

## CHAPTER 12

### ECONOMIES OF SCALE AND SCOPE, CONTESTABILITY, WINDFALL PROFITS AND REGULATORY RISK

#### 1. Introduction

Norman and Thisse (1997) have shown how the work of Eaton and Wooders (1985) and others as well as the contestability ideas of Baumol et al (1982) can be incorporated into a single period multiregional oligopolistic analysis. But, apart from implications for regulation based on contestability considerations, none of these approaches explicitly includes regulatory objectives and constraints. A fortiori none of these approaches allows for the modelling of post privatization windfall profits.

The primary purpose of this chapter is to address this issue and to develop an explicitly regulation related intertemporal optimization model within a regulated contestability based state preference framework. It will follow that, even though there may be significant switches in regulatory regime, because regulatory risk would then be fully anticipated, there would be no injustice (in the sense of retrospective legislation) in levying retrospective “windfall profit” taxes.

Before developing a state preference structure and with the motivation of regionally based utilities, I will briefly review work by Baumol et al (1982), Baumol and Willig (1986) and others on contestability with contexts of monopoly and oligopoly and its regulation. I will also introduce definitions of economies of scale and scope which draw on ideas in Ryan (1980), and Charnes et al (1980),(1987) with reference to the More for Less (Nothing) Paradox in linear programming. In these latter papers our results follow from the general optimality principle that a relatively less restricted constraint set may lead to lower (or equal) optimal costs, even though more has been produced and/or shipped. Extensions of that principle here yield new definitions for economies of scale (stemming from relaxations of restrictions on plant output) and economies of scope (stemming from reductions in costs of intermarket

shipments) for intertemporal as well as spatial cases. I will use those definitions and underlying more for less (nothing) ideas as natural ways of motivating the existence, or coexistence, of regional monopoly and oligopoly and associated contestability based regulatory frameworks under both profit and nonprofit maximizing conjectures and behaviours.

The structure of the chapter is as follows: After more background on contest-ability and regulation in Section 2, in Section 3 goal oriented regulation is introduced more formally via a multiregional goal programming generalization of Littlechild’s (1970) peak load pricing model. In sections 3 and 4 I use that model to give multiperiod and multiregional definitions of plant related (regional) economies of scale and firm related (interregional) economies of scope. In Sections 5 through 9 I consider implications of a multiperiod goal programming approach for concentration, for contestability and for tax and subsidy based regulation for firms with single or multiple production technologies. In Section 10 the preceding analyses are extended to a state preference framework in which outputs and profits, as well as associated oligopolistic conjectures, entry conditions and regulatory frameworks, all become state contingent. This approach leads directly to state contingent implications for profits and for industrial and market concentration. It also provides a natural way of modelling regulatory risk and associated windfall gains (or losses).

Finally, although that is not its primary purpose, the paper provides a means of analyzing the fact that, to the extent that interregional barriers to entry are effective, regionally monopolistic or oligopolistic enterprises will have opportunities to pursue a *variety* of objectives -including nonprofit maximizing objectives - which may coexist and be different for different firms in different regions.

## 2. Background

For Baumol et al a contestable market has the following properties:

First, the potential entrants can, without restriction, serve the same market demands and use the same productive techniques as those available to the incumbent firms...Second, the potential entrants evaluate the profitability of entry at the incumbents' pre-entry prices. (Baumol et al (1982) p.5)

Shepherd (1984) terms the first condition, together with conditions of perfect reversability and that the entrant could establish itself before any price response by incumbents, "ultra free entry". He goes on to state: "the premier question is whether the ultra-free entry results apply when entry is not ultra free". (p.573). In that context he maintains that Baumol et al's two properties are inconsistent, since if entry is ultra free there could be no contest and the second kind of calculation would be redundant. But in fact Baumol et al's conditions may be sustainable for an "entrant" supplying perfect substitutes by means of imports. Indeed owners of firms making only normal profits in a market would be indifferent at the margin between supply by production or supply by import for such a market - *and that market might be appropriately regulated to bring about such conditions*. (This spatial example has a particular resonance here since Baumol et al's own example was essentially spatial namely that of entry by a new operator onto an existing airline route.)

In the wider context of potentially regulated markets Morrison and Winston (1987), who set out to test the airline-related contestability hypothesis, made a distinction between perfect (i.e. B-P-W) and imperfect contestability and found that airline markets were imperfectly contestable according to their definition. (See also Shwartz (1986) on Morrison and Winston (1985)). Specifically they found that the difference between measures of actual and optimal welfare was responsive to numbers of *potential* competitors. They argued that, for that reason, the contestability hypothesis remained important for regulation. Underlying this: by then Baumol and Willig (1986) had accepted that the main contribution of contestability may be "as a guide

for regulation rather than as an argument for its elimination" (Baumol and Willig p.27), a point emphasised by Armstrong et al (1994) using this quotation.

The fact that contestability may be purposively induced by means of a contingently appropriate regime of taxes or subsidies to producers and/or to importers in designated regions, is a key idea in the present paper. A second key idea here is that in practice regulatory regimes will *change* and that it is possible to model implications for realized economies of scale and scope and for contingent profits of contestability based regulatory regimes. [In this respect the present paper differs with Cairns 1996 on contestability and risk. Cairns maintains that both contestability and risk can be modelled appropriately but not simultaneously. Briefly, Cairns introduces asymmetric information related uncertainty such that exit is not "free" since potentially entering firms may then realize unanticipated losses relative to incumbents. By contrast the state contingent approach to risk in the present chapter is based on full information assumptions and is potentially wholly consistent with the existence of *contingent* gains and losses, including those stemming from changes, if any, in regulatory regimes relating to spatially defined monopolies and oligopolies and associated entry conditions. (In such a framework both new and existing firms could in principle always finance their activities so as to realize only normal profits.)]

Spatial monopoly or oligopoly can be motivated on grounds of barriers to entry due to plant economies of scale, to tariffs, and/or due to high transport costs which effectively prevent competition from other regions. In any of these circumstances a regulator charged with introducing effective competition into such markets may consider a *regulated* contestability based policy by reducing interregional transport costs and reducing or removing impediments to imports, such as import tariffs or export subsidies. They may also consider contestability based legislation and regulation giving potential rivals access to existing capacity and distribution systems. [An example here is electricity in the UK. Prior to privatization electricity companies were nationally owned, but regionally run, with substantial interregional interconnections, among

other things to allow for the economical accommodation of demand peaks. Electricity was privatised in the form of a number of suppliers and a separately owned company controlling the interregional system (“the supergrid”). The industry also has a regulator with control over pricing and entry conditions and one form which regulation has taken is to control regional companies’ local supply prices. More recently another form of regulation has been contestability based regulation of prices at which regional companies can have access to distribution systems and thence potentially have access to consumers in other regions.]

Because it has implications for potential entry by existing firms into others’ markets, clearly contestability based regulation of this kind would have implications for firms’ level of operations stemming from reductions in intermarket shipment costs (*economies of scope*). But in an inter-regionally linked system *economies of scale* can arise even given constant intermarket shipment costs. If local demand and local supply both increase by the same absolute amount it may be the case that overall supply cost reduces due to the possibility of achieving a less costly overall pattern of distribution. In any case the scale at which a regionally based company’s plant can be operated and the price which can be gained for its output both have implications for profits. Indeed in the UK primary purposes of regulation in general, and contestability based regulation in particular, are to curb superprofits while at the same time ensuring sufficient supplies to all regions. (See Beesley (1996).).

More subtly a *change* in regulatory regime may have substantial implications for *changes* in profits. An example here is the radical change in regulatory policy and practice that commonly follows privatization of nationalized assets. Less radical policy changes may also follow a change of party control after a national election. For instance after the UK election of May 1997 recently privatized utilities were subjected to a windfall tax levied retrospectively on post privatization profits. In the following Sections, starting with a simple peak load pricing model, I will develop the contestability idea to show how, if contingent changes in entry conditions and/or in regulatory regimes are foreseeable in distribution,

then switches in regulatory policy and their implications for windfall gains (and losses), as well implications of cost and regulatory changes for entry conditions, can all be modelled as parts of an overall intertemporal optimization within a contestability based state preference framework.

### 3. A simplified n period peak load pricing model

Assume that  $d_{rt}$  and  $s_{rt}$  are demands and supplies for a product in region  $r$  and period  $t$  and that local production is generated either from existing capacity  $\kappa_{rtt-h}$  of ages  $h=1..H$  or from new capacity  $\kappa_{rtt-0}$ . If  $s_{rkt} \geq 0$  are shipments from region  $r$  to region  $k$  in period  $t$  and  $a_{rtt-h}$  allows for changing physical productivity of capacity with age then  $d_{rt} \leq s_{rt}$  and  $s_{rt} = \sum s_{rtt-h} + \sum s_{krt} - \sum s_{rkt}$  with  $s_{rtt-h} \leq a_{rtt-h} \kappa_{rtt-h}$   $h=0..H$ . (While allowing sales of capacity  $\kappa_{rtt-h}$ , for simplicity I assume that initial stocks of capacity are zero in every region. Extensions to include initial stocks would not substantively affect the results.)

If  $v_{rt}$  are acquisition costs of new capacity,  $v_{rtt-h}$ ,  $m_{rtt-h}$ , are sale values and maintenance costs of preexisting capacity and  $c_{rtt-h}$  and  $c_{rkt}$  are variable production costs of output and interregional transmission costs and if the overall objective of the multiregional enterprise is to maximize a measure of net return to its production capacity over an interval  $t=1,2..T$ , the problem for this enterprise can be expressed as in (I). (In that system  $1/(1+v)^t$  is an appropriate discount factor.)

With  $f_{rt}(d_{rt}) = \text{def } p_{rt}(d_{rt}) \delta d_{rt}$ , (I) is a multiregion variant of the consumers plus producers surplus based peak load pricing model used by Littlechild (1970). Or, with  $f_{rt}(d_{rt}) = \text{def } p_{rt}(d_{rt}) d_{rt}$ , (I) is a multiperiod, multiregion variant of standard multi-plant monopoly or perfectly collusive oligopoly models. In either case the objective is concave and the Kuhn Tucker conditions are necessary and sufficient for an optimum.

Since optima for (I) and subsequent models can always be found using appropriate linear or nonlinear programming computation packages, I will focus on economic interpretations of the Kuhn Tucker conditions at an optimum rather

than on processes of optimization per se.

Accordingly, with the following dual variables

$1/(1+\nu)^t \varphi_{rt}, 1/(1+\nu)^t \xi_{rt}, 1/(1+\nu)^t \mu_{rt-h}, 1/(1+\nu)^t \psi_{rt-1-t-h+1}$   
the Kuhn Tucker conditions for (I) are as in (I)'.

$$\begin{aligned}
 & \text{Max } \sum 1/(1+\nu)^t \sum [f_{rt}(d_{rt}) + v_{rt-h} \bar{\kappa}_{rtt-h} - m_{rtt-h} \kappa_{rtt-h} - v_{rt} \kappa_{rtt-0} - c_{rtt-h} s_{rtt-h} - c_{rkt} s_{rkt}] \\
 & \text{st} \quad d_{rt} \leq s_{rt} \\
 & \quad s_{rt} = \sum s_{rtt-h} + \sum s_{krt} - \sum s_{rkt} \\
 & \quad s_{rtt-h} \leq a_{rtt-h} \kappa_{rtt-h} \\
 & \quad \kappa_{rtt-h} + \kappa_{rtt-h} = \kappa_{rtt-(h-1)} \\
 & \quad \text{All variables nonnegative} \\
 & \quad \varphi_{rt} \geq 1/(1+\nu) \delta f(d_{rt}) / \delta d_{rt} \\
 & \quad \varphi_{rt} \leq \xi_{rt} \\
 & \quad \xi_{rt} \leq c_{rtt-h} + \mu_{rtt-h} \\
 & \quad \xi_{rt} \leq \xi_{kt} + c_{krt} \\
 & \quad \xi_{rt} \geq \xi_{kt} - c_{rkt} \\
 & \quad \psi_{irtt-0} \leq v_{rt} \\
 & \quad a_{rtt-h} \mu_{rtt-h} \leq \psi_{rtt-h} - 1/(1+\nu) \psi_{rt+1-t-h-1} + m_{rtt-h} \\
 & \quad \psi_{rtt-h} \geq v_{rtt-h} \\
 & \quad \varphi_{rt}, \mu_{kt} \geq 0, \xi_{rt} \text{ unrestricted}
 \end{aligned} \tag{I}$$

From the first conditions output will optimally be supplied in region  $r$  in period  $t$ , if at all, then only to the point where the discounted marginal return  $1/(1+\nu) \delta f(d_{rt}) / \delta d_{rt}$  equals marginal supply price  $\varphi_{rt}$ . By complementary slackness, if such supplies are positive, supply price  $\varphi_{rt}$  will be equated to supply cost  $\xi_{rt}$  (from the second constraint) and, by the third and fourth constraints, supplies in region  $r$  will be from the cost minimizing set of ages of capacity in that region and/or from imports to that region. (The third constraints also imply that supply cost will be made up of variable cost and capacity cost components  $c_{rtt-h}$ ,  $\mu_{rtt-h}$  for each age of capacity used.)

The fifth constraint of (I)' requires that output will be exported to region  $k$  from region  $r$ , if at all, then to the point where supply cost in region  $k$  equals supply cost in region  $r$  plus interregional transmission cost  $c_{rkt}$ . The last three constraints give optimal rules for acquiring, retaining and disposing of capacity. With complementary slackness they require that capacity would be: i) acquired only if its internal valuation is sufficient to recoup the acquisition cost; ii) retained only as long as current and anticipated incomes from production (if any)

$a_{rtt-h} \mu_{rtt-h}$  and internally evaluated anticipated capital accumulation  $1/(1+\nu) \psi_{rt+1-t-h+1}$  is sufficient to recoup the end period internal valuation  $\psi_{rtt-h}$  plus maintenance cost  $m_{rtt-h}$ , and; iii) relinquished otherwise.

For elements of capacity optimally acquired in region  $r$  and period  $t$  and relinquished in period  $(t+h_1)$ , from the last three constraints and complementary slackness:

$$\kappa_{rtt-0} > 0 \Rightarrow \psi_{irtt-0} = v_{rt} \tag{1}$$

$$\kappa_{rt+ht-0} > 0 \Rightarrow a_{rt+ht-0} \mu_{rt+ht-0} = \psi_{rt+ht-0} - 1/(1+\nu) \psi_{rt+h+1-t-0} + m_{rt+ht-0} \quad h=0..h_1 \tag{2}$$

$$\kappa_{rt+h_1-t-0} > 0 \Rightarrow \psi_{rt+h_1-t-0} = v_{rt+h_1-t-0} \tag{3}$$

Conditions (1)-(3) together imply:

$$v_{rt} = \sum_{h=0}^{h=h_1} 1/(1+\nu)^h [a_{rt+ht-0} \mu_{rt+ht-0} - m_{rt+ht-0}] + v_{rt+h_1-t-0} / (1+\nu)^{h_1} \tag{4}$$

So at an optimum with tariffs as in (I)', acquisition costs would be exactly recouped from rentals net of maintenance costs together with terminal valuations over the intervals during which

elements of capacity are retained.

In that way  $a_{rt,h}\mu_{rtt-h}$  become optimal amortization allowances analogous to those derived for nonspatial models by Turvey (1969) and Littlechild (1970) and, in an explicitly spatial context, by Ryan (1978),(1992).

[Notice that optimal retention intervals for capacity are endogenous, with newer capacity potentially replacing old due to ultimately increasing maintenance cost  $m_{rtt-h}$  and/or reducing productivity  $a_{rtt-h}$ . Notice, too, that acquisition and selling prices of capacity need not be positive. Indeed for elements of nuclear generating

capacity terminal selling prices might quite plausibly be negative.]

Program (I) can be extended to include demand ( $d_{rt}$ ), supply ( $s_{rt}$ ) and capacity ( $\kappa_{rtt-h}$ ) related regulatory goals  $d_{rt}^*, s_{rtt-h}^*, \kappa_{rtt-h}^*$  with penalties  $d_{rt}^+, d_{rt}^-, e_{rtt-h}^+, e_{rtt-h}^-, m_{rtt-h}^+, m_{rtt-h}^-$  for exceeding or falling short of them as follows as in (II). Then associating dual variables  $1/(1+\nu)^t \Delta p_{rt}$ ,  $1/(1+\nu)^t \Delta \tau_{rtt-h}$ ,  $1/(1+\nu)^t \Delta \omega_{rtt-h}$  with the last three constraints of (II), the Kuhn-Tucker conditions (I)' are modified to give (II)':

$$\begin{aligned}
 \text{Max } \sum_t 1/(1+\nu)^t \sum [f_{rt}(d_{rt}) + v_{rtt-h} \bar{\kappa}_{rtt-h} - m_{rtt-h} \kappa_{rtt-h} - v_{rt} \kappa_{rtt-0} - c_{rtt-h} s_{rtt-h} - c_{rkt} s_{rkt} \\
 - c_{rt}^+ d_{rt}^+ - c_{rt}^- d_{rt}^- - e_{rtt-h}^+ s_{rtt-h}^+ - e_{rtt-h}^- s_{rtt-h}^- - m_{rtt-h}^+ \kappa_{rtt-h}^+ - m_{rtt-h}^- \kappa_{rtt-h}^-] \\
 \text{st} \quad d_{rt} \leq s_{rt} \\
 s_{rt} = \sum s_{rtt-h} + \sum s_{krt} - \sum s_{rkt} \\
 s_{rtt-h} \leq a_{rtt-h} \kappa_{rtt-h} \\
 \kappa_{irtt-h} + \kappa_{irtt-h}^- = \kappa_{irt-1t-h+1} \\
 d_{rt} + d_{rt}^+ - d_{rt}^- = d_{rt}^* \\
 s_{rtt-h} + s_{rtt-h}^+ - s_{rtt-h}^- = s_{rtt-h}^* \\
 \kappa_{rtt-h} + \kappa_{rtt-h}^+ - \kappa_{rtt-h}^- = \kappa_{rtt-h}^* \\
 \text{All variables nonnegative} \\
 \varphi_{rt} + \Delta p_{rt} \geq 1/(1+\nu) \delta f(d_{rt}) / \delta d_{rt} \\
 \varphi_{rt} \leq \xi_{rt-1} \\
 \xi_{rt} \leq c_{rtt-h} + \mu_{rtt-h} + \Delta \tau_{rtt-h} \\
 \xi_{rt} \leq \xi_{kt} + c_{krt} \\
 \xi_{rt} \geq \xi_{kt} - c_{rkt} \\
 \psi_{irtt-0} \leq v_{rt} \\
 a_{rtt-h} \mu_{rtt-h} \leq \psi_{rtt-h} - 1/(1+\nu) \psi_{rt+1t-h-1} + m_{rtt-h} + \Delta \omega_{rtt-h} \\
 \psi_{irtt-h} \geq v_{rtt-h}^- \\
 -c_{rt} \leq \Delta p_{rt} \leq c_{rt}^+ \\
 -e_{rtt-h}^- \leq \Delta \omega_{rtt-h} \leq e_{rtt-h}^+ \\
 -m_{rtt-h}^- \leq \Delta \tau_{rtt-h} \leq m_{rtt-h}^+ \\
 \varphi_{rt}, \mu_{kt} \geq 0, \xi_{rt} \text{ unrestricted}
 \end{aligned} \tag{II}'$$

Since (II) is a goal programming extension of (I) not surprisingly at an optimum interpretations of (II)' are as for (I)', except, by complementary slackness,  $s_{rt}^+ > 0 \Rightarrow \Delta p_{rt} = c_{rt}^+$  and  $s_{rt}^- > 0 \Rightarrow \Delta p_{rt} = -c_{rt}$  so that  $\Delta p_{rt}$  takes on the interpretation of an optimal tax or subsidy.

Similar output and capacity related tax and subsidy interpretations arise via  $e_{rtt-h}^+$ ,  $-e_{rtt-h}^-$  (resp  $m_{rtt-h}^+$ ,  $-m_{rtt-h}^-$ ). Before considering these interpretations and associated issues in relation to regulatory principles and policies in more detail in Section 6, I first explicitly define economies of scope and economies of scale.

#### 4. Economies of scope

Economies of scope arise in a multiregion economy when hitherto unavailable (or untaken) shipping opportunities become available via reductions in one or more relevant shipping costs. Specifically: Associate penalties  $M$  arbitrarily larger than  $c_{rkt}'$  with initially unavailable shipping opportunities  $s_{rkt}'$  potentially connecting subsets of

regions  $r_2, k_2 \in R_2$  to initially connected subsets  $r_1, k_1 \in R_1$  to give:

#### THEOREM 1 (*Economies of Scope*)

Assume two alternative cost regimes  $\{c_{rkt}, M\}$  and  $\{c_{rkt}, c_{rkt}'\}$  for potential shipments between regions  $r_1, k_1 \in R_1$  and  $r_2, k_2 \in R_2$  then, if a feasible solution exists for (IIa\*):

$$\begin{aligned} \text{Max } \sum 1/(1+v)^t [f_{rt}(d_{rt}) + v_{rtt-h} \bar{k}_{rtt-h} - m_{rtt-h} k_{rtt-h} - v_{rt} k_{rtt-0} - c_{rtt-h} s_{rtt-h} - \sum_{R_1} c_{rkt} s_{rkt} - \sum_{R_2} M s_{rkt}' \\ - c_{rt}^+ d_{rt}^+ - c_{rt} \bar{d}_{rt} - e_{rtt-h}^+ s_{rt-h}^+ - e_{rtt-h} \bar{s}_{rtt-h} - m_{rtt-h}^+ k_{rtt-h}^+ - m_{rtt-h} \bar{k}_{rtt-h}] \\ \text{st constraints of (II)} \end{aligned} \quad (\text{IIa})*$$

$$\leq$$

$$\begin{aligned} \text{Max } \sum 1/(1+v)^t [f_{rt}(d_{rt}) + v_{rtt-h} \bar{k}_{rtt-h} - m_{rtt-h} k_{rtt-h} - v_{rt} k_{rtt-0} - c_{rtt-h} s_{rtt-h} - \sum_{R_1} c_{rkt} s_{rkt} - \sum_{R_2} c_{rkt}' s_{rkt}' \\ - c_{rt}^+ d_{rt}^+ - c_{rt} \bar{d}_{rt} - e_{rtt-h}^+ s_{rt-h}^+ - e_{rtt-h} \bar{s}_{rtt-h} - m_{rtt-h}^+ k_{rtt-h}^+ - m_{rtt-h} \bar{k}_{rtt-h}] \\ \text{st constraints of (II)} \end{aligned} \quad (\text{IIa})**$$

#### PROOF

Any feasible solution to (IIa)\* is a feasible solution to (IIa)\*\* and conversely. But an optimal solution to (IIa)\* is a feasible but not necessarily an optimal solution to (IIa)\*\*. It follows that there may exist optimal solutions to (IIa)\*\* such that  $z < z'$  or  $z = z'$  with  $s_{rkt}' > 0$   $r_2, k_2 \in R_2$ .

That is: with weights  $M$  as in (IIa)\* there will be no interregional shipments connecting subsets  $r_2, s_2$  to  $r_1, s_1$ . But, if one or more weights  $M$  becomes non-preemptive, *economies of scope* may become available and may be exploited via a relatively higher optimum in (IIa)\*\* via nonpreemptive weights  $c_{srkt}'$  vis a vis the preemptive weights  $M$  in (IIa)\*.

#### 5. Economies of scale

Economies of scale are defined as arising in a

multiregion economy when it is possible to increase total amount produced in at least one region and for at least one market so that average production costs are reduced, even when (increased) economies of scope are not available. Formally, if  $d_{rt} > 0$ ,  $s_{rtt-h} > 0$  at least one  $r, t, h$  in an optimal solution to (II), economies of scale are attainable relative to the following system which associates preemptively large penalties  $M$  with at least one relevant pair  $d_{rt}, s_{rtt-h}$ :

#### THEOREM 2 (*Economies of Scale*)

Consider two distinct regulatory regimes, one associating prohibitive penalties  $M$  and the other non prohibitive penalties  $(a_{rt}, e_{rtt-h})$  with potentially marginal increases in sub-market demand and supply levels  $d_{rt}, s_{rtt-h}$  in (II). Then, if a feasible solution exists for (IIb)

$$\begin{aligned} \text{Max } \sum 1/(1+v)^t [f_{rt}(d_{rt}) + v_{rtt-h} \bar{k}_{rtt-h} - m_{rtt-h} k_{rtt-h} - v_{rt} k_{rtt-0} - c_{rtt-h} s_{rtt-h} - \sum_{R_1} c_{rkt} s_{rkt} - \sum_{R_2} M s_{rkt}' \\ - c_{rt}^+ d_{rt}^+ - M d_{rt} - e_{rtt-h}^+ s_{rt-h}^+ - M s_{rtt-h} - m_{rtt-h}^+ k_{rtt-h}^+ - m_{rtt-h} \bar{k}_{rtt-h}] \\ \text{st constraints of (II)} \end{aligned} \quad (\text{IIb})*$$

$$\leq$$

$$\begin{aligned} \text{Max } \sum 1/(1+v)^t [f_{rt}(d_{rt}) + v_{rtt-h} \bar{k}_{rtt-h} - m_{rtt-h} k_{rtt-h} - v_{rt} k_{rtt-0} - c_{rtt-h} s_{rtt-h} - \sum_{R_1} c_{rkt} s_{rkt} - \sum_{R_2} M s_{rkt}' \\ - c_{rt}^+ d_{rt}^+ - c_{rt} \bar{d}_{rt} - e_{rtt-h}^+ s_{rt-h}^+ - e_{rtt-h} \bar{s}_{rtt-h} - m_{rtt-h}^+ k_{rtt-h}^+ - m_{rtt-h} \bar{k}_{rtt-h}] \\ \text{st constraints of (II)} \end{aligned} \quad (\text{IIb})**$$

## PROOF

Any feasible solution to (IIb)\* is a feasible solution to (IIb)\*\* and conversely. But an optimal solution to (IIb)\* is a feasible but not necessarily an optimal solution to (IIb)\*\*. It follows that there may exist optimal solutions to (IIa)\*\* such that  $z < z'$  or  $z = z'$   $d_{rt} > 0$ ,  $s_{rtt-h} > 0$  at least one  $r, t, h$ .

If marginal revenue  $\delta f_{rt}(d_{rt})/\delta d_{rt}$  is constant and identical to  $p_{rt}^*$  in the relevant range of output, *cost economies of scale* may arise when demand and supply in a region  $r$  and time period  $t$  can be increased so that optimally  $d_{rt} > 0$ ,  $s_{rtt-h} > 0$ , and yet the marginal cost is zero or negative. (That is: when (IIb)\*\*, with nonpreemptive weights  $c_{rt}$ ,  $e_{rtt-h}$ , yields a lower optimal cost than (IIb)\*.) This zero or negative incremental cost may come about for either or both of two reasons: First because of increased use of relatively lower cost capacity  $s$  in region  $r$  to generate the increased supply  $s_{rtt-h} > 0$  and, secondly, due to changes in shipments to or from other markets in response to increasing demands and supplies  $d_{rt} > 0$ ,  $s_{rtt-h} > 0$  at market  $r$ . In effect the age mix of locally used capacity and the regional mix of sources of supply of product  $i$  to region  $r$  are both scale dependent - even with a given set of potential interregional shipping routes and constant unit shipping costs over those routes.

## 6. Generalised economies of scale and scope and implications for concentration and regulation

Clearly in general a potentially connected set of markets may exhibit economies of scale and then of scope or, conversely, of scope and then of scale. In each case the resulting configuration will conform to an optimal solution to an overall model of the form of (II) with the appropriate parameters. That is: with the appropriate parameters, (II) includes all of (IIa)\*, (IIa)\*\*, (IIb)\* and (IIb)\*\* as special cases. Further, insofar as  $c_{rkt}$ ,  $d_{rt}$ ,  $c_{rtt-h}$  may take on arbitrarily large values, (II) already incorporates interpretations in relation to economies of scale, as in (IIa)\*\*, and in relation to economies of scope, as in (IIb)\*\*. With that context the goal oriented parameters  $d_{rt}$ ,  $c_{rtt-h}$  take on interpretations as intra-regional taxes (if negative)

and subsidies (if positive). Also, by means of changes in effectively product specific interregional or international tariff or subsidy related differentials via  $c_{rkt}$ , potential for economies of scope, and thence potentially for scale, can be correspondingly enhanced or reduced.

Such potential to use taxes or subsidies to exploit (or inhibit) economies of scale and of scope has implications for industrial and market concentration. If interregional shipping costs are sufficiently large, markets will be effectively isolated and correspond to conditions of *regional monopoly*. In turn such conditions, by effectively precluding imports from other regions, would impose constraints on opportunities for them to exploit economies of scale vis a vis the remainder. Conversely, use of taxes or subsidies to remove impediments to economies of scope (and thence possibly to economies of scale) for subsets of regions evidently has implications for industrial and market concentration.

More generally taxes and subsidies may be used to regulate regional demands, outputs and capacities, as well as interregional shipments, and all of these types of regulation have implications for regional industrial and market structures. For example a sufficiently large reduction in entry costs may reduce market concentration by generating supplies from an additional source. But market concentration would not inevitably be reduced since entry costs may fall so far that one or all existing suppliers is displaced from the entered market. At the same time, diversion of supplies from one or more preexisting markets to supply this now more accessible market may have the consequence of an optimally increased market concentration in the region in which those incremental supplies are produced. This is an empirically important observation since under the antitrust laws of both the EU and of the US, regulation of entry conditions, of mergers and takeovers, as well as of output prices and processes, all stem from a preoccupation with collusive arrangements and/or entry conditions.

## 7. Multiple technologies and regulation

One way to model a number of firms in a region is to consider different ages of plant as relating to different firms. In that way (II) might include up to  $T$  firms in each region over a horizon  $T$ ,  $t=0,1\dots T$ . A more general approach is to assume that in each region a product (e.g. electricity) may be supplied by one or all of a variety of technologies  $i$  (e.g. oil, natural gas, coal, nuclear) each potentially with a variety of ages of capacity. This extension - and with it the extension to  $i=1,2\dots N_r$  distinct types of enterprise in region  $r$ , is made in (III) below by assuming constraints as in (II) for each of  $N_r$  technologies. In this way in region  $r$  the total output available to supply local demand or for export will be the sum of outputs from technologies  $i$ . Then supply in a region can be seen as potentially supplied by up to  $N_r$

technologically differentiated oligopolists each potentially operating multiple plants (plants of multiple ages). As another extension (III) also incorporates targets  $s_{rkt}^*$  for interregional shipments in addition to and analogous to those for demand, output and capacity in (II).

The objective of (III) is concave so the Kuhn-Tucker conditions are necessary and sufficient for an optimum. Making dual variables specific to technologies  $i$  where appropriate and associating dual variables  $1/(1+\nu)^t \Delta \sigma_{irkt}$   $1/(1+\nu)^t \Delta \sigma_{ikrt}$  with the final interregional shipment targets  $s_{rkt}^*$ ,  $s_{krt}^*$  (III) yields technology dependent optimality conditions with interpretations corresponding to those in (II) as in (III)'.

$$\begin{aligned}
 & \text{Max } \sum 1/(1+\nu)^t [f_{rt}(d_{rt}) + v_{irtt-h} \bar{K}_{irtt-h} - c_{irtt-h} S_{irtt-h} - m_{irtt-h} \bar{K}_{irtt-h} - v_{irt} K_{irtt-0} - c_{rkt} s_{rkt} - c_{rt}^+ d_{rt}^+ \\
 & \quad - c_{rt}^- d_{rt}^- - e_{irtt-h}^+ S_{irtt-h}^+ - e_{irtt-h}^- S_{irtt-h}^- - m_{irtt-h}^+ \bar{K}_{irtt-h}^+ - m_{irtt-h}^- \bar{K}_{irtt-h}^- - c_{rkt}^+ s_{rkt}^+ - c_{rkt}^- s_{rkt}^- - c_{krt}^+ s_{krt}^+ - c_{krt}^- s_{krt}^-] \\
 & \quad \text{st} \quad d_{rt} \leq s_{rt} \\
 & \quad s_{rt} = \sum S_{irtt-h} + \sum s_{krt} - \sum s_{rkt} \\
 & \quad S_{irtt-h} \leq a_{irtt-h} K_{irtt-h} \\
 & \quad K_{irtt-h} + \bar{K}_{irtt-h} = K_{irt-1t-h+1} \\
 & \quad d_{irt} + d_{rt}^+ - d_{rt}^- = d_{rt}^* \\
 & \quad S_{irtt-h} + S_{irtt-h}^+ - S_{irtt-h}^- = S_{irtt-h}^* \\
 & \quad K_{irtt-h} + \bar{K}_{irtt-h}^+ - \bar{K}_{irtt-h}^- = K_{irtt-h}^* \\
 & \quad s_{rkt}^* + s_{rkt}^+ - s_{rkt}^- = s_{rkt}^* \\
 & \quad s_{krt}^* + s_{krt}^+ - s_{krt}^- = s_{krt}^* \\
 & \quad \text{All variables nonnegative} \\
 & \quad (1+\nu)\varphi_{rt} + \Delta p_{rt} \geq \delta f(d_{rt})/\delta d_{rt} \\
 & \quad \varphi_{rt} \leq \xi_{irt-1} \\
 & \quad \xi_{irt} \leq c_{irtt-h} + \mu_{irtt-h} + \Delta \tau_{irtt-h} \\
 & \quad \xi_{irt} \leq \xi_{ikt} + c_{ikrt} + \Delta \sigma_{ikrt} \\
 & \quad \xi_{irt} \geq \xi_{ikt} - c_{irkt} + \Delta \sigma_{irkt} \\
 & \quad \psi_{irtt-0} \leq v_{irt} \\
 & \quad a_{irtt-h} \mu_{irtt-h} \leq \psi_{irtt-h} - 1/(1+\nu) \psi_{irt+1t-h-1} + m_{irtt-h} + \Delta \omega_{irtt-h} \\
 & \quad \psi_{irtt-h} \geq v_{irtt-h}^- \\
 & \quad -c_{rt}^- \leq \Delta p_{rt} \leq c_{rt}^+ \\
 & \quad -e_{irtt-h}^- \leq \Delta \omega_{irtt-h} \leq e_{irtt-h}^+ \\
 & \quad -m_{irtt-h}^- \leq \Delta \tau_{irtt-h} \leq m_{irtt-h}^+ \\
 & \quad -c_{irkt}^- \leq \Delta \sigma_{irkt} \leq c_{irkt}^+ \\
 & \quad \varphi_{rt}, \mu_{irtt-h}, \xi_{irt} \text{ unrestricted}
 \end{aligned} \tag{III}$$

By contrast with (II)', policy oriented interregional tax and subsidy variables  $\Delta p_{rt}, \Delta \omega_{irtt-h}, \Delta \tau_{irtt-h}, \Delta \sigma_{irkt}$  are explicitly distinguished from

interregional shipping costs in (III)'. At the same time the targets  $s_{rkt}^*$  and dual relations have potential interpretations in relation to conditions



of entry and thence to contestability - regulated or otherwise.

Consider the latter possibilities first: i) If in (III) all markets were consistent with conditions of competitive free entry, optimal allocations of outputs and capacities would be such that price equals marginal supply cost in each market with products being supplied from the least cost source (or sources), be those from local production and/or from imports. And; ii) costs of acquiring and maintaining capacity would be such that only normal profits would be made. Together i) and ii) are equivalent to conditions as if  $\delta f_{rt}(d_{rt})/\delta d_{rt} = \text{price} = \varphi_{rt}(1+v) = \text{marginal supply cost} = \min[c_{irtt-h} + \mu_{irtt-h} + \Delta\tau_{irtt-h}]$  for all products, production technologies and markets associated with positive outputs in an optimum for (III). That is, they are equivalent to conditions in which all taxes and subsidies are *zero* at an optimum for (III) and so with a no regulation optimum for that system. In that way optimal solutions to (III) may correspond to a special class of cases consistent with a “first best” optimum generated by conditions of free entry and free exit and so of contestability for every market.

[For any optimal solution to (III) all revenues, which may include supernormal profits in some circumstances, are imputed to those with property rights in resources. Here these include rights to specify regulatory goals as well as rights associated with elements of production capacity. This focus on income to property rights suggests extensions of (III) to comprehend budget constrained consumption and thence general equilibria for non regulated competitive cases, and where appropriate, for regulated second best noncompetitive cases.]

## 8. Capitalization, regulation and contestability

Consider a collectively profit maximizing collusive optimum with reference to (III). In that case in (III)' superprofits may be implicit in the quantities (marginal revenues)  $\delta f_{rt}(d_{rt})/\delta d_{rt}$  and so in supply costs  $[c_{irtt-h} + \mu_{irtt-h} + \Delta\tau_{irtt-h}]$  and, via  $a_{irtt-h}\mu_{irtt-h}$ , in capital values  $\psi_{irtt-h}$ . In such cases, even though superprofit may be implicit in current prices, such profit will nevertheless be capitalized via  $\psi_{irtt-h}$ . And, by arguments analogous to the derivation of (4) from (1)..(3), the quantities  $\psi_{irtt-h}$  are themselves consistent with regulation related

optimal amortization interpretations for  $a_{irtt-h}\mu_{irtt-h}$ .

At an optimum:

$$S_{irtt-h} > 0 \Rightarrow \xi_{irt} = c_{irtt-h} + \mu_{irtt-h} + \Delta\tau_{irtt-h} \quad (5)$$

These conditions together with the remaining conditions (III)' and complementary slackness for a retention interval  $h=0\dots h_1$  yield:

$$V_{irt} = \sum_{h=0}^{h=h_1} [a_{irtt-h} \mu_{irtt-h} - m_{irtt-h} - \Delta\omega_{irtt-h}] 1/(1+v)^h + V_{irt+h_1} / (1+v)^{h_1} \quad (6)$$

In (6) supernormal profits (if any) net of regulatory taxes or subsidies (if any) are capitalized in the value of acquired and retained capacity and so in its ownership. With the perfect capital markets implicit in the adoption of a common discount rate  $v$  throughout the preceding analysis, no superprofit would be attributable as transfer earnings to new investors in elements of production capacity. And since purchases and sales of elements of production capacity would then be free of transfer earnings, markets would always be *contestable* by potentially new owners of capacities  $Z_{irtt-h}$  in (III). That is: such investors or disinvestors would earn a normal rate of return on their investments (disinvestments) and both entry and exit would be free for such potentially entering or leaving owners.

## 9. Industrial vs market contestability

In general contestability may refer to conditions of potential entry and exit by new producers and not just to changes of ownership for existing producers in a region (as in the previous Section). In that context it is useful to make a further distinction between *industrial contestability* and *market contestability*, referring respectively to conditions of entry and exit for the marginal potential producer and for the marginal potential supplier in a region.

Conditions of *industrial* contestability are equivalent to whether or not, for the marginal potential entrant  $j$ , the condition  $\xi_{jrt} \leq c_{jrtt-h} + \mu_{jrtt-h} + \Delta\tau_{jrtt-h}$  in (III)' holds with equality *and* whether for such a potentially entering producer profits would be normal.

Notice that, via  $\Delta\tau_{irtt-h}$ , (III) incorporates opportunities for regulators to select contingent taxes

and subsidies - and so to vary conditions determining market contestability - with reference both to existing firms and to potential entrants. (In general existing suppliers may be making supernormal profit even if the relevant market is contestable by a new firm with a different technology. This emphasizes that (III) is consistent with existing suppliers, possibly using a different technology, making supernormal profits - or losses - even if the industry is contestable by a new firm.)

Conditions of *market* contestability are equivalent to whether or not inter-market (potential) entry conditions  $\xi_{jrt} \leq \xi_{jkt} + c_{jkrt} + \Delta\omega_{irt-h}$  in (III)' hold with equality for at least one supplier  $j$  at present not supplying region  $r$ . Again (III) implies that existing suppliers, possibly using a different technology, may make supernormal profits, even if the industry in region  $r$  is contestable in this way by a new firm from outside it. Further, as in the industrial contestability case, (III) implicitly incorporates opportunities for regulators to select values of contingent taxes and subsidies, or otherwise to affect contingent costs (such as interregional transmission costs) and to vary conditions of industrial contestability - both with reference to existing firms and with reference to marginal potential entrants.

In these ways, even under second best conditions, regulatory authorities may use the approach implicit in (III) to synthesize conditions of potential entry in general and for regulation via a synthesized form of market contestability in particular. (Notice that competition and collusive profit maximization could coexist in an optimal solution to (III). Depending on the magnitudes of its regulatory parameters, optimal solutions to (III) may be consistent at the same time with competition and/or contestability in one subset of regions and collusion and noncontestability in another.)

Two more classes of cases consistent with (III) are Cournot and market sharing cases. To illustrate this assume that in (III) one producer is associated with each region  $r$  and targets  $s_{rkt}^*$  relate to potential rivals' interregional shipments. If such targets are common knowledge, in principle each potential producer may determine a relatively decentralized optimum by maximizing

regional returns contingent on knowledge of the others' shipments (if any) to that region. In the (unlikely) event that such conjectures for the relevant subregion(s) correctly correspond to those required for a Cournot equilibrium that equilibrium would be attained. Similarly, if such conjectures for the relevant subregion(s) correspond to those required for a market sharing equilibrium *that* equilibrium would be attained. Indeed in principle both of these distinct types of oligopolistic conjecture may be simultaneously appropriate in different regions as parts of an overall optimal solution for (III). This underlines the fact that (III) potentially comprehends a variety of behaviours and entry conditions and that an optimal solution may be simultaneously consistent with Cournot and market sharing behaviours for effectively disjoint subsets of markets.

As one application developments in the previous section can be related to electricity generation. Economies of scope might then refer to potential for inclusion of as yet unincluded regions into a distribution network by sufficiently reducing prohibitively high interregional transmission costs to nonprohibitive levels. (An example here would be a further supergrid link from the UK to France.) Where appropriate such links could themselves be taxed or subsidized as means of executing a contestability related regulatory policy. In any case such links directly affect conditions of entry and thence of market contestability and would have implications for industrial and market concentration in electricity supply in both countries.

Economies of scope, if realized, would be the means of raising effective demand and supply for at least one interregional exporter. It may be that such an exporter would thereby be able to secure hitherto unattainable economies of scale. But economies of scale could be secured from proportionately increased demands and supplies (and non-proportionately increased costs) from existing markets in a manner consistent with definitions in Section 3, too.

# 10. Regulatory risk, windfall profits and state preference extensions.

Apart from its implications for economies of scale and scope and contestability in general, (III) also incorporates specific implications for regulation via the explicitly goal oriented regulatory variables  $\Delta p_{rt}, \Delta \omega_{irtt-h}, \Delta \tau_{irtt-h}, \Delta \omega_{irtt-h}$  relating respectively to price regulation, output process related regulation (e.g. coal vs oil generated electricity), capacity related regulation and to import related regulation.

In practice, however, economic activity takes place in an environment of risk and uncertainty with the actual regulatory framework, as well as actual sales of output and of capacity within and between regions, being state contingent. For example electricity demand will generally vary with the weather and the regulatory regime, as well as with the time of day and the cost of generating power. And, in general, both weather patterns and contingent regulatory regimes will, at best, be known in distribution. To model this assume first that the future is known in

distribution with  $\theta_{vt|t-1}$  being the (conditional) probability with which a regulatory and weather dependent state  $vt$  occurs in period  $t$  given that state  $vt-1$  occurs in period  $t-1$ .

Thus  $\theta_{vt|t-1} > 0$  implies that state  $vt$  is *accessible* from state  $vt-1$  and  $\theta_{vt|t-1} = 0$  implies  $vt$  is *inaccessible* from state  $vt-1$ . More generally,  $vtAvt-1$  will denote states  $vt$  in period  $t$  accessible from a particular state  $vt-1$  in period  $t-1$ .

Next define  $v_{vt|t-1}$  as the contingent discount rate applicable to transitions from state  $vt-1$  in period  $t-1$  to an accessible state  $vt$  in period  $t$ . Assuming for simplicity that relevant enterprises maximize a measure of expected returns and that it is understood that interperiod conservation conditions accord with accessibility for relevant state contingent transformations, (III) becomes:

$$\begin{aligned} \text{Max} \sum \Pi \theta_{vt|t-1} (1 + v_{vt|t-1})^t & [f_{rvt}(d_{rvt}) + v_{irvtt-h} \bar{K}_{irvtt-h} - c_{irvtt-h} S_{irvtt-h} - m_{irvtt-h} \bar{K}_{irvtt-h} - v_{irvt} \bar{K}_{irvtt-0} - c_{rvt}^+ d_{rvt}^+ - c_{rvt}^- d_{rvt}^- \\ & - e_{irvtt-h}^+ S_{irvtt-h}^+ - e_{irvtt-h}^- S_{irvtt-h}^- - m_{irvtt-h}^+ \bar{K}_{irvtt-h}^+ - m_{irvtt-h}^- \bar{K}_{irvtt-h}^- - c_{rkvt}^+ S_{rkvt}^+ - c_{rkvt}^- S_{rkvt}^- - c_{krvt}^+ S_{krvt}^+ - c_{krvt}^- S_{krvt}^-] \\ \text{st} \quad & d_{rvt} \leq S_{rvt-1} \\ S_{rvt} = & \sum S_{irvtt-h} + \sum S_{vkrt} - \sum S_{vrkt} \\ S_{irvtt-h} \leq & a_{irvtt-h} \bar{K}_{irvtt-h} \\ \bar{K}_{irvtt-h} + \bar{K}_{irvtt-h} = & \bar{K}_{irvt-1t-h+1} \\ d_{rvt} + d_{rvt}^+ - d_{rvt}^- = & d_{rvt}^* \\ S_{irvtt-h} + S_{irvtt-h}^+ - S_{irvtt-h}^- = & S_{irvtt-h}^* \\ \bar{K}_{irvtt-h} + \bar{K}_{irvtt-h}^+ - \bar{K}_{irvtt-h}^- = & \bar{K}_{irvtt-h}^* \\ S_{rkvt} + S_{rkvt}^+ - S_{rkvt}^- = & S_{rkvt}^* \\ S_{krvt} + S_{krvt}^+ - S_{krvt}^- = & S_{krvt}^* \\ \text{All variables nonnegative} \end{aligned} \tag{IV}$$

If  $f_{rvt}(d_{rvt}) = \int p_{rvt}(d_{rvt}) \delta d_{rvt}$  this is a multiregion and explicitly regulated variant of the model in the appendix of Littlechild (1970) with dual relations analogous to those of (III). More generally, elements of the objective and constraints of (IV) may correspond to variously collusive, Cournot, or market sharing conditions depending on the contingency which is forthcoming.

With  $f_{rvt}(d_{rvt}) = \int p_{rvt}(d_{rvt}) \delta d_{rvt}$  or  $f_{rvt}(d_{rvt}) = p_{rvt} d_{rvt}$  the objective of (IV) is concave and the Kuhn Tucker conditions are necessary and sufficient for an optimum. Associating transition contingent dual variables  $\varphi_{rvt}/\Pi(1 + v_{vt|t-1})$  etc with the constraints of (IV) in a manner analogous to the association of dual variables  $\varphi_{rt}/(1 + v)^t$  etc with the constraints of (III), the Kuhn Tucker conditions are as in (IV)' below.

$$\begin{aligned}
(1+u_{vt-1})\varphi_{rvt} + \Delta p_{rvt} &\geq \delta f(d_{rvt})/\delta d_{rvt} \\
\sum \theta_{vt} \varphi_{rvt} &\leq \xi_{irvt-1} \\
\xi_{irvt} &\leq c_{irvtt-h} + u_{irvtt-h} + \Delta \tau_{irvtt-h} \\
\xi_{irvt} &\leq \xi_{ikvt} + c_{ikrvt} + \Delta \sigma_{ikrvt} \\
\xi_{irvt} &\geq \xi_{ikvt} - c_{ikrvt} + \Delta \sigma_{ikrvt} \\
\psi_{irvtt-0} &\leq v_{irvt} \\
a_{irvtt-h} u_{irvtt-h} &\leq \psi_{irvtt-h} - \sum \theta_{vt+1}/(1+u_{vt+1}) \psi_{irvt+1t-h-1} + m_{irvtt-h} + \Delta \omega_{irvtt-h} \\
\psi_{irvtt-h} &\geq v_{irvtt-h} \\
-c_{rvt} &\leq \Delta p_{rvt} \leq c_{rvt} \\
-e_{irvtt-h} &\leq \Delta \omega_{irvtt-h} \leq e_{irvtt-h} \\
-m_{irvtt-h} &\leq \Delta \tau_{irvtt-h} \leq m_{irvtt-h} \\
-c_{irkt} &\leq \Delta \sigma_{irkt} \leq c_{irkt} \\
\varphi_{rvt}, u_{irvtt-h} &\geq 0, \quad \xi_{irvt} \text{ unrestricted}
\end{aligned} \tag{IV}'$$

Conditions (IV)' have interpretations analogous to those associated with (III)' and can be used to calculate transition dependent contingent gains or losses. To illustrate this consider one such transition: Assume that optimally demand  $d_{rvt}$  in region  $r$  and state  $vt$  is positive and is in part supplied locally from previous production  $s_{rvt-1} > 0$  in that region and in part supplied by imports. Then, by complementary slackness, the first constraints of (IV)' give:

$$d_{rvt} > 0 \Rightarrow (1+u_{vt-1})\varphi_{rvt} + \Delta p_{rvt} = \delta f(d_{rvt})/\delta d_{rvt} \tag{7}$$

$$s_{rvt-1} > 0 \Rightarrow \sum \theta_{vt} \varphi_{rvt} = \xi_{irvt-1} \tag{8}$$

That is: output is optimally supplied in period  $vt$  and region  $r$ , if at all, then to the point where marginal revenue is sufficient to recoup the transition contingent imputation to supply costs  $(1+u_{vt-1})\varphi_{rvt}$  plus consumption tax or subsidy  $\Delta p_{rvt}$ . From the second constraints, at an optimum supply will be undertaken (if at all) up to the point where the expected imputation to supply costs equals the marginal production or import cost  $\xi_{irvt-1}$  for each process  $i$  used in time state  $vt-1$ . (Notice that an alternative source of supply may be from imports via conditions implicit in the fourth constraint.)

If output is optimally positive in time state  $vt$  in region  $r$  then output must be positive for at least one technology  $i$  and age of capacity in that time

state so that, by complementary slackness:

$$s_{irvtt-h} > 0 \Rightarrow \xi_{irvt} = c_{irvtt-h} + u_{irvtt-h} + \Delta \tau_{irvtt-h} \tag{9}$$

$$\begin{aligned}
s_{irvtt-h} > 0 \Rightarrow \\
a_{irvtt-h} u_{irvtt-h} &= \psi_{irvt-1t-h} - \sum \theta_{vt+1}/(1+u_{vt}) \psi_{irvt+1t-h-1} + \\
&\quad m_{irvtt-h} + \Delta \omega_{irvtt-h}
\end{aligned} \tag{10}$$

From (9) the marginal supply cost is made up of a marginal variable cost, a marginal capacity cost and (potentially) an output specific regulatory tax or subsidy. And, from (10), at an optimum capacity of type  $i$  and age  $h$  will be maintained in time state  $vt$  in region  $r$  to the point where the marginal imputation (if any)  $a_{irvtt-h} u_{irvtt-h}$  charged to output is sufficient to meet the sum of the maintenance cost, and expected capital depreciation (or appreciation)  $\psi_{irvt-1t-h} - \sum \theta_{vt+1}/(1+u_{vt}) \psi_{irvt+1t-h-1}$  *net of* any taxes or subsidies  $\Delta \omega_{irvtt-h}$  to that element of capacity.

Finally: if supply in region  $r$  and time state  $vt$  is met in part by imports, then  $s_{krvt} > 0$  some  $k$  and, by complementary slackness:

$$s_{krvt} > 0 \Rightarrow \xi_{irvt} = \xi_{ikvt} + c_{ikrvt} + \Delta \sigma_{ikrvt} \tag{11}$$

(Of course, if any of (7)..(11) optimally corresponds to a strict inequality then by complementary slackness the optimal decision is respectively *not* to sell, produce output, maintain that type and age of capacity, or import in the relevant time state.)

If time state  $vt$  follows time state  $vt-1$  in region  $r$  (7).(11) can be used to generate output and capacity cost related variances  $\Delta(\ )$  as follows:

If  $\Delta(\varphi_{rvt}) =_{\text{def}} (\sum \theta_{vt} \varphi_{rvt} - \varphi_{rvt})$  then rearranging (7):

$$d_{rvt} > 0 \Rightarrow (1 + v_{vt-1}) [\sum \theta_{vt} \varphi_{rvt} - \Delta(\varphi_{rvt})] + \Delta p_{rvt} = \delta f(d_{rvt}) / \delta d_{rvt} \quad (12)$$

$$\text{If } \Delta(\psi_{irvt+1t-h-1}) =_{\text{def}} (\sum \theta_{vt+1} / (1 + v_{vt}) \psi_{irvt+1t-h-1} - 1 / (1 + v_{vt}) \psi_{irvt+1t-h-1}) \quad (13)$$

Then (10) gives:

$$s_{irvt-h} > 0 \Rightarrow a_{irvt-h} \mu_{irvt-h} = \psi_{irvt-1t-h} - [1 / (1 + v_{vt}) \psi_{irvt+1t-h-1} - \Delta(\psi_{irvt+1t-h-1})] + m_{irvt-h} + \Delta \omega_{irvt-h} \quad (14)$$

At an optimum (12) has the interpretation: Set sales (if any) where marginal revenue equals marginal production cost plus a state contingent tax or subsidy, *less* the transition contingent loss (or gain)  $\Delta(\varphi_{rvt})$ . And, from (13): Choose capacity of type  $i$  and age  $h$  where marginal rentals imputed to it will equal its maintenance cost plus a transition contingent capital gain or loss, plus a state contingent tax or subsidy *less* a transition specific variation  $\Delta(\psi_{irvt+1t-h-1})$ . That is: in every case, state contingent prices and valuations will include both state contingent regulatory taxes and/or subsidies and risk related and transition contingent price and capital variations  $\Delta(\varphi_{rvt})$ ,  $\Delta(\psi_{irvt+1t-h-1})$ . And, contingent regulatory variables  $\Delta p_{rvt}$ ,  $\Delta \omega_{irvt-h}$  may reinforce or mitigate the relative gains or losses  $\Delta(\varphi_{rvt})$ ,  $\Delta(\psi_{irvt+1t-h-1})$  either directly or indirectly e.g. via relatively favourable or unfavourable treatment of intraregional prices or capital allowances vis a vis prices of exports and imports. (These may also be taxed or subsidized at an optimum.)

In these ways sale price and capacity valuations will interact both with intraregional and with interregional regulatory regimes to determine transition dependent and state contingent net revenue gains or losses and net capital gains or losses. In a state contingent framework such as (IV), (IV)' such potential variations in net returns will be reflected in output and capital investment

decisions via relatively increased or decreased levels of production of output and/or of investment in capacity.

Summarizing: contingent gains or losses may stem both from relatively profit-able vs relatively unprofitable market states and from intraregional and/or interregional regulatory regimes relatively favourable or unfavourable to incumbent firms.

Of course these variously market based and regulation based effects may reinforce or weaken each other. But the important point here is that, as the state contingent framework of (IV) and (IV)' demonstrates, in such a framework each of these contingent effects would be fully anticipated. These are the main results of the paper: in a state preference based industrial and regulatory framework, regulatory risk in general, and regulatory induced windfall gains or losses, would be fully anticipated. It follows immediately that “windfall” taxes, e.g. on <sup>(13)</sup>privatization based returns to assets, even though at first seemingly retrospective in nature, would also be fully anticipated. In particular in a fully specified state preference analysis windfall profit taxes of the kind recently introduced by the UK government with reference to privatized utilities would have been fully anticipated both in magnitude and effect - albeit on a contingent basis.

More generally by endogenising contingent regulatory regimes (IV) and (IV)' also demonstrate that, if enterprises optimise their pricing and output and investment decisions in a manner analogous to that of such a regime, then finite probabilities of the exaction of state contingent “windfall” taxes will affect output and investment and export decisions, *whether or not those taxes are in fact exacted*.

## 11. Conclusion

In this chapter I have introduced new results on economies of scale and scope and have developed implications of these results for contestability and for regulation. In those contexts I have also shown how useful distinctions can be made between industrial and market contestability. Finally I have used a state preference argument to show how the contestability idea can be extended to a state preference framework and, in particular how in such a framework windfall profits or losses would

not only be fully anticipated but so would any associated contingent regulatory actions, including windfall taxes.

I conclude with three observations: Firstly, the preceding developments and interpretations are not dependent in any fundamental way on the linear form of the constraints of the models which have been used. Goal programming approaches and associated interpretations in relation to economies of scale and scope and/or in relation to regulatory prohibitions, taxes or subsidies can also be applied in entirely analogous ways to models with nonlinear production related constraints. Secondly, here quantity related goals have been related to tax or subsidy related regulatory interpretations. But these same variables also have potential interpretations as elements of decentralized solutions in general, and of contingently oligopolistic interpretations such as Cournot or market sharing interpretations in particular. For example, with reference to electricity generation (IV),(IV)' could be interpreted as potentially decentralized via appropriately output and transmission specific goals (e.g. power generating plant and distribution grid specific goals) with each enterprise potentially subject to contestability based regulation contingent on the forthcoming state. But such state contingent decentralization goals may in principle simultaneously relate to oligopolistic conjectures such as Cournot, market sharing or collusion conjectures. In such cases output and transmission variables can be directly susceptible to regulatory interventions also designed to bring about policy based changes in regional industrial and market structures.

Finally, throughout the paper the emphasis has been on regulated contest-ability and related transition processes in a state preference framework. In practice outcomes would not all be fully anticipated as a state preference structure implicitly assumes. From that perspective the present analysis and results, which have explicitly considered implications of shifts between contestability based policy regimes, can be seen as one step towards a wider framework incorporating uncertainty as well as state contingent regulatory risks into a still more comprehensive policy oriented analysis.

## References

- Armstrong, M, Cowen C. and J.Vickers, 1994, *Regulatory Reform: Economic Analysis and British Experience*, MIT Press.
- Baumol, W.J, Panzer J.C. and R.D.Willig, 1982, *Contestable Markets and the Theory of Industrial Structure*, Harcourt Brace Jovanovich.
- Baumol, W.J and R.D.Willig, 1986, "Contestability: Developments Since the Book", *Oxford Economic Papers* (supplement), 38: 9-36.
- Beesley, M., 1996, *Regulating Utilities: A Time for Change?* IEA Readings 44, Institute of Economic Affairs, London.
- Cairns, R.D., 1996, "Uncertain Contestability", *Journal of Economic Behaviour and Organization*, 30, 1, 125-131.
- Charnes, A, Duffuaa, S. and Ryan M.J., 1980, "Degeneracy and the More for Less Paradox", *Journal of Information and Optimization Sciences*, 1, 52-56.
- Charnes, A, Duffuaa, S., and Ryan, M.J., 1987, "The More for Less Paradox in Linear Programming", *European Journal of Operational Research*, 31, 194-197.
- Eaton, B.C. and Wooders, M.H., 1985, "Sophisticated Entry in a Model of Spatial Competition", *Rand Journal of Economics*, 16, 282-297.
- Littlechild, S.C. (1970). "Marginal Cost Pricing with Joint Costs", *Economic Journal*, 80, 222-235.
- Morrison S. and Winston. C., 1987, "Empirical Implications and Tests of the Contestability Hypothesis." *Journal; of Law and Economics*, 30, 53-66.
- Morrison S. and Winston, C., 1985, "Empirical Implications and Tests of the Contestability Hypothesis." Manuscript. Brookings Institution.
- Norman G. and Thisse, J-F., 1997, "Product Variety and Welfare Under Tough and Soft Pricing Regimes", *Economic Journal*, 106, 76-91.
- Ryan, M.J., 1978, "Qualitative Interpretations of Urban Phenomena", *Journal of Regional Science*, 18, 3, 383-394.
- Ryan, M.J., 1980, "More on the More for Less Paradox in the Distribution Model". In A.V.Fiacco and K.Kortanek, eds, *Lecture Notes in Mathem-atical Systems: Extremal Methods and Systems Analysis*, 275-303. Springer.
- Ryan, M.J., 1992, *Contradiction, Self Contradiction and Collective Choice*, Avebury.