STRATHCLYDE

DISCUSSION PAPERS IN ECONOMICS

THE WELFARE IMPACTS OF DISCRIMINATORY PRICE TARIFFS

BY

NIKOLAOS DANIAS AND J KIM SWALES

NO 16-09

DEPARTMENT OF ECONOMICS
UNIVERSITY OF STRATHCLYDE
GLASGOW

The welfare impacts of discriminatory price tariffs

Nikolaos Danias[a]

J. Kim Swales[b]

[a]Department of Economics, University of Strathclyde, Glasgow, UK, G4 0QU. E-mail: nikolaos.danias@strath.ac.uk

[b]Department of Economics, University of Strathclyde, Glasgow, UK, G4 0QU. E-mail:

j.k.swales@strath.ac.uk

June 28, 2016

Abstract

This paper looks at the use of asymmetric tariffs as a regulatory instrument. We use a monopolistic market setup with two markets and we introduce price controls in one of the two. The purpose of the regulator is to maximise consumer welfare through this price discriminatory practice. We consider cases where the welfare of the consumers in the two markets is weighted equally and cases where it is not. In some cases we allow for the two markets to be linked through a monopsonistic input market. The paper focuses on the welfare implications of this regulatory approach, with the firm operating under a profit restriction. Results suggest that having only one price-controlled market is in certain cases a good option from a welfare perspective.

Keywords: monopoly, asymmetric regulation, tariffs, welfare

JEL classification: D42, D61, I31, I38, L12, L51

1. Motivation

The motivation behind this paper is to shed some light on, and to help understand, the mechanisms of asymmetric regulation. Policy makers are faced with questions of how and to what extent to regulate markets. The dilemma is to balance the increased levels of control offered by stronger and tighter regulation against the outcomes that result when markets are allowed to operate based on their own devices. Relying on market operation allows profit-maximising firms to address the challenges in the business environment either using a long or short-term perspective, which can have advantages, since they have the opportunity to base their decision-making on their knowledge of the market and cost structures. Additionally, if the firm is properly incentivised to act in a way that serves its profit-maximizing interests, then enforcement of regulation may be less expensive and less complicated.

This paper addresses the issue of asymmetric regulation. In a simple model where a monopolistic/monopsonistic firm purchases in one wholesale market and sells in two retail markets, we examine the outcomes when one of the retail markets operates under price controls, whereas the other faces no regulation. An important restriction assumed for both markets is that the firm bears an obligation to meet the market demand in the controlled market, meaning that quantity rationing is not allowed in that market. The asymmetric regulatory approach (where consumers face price-discriminatory tariffs) presented in this paper could be a realistic and sensible option with real-world applications if we consider the possibility of it's being adopted : where there is lack of knowledge about costs; where there is not enough political will to introduce full tariffs; and where social policies targeted on specific populations are introduced.

In this paper we assume that the aim of the regulator is to maximize consumer welfare subject to constraints. We focus specifically on the case where there is discriminatory tariff setting. By separating markets in price-controlled and non-price-controlled ones, the regulator can engage in activities that result in indirect income redistribution, as an implied cross-subsidization probably takes place between the markets. In our paper these markets are modelled using two markets. Market 1 is a price-controlled market which can represent a market serving vulnerable consumers needing protection from high prices, and Market 2 is a non-price-controlled market which can represent a market serving the remainder of the market.

An important element of our model is that it allows us to consider the possibility that the two markets are not only served by the one firm but they are also inter-connected through their input prices. This is because the output sold to each market is supplied using input purchased in one common wholesale market. With ascending wholesale input prices, the two retail markets are interrelated as larger activity in one of them increases costs for the other one.

The model presented in this paper and the understandings that it provides are applicable in cases such as tariff-setting for transportation services (regulated as against unregulated fares), price capped finance costs for mortgages for certain groups, such as first time buyers) and control of energy prices for certain users. In these cases, asymmetric or partial tariffs can be set to serve income redistributing purposes through price discrimination driven by consumer welfare concerns for protected customer groups.

Regulation is a very popular topic in the economics literature (Bradley and Price, 1988; Vogelsang, 2002; Dobbs, 2004; Stigler and Friedland, 1962; Laffont and Tirole, 1986; Averch and Johnson, 1962; Baldwin and Cave, 1999; Newbery, 2002). Price-cap regulation is a well-known and widely used instrument of regulation which has been examined previously in the literature (Cowan, 2002; Parker, 1997; Braeutigam and Panzar, 1993).This paper takes the approach of the regulator into account and therefore the latter sections examine consumer welfare, aiming to appreciate impacts of regulation for society as a whole. The effect of regulation on consumer welfare has been examined in papers in the published literature (Clemenz, 1991; Kang et al., 2000; Sappington and Sibley, 1992; Eaton and Grossman, 1986).

Section 2 outlines the model with no regulation. Section 3 introduces a tariff in a unified retail market and discusses interventions that optimize the subsequent welfare impact. Section 4 discusses the introduction of differential tariffs in the two separate retail markets, first where unit wholesale prices are fixed and then where they rise as total output rises. Section 5 is the conclusion.

2. The basic model with no regulation

In order to ease the analysis, a number of simplifying assumptions are made. The monopoly electricity supplier faces two identical markets, each characterized by the linear inverse demand function:

(1) $$p_i = a - bq_i \qquad i = 1, 2$$

where p_i and q_i are the prices and quantities in market i, and a and b are parameters which take positive values. For heuristic reasons, that the two markets are initially assumed to be identical, so that the values of a and b do not vary across the two markets. The value taken by the parameter a is the maximum price that the monopolist can charge in either market and have non-negative sales. Therefore in all the analysis the tariff is never set above a. The parameter b is the (negative) slope of the inverse demand curve. The firm's total cost, C_T, is made up of a fixed cost, Γ, and the cost of purchasing electricity in the wholesale market, in which it acts as a monopsonist. The wholesale price of electricity, p_W , is again assumed to be a linear function of total electricity supply, q_T . This implies:

(2) $$C_T = \Gamma + p_W q_T$$

where

(3) $$p_W = c + dq_T$$

and

(4) $$q_T = \sum_i q_i = q_1 + q_2$$

Again c and d are parameters which take positive values; c is the minimum price in the wholesale market that would generate a non-negative supply and d is the slope of the wholesale electricity supply curve. We assume that both c and d are non-negative. Where d is zero, there is a constant wholesale price for electricity equal to c.

Before the introduction of any price caps, the first order conditions for profit maximization for the monopolist supplier is to set marginal revenue equal to marginal cost in each market. Using equation (1), marginal revenue in market i (MR_i) is given as:

(5) $$MR_i = \frac{d(p_i q_i)}{dq_i} = a - 2bq_i$$

From the standard monopolistic argument, $p_i > MR_i$. The linear marginal revenue curve has the familiar characteristic that it meets the price axis at the same point as the demand curve, but has a negative slope that is twice as steep.

Using equations (2), (3) and (4), the marginal cost in market i (MC_i) is derived as:

(6) $$MC_i = \frac{dC}{dq_T}\frac{\partial q_T}{\partial q_i} = MC_T = c + 2dq_T$$

Equation (6) is important for three reasons. First, the marginal cost is above the average variable cost, reflecting the monopsonistic position of the single buyer. As a result, again the marginal cost curve cuts the price axis at the same point as the wholesale electricity supply curve but is twice as steep. Second, the marginal cost is the same in each market. Third, the level of the marginal cost is dependent on the total supply of electricity. This implies that the marginal cost in one market is dependent on the output levels in both markets. For subsequent analysis this means that the marginal cost in market 2 is a function of the output in market 1 and vice versa. Through this link interventions in one market affect activity in the second.

In the case with no regulation, each market faces the same marginal cost and has an identical marginal revenue curve. Using equations (5) and (6), this implies that the unconstrained profit-maximizing outputs in markets one and two are given as:

$$q_i^U = \frac{a-c}{2(b+2d)} \qquad i = 1,2 \tag{7}$$

These quantities are positive as long as production is financially viable at all (that is as long as a > c). The corresponding prices in the individual markets can be obtained from substituting equation (7) into equation (1). This gives:

$$p_i^U = \frac{4ad+ab+bc}{2(b+2d)} \tag{8}$$

The general expression for total profits for the company takes the form

$$\Pi_T = \sum_i p_i q_i - p_W q_T - \Gamma \tag{9}$$

where the T subscripts represents the firm total summing activity in both retail markets. In the specific unconstrained case, substituting equations (3), (7) and (8) into equation (9) produces, with some manipulation:

$$\Pi_T^U = \frac{(a-c)^2}{2(b+2d)} - \Gamma \tag{10}$$

In the short run production is profitable as long as a > c. In the long-run revenue earned needs to be enough to cover the fixed costs Γ in order that non-negative profits can be made. Each market makes half of the revenue, so that the short-run profits made in each market, Π_i^U, can be identified as:

(11) $$\Pi_i^U = \frac{(a-c)^2}{4(b+2d)}$$

This gives the current profit in each market but does not allocate the fixed costs across markets.

3. Introduction of a tariff in a single, unified market

Primarily for pedagogic reasons, it is useful to begin by analyzing the impact of the introduction of a price tariff $\overline{p}_1^F$ in a single unified market and to trace the welfare implications. To be clear, in this case there is only one retail market which, for convenience, is identified as market 1. From equations (1), (2) and (3), we can determine the output and cost as linear functions of the price tariff. Applying equation (9) and setting the fixed cost, Γ, equal to zero gives profits as:

(12) $$\Pi_1^F = (\overline{p}_1^F - p_W)q_1^F = \frac{((b+d)\overline{p}_1^F - (ad+bc))(a-\overline{p}_1^F)}{b^2}$$

where $\frac{d\Pi_1^F}{d\overline{p}_1^F} = \frac{b(a+c)+2ad-2(b+d)\overline{p}_1^F}{b^2}$ and $\frac{d^2\Pi_1^F}{d(\overline{p}_1^F)^2} = \frac{-2(b+d)}{b^2} < 0$

Where $\overline{p}_1^F \geq a$, quantity demanded is zero, so that profits are therefore also zero for tariffs in this range. Similarly, where $\overline{p}_1^F = p_W$, so that the market price equals average variable cost, again profits fall to zero. From equation (12), this tariff is $\frac{ad+bc}{b+d}$. Given that equation (12) is quadratic, the tariff level which maximizes the profits (and therefore also the price level which maximizes profits in a single unified market) is the mid-point of these two tariff values and is calculated as $\overline{p}_1^F = \frac{2ad+b(a+c)}{2(b+d)}$, with a corresponding output of $\overline{q}_1^F = \frac{a-c}{2(b+d)}$.

The full welfare implications of imposing the tariff can be analysed as three elements: the consumer surplus; the producer surplus to the competitive wholesale suppliers; and the profits to the retail monopolist/monopsonist.

The consumer surplus (S_1^F) can be derived, using equation (1), as:

$$(13)\ S_1^F = \frac{(a-\overline{p}_1^F)^2}{2b}$$

where:

$$(14)\ \frac{dS_1^F}{d\overline{p}_1^F} = \frac{-(a-\overline{p}_1^F)}{b} < 0, \quad \frac{d^2 S_1^F}{d(\overline{p}_1^F)^2} = \frac{1}{b} > 0$$

Similarly, from equations (1), (3) and (4) the producer surplus in the wholesale market equals:

$$(15)\ P_1^F = \frac{d(a-\overline{p}_1^F)^2}{2b^2}$$

so that

$$(16)\ \frac{dP_1^F}{d\overline{p}_1^F} = \frac{-d(a-\overline{p}_1^F)}{b^2} < 0$$

Equation (12) gives the profits to the retailer, with

$$(17)\ \frac{d\Pi_1^F}{d\overline{p}_1^F} = \frac{2ad + b(a+c) - 2(b+d)\overline{p}_1^F}{b^2}$$

If weights ω_S, ω_P and ω_Π are attached to the consumer surplus, producer surplus and profits, the change in welfare as the tariff changes is given as:

$$(18)\ \frac{dW_1^F}{d\overline{p}_1^F} = \omega_S \frac{dS_1^F}{d\overline{p}_1^F} + \omega_P \frac{dP_1^F}{d\overline{p}_1^F} + \omega_\Pi \frac{d\Pi_1^F}{d\overline{p}_1^F}$$

Substituting expressions (14), (16) and (17) into equation (18) gives:

(19) $$\frac{dW_1^F}{d\overline{p}_1^F} = \frac{-(\omega_P d + \omega_S b)(a - \overline{p}_1^F) + \omega_\Pi (2ad + b(a+c) - 2(b+d)\overline{p}_1^F)}{b^2}$$ where

$$\frac{d^2 W_1^F}{d(\overline{p}_1^F)^2} = \frac{(\omega_P d + \omega_S b) - 2\omega_\Pi (b+d)}{b^2}.$$

Applying the first and second order conditions implies that welfare is maximized where the tariff is set at the optimal level, indicated by a star superscript

(20) $$\overline{p}_1^{F*} = \frac{\omega_\Pi (2ad + b(a+c) - a(\omega_P d + \omega_S b)}{2\omega_\Pi (b+d) - \omega_P d - \omega_S b}$$

as long as the second order conditions hold, that is:

(21) $$\omega_\Pi > \frac{\omega_P d + \omega_S b}{2(b+d)}.$$

Expression (21) illustrates an important problem for the analysis. A reduction in the tariff will always increase the producer and consumer surplus. But where $\overline{p}_1^F$ is less than the profit maximising level, a reduction in the tariff will reduce profits. If the welfare weight on profits is so low that expression (21) does not hold, the first order condition is a welfare minimum and welfare will be increased without limit as the tariff falls. Similarly, if

(22) $$\frac{a(\omega_P d + \omega_S b)}{2ad + (a+c)} > \omega_\Pi > \frac{\omega_P d + \omega_S b}{2(b+d)}.$$

then although there is a determinate welfare maximizing price tariff, it is negative.

The sensitivity of the welfare optimal tariff to the weights gives:

(23) $$\frac{\partial \overline{p}_1^{F*}}{\partial \omega_P} = \frac{-\omega_\Pi bd(a-c)}{(2\omega_\Pi (b+d) - \omega_P d - \omega_S b)^2} < 0$$

(24) $$\frac{\partial \overline{p}_1^{F*}}{\partial \omega_S} = \frac{-\omega_\Pi b^2(a-c)}{(2\omega_\Pi (b+d) - \omega_P d - \omega_S b)^2} < 0$$

Equations (23) and (24) imply that the greater the weight on consumer surplus and the producer surplus in the wholesale market, the lower the welfare maximising tariff.

Some special cases produce familiar results. Maximising welfare where $\omega_S, \omega_P = 0$, so that no concern is paid to the welfare of consumers or wholesale producers, generates a tariff set at the profit maximizing price discussed in Section 3:

$$(25)\ \overline{p}_1^{F*} = \frac{2ad + b(a+c)}{2(b+d)}$$

When $\omega_P = 0$ and $\omega_S = \omega_\Pi$, the welfare maximizing point is set where the marginal cost curve for the monopsonistic retailer cuts the demand curve, so that

$$(26)\ \overline{p}_1^{F*} = \frac{2ad + bc}{b+2d}\ .$$

Alternatively, if $\omega_S = 0$, the welfare maximizing point is where the marginal revenue curve cuts the wholesale supply curve, so that:

$$(27)\ \overline{p}_1^{F*} = \frac{ab + ad + bc}{2b+d}\ .$$

Finally, if the welfare of all actors is weighted equally, so that $\omega_S, \omega_P, \omega_\Pi = 1$, the welfare maximizing tariff is the competitive price, where the consumer demand curve cuts the wholesale competitive supply curve, so that:

$$(28)\ \overline{p}_1^{F*} = \frac{ad + bc}{b+d}\ .$$

At this point retail profits are zero.

This implies that in a unified market, if the welfare of consumers and wholesale (competitive) producers are weighted higher than the welfare of the monopolist, then the welfare optimizing price tariff leaves the monopolist making a loss. This raises the question of how this is sustained. How does the government force the profit maximizing firm to maintain production? This is another aspect of the issues raised in the discussion around expressions (21) and (22). In the analysis in subsequent sections, we therefore take the non-negative profits as a constraint in order to maintain the continuing existence of the firm. Also for pedagogic reasons we also typically give zero weight to the producer surplus in the wholesale market.

Although the aim of this section is mainly introductory, where tariffs are introduced in one market but the firm sells in two, there are conditions under which output in the second market falls to zero. The second market collapses. This is a topic we consider in greater detail later in the paper. When this occurs and only the controlled market remains operative, under subsequent reductions in the tariff rate, equations outlined in the section above will apply to the remaining single controlled market.

4. Impact of separate tariffs in each of the two markets

Where there are two retail markets, markets 1 and 2, the regulator can discriminate and differentiate the tariffs set in each. In general it will prove useful to define all the relevant variables as functions of the two price tariffs and the demand and cost parameters. We begin with the consumer surplus. Using equation (13) and defining the total consumer welfare, S_T^F , as the weighted sum of the consumer surplus in markets 1 and 2, for values of $\overline{p}_1^F, \overline{p}_2^F \leq a$:

$$S_T^F = \sum_{i=1,2} \omega_S^i S_i^F = \frac{1}{2b} \sum_{i=1,2} \omega_S^i (a - \overline{p}_i^F)^2 \tag{29}$$

where:

$$\frac{\partial S_T^F}{\partial \overline{p}_i^F} = \frac{-\omega_S^i (a - \overline{p}_i^F)}{b} \leq 0$$

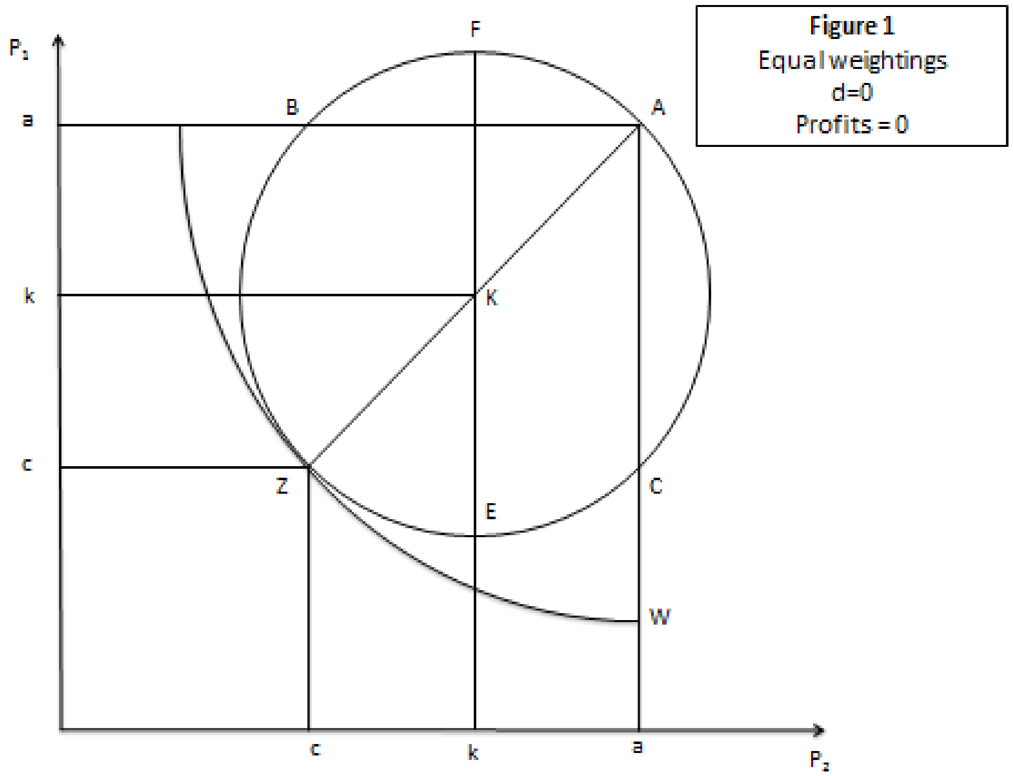
and ω_S^i is the weight put on the consumer surplus in market i.

The consumer's welfare is minimized at the value zero where $\overline{p}_1^F, \overline{p}_2^F = a$. Where the consumer's welfare is fixed at some positive level, $\overline{S}_T^F$, equation (29) can be interpreted as the associated iso-consumer welfare function, which for convenience, we refer to subsequently as the iso-welfare function. Diagrammatically, these functions take the form of segments of an ellipse which is centred on point A in $\overline{p}_1^F, \overline{p}_2^F$ space. Point A has co-ordinates (a,a) and the segments are values where $\overline{p}_1^F, \overline{p}_2^F \leq a$. One of the iso-welfare curves (where $\omega_S^1 = \omega_S^2 = 1$) is shown as W in Figure 1. In general the slope of the iso-consumer's welfare function is given as:

$$\text{(30)} \quad \frac{d\overline{p}_1^F}{d\overline{p}_2^F} = \frac{-\partial S_T^F / \partial \overline{p}_2^F}{\partial S_T^F / \partial \overline{p}_1^F} = \frac{-\omega_S^2 (a - \overline{p}_2^F)}{\omega_S^1 (a - \overline{p}_1^F)}$$

Broadly, the lower are the values of the tariffs below a, the higher the level of consumer welfare. For all iso-welfare functions, where $\overline{p}_1^F = a$, the function is vertical and where $\overline{p}_2^F = a$ it is horizontal.

Where the weights for each market are equal, the iso-welfare functions are segments of circles, as shown in Figure 1. This is illustrated by W and, from equation (30), all iso-welfare functions where the consumer surplus in each market is weighted equally have a slope equal to -1 where $\overline{p}_1^F = \overline{p}_2^F$. If the consumer's surplus in market 1 is weighted more heavily than in market 2, the iso-consumer's welfare functions have a shape similar to W_1 and W_2 in Figure 2. Whilst the curves will still be vertical and horizontal where the relevant prices take their maximum value, a, the slope where $\overline{p}_1^F = \overline{p}_2^F$ now has a lower absolute value equal to ω_S^2 / ω_S^1.



Adapting equation (15) to the case where there are two retail markets, the producer surplus, P_T^F is given as

$$(31)\quad P_T^F = \frac{d(2a - \overline{p}_1^F - \overline{p}_2^F)^2}{2b^2}$$

so that

$$(32)\quad \frac{\partial P_T^F}{\partial \overline{p}_i^F} = \frac{-d(2a - \overline{p}_1^F - \overline{p}_2^F)}{b^2} < 0$$

Any increase in price in either retail market reduces total output and therefore reduces producer surplus and at point A, where $\overline{p}_1^F = \overline{p}_2^F = a$, the producer surplus is zero, given that output is zero. The slope of the iso-producer surplus lines is given as:

$$(33)\quad \frac{\partial \overline{p}_1^F}{\partial \overline{p}_2^F} = -\frac{\partial P_T^F / \partial \overline{p}_2^F}{\partial P_T^F / \partial \overline{p}_1^F} = -1$$

The subsequent iso-producer surplus functions are therefore negative 45 degree straight lines, where further to the SW of point A, the higher the producer surplus.

Finally, to calculate the monopolist/monopsonist's profit, output in each market is determined by equation (1), average wholesale price by equation (2) and the monopolist's profits by equation (9), where in each case a tariff is imposed as the price.

(34)

$$\Pi_T^F = \frac{(b\overline{p}_1^F - bc - d(2a + \overline{p}_1^F + \overline{p}_2^F))(a - \overline{p}_1^F)}{b^2} + \frac{(b\overline{p}_2^F - bc - d(2a + \overline{p}_1^F + \overline{p}_2^F))(a - \overline{p}_2^F)}{b^2}$$

which gives

(35)

$$\Pi_T^F = \frac{-2a(bc + d(2a - (\overline{p}_1^F + \overline{p}_2^F)))}{b^2} + \frac{(bc + ab + d(2a - (\overline{p}_1^F + \overline{p}_2^F)))(\overline{p}_1^F + \overline{p}_2^F)}{b^2} - \frac{b((\overline{p}_1^F)^2 + (\overline{p}_2^F)^2)}{b^2}$$

This implies that

(36) $$\frac{\partial \Pi_T^F}{\partial \overline{p}_i^F} = \frac{4ad + bc + ab - 2d(\overline{p}_1^F + \overline{p}_2^F) - 2b\overline{p}_i^F}{b^2}, \quad \frac{\partial^2 \Pi_T^F}{\partial (\overline{p}_i^F)^2} = \frac{-2(d+b)}{b^2} < 0$$

Using expression (36), the first and second order conditions imply that profits are maximized where:

(37) $$\overline{p}_1^F = \overline{p}_2^F = \frac{4ad + bc + ab}{2b + 4d}$$

This set of tariffs equal the unconstrained monopoly prices where there are two retail markets.

Iso-profit curves for the firm are the locus of values for $\overline{p}_1^F$ and $\overline{p}_2^F$ which produce a constant profit for the firm. These are implied by equation (35). In general, this takes the form of an elipse. Initially the no-fixed-cost zero-profit iso-profit case is taken as a benchmark. Where the markets are treated identically, so that the price tariff is the same in both, zero profits occur where: $\overline{p}_1, \overline{p}_2 = a$ or $\frac{bc+2ad}{b+2d}$. Where both prices equal a, there is zero demand, therefore zero profits. Again note that *a* is an effective upper bound for the price tariff, as higher values imply negative demand. Where the two tariffs take the lower value, $\frac{bc+2ad}{b+2d}$, which is subsequently denoted as z, the tariff is set at the wholesale price. Note that with identical markets and tariffs set in both markets, their identical profit maximizing value will be denoted as k, the mean value between the two zero profit values.

4.1 The introduction of tariffs where the wholesale price is constant (d = 0)

It is useful to initially take the special case where the wholesale price is fixed at the level c and is invariant to the output level so that d = 0. From equation (35) the total profits are now given as:

$$\text{(38)} \qquad \Pi_T^F = \frac{(\overline{p}_1^F - c)(a - \overline{p}_1^F)}{b} + \frac{(\overline{p}_2^F - c)(a - \overline{p}_2^F)}{b}$$

Equation (38) can be reformulated as:

$$\text{(39)} \qquad \left[\overline{p}_1^F - \left[\frac{a+c}{2}\right]\right]^2 + \left[\overline{p}_2^F - \left[\frac{a+c}{2}\right]\right]^2 = \frac{(a-c)^2}{2} - b\Pi_T^F$$

This is the equation for the iso-profit curve where profits equal Π_T^F. In this special case it represents a circle whose center is $\frac{a+c}{2}$ and whose radius, r, is $\sqrt{\frac{(a-c)^2}{2} - b\Pi_T^F}$.

Imposing equal tariffs, the lowest value on the zero iso-profit curve, which we call more generally z, is equal to c and from equation (37) the profit maximizing prices are:

$$(40) \qquad \overline{p}_1^F = \overline{p}_2^F = \frac{a+c}{2}$$

Figure 1 shows the zero iso-profit curve in this case. Given the constraint that output cannot take negative values. The zero iso-profit curve therefore comprises an isolated point, A, and the arc BZC. At points A and Z, the tariffs in both markets take the value a and z respectively. In this case the segment BZC comprises one half of a circle whose center is the profit maximizing tariffs (k,k). This zero iso-profit curve goes through the points (c,c), (c,a) and (a,c) and has a radius equal to √2f, where $f = \frac{a-c}{2}$ so that the radius can also be expressed as $\frac{a-c}{\sqrt{2}}$.

One key element of the analysis is the locus of tariff values where the iso-profit curves are horizontal or vertical. These points identify the profit-maximising price in one market, given a specific tariff in the second market. However, a second interpretation is that these points show the minimum values of each tariff that will support a given profit level. The minimum level of $\overline{p}_1$ is shown where the iso-profit curve is horizontal, the minimum level of $\overline{p}_2$ where it is vertical.

The general slope of the iso-profit curve, not restricted to the case where d = 0, is given as:

$$(41) \qquad \frac{d\overline{p}_1^F}{d\overline{p}_2^F} = \frac{-\partial\Pi_T^F / \partial\overline{p}_2^F}{\partial\Pi_T^F / \partial\overline{p}_1^F} = -\frac{4ad + bc + ab - 2d(\overline{p}_1^F + \overline{p}_2^F) - 2b\overline{p}_2^F}{4ad + bc + ab - 2d(\overline{p}_1^F + \overline{p}_2^F) - 2b\overline{p}_1^F}$$

Where the tariffs are equal in both markets, the iso-profit curve will have a -45% slope. This reflects the symmetry in the model. It implies that the impact on profits of a change in the either price would be just offset by an equal and opposite change in the other price. The iso-profit curve is horizontal where $d\overline{p}_1^P / d\overline{p}_2^P = 0$. From equation (41) this requires that $\partial \Pi_T^F / \partial \overline{p}_2^F = 0$ which occurs where

(42) $$4ad + bc + ab - 2d(\overline{p}_1^F + \overline{p}_2^F) - 2b\overline{p}_2^F = 0.$$

Rearranging equation (42) produces the profit-maximising value of $\overline{p}_2^P$, given a tariff $\overline{p}_1^P$.

(43) $$\overline{p}_2^F = \frac{4ad + bc + ab}{2(b+d)} - \frac{d}{b+d}\overline{p}_1^F$$

Equation (43) can be seen as a reaction function: it is the price set in the unregulated market 2, $\overline{p}_2^P$, that will be the best response to the regulator setting a specific tariff, $\overline{p}_1^P$, in market 1. In the special case where d = 0, this reaction function reduces to $\overline{p}_2^F = \frac{a+c}{2}$. This simply means that if a tariff is set in market 1, the profit maximizing tariff in market 2 is the initial monopoly price and that as the tariff in market 1 falls, the maximum profit price for market 2 remains fixed at $\frac{a+c}{2}$. This is represented by the vertical line FKE in Figure 1.

With no discretionary weights (that is, setting all weights to unity) the first order conditions for maximizing welfare are that:

(44) $$\frac{\partial W_i^F}{\partial \overline{p}_i^F} = \frac{\partial S_i^F}{\partial \overline{p}_i^F} + \frac{\partial P_T^F}{\partial \overline{p}_i^F} + \frac{\partial \Pi_T^F}{\partial \overline{p}_i^F} = 0 \quad \forall_i$$

In this case, using equations (29) and (36) and noting that setting d =0 implies that the producer surplus is zero, equation (44) can be expressed as:

(45) $$\frac{\partial W_i^F}{\partial \overline{p}_i^F} = \frac{-(a - \overline{p}_i^F)}{b} + \frac{\left[(a - \overline{p}_i^F) - (\overline{p}_i^F - c)\right]}{b} = 0 \quad \forall_i$$

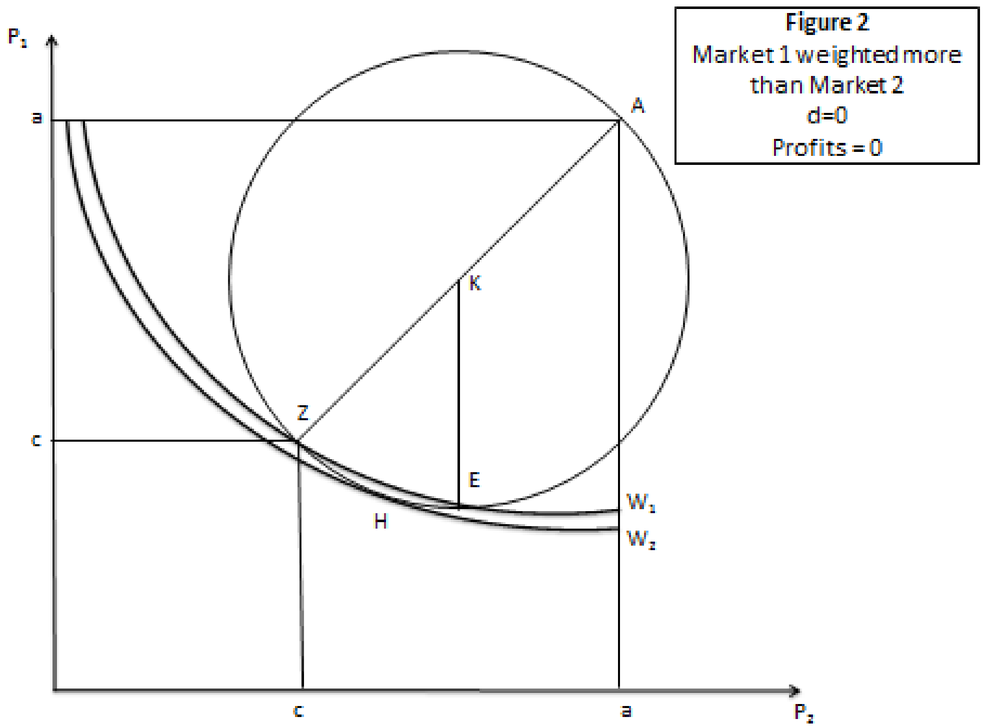
which implies $\overline{p}_1^F, \overline{p}_2^F = c$. In this case the welfare maximizing position is clear. The regulator should impose a tariff equal to c in each market and the welfare maximizing point is at Z. The monopolistic firm makes zero profits.

Figure 1 shows that the zero iso-profit curve as the half circle, center K, which goes through Z. If the welfare of consumers in each market is given the same weight, the highest attainable consumer iso-welfare function, consistent with non-negative profits is the quarter circle, center A, which again passes through Z. This is the optimal position in this case.

If the consumer welfare weights differ across the two retail markets, the optimal outcome is changed. Increasing the weight on market 1 shifts the consumer iso-welfare function so that it is tangent to the zero iso-profit curve along the segment ZE. This is illustrated in Figure 2. The iso-welfare curve W_1 that passes through point Z is no longer tangent to the zero iso-profit curve. The highest (weighted) consumer surplus is now found at point H on the iso-welfare curve W_2.

The most extreme case is where the welfare of consumers in market two is given zero weight by the regulator. In this case, the iso-welfare curves become straight horizontal lines, with lower lines producing higher consumer welfare. This would lead to an optimal outcome given by point E. The monopoly price is still charged in market 2 and the profits generated in that market are wholly used to subsidise the consumers in market 1.

The segment ZE in both Figures 1 and 2 might include a range where the tariff in market 1 is negative. Take, for example, the situation where there are no wholesale cost so that c = 0. In this case, all the points on the zero iso-profit line, apart from points A and Z, have one negative price tariff. In principle the existence of negative prices is not problematic (although negative output clearly is). Negative prices could represent the firm paying subsidies to consumers in one market, rather than charging a positive price.



4.1.1 A tariff in only one of the two markets, zero profit constraint

This paper focusses on situations where there are restrictions on the extent to which the separate markets can be regulated and, in particular, the extreme case where tariffs can only be set in one market. A specific concern is the size of the loss in efficiency that such a constraint would imply, compared to the outcome where tariffs are optimally set in both markets.

From equation (29), where consumers in each market are given a weight of 1, the iso-consumer surplus curve is given by the formula:

(46) $\bar{S}_T^F = \frac{1}{2b}\left[(a-\bar{p}_1^F)^2 + (a-\bar{p}_2^F)^2\right]$

Expression (46) is the equation of a circle whose center is point A (remembering that this only operates for values of $\bar{p}_i^F \leq a$). If R is the radius of that circle, equation (46) can be replaced by:

(47) $\bar{S}_T^F = \frac{R^2}{2b}$

Using Pythagoras's theorem, the iso-welfare function that passes through the point that maximizes the monopoly profit has a radius $\sqrt{2}f$ (recalling that we defined f, in the discussion following equation (40), as $\frac{a-c}{2}$). From equation (47) the associated consumer surplus is therefore $\frac{f^2}{b}$. With optimal regulation, given the zero profit constraint, tariffs are set at Z. The radius of the iso-welfare function passing through that point is $2\sqrt{2}f$, so that the associated consumer surplus is $\frac{4f^2}{b}$. Moving to the optimal regulation improves consumer welfare, as measured by the consumer surplus, by 400%. However, if the regulator is restricted to setting a tariff in only one market, how does this affect the resulting welfare?

With only one tariff in operation, consumer welfare is optimized at point E, and with equal consumer weights in each market the radius of the iso-welfare function that goes through this point is given as:

(48) $R^2 = (\sqrt{2}f + f)^2 + f^2 = 2(2+\sqrt{2})f^2$

Substituting this result into equation (47) gives a consumers surplus of $\frac{(2+\sqrt{2})f^2}{b}$.

This value, expressed as a ratio of the optimal welfare is $\frac{2+\sqrt{2}}{4} = 0.85$. Alternatively, if the changes in welfare, rather than the absolute values, are compared the ratio is $\frac{1+\sqrt{2}}{3} = 0.80$. Therefore although there is a loss in effectiveness in only being able to target one market, the increase in welfare is only reduced by 20% and the consumer surplus in the market not receiving the tariff is unchanged.

Moreover, in the case where d = 0, if the regulator can decide which market to target, 20% represents the maximum lost potential welfare. Where there is greater weight placed on the benefits to one set of consumers, the resulting loss in effectiveness, as against a situation where the same tariff is applied to both sets of consumers, will be lower. In fact we can use the analysis to find the relative weights that have to be placed on the consumer surplus in the two markets that would lead the regulator to actually prefer imposing a tariff in only one market, rather than the same tariff in both.

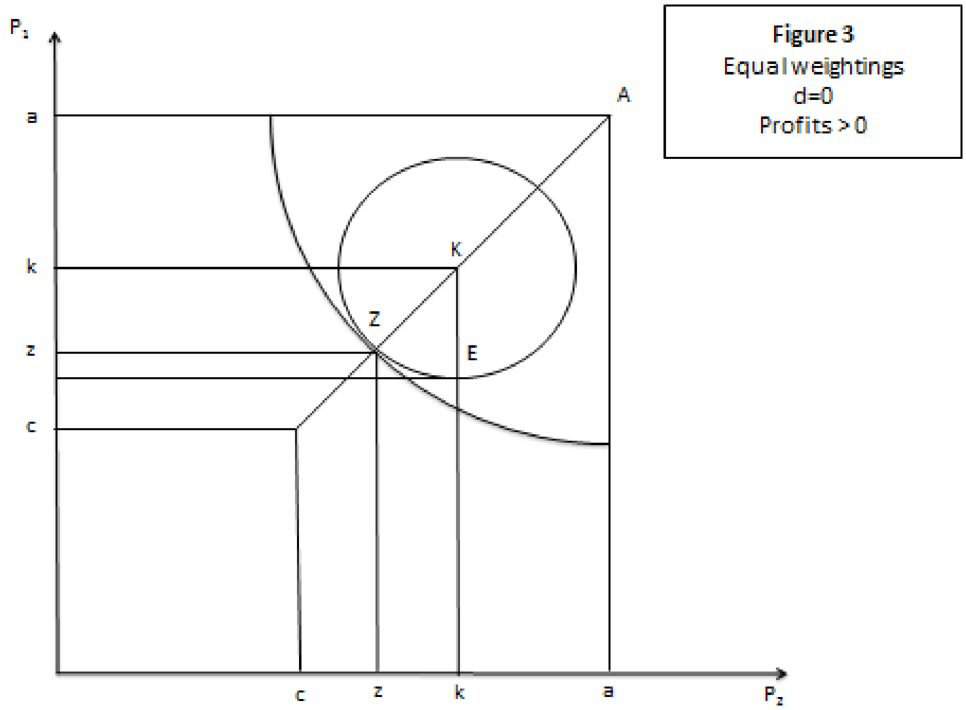
Where the two markets are weighted unequally, the maximum consumer surplus where both tariffs are constrained to take the same value is $f^2\left[\frac{2\omega_1 + 2\omega_2}{b}\right]$. Where the tariff can only apply in market 1, the maximum consumer surplus is $f^2\left[\frac{\omega_1(3+2\sqrt{2})+\omega_2}{2b}\right]$. Therefore if the regulator is faced with this choice, the optimal decision would be to apply the tariff only in market 1 if:

$$f^2\left[\frac{\omega_1(3+2\sqrt{2})+\omega_2}{2b}\right] > f^2\left[\frac{2\omega_1 + 2\omega_2}{b}\right] \rightarrow \frac{\omega_1}{\omega_2} > \frac{3}{2\sqrt{2}-1} = 1.64 .$$

Therefore as long as consumer surplus in market 1 is weighted more than 64% greater than in market 2, it is better to place an optimal tariff only in market 1 than have to impose equal tariffs in both markets.

4.1.2 A tariff in only one of the two markets, positive profit constraint

Up to now we have only considered the case where the constraint operates with zero profits. However, if fixed costs are positive, so that $\Gamma > 0$, or if the company has some kind of power with which it can push back against the regulator, the profit constraint will be positive.



The maximum, unregulated, profits, Π_{MAX}^{F}, are determined by substituting the value $\overline{p}_1^F, \overline{p}_2^F = \dfrac{a+c}{2}$ into equation (38). This gives the result:

(49) $$\Pi_{MAX}^{F} = \frac{2}{b}\left[\frac{a-c}{2}\right]^2 = \frac{2f^2}{b}$$

Using equation (49), it is convenient to express the actual profit constraint, $\overline{\Pi}_T^F$, as

(50) $$\overline{\Pi}_T^F = \overline{\rho}\Pi_{MAX}^{F} = \frac{2\overline{\rho} f^2}{b}$$

where $1 \geq \overline{\rho} \geq 0$. Substituting this into the expression associated with equation (38) for the radius of the iso-profit curve, r, gives the result that[1]

(51) $$r^2 = 2(1-\overline{\rho})f^2 \rightarrow r = \sqrt{2(1-\overline{\rho})}\, f\ .$$

Therefore if $\overline{\rho} = 0$ we have the zero profit case where r = √2f, whereas $\overline{\rho} = 1$ produces the profit maximizing case where r = 0. For values of $\overline{\pi}$ between 1 and 0, the radius of the iso-profit line lies between 0 and √2f . As the profit constraint increases the circular iso-profit curve moves closer to the profit-maximising point (k,k). Note that now the profit constraint no longer goes through point A (representing prices a,a). Also the point Z which minimizes prices in both markets, consistent with the profits constraint and price equality, now has prices greater than c.

We can use equations (46) and (50) to calculate the new maximum welfare where there is a positive profit identified by $\overline{\rho}$. The maximum value of R is now $\left[\sqrt{2} + \sqrt{2(1-\overline{\rho})}\right] f$. The maximum consumer surplus, where consumers in both markets are given a weight of unity equals:

(52) $$\overline{S}_T^F = \frac{\left[\sqrt{2} + \sqrt{2(1-\overline{\rho})}\right]^2 f^2}{2b} = \frac{\left[2 - \overline{\rho} + 2\sqrt{(1-\overline{\rho})}\right] f^2}{b}$$

where $\dfrac{\partial \overline{S}_T^F}{\partial \overline{\rho}} = \dfrac{-\left[1 + (1-\overline{\rho})^{-\frac{1}{2}}\right] f^2}{b} < 0$.

We can calculate the implication of introducing the tariff in only one market and imposing the profit constraint. Again if the tariff is imposed in market 1, the output in market 2 remains at the monopoly level and the excess profits in that market subsidise output in market 1. The maximum consumer surplus (both markets with a unitary weight) is now:

[1] Note that we use the upper-case, R, for the radius of the iso-welfare function and the lower case, r, for the iso-profit function.

(53) $$\overline{S}_T^F = \frac{\left[1+\left[1+\sqrt{2(1-\overline{\rho})}\right]^2\right]f^2}{2b} = \frac{\left[2-\overline{\rho}+\sqrt{2}\sqrt{(1-\overline{\rho})}\right]f^2}{b}$$

With $\dfrac{\partial \overline{S}_T^F}{\partial \overline{\rho}} = \dfrac{-\left[1+(2(1-\overline{\rho}))^{-\frac{1}{2}}\right]f^2}{b}$. Similarly, the ratio, Ω, of the consumer welfare where one tariff is imposed, as against a common tariff in both markets, is:

(54) $$\Omega = \frac{2-\overline{\rho}+\sqrt{2}\sqrt{(1-\overline{\rho})}}{2-\overline{\rho}+2\sqrt{(1-\overline{\rho})}} < 1.$$

However, $\dfrac{\partial \Omega}{\partial \overline{\rho}} > 0$, so that this ratio gets higher, and approaches 1 as the value of $\overline{\rho}$ approaches 1. That is to say, the proportionate loss in welfare in an optimal single market tariff, as against a common tariff in both markets falls as the profit constraint tightens.

Where the profit constraint is positive, there is an increase in welfare where the weights for market 1 and 2 are such that:

(55) $$\frac{\omega_1\left[3-2\overline{\rho}+2\sqrt{2}\sqrt{(1-\overline{\rho})}\right]f^2}{2b} + \frac{\omega_2 f^2}{2b} > \frac{\omega_1\left[2-\overline{\rho}+2\sqrt{(1-\overline{\rho})}\right]f^2}{2b} + \frac{\omega_2\left[2-\overline{\rho}+2\sqrt{(1-\overline{\rho})}\right]f^2}{2b}$$

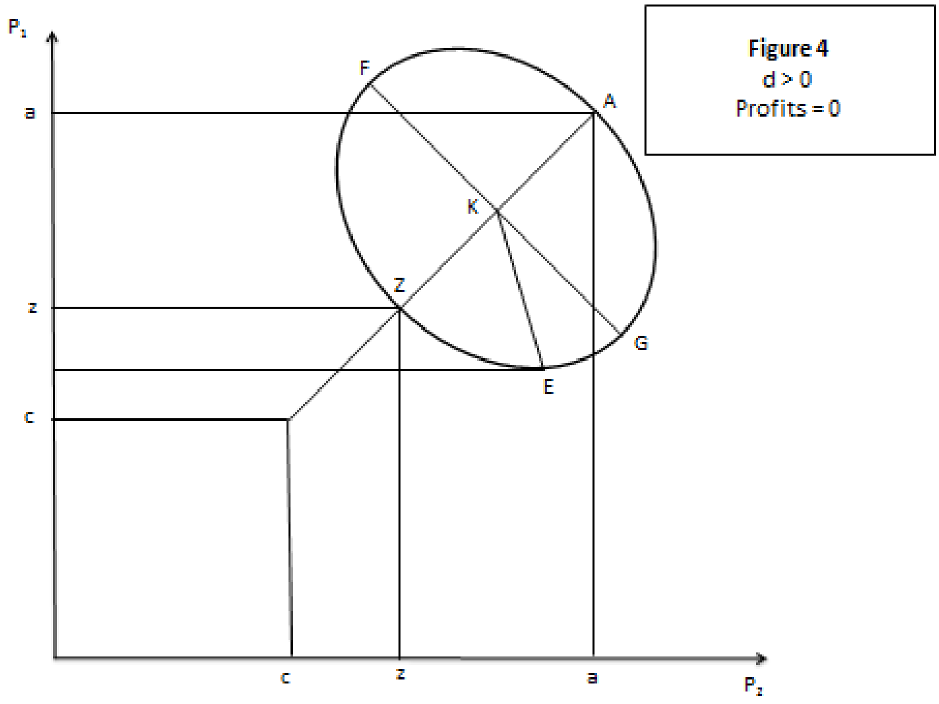
which implies $\dfrac{\omega_1}{\omega_2} > \dfrac{\left[1-\overline{\rho}+\sqrt{2}\sqrt{(1-\overline{\rho})}\right]}{\left[1-\overline{\rho}+2(\sqrt{2}-1)\sqrt{(1-\overline{\rho})}\right]}$.

This means that, for example, if ρ equals 0.5, then if $\dfrac{\omega_1}{\omega_2} > 1.38$, then it is better for the regulator to impose a tariff solely in market 1, rather than impose a common tariff in both markets.

This section essentially analyses the situation where the positive profits made in the unregulated retail market (market 2) are used to subsidise consumers in the regulated market. There is no detrimental impact on consumers in market 2. However, on the other hand, market 2 consumers receive no benefit from the regulation. In the next section, where we introduce a positively sloping wholesale supply curve, the analysis becomes more complex.

4.2 The introduction of tariffs where the wholesale price increases with output (d > 0)

In the general case, where the wholesale cost rises as total output increases the iso-profit curves are given by the expression (35). The adoption of a positive value for the wholesale supply parameter d has a number of important implications. Consider the zero iso-profit curve, and compare it to the circular curve, where d = 0, discussed in detail in Section 4.1. The new iso-profit curve, where d > 0, is illustrated in Figure 4.



The zero iso-profit curve now has the form of an oval. The point A (a,a) is still on the curve but the point Z (z,z), where wholesale cost equals retail price, has moved further up the 45 degree line through the origin: that is to say, z > c. The unconstrained profit maximizing point, K (k,k), is still at the mid-point between A and Z. The zero iso-profit curve is symmetric around the two perpendicular axes that have negative and positive 45 degree slopes and pass through the point K. The width on the 45 degree line through the origin, AZ, is now shorter than the distance along the other axis, FG. This is also verified by the fact that the points (z,a) and (a,z) are no longer on the zero iso-profit curve, but generate positive profits.

From equation (43), if the tariff is imposed only in market 1, and the firm sets the profit maximizing price in market 2, the market prices lie on a line with a slope equal to $-\frac{b+d}{d}$. This is shown as KE in Figure 4 for values of $\overline{p}_1^F \leq k$. That is to say, if the tariff in market 1 falls by 1 unit, the profit maximizing price in market 2 will increase by $\frac{d}{b+d} \leq 1$. There is now a clearer trade-off between the consumers in the separate markets. A tariff imposed solely in market 1 in a previously unregulated system means that the consumer surplus in market 2 now falls as the consumer surplus in market 1 rises,

One potential issue is whether the restriction that the price in market 2 cannot rise above the value a acts as a constraint in this case. Essentially this is asking the question: can the regulator set a tariff in market 1 such that the profit-maximising response would be to produce zero output in market 2, yet the firm would still be make positive profits?

For this to occur implies the following dual restrictions. First, the tariff in market 1 must be greater than the wholesale price for supplying the output of market 1. Simultaneously, the marginal wholesale cost at that output must be greater than a, and therefore greater than the highest marginal revenue in market 2. We proceed by finding the highest tariff, $\overline{p}_1^{F*}$, where this applies. This requires that the marginal wholesale cost just equals a, so that using equation (6):

$$(56) \qquad a = c + 2d\overline{q}_1^{F*}$$

This rearranges to

$$(57) \qquad \overline{q}_1^{F*} = \frac{a-c}{2d}$$

The tariff at which this occurs is given as:

$$(58) \qquad \overline{p}_1^{F*} = a - b\overline{q}_1^{F*} = \frac{2ad - b(a-c)}{2d}$$

This is therefore the highest tariff in market 1 at which the firm will voluntarily leave market 2. The key issue is whether the company is profitable at that tariff. This requires the tariff in market 1 to be greater than the wholesale price, p_W^{F*}. The wholesale price is given as:

$$(59) \qquad p_W^{F*} = c + d\overline{q}_1^{F*} = c + \frac{a-c}{2} = \frac{a+c}{2}$$

Therefore, for profitability:

$$(60) \qquad \overline{p}_1^{F*} \geq p_W^* \rightarrow \frac{2ad - b(a-c)}{2d} \geq \frac{a+c}{2} \rightarrow d \geq b$$

This is simply the condition that the absolute slope of the supply function is greater than the slope of the demand function. If this expression holds, the minimum price tariff $\overline{p}_1^{F**}$ consistent with zero profits and zero profit-maximising output in market 2 would be where the tariff in market 1 just equals the wholesale price. This implies that:

$$(61) \qquad \overline{p}_1^{F**} = p_W^{F**} = c + d\overline{q}_1^{F*} = c + d\left[\frac{a - \overline{p}_1^{F*}}{b}\right] = \frac{ad+bc}{b+d}$$

Therefore, where $d \geq b$, we can define a tariff range for market 1 where the profit maximising output for market 2 is zero. This range is given by:

(62) $$\frac{2ad - b(a-c)}{2d} \geq \overline{p}_1^F \geq \frac{ad + bc}{b+d}$$

If the tariff is set in this range, the analysis reverts to that outlined in Section 3. The firm will only operate in the regulated market.

It is interesting to consider the special case where b = d. In this case the zero iso-profit is horizontal at E where it cuts the vertical line through A. This implies that the minimum price tariff in market 1 consistent with non-negative profits just leads to output in market 2 becoming unprofitable. This is illustrated in Figure 5. In this case the line KE is the line given in equation (43) where in this case the slope equals – 2.

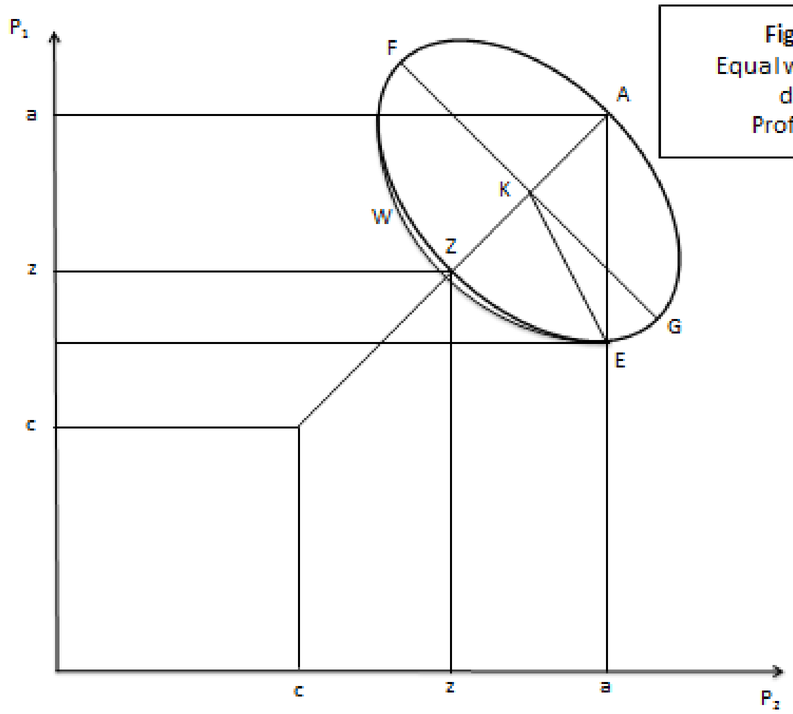


Figure 5
Equal weightings
d = b
Profits = 0

Using the expressions around equation (37), in this case the values of z and k are (2a+c)/3 and (5a +c)/6. It is useful to adopt the following notation:

(63) $$g = \frac{a-c}{6} = \frac{f}{3}$$

This means that the distances AK, AZ and AE take the values √2g, 2√2g and 3g respectively. If the only concern is the consumer surplus with differential weighting between markets, subject to a profit constraint, then it is straightforward to analyse this as in sub-section 4.1 where the wholesale price was fixed. Where there is no intervention, the outcome is at K and, using equation (47), the consumer surplus is g^2/b. Where the tariff is introduced with the same value in both markets, the outcome is at z and the consumer surplus equals $4g^2/d$. As in the case where d = 0, introducing the uniform minimum tariff increases consumer surplus to 4 times its original value. However, if the minimum tariff is only imposed in one market, in this case the consumer surplus is given as $9g^2/2b$. This gives a higher value than either of the other options. Of course, this implies that if the consumer surplus in market 1 were given a greater weight than those in market 2, this would furnish an even stronger argument for favouring the introduction of a single tariff in only that market.

We know that for $d \geq b$, then the zero profit outcome means setting price equal to the wholesale price in market 1. In that case the consumer surplus, $\overline{S}_1^F$, is given as:

(64) $$\overline{S}_1^F = \frac{b(a-c)^2}{2(b+d)^2}$$

However, the consumer surplus where the tariff is introduced in both markets simultaneously, $\overline{S}_2^F$, equals:

(65) $$\overline{S}_2^F = \frac{b(a-c)^2}{(b+2d)^2}$$

For $\overline{S}_1^F \geq \overline{S}_2^F$, then $d > \frac{b}{\sqrt{2}}$. But where market 2 is reduced to an output of zero, then $d \geq b > \frac{b}{\sqrt{2}}$, so that wherever this occurs, the choice between equal minimum tariffs in both markets and concentrating in just one market means that the consumer surplus would be maximised by just applying the tariff to one market. [2]

[2] It also means that there is a range of values for d given by $\frac{b}{\sqrt{2}} \leq d \leq b$ where a greater consumer surplus would be generated, with a zero profit constraint, by setting a tariff in market equal to the

5. Conclusions

It appears often to be the case that in markets with a degree of monopoly power, regulators wish to impose price controls in only part of. This is typically for redistributive reasons, with the welfare of one set of consumers weighted more heavily than others. In the case of the electricity market this could be motivated by concerns over fuel poverty. In this paper we formally analyse this behaviour using a very stylised model in which a monopolist serving two separate retail markets is a monopsonist in an otherwise competitive wholesale market. The paper focuses on the consumer welfare implications of imposing a price tariff in only one market, against having to impose a uniform tariff in both markets.

The analytical results suggest that the welfare costs of imposing price constraints in only one market are relatively low, and this is especially the case where the benefits to the favoured market are weighted more heavily. Where the minimum profit constraint is increased, the relative welfare loss from price control only operating in one market is further reduced. Moreover, with a degree of scarcity in the wholesale market we get the potential for a counterintuitive result. This is that there are conditions where it is better to only control one tariff, rather than impose a uniform tariff in both markets, even with neutral consumer welfare weights.

average wholesale price and a tariff higher than the marginal wholesale cost in the second market, ather than a common tariff in both markets.

References

Averch, Harvey, Leland L. Johnson, (1962), “Behavior of the firm under regulatory constraint”, The American Economic Review, Vol. 52, No. 5, pp. 1052-1069.

Baldwin, Robert, Martin Cave, (1999), “Understanding Regulation: Theory, Strategy, and Practice”, Oxford University Press.

Bradley, Ian, Catherine Price, (1988), “The economic regulation of private industries by price constraints”, The Journal of Industrial Economics, Vol. 37, No. 1, pp. 99-106.

Braeutigam, Ronald R., John C. Panzar, (1993), “Effects of the Change from Rate-of-Return to Price-Cap Regulation”, The American Economic Review, Vol. 83, No. 2, pp. 191-198.

Clemenz, Gerhard, (1991), “Optimal Price-Cap Regulation”, The Journal of Industrial Economics, Vol. 39, No. 4, pp. 391-408.

Cowan, Simon, (2002), “Price-cap regulation”, Swedish Economic Policy review, Vol. 9, pp. 167-188.

Dobbs, Ian M., (2004), “Intertemporal price cap regulation under uncertainty”, The Economic Journal, 114, pp. 421-440.

Eaton, Jonathan, Gene M. Grossman, (1986), “Optimal Trade and Industrial Policy under Oligopoly”, The Quarterly Journal of Economics, Vol. 101, No. 2, pp. 383-406.

Kang, Jaesung, Dennis L. Weisman, Mingyuan Zhang, (2000), “Do consumers benefit from tighter price cap regulation?”, Economic Letters, Vol. 67, pp. 113-119.

Laffont, Jean-Jacques, Jean Tirole (1986), “Using Cost Information to Regulate Firms”, Journal of Political Economy, Vol. 94, No. 3, pp. 614-641.

Newbery, David M., (2002), “Privatization, Restructuring and Regulation of Network Utilities”, MIT Press.

Sappington, David E.M., David S. Sibley, (1992), “Strategic Nonlinear Pricing under Price-Cap Regulation”, The RAND Journal of Economics, Vol. 23, No. 1, pp. 1-19.

Sibley, David, (1989), “Asymmetric information, incentives and price-cap regulation”, RAND Journal of Economics, Vol. 20, No. 3, pp. 392-404.

Stigler, George J., Claire Friedland, (1962), “What can regulators regulate? The case of electricity”, The Journal of Law and Economics, Volume V, pp. 1-16.

Parker, David, (1997), “Price cap regulation, profitability and returns to investors in the UK regulated industries”, Utilities Policy, Vol. 6, No. 4, pp. 303-315.

Vogelsang, Ingo, (2002), “Incentive Regulation and Competition in Public Utility Markets: A 20-Year Perspective”, Journal of Regulatory Economics, Vol. 22, Issue 1, pp. 5-27.