symmetric Cournot-Nash equilibria, following the same reasoning as before we arrive at equilibrium prices:

$$\begin{aligned} P_1 &= \frac{c + a}{1 - \frac{1}{\epsilon N}} && (37) \\ P_2 &= P_3 = \frac{c + \frac{a}{2}}{1 - \frac{1}{\epsilon N}} && (38) \end{aligned}$$

Thus we can see that prices are lower in markets where spectrum can be shared. In a sector characterized by two-edged graphs prices will be lower in some markets than in sectors with complete graphs but profits will be higher (again by appeal to Proposition 1). It follows that in the less congested sector (in the radio interference sense) higher profits will encourage more entry and markets will be more competitive.

5.3 A Graph with One Edge

In market 1 (node A) all channels are available so $z_{k1} = Z_k$. In markets 2 and 3 channels can be shared and $z_{k2} = z_{k3} = \frac{Z}{2}$. Now firm k maximizes a Lagrangian

$$\mathcal{L}_k = \Pi_k - \lambda_1(q_{k1} - Z_k) - \lambda_2(q_{k2} - \frac{Z_k}{2}) - \lambda_3(q_{k3} - \frac{Z_k}{2}) \quad (39)$$

with respect to q_{kj}, Z_k, λ_k given the corresponding decisions of other firms. The first order condition are for $j = 1, 2, 3$

$$
\begin{aligned}
q_{kj} &: (P_j - c) + q_j P_j' - \lambda_j = 0 \qquad (40) \\
Z_k &: -a + \lambda_1 + \frac{1}{2}(\lambda_2 + \lambda_3) = 0 \qquad (41) \\
\text{CS} &: \lambda_{kj}(q_{kj} - z_{kj}) = 0 \qquad (42)
\end{aligned}
$$

We first solve for a *Type I equilibrium where all constraints bind* ($\lambda_j > 0$, j=1,2,3). By symmetry we know that $\lambda_2 = \lambda_3$. From the first order conditions, the solution then satisfies

$$
\begin{aligned}
P_1 &= \frac{c + \lambda_1}{1 - \frac{1}{\epsilon N}} \qquad (43) \\
P_2 = P_3 &= \frac{c + a - \lambda_1}{1 - \frac{1}{\epsilon N}} \qquad (44) \\
D(P_1) &= 2D(P_2) \text{ where } D(P) = AP^{-\epsilon} \qquad (45) \\
\lambda_2 = \lambda_3 &= a - \lambda_1 \qquad (46)
\end{aligned}
$$

Solving this set of equations for P_1, P_2, λ_1 and λ_2, a little algebra gives

$$
\begin{aligned}
P_1 &= \frac{1}{(1 + 2^\epsilon)} \left[\frac{2c + a}{1 - \frac{1}{\epsilon N}} \right] \qquad (47) \\
P_2 = P_3 &= \frac{2^\epsilon}{(1 + 2^\epsilon)} \left[\frac{2c + a}{1 - \frac{1}{\epsilon N}} \right] \qquad (48) \\
\lambda_1 &= \frac{c(1 - 2^\epsilon) + a}{1 + 2^\epsilon} \qquad (49) \\
\lambda_2 = \lambda_3 &= \frac{(2^\epsilon - 1)(a + c)}{(1 + 2^\epsilon)} \qquad (50)
\end{aligned}
$$

It follows from this solution that $\lambda_2 = \lambda_3 > 0$ if $\epsilon > 1$ which we have already imposed, but that $\lambda_1 > 0$ imposes a condition on the licence price

$$a > c(2^\epsilon - 1) \qquad (51)$$

If (51) does not hold then we must have a *Type II equilibrium where the capacity constraint in market 1 does not bind* and there is are spare radio channels. Then

$\lambda_1 = 0$ and

$$P_1 = \frac{c}{1 - \frac{1}{\epsilon N}} \tag{52}$$

$$P_2 = P_3 = \frac{c + a}{1 - \frac{1}{\epsilon N}} \tag{53}$$

$$\lambda_2 = \lambda_3 = a \tag{54}$$

These results for a single edge graph in particular highlight the subtle relationship between the interference constraints and the pricing and channel use decisions of network operators. Iff the regulator sets a sufficiently high licence price such that (51) holds all channels will be fully utilized in each market. The drawback is that prices will be higher directly through the effect of the licence price on the retail Lerner index and indirectly through higher concentration in a free-entry equilibrium.

The mathematical framework developed for analyzing these examples can be used to develop software capable of handling much larger problems. However the small node-number examples considered here are sufficient to demonstrate that the spatially distributed aspects of channel assignment problems provide new challenges for analysis that go beyond standard economic treatments.

5.4 The Optimal Licence Price

We now turn to the regulator's choice of an optimal licence price for a particular sector i (recall we have suppressed the sector i subscript in the preceding analysis). We consider a free entry equilibrium where firms enter until profits are driven down to zero. For analytical convenience, we confine ourselves to homogeneous graphs, but unlike the previous section we now generalize the analysis to any number of local markets ℓ_i in sector i. Furthermore as in the previous section we continue to consider network operators providing a homogeneous service across ℓ_i local markets in sector i, but with homogeneous graphs the regulator's problem is identical if we assume local operators.[7] N_i network operators or firms are identical in each

[7] Let $\underline{a}^L$ be the licence price for local operators. Putting $\underline{a} = \ell \underline{a}^L$ we then arrive at an identical optimization problem described below.

sector i and demand Z_i radio channels at a licence price a_i to now be determined. The revenue in sector i is therefore $N_iZ_ia_i$ and since graphs are homogeneous, retail prices and identical in a particular sector across local markets; i.e., $P_{ij} = P_i$. Putting $S_{ij} = S_i = \int_{P_i}^{\infty} D_i(P)dP$ the regulator's problem set out in general form in section 3 now becomes one to maximize with respect to $\underline{a} = (a_1, a_2, \cdots, a_p)$ the social welfare function

$$W = W(\underline{a}) = \sum_{i=1}^{p} \ell_i \int_{P_i}^{\infty} D_i(P)dP + (1+\Lambda) \sum_{i=1}^{p} N_iZ_ia_i \tag{55}$$

subject to

$$\sum_{i=1}^{p} N_iZ_i \leq Z^{max} \tag{56}$$

and the interference constraints implied by the graphs.

Confining ourselves to homogeneous graphs, there are two types of sectors to consider. Those with graphs consisting of ℓ_i nodes all connected to each other are referred to as sectors *without spectrum re-use.* Then the demand for spectrum is $N_iZ_i = \ell_iD_i(P_i)$ and from the previous sub-section the retail price set by the firm is given by $P_i = \frac{c_i+a_i}{1-\frac{1}{\epsilon_iN_i}}$. Graphs consisting of ℓ nodes without any edges are referred to as sectors *with spectrum re-use.* Then the demand for spectrum is $N_iZ_i = D_i(P_i)$ and from the previous sub-section the retail price set by the firm is given by $P_i = \frac{c_i+\frac{a_i}{\ell_i}}{1-\frac{1}{\epsilon_iN_i}}$. Define $k_i = 1$ for sectors without spectrum re-use and $k_i = \frac{1}{\ell_i}$ for sectors with spectrum re-use. Then (55) can be written as

$$W = W(\underline{a}) = \sum_{i=1}^{p} \ell_i \left[\int_{P_i}^{\infty} D_i(P)dP + (1+\Lambda)k_iD_i(P_i)a_i \right] \tag{57}$$

where

$$P_i = \frac{c_i + k_ia_i}{1 - \frac{1}{\epsilon_iN_i}} \tag{58}$$

The regulator maximizes (57) with respect to $\underline{a}$ subject to the spectrum resource constraint

$$\sum_{i=1}^{p} k_i\ell_iD_i(P_i) \leq Z^{max} \tag{59}$$

and the interference constraints. Within sectors these are that channels cannot be shared between firms in a local market but within the firm, as a network operator, it can share between markets in sectors with spectrum re-use ($k_i = \frac{1}{\ell_i}$), but not in sectors without spectrum re-use ($k_i = 1$). Within sectors the regulator then assigns different channels to each firm and imposes constraints across markets if appropriate (though firms, not wanting to cause interference between their own sites, might self-impose such constraints). Given the licence price firms then compete making entry (or exit) decisions and resulting in a retail price in each sector given by (58). Between sectors we assume that harmonisation agreements prevent the possibilities of sharing spectrum.

To carry out the regulator's optimization problem define a Lagrangian

$$\mathcal{L} = W(\underline{a}) - \mu(\sum_{i=1}^{p} k_i \ell_i D_i(P_i) - Z^{max}) \tag{60}$$

where μ is a Lagrange multiplier associated with the spectrum resource constraint (i.e., the shadow price of spectrum). Writing (58) $P_i = P_i(a_i, N_i(a_i))$ as in 3.2.1, the first order condition with respect to a_i is

$$-D_i(P_i)\frac{dP_i}{da_i} + (1+\Lambda)k_i\left(D_i(P_i) + a_i D'(P_i)\frac{dP_i}{da_i}\right) = \mu k_i D'(P_i)\frac{dP_i}{da_i} \tag{61}$$

The left-hand-side of (61) is the marginal benefit from a marginal increase in the spectrum price from increased revenue (the second term) minus the marginal cost from a drop in consumer surplus (the first term). The right-hand-side is the marginal cost of spectrum evaluated at its shadow price μ.[8]

To proceed with the analysis we need to evaluate $\frac{dP_i}{da_i}$ using the free-entry condition $\Pi_i = \frac{\ell_i P_i D_i(P_i}{\epsilon_i N_i^2} - F_i = 0$ (using (A.1) from the appendix). Differentiating $P_i = P_i(a_i, N_i(a_i))$ we have

$$\frac{dP_i}{da_i} = \frac{\partial P_i}{\partial a_i} + \frac{\partial P_i}{\partial N_i}\frac{dN_i}{da_i} \tag{62}$$

The first term on the right-hand-side of (62) is the direct effect of an increase in the licence price on the retail price. The second term is an indirect effect arising from

[8]Notice that the expression for the optimal licence price in each sector depends only on demand and supply conditions in that sector and the shadow price of spectrum, not on conditions in other sectors. The reason for this convenient decomposition is the assumed absence of any substitutability or complementarity of services between sectors.

the exit of firms as the cost spectrum rises. Differentiating the free-entry condition a little algebra gives

$$\frac{dN_i}{da_i} = -\frac{(\epsilon_i - 1)(\epsilon_i N_i - 1)}{(2\epsilon_i N_i - 1 - \epsilon_i)} \frac{N_i \partial P_i}{P_i \partial a_i} < 0 \tag{63}$$

for $\epsilon_i > 1$ which we have assumed throughout. Substituting into (62) we can now write

$$\frac{dP_i}{da_i} = \rho_i \frac{\partial P_i}{\partial a_i} \tag{64}$$

where we have defined

$$\rho_i = \rho_i(\epsilon_i, N_i) = \frac{2(\epsilon_i N_i - 1)}{(2\epsilon_i N_i - 1 - \epsilon_i)} \tag{65}$$

Note that as $N_i \rightarrow \infty$ and markets become competitive, $\rho_i \rightarrow 1$. Given N_i, as ϵ_i varies in the range $\epsilon_i \in [1, \infty]$ then $\rho_i \in \left[1, \frac{2N_i}{2N_i - 1}\right]$.

Dividing (61) by $D_i(P_i)$ and using $\epsilon_i = -\frac{P_i dD_i}{D_i dP_i}$ and (64) we can, after some algebra, write the first-order condition as

$$a_i = \frac{\mu \epsilon_i \rho_i \left(1 - \frac{1}{\epsilon_i N_i}\right) + \frac{c_i}{k_i}\left((1+\Lambda)\left(1 - \frac{1}{\epsilon_i N_i}\right) - \rho_i\right)}{\rho_i + (1+\Lambda)(\epsilon_i \rho_i - 1)\left(1 - \frac{1}{\epsilon_i N_i}\right)} \tag{66}$$

This relationship along with the free-entry condition

$$\Pi_i = \frac{\ell_i P_i D_i(P_i)}{\epsilon_i N_i^2} - F_i = 0 \tag{67}$$

the Kuhn-Tucker complementary slackness condition

$$\mu\left(\sum_{i=1}^{p} k_i D_i(P_i) - Z^{max}\right) = 0 \tag{68}$$

and the retail price equation (58) gives four equations for N_i, a_i, P_i and μ at the optimum, given parameters ϵ_i, c_i, k_i, Λ and A_i. This completes the solution for the optimal licence price.

From (66) an important result linking the optimal licence price to the extent of congestion captured by our parameter k_i is apparent. Given the number of firms N_i (66) says that the optimal price *falls* as congestion increases and the $a_i(N_i)$ shifts downwards Furthermore since from (58), congestion effectively increases the spectrum price per unit of output, so the downward-sloping free-entry relationship $N_i(a)$ shifts to the left. These two effects are illustrated in figure 4. We summarize this result as:

Proposition 3.

Assuming linear technology for spectrum and homogeneous graphs, ceteris paribus the optimal licence price in sectors without spectrum re-use is lower than that in sectors with spectrum re-use.

The result that the inability to re-use spectrum is not accompanied by a 'congestion tax' consisting of a higher spectrum licence price may seem counter-intuitive. The reason why a higher licence price is not necessary to reduce demand for spectrum in sectors without spectrum re-use is that network operators perform this function by raising the retail price. This is illustrated in figures 5 and 6 for the cases where spectrum is scarce ($\mu > 0$) and where it is not ($\mu = 0$) respectively.[9] If spectrum is scarce the retail price rises as k_i rises from $k_i = \frac{1}{\ell_i}$ for homogeneous graphs without edges (where spectrum can be shared across local markets) to $k_i = 1$ for homogeneous graphs with every node connected where no channel sharing is possible. For example if $\ell_i = 5$ we need to compare $k_i = \frac{1}{5}$ and $k_i = 1$ for cases of sectors with spectrum re-use and without spectrum re-use respectively. As we move from these two cases, if $\mu > 0$ and spectrum is scarce then in our illustrative example the retail price more than doubles, the number of firms drops from 5 to 3. The regulator who is equally concerned for the welfare of consumers in sectors with and without spectrum re-use, cushions the effect of the latter by *lowering and not raising* the licence price in that sector relative to the sector where spectrum re-use is possible (a drop from around 2.5 to 2.2 in our example). If spectrum is plentiful ($\mu = 0$) we can see from (66) that $a_i k_i$ and therefore from (58) the retail price is independent of the parameter k_i. It follows from the free-entry condition that this must also be true of N_i and therefore a_i is simply proportional to $\frac{1}{k_i}$, as in apparent in figure 6.

[9]Baseline parameter values are $c_i = A_i = 1$, $\Lambda = 0.3$, $\epsilon_i = 2$ and $\ell_i = 3$. Fixed entry costs F_i are chosen so that at baseline parameter values in sectors without spectrum re-use ($k_i = 3$), $N_i = 4$.

A number of simpler expression for special cases provide useful insights. First let N_i become large. Then for this competitive case since $\rho \to 1$, (66) becomes

$$a_i \to \frac{\mu\epsilon_i + \frac{c_i}{k_i}\Lambda}{(1+\Lambda)\epsilon_i - \Lambda} \tag{69}$$

From this result it follows that for $\Lambda > 0$, $\frac{da_i}{d\epsilon_i} < 0$ a result that is confirmed numerically in figure 7 for any N_i. The result that the licence price is inversely related to the elasticity of demand is a familiar property of a *Ramsey price* from the regulation literature that applies to optimal spectrum licence pricing as long a taxes are distortionary ($\Lambda > 0$). To summarize:

Proposition 4.

For large N_i the inverse elasticities rule of Ramsey prices and taxes apply: the optimal license price in a sector is inversely related to the elasticity of demand. Numerical results suggest this may hold for small N_i.

Finally if in addition to N_i being large, $\Lambda = 0$ and there are no distortionary effects from taxation, then $a_i = \mu$ and the optimal licence price is simply equal to the shadow price of spectrum and independent of all the characteristics that distinguish the sectors.

Up to this point we have reported results for the case where the regulator maximizes a social welfare function with $\Lambda > 0$ that (correctly) assumes there are welfare benefits from the revenue from licence fees. We now assume that the regulator is constrained by law to ignore these benefits. Our general framework handles this case if we put $\Lambda = -1$. The competitive case where N_i is large is straightforward to analyze. Then (69) becomes

$$a_i \to \mu\epsilon_i - \frac{c_i}{k_i} \tag{70}$$

It is now apparent that the relationship between the optimal licence price under this constraint and both the spectrum re-use parameter k_i and the elasticity of demand ϵ_i is now the opposite of that reported in propositions 3 and 4. In other words these results depend critically on revenue generation considerations. Thus we have the proposition:

Proposition 5

If the regulator adopts an objective function that ignores the welfare benefits of revenue from licence fees the the constrained optimal licence price in sectors without spectrum re-use is now *greater*. For large N_i, the constrained optimal licence price in sectors with spectrum re-use is *positively* related to the elasticity of demand.

Figure 8 for $\Lambda = -1$ demonstrates the first part of this result and corresponds to figure 4 for $\Lambda = 0.3$. Figure 9 for $\Lambda = -1$ that the relationship between he constrained optimal licence and the elasticity of demand becomes ambiguous for the oligopoly case.

5.5 The Incorporation of Costs of Adjusting Licence Prices

A feature of the administered incentive pricing (AIP) scheme discussed and compared with our optimal pricing scheme in a later section is that it is incremental, gradually adjusting towards a more efficient allocation of the spectrum. This feature can be introduced into optimal pricing by incorporating *adjustment costs* of changing the licence price. The optimization problem now is no longer static. Let $\underline{a}^*$ be the vector of optimal prices that solves the static problem set out in the previous section. Consider the *intertemporal* welfare loss function

$$\Omega = \sum_{t=1}^{\infty} \beta^t \sum_{i=1}^{p} [(a_i^* - a_i)^2 + \Phi(a_i(t) - a_i(t-1))^2] \tag{71}$$

where $\beta \in (0, 1]$ is a discount factor. This welfare loss expresses the idea that the regulator would prefer to be at the static optimal set of prices but is subjected to costs of change prices proportion to $(\Delta a(t))^2$ where $\Delta a_i(t) = a_i(t) - a_i(t-1)$ is the change in the licence price over the interval $[t-1, t]$ in sector i. As the cost of adjustment parameter $\Phi \to 0$ we approach the previous problem where the regulator instantaneously jumps to the static optimum $\underline{a} = \underline{a}^*$.

At time $t = 1$, the regulator now minimizes Ω with respect to $\underline{a}$ given historical prices $\underline{a}(0)$. The first-order condition for a minimum in sector i is given by

$$-(a_i^* - a_i(t)) + \Phi(a_i(t) - a_i(t-1) - \Phi\beta(a_i(t+1) - a_i(t)) = 0 \tag{72}$$

Let $\hat{a}_i(t) = a_i(t) - a_i^*$ be the deviation of the licence price about the static optimum. Then the first-order condition can be written as

$$\beta\Phi\hat{a}_i(t+1) - (1 + \Phi(1+\beta))\hat{a}_i(t) + \Phi\hat{a}_i(t-1) = 0 \tag{73}$$

To solve this second-order difference equation take z-transforms to give the characteristic equation

$$\beta\Phi z^2 - (1 + \Phi(1+\beta))z + \Phi = 0 \tag{74}$$

It is straightforward to show that this has two positive roots $z_1 < 1$ and $z_2 > 1$. The system is therefore saddle-path stable with solution

$$\hat{a}_i(t) = z_1^t \hat{a}_i(0) \tag{75}$$

Figure 10 illustrates this adjustment process with costs of adjustment by plotting the solution (75) for a low cost, medium cost and high cost cases. For the sector in question the optimal static spectrum price as assumed to be unity and the initial price is zero. As one would expect as costs of adjustment are lowered, the speed of adjustment to the static optimum increases. More generally this section shows how a rational regulator facing adjustment costs would act in an incremental fashion moving quickly or slowly toward the static optimum, depending on the size of the costs of adjustment, so mimicking the adjustment process advocated by AIP.

6 Optimal Pricing with General Technology

In the previous subsection for reasons of tractability we confined ourselves to linear technology where there was no scope for substituting spectrum for other factors of production. Now we sketch out how optimal pricing may be formulated for more general production technologies which do allow for alternative spectrum saving ways of producing services such as those expressed by production functions (1) or (2). The total costs of producing q_i output in sector i with spectrum costing a_i now becomes

$$C_i(q_i, a_i) = F_i + c_i(w_i, r_i, k_i a_i)q_i \tag{76}$$

where $c_i(w_i, r_i, k_i a_i)$ is the cost per unit of output and $k_i a_i$ is the effective cost of spectrum for network operators. By Shephard's Lemma we have that the demand for spectrum in sector i is given by

$$Z_i = \frac{\partial C_i}{\partial a_i} = \ell_i k_i c_{3i} \frac{D_i(P_i)}{N_i} \tag{77}$$

where $c_{3i} = \frac{\partial c_i(w_i, r_i, x)}{\partial x}$.

We can now generalize the regulator's problem expressed by (57) to (59) for linear technology to the maximization with respect to $\underline{a}$ of

$$W = W(\underline{a}) = \sum_{i=1}^{p} \ell_i \left[\int_{P_i}^{\infty} D_i(P)dP + (1+\Lambda) k_i c_{3i}(w_i, r_i, k_i a_i) D_i(P_i) a_i \right] \tag{78}$$

where

$$P_i = \frac{c_i(w_i, r_i, k_i a_i)}{1 - \frac{1}{\epsilon_i N_i}} \tag{79}$$

subject to the spectrum resource constraint

$$\sum_{i=1}^{p} k_i c_{3i}(w_i, r_i, k_i a_i) D_i(P_i) \leq Z^{max} \tag{80}$$

and the interference constraints.

To implement this procedure we require either a production function for the service using spectrum is required or the cost function (76). Empirically the latter is usually the preferred method of estimating factor demands using, for example, translog functional forms. Unlike the previous case of linear technology the first order conditions are not analytically tractable and would require numerical solution using an estimated or calibrated production or cost function.[10]

7 AIP versus Optimal Prices

We are now in a position to compare the 'Administered Incentive Prices' (AIP) proposed by Smith-NERA (1996) with optimal prices. To examine the latter assume

[10]Even for the simplest non-linear technology, a Cobb-Douglas production function, the first-order condition corresponding to (66) is a high-order polynomial in the licence price a_i.

there are 2 sectors ($p = 2$) with cost functions per unit of output (in value terms) for the representative firm:

$$c_i(w, r, a) = wL_i(w, r, a) + rK_i(w, r, a) + aZ_i(w, r, a)\,;\; i = 1, 2 \tag{81}$$

If firms are unconstrained in the choice of inputs then $c_i(w, r, a)$ are the *minimum* costs chosen by the firms given factor prices (w, r, a). Suppose however that the licence price is too low to clear the market so at least one of the firms is short of the spectrum it needs to achieve this minimum cost. Suppose firm 1 is short of spectrum but firm 2 is unconstrained. Consider a small incremental increase in spectrum for firm 1, ΔZ_1 which it substitutes for ΔL_1 of labour and ΔK_1 of capital. i.e., the change in cost is

$$\Delta c_1 = -w\Delta L_1 - r\Delta K_1 + a\Delta Z_1 < 0 \tag{82}$$

whereas for the unconstrained firm 2 the change in cost is

$$\Delta c_2 = w\Delta L_2 + r\Delta K_2 - a\Delta Z_1 > 0 \tag{83}$$

since $\Delta Z_2 = -\Delta Z_1$ and firm 2 has minimized costs before the reallocation. $\mid \Delta c_1 \mid$ is the Smith-NERA 'opportunity cost' or what society forgoes by allocating spectrum to sector 2 rather than sector 1.

According to the Smith-NERA methodology, licence prices are based on available estimates of the costs of alternative uses of the radio spectrum. Initial valuations and the subsequent modifications take qualitative factors affecting spectrum use into account. With relevant data largely unavailable for estimating marginal profit values, the Smith-NERA (1996) study turned to an approach based marginal values of the costs of alternatives. One of the more straight forward cases is that of terrestrial fixed links. For this sector the a number of methods of relieving excess demand for spectrum in fixed links bands including: installing narrowband equipment, moving to higher frequencies, moving to cable and releasing spectrum from other uses sharing the same bands.

For terrestrial fixed links, using the spectrum valuation model, Smith-NERA used the following three-step procedure :

1. Establish what or who is the next best use/user of the spectrum: In lower bands (e.g. below 2GHz) this might be mobile radio. In higher bands this might be other fixed link users, satellite links etc.

2. Determine what can the next best use/user do if access to spectrum is denied

   (a) **Mobile**: The options include move to different frequency band, use narrowband equipment, implement more cells, change to different operator.

   (b) **Fixed Link**: If all access to spectrum is denied then fixed link users will be forced to move to cable. Otherwise, users can install narrowband equipment or, where this is not feasible, move to a different frequency band.

   (c) **Satellite Earth Station**: Move to a different geographical area and possibly a different country.

3. Take the minimum of these costs. Then the change in costs is our equation (83).

This description of the theory and practice of AIP has two fundamental features: first it is incremental based on observations of inefficient allocations. Second it is solely concerned with the input side of efficiency and ignores the consumer willingness to pay and the benefits of reducing distortionary taxes by other means. An alternative suggested by our analysis is to use consumer surplus plus tax revenue calculated using the shadow price of taxation forgone in sector 1 by allocating a unit of spectrum to sector 2. Even if the revenue effects are ignored by putting $\Lambda = -1$, the forgone benefit in terms of the consumer surplus will depend not only on the lowering of cost but also its subsequent impact on the price through the mark-up formula

$$P_1 = \frac{c_1(w, r, a)}{1 - \frac{1}{\epsilon_1 N_1}} = P_1(w, r, a, N_1) \tag{84}$$

which depends on cost, the elasticity of demand ϵ_1 and market structure through the effect on N_1. Writing the consumer surplus in sector 1 as $S(P_1)$ we have that

$S'(P_1) = -D(P_1)$. Thus the effect of a lowering of costs by Δc_1 and an increase in the number of firms by ΔN_1 is given by

$$\Delta S_1 = S'(P_1)\Delta P_1 = \underbrace{D(P_1)}_{\text{demand effect}} \left[\underbrace{-\frac{\Delta c_1}{\left(1 - \frac{1}{\epsilon_1 N_1}\right)}}_{\text{cost effect}} + \underbrace{\frac{c_1 \Delta N_1}{\epsilon_1 N_1^2 \left(1 - \frac{1}{\epsilon_1 N_1}\right)^2}}_{\text{market structure effect}} \right] \tag{85}$$

Since $\Delta c_1 < 0$ and $\Delta N_1 > 0$ both the cost effect and the market structure effect on the consumer surplus based notion of opportunity cost are positive. Note that (85) corresponds to the objective function used to compute constrained optimal licence prices by putting $\Lambda = -1$ in (16).

Comparing the Smith-NERA AIP given earlier with (85) as criteria for making incremental changes to licence prices it is now apparent that of the three effects present in the latter – demand, cost and market structure – only the cost effect is taken into account by AIP.[11] Thus we have developed a natural extension of the Smith-NERA methodology that also takes into account imperfect competition, the revenue effect (if so desired by the regulator), the willingness of consumers to pay, the ability to substitute other factors of production for spectrum and costs of changing spectrum prices.

8 Conclusions and Future Research

The objective of this paper has been to explore whether valuable light can be shed on policy questions in the area of spectrum allocation by the combination of models from information theory and economics. We have focused on several issues that

[11] Indepen (2004) acknowledge that in principle costs and benefits of reallocation of spectrum through relaxing constraints imposed by harmonisation should be measured with respect to their impact on consumer and producer surplus. Our free-entry condition drives the latter to zero and so does not feature in the calculations. In that study costs are used as a proxy for the impact on welfare only where suitable data on consumer surplus is lacking.

we believe are central to existing work by economists in this area. Our approach complements this literature by offering rigorous modelling of the issues identified, whilst advancing it by the multidisciplinary approach adopted. We believe this approach has yielded insights both for policymakers and for how the literature might be developed in order to strengthen these.

Perhaps one way to test this claim is to consider whether a useful research agenda has emerged from our work. We believe that this is the case. First, there are important questions of implementation—as we believe is also true of the existing AIP methodology. In particular, the introduction of consumer surplus and foregone distortionary taxes leading to Ramsey-type licence prices places additional emphasis on demand studies. The full implementation of our proposed procedure would require estimation of cost functions sector-by-sector.

Second, an important development would follow naturally from our framework. We focus on allocative and productive efficiency, thus ignoring dynamic (intertemporal) incentives/efficiency: in particular, for investment in spectrum-efficient technologies. In such a development, it would be socially optimal for investment to take place in settings where interference graphs are dense—so as to free up spectrum for re-use around the graph. Yet it seems almost certain that firms will not be able to appropriate all of the gains from such investment, thus rendering their investment decisions suboptimal. A natural way to encourage such investment would be to charge high spectrum prices where interference is greatest but this conflicts with the optimal static AIPs, as described above—recall Figure 7. Thus, the question of encouraging static and dynamic efficiency needs significant future research, within a model that explicitly incorporates aspects of the underlying engineering problem.

References

Indepen (2004). An economic study to review spectrum pricing: Final report. Mimeo, Indepen, Aegis Systems and Warwick Business School.

Jones, S., Levine, P., and Rickman, N. (2004). The economics of radio spectrum. Report to Ofcom, London.

Leese, R. (1998). Mathematical methods in radio frequency aassignment. In *Proceedings of the 14th International Wroclaw Symposium on Electromagnetic Compatibility, Poland*, pages 644–647.

Leese, R., Levine, P., and Rickman, N. (2000). Exploration into the relationship between mathematical and economic problems of spectrum allocation. Report to the Radiocommunications Agency, London.

Levin, H. J. (1970). Spectrum allocation without market. *The American Economic Review*, **60**(2), 209–218.

Perry, M. (1984). Scale economies, imperfect competition and public policy. *Journal of Industrial Economics*, **32**, 293–312.

A Proof of Proposition 1

Using (7) we may write (11) as

$$\Pi = \Pi(a, N) = \frac{P(a, N)D(P(a, N)}{\epsilon N^2} - F \tag{A.1}$$

Then, differentiating (A.1) partially with respect to a, we arrive at

$$\frac{\partial \Pi}{\partial a} = \frac{[PD' + D]}{\epsilon N^2} \tag{A.2}$$

Hence $\frac{\partial \Pi}{\partial a} < 0$ if $PD' + D < 0$ ie $\epsilon > 1$.

Similarly differentiating (A.1) partially with respect to N:

$$\frac{\partial \Pi}{\partial N} = \frac{\partial P}{\partial N}\frac{1}{\epsilon N^2}\left[D + PD'\right] - \frac{2PD}{\epsilon N^3} \tag{A.3}$$

Again from (7) we have

$$\frac{\partial P}{\partial N} = -\frac{P}{N(\epsilon N - 1)} \tag{A.4}$$

Combining (A.3) and (A.4) we arrive at

$$\frac{\partial \Pi}{\partial N} = -\frac{P[(2N-1)\epsilon - 1]}{N^3(\epsilon N - 1)} \tag{A.5}$$

Since we assume that $\epsilon > 1$ we have that $(2N-1)\epsilon > N(\epsilon-1) > 0$ for $N \geq 1$. It follows that $\frac{\partial \Pi}{\partial N} < 0$. $\square$

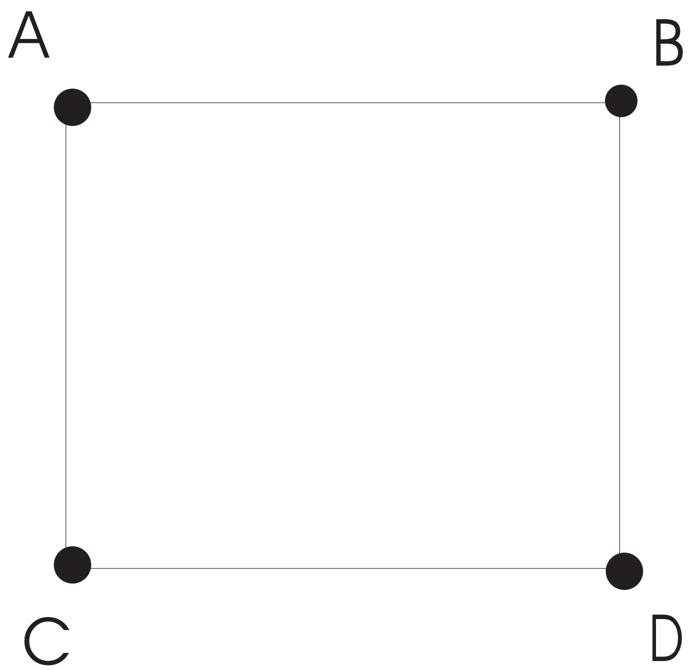

Figure 1: A 4-Node Graph of 4 Markets

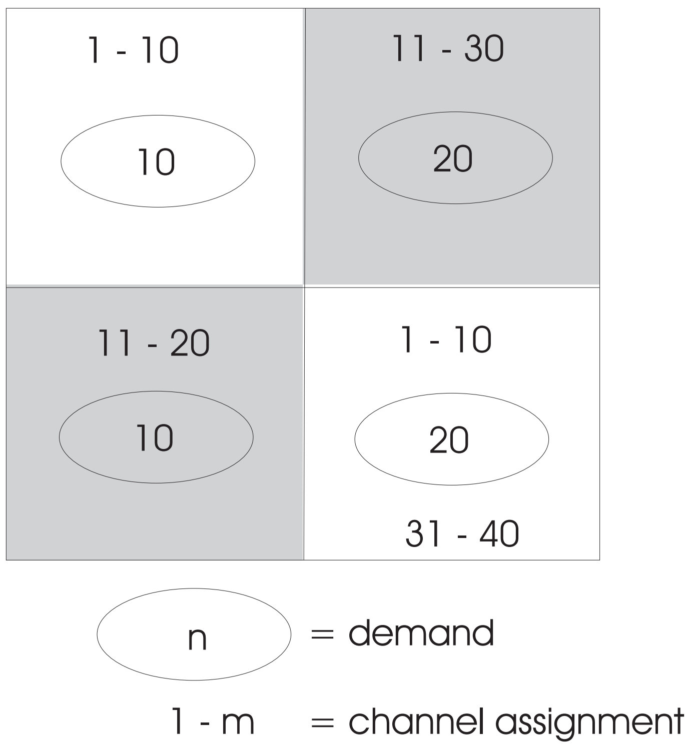

Figure 2: A Coloured Map of 4 Markets

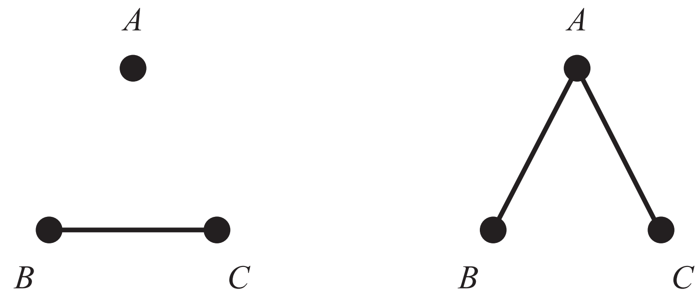

Figure 3: Single Edge and Two Edge Graphs

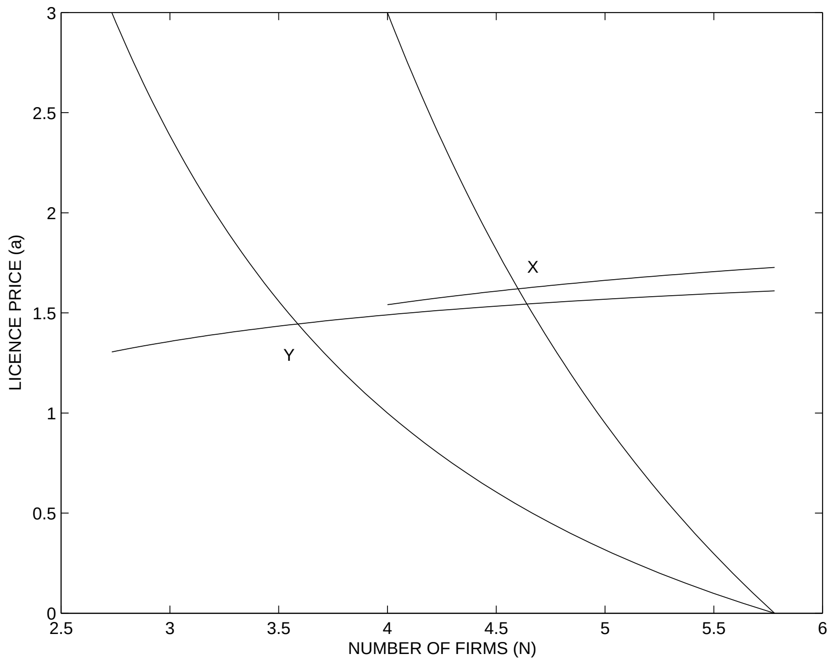

Figure 4: **The Optimal Licence Price and Firm Numbers:** $\Lambda = 0.3$, $\ell_i = 3$, $\mu = 3c$. **X=sector with spectrum re-use. Y=sector without spectrum re-use.**

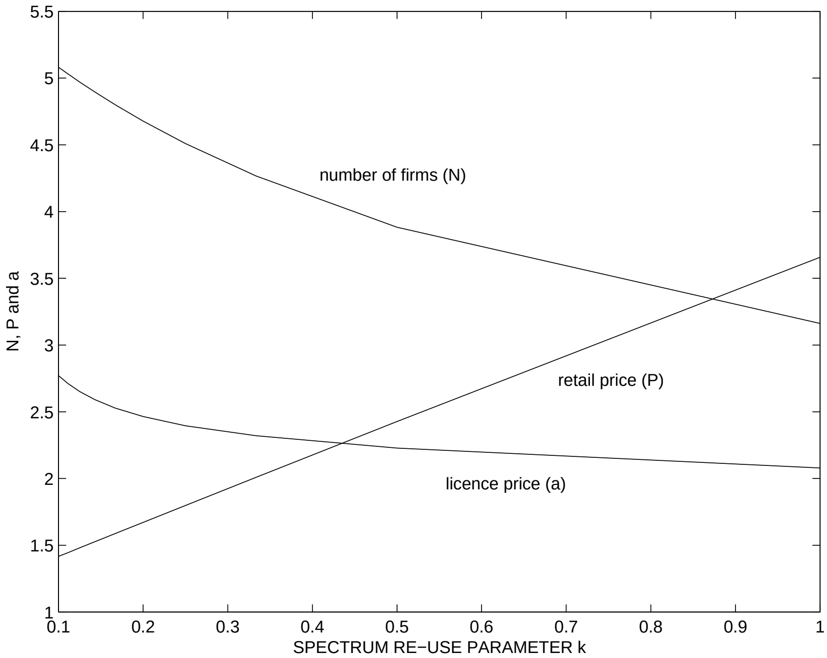

Figure 5: **The Optimal Licence Price and Spectrum Re-Use:** $\Lambda = 0.3$**,**$\mu = 3c$**.**

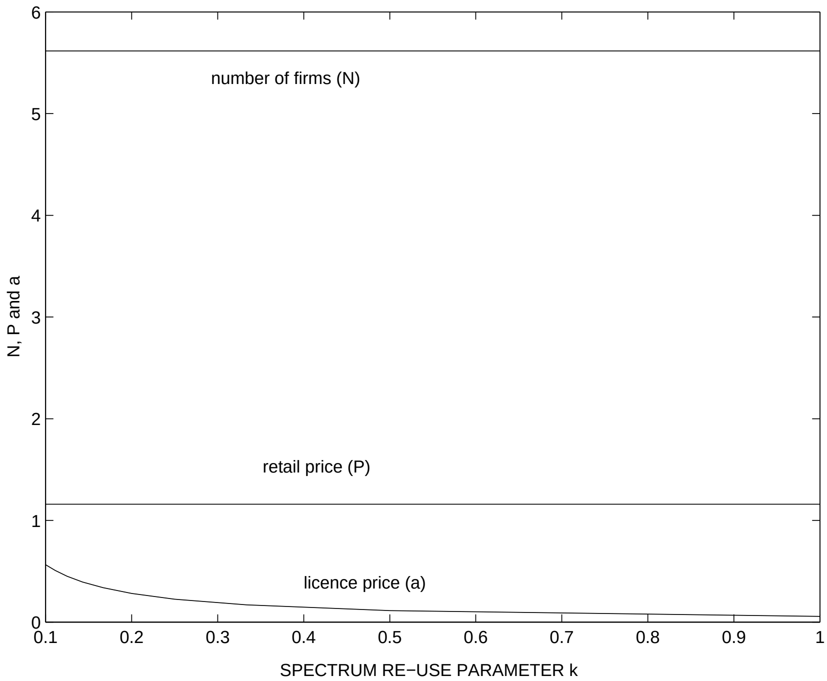

Figure 6: **The Optimal Licence Price and Spectrum Re-Use:** $\Lambda = 0.3$**,** $\mu = 0$**.**

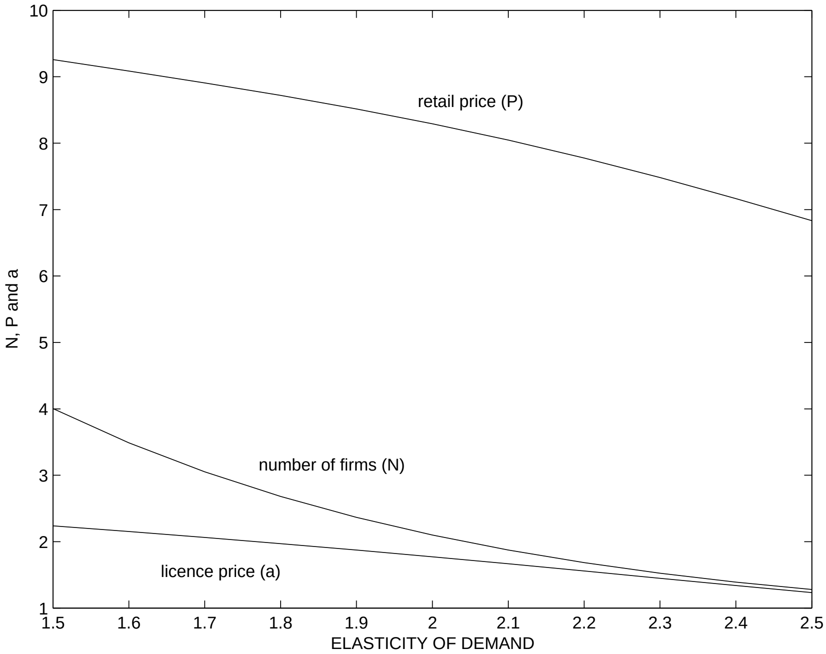

Figure 7: **The Optimal Licence Price and Elasticity of Demand (ϵ): $\Lambda = 0.3$, $\mu = 3c$, $\ell_i = 3$.**

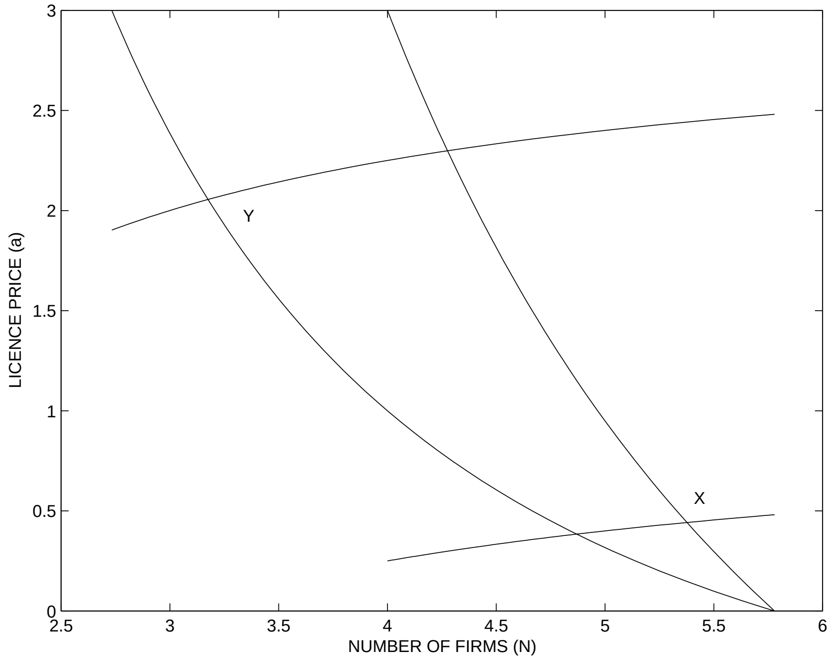

Figure 8: **The Constrained Optimal Licence Price and Firm Numbers: $\Lambda = -1$, $\ell_i = 3$. X=sector with spectrum re-use. Y=sector without spectrum re-use.**

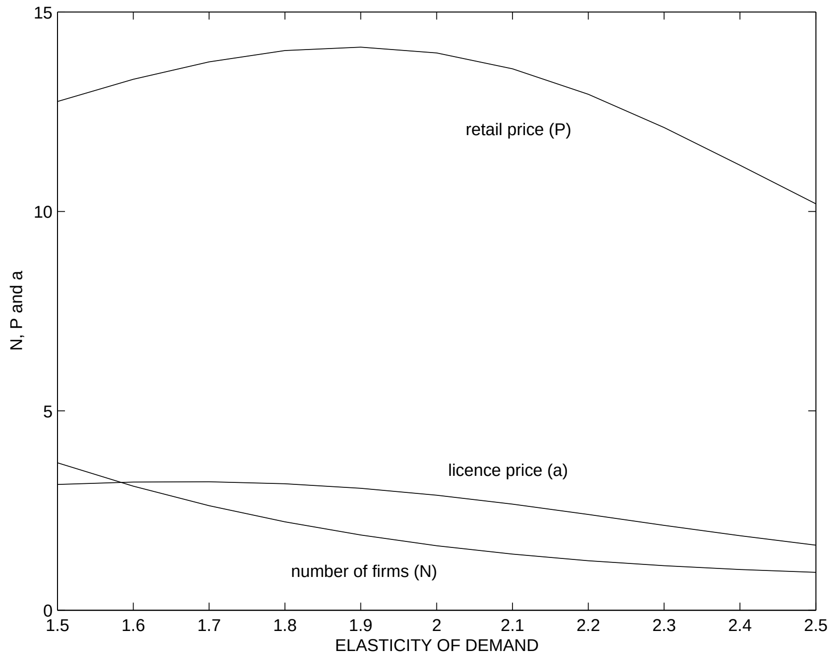

Figure 9: **The Constrained Optimal Licence Price Elasticity of Demand (ϵ): $\Lambda = -1$, $\mu = 3c$, $\ell_i = 3$.**

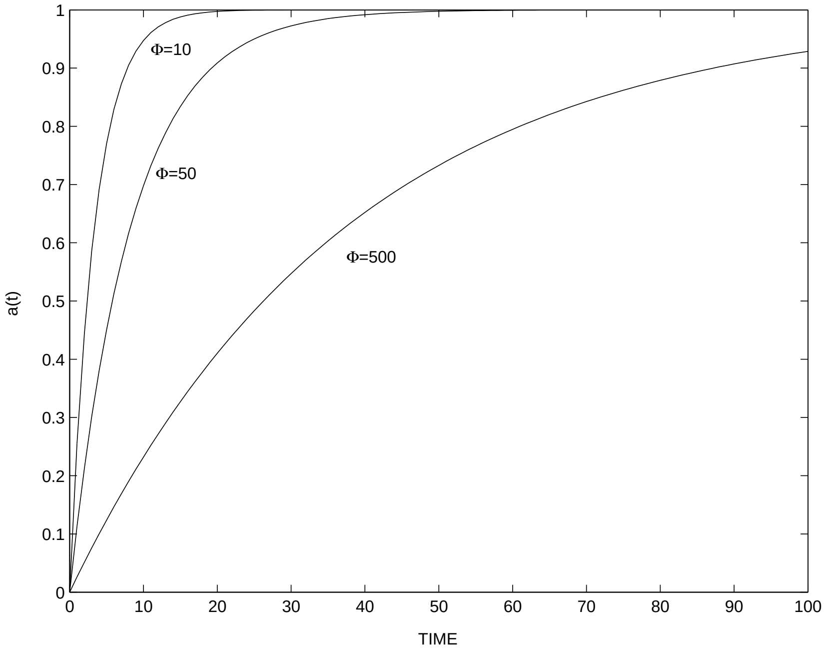

Figure 10: **Dynamic Adjustment.** $a_i^* = 1$**;** $a_i(0) = 0$**.**