Optimal Concentration and Deadweight Losses in Canadian Manufacturing

by

Vaughan Dickson and Jian He

Working Paper Series 95-01



DEPARTMENT OF ECONOMICS
THE UNIVERSITY OF NEW BRUNSWICK
FREDERICTON, CANADA

Optimal Concentration and Deadweight Losses in Canadian Manufacturing

Vaughan Dickson, University of New Brunswick
Fredericton, N.B., Canada E3B 5A3

Jian He, Simon Fraser University
Burnaby, B.C., Canada V5A 1S6

May 1995

Optimal Concentration and Deadweight Losses in Canadian Manufacturing

A recurring theme in Canadian industrial organization has been that the relatively small Canadian market produces industries with too many firms too small to exhaust economies of scale.[1] More recently the possibility that industries can exhibit excessive entry has been formally analyzed by, among others, von Weizsacker (1980), Perry (1984) and Mankiw and Whinston (1986). In these models entry reduces average firm size and so thwarts the exploitation of economies of scale. Although more firms imply reduced market power, this benefit may not offset the adverse effect of higher unit cost. From a policy perspective, one solution implied by this trade-off is price regulation of a monopolist or, depending on the importance of economies of scale relative to market size, regulation of a small numbers natural oligoply. However because firms have incentives to avoid regulation and information is impacted, even the best-intentioned regulation of pricing conduct is difficult and produces its own inefficiencies. Given the difficulty of conduct regulation, a preferred direction may be to regulate structure. This leads naturally to the concept of the optimal level of industry concentration which for given firm conduct maximizes industry welfare (the sum of producers and consumers' surplus) by trading off economies of scale and market power effects.

In fact one of our principal concerns will be to estimate optimal levels of concentration in Canadian industries. Specifically we shall calculate optimal concentration for each of 107 Canadian manufacturing industries where concentration is measured by the Herfindahl index (H). This will allow us to compare optimal H to actual H and thereby define and calculate a deadweight loss as the difference between total surplus at optimal H and total surplus at actual

H. For purpose of comparison we shall also calculate a deadweight loss based on the first-best premise that effective regulation of pricing conduct is feasible. To compute optimal concentration and accompanying welfare losses, we introduce simulations of industry equilibium based on estimated industry cost functions and assumed elasticities of demand. The industry equilibrium is initially taken to be Cournot-Nash after which we also examine the impact of more collusive behaviour. Estimates of optimal concentration emerge because we are able to express both industry price and cost as functions of H.

Our closest antecedent is Perry (1984) who examines the differences in welfare produced by the number of firms in a free-entry zero-profit Cournot equilibrium and the welfare-maximizing optimal number of firms. Perry does this with simulations of hypothetical markets for a variety of assumptions about demand elasticities, scale elasticities and interactions among firms. Our analysis uses real industries, real H values and actual estimated scale elasticities.

To briefly summarize our results we find for our base case, i.e., Cournot equilibrium and unitary demand elasticities, that optimal concentration typically exceeds existing concentration. The average optimal H in our sample is .147 whereas average H is .117. In fact in 79 of the 107 industies optimal H is larger than current H. The failure to be at optimal H, either because existing H is too low or too high, produces on average a deadweight loss of 2.75 percent of shipments. We also find that higher elasticities of demand and greater scale economies lead to higher estimates of optimal H and higher deadweight losses. Finally we find that if one assumes that control of conduct is feasible then deadweight losses are larger. For example for our base case deadweight losses average 4.57 of industry shipments.

Further elaboration of results and how they compare to the more traditional

deadweight loss literature will follow after the introduction of our model and presentation of results.

The Model

The model introduced will be used to measure optimal concentration and welfare losses in 107 Canadian manufacturing industries for 1980. The model contains a representation of an industry cost function, an industry demand function and an equilibrium condition. We begin with the industry cost function.

To construct the industry cost function we first introduce a Cobb-Douglas cost function for any firm i in a particular industry. (The Cobb-Douglas form is chosen because the economies of scale parameter will be based on Cobb-Douglas production functions estimated by Baldwin and Gorecki (1986)). Accordingly:

$$(1) \qquad c_i = Aq_i^w$$

In equation (1) c_i is total cost for firm i, q_i is firm output and A is a parameter representing exogenous input prices and technology. To move from the firm to the industry we sum firm costs to obtain industry costs:

$$(2) \qquad (\Sigma c_i = C) = A\Sigma q_i^w$$

The difficulty with equation (2) is that industry cost, C, is expressed in terms of unobserved firm output instead of observed industry output. To resolve the problem we first define firm i's share of industry output (Q) as s_i

$= q_i/Q$ and rewrite equation (2) as:

$$(3) \qquad C = AQ^w \Sigma s_i^w$$

Although industry output appears in equation (3), unavailable firm data is still required to measure firm shares. To deal with this we use the approximation $\Sigma s_i^w = H^{w-1}$. This approximation is based on the Herfindahl numbers equivalent (n_e) which is the number of equal-size firms associated with a corresponding H value, i.e., $n_e = 1/H$. Thus if all firms are equal, $s_i = 1/n_e = H$ and $\Sigma s_i^w = \Sigma H^w = H^{w-1}$.[2] With this substitution, industry total and marginal cost (MC) can be expressed with industry level variables only:

$$(4) \qquad C = AQ^w H^{w-1}$$

$$(5) \qquad MC = AwQ^{w-1}H^{w-1}$$

From equation (5) it is clear that with given concentration larger industry output will lead to lower marginal cost and with given output larger concentration will produce lower marginal cost whenever there are economies of scale, i.e., $0 < w < 1$. The trade-off problem is that higher concentration also implies higher market power.

To complete the model we need a demand curve and an equilibrium condition. For the demand curve we use the iso-elastic form:

$$(6) \qquad Q_d = B/P^e$$

In equation (6) Q_d, P and e are respectively quantity demanded, price and the price elasticity of demand.

Finally for the equilibrium condition we initially use the Cournot-Nash solution. As is well-known the industry Lerner index in this circumstance is:[3]

(7) $(P-MC)/P = H/e$

Given the cost, demand and equilibrium specifications, we can derive optimal concentration and concomitant welfare losses. To this end we calibrate the model for a given industry by setting P = 1 and Q = 100 when best possible industry MC intersects industry demand. Since the MC curve plotted against industry output shifts down as H increases, the best possible level occurs at H = 1. Thus calibrated and substituting v = w-1 we can summarize the model as:

(8) $Q = 100/P^e$

(9) $MC = (Q/100)^v H^v$

(10) $P = MC\ e/(e-H)$

To find optimal H we first find equilibrium output and the associated total surplus both expressed as a function of H . We then vary H to produce alternative outputs and surpluses with optimal H occurring when surplus is maximized. Accordingly substituting equation (9) into (10) defines equilibrium price as:

(11) $P = (Q/100)^v H^v e/(e-H)$

To find associated output for a given H, H_o, we substitute (11) into the demand function:

$$(12) \qquad Q_0 = 100\,[(e-H_o)/e]^{e/(1+ev)} H_o^{-ev/(1+ev)}$$

Given (12) total surplus (TS) in the Cournot equilibrium may be evaluated as the area underneath the demand curve minus total cost:

$$(13) \qquad TS = \int_0^{Q0} (100/Q)^{1/e} dQ - [Q_o^{v+1}/(v+1)100^v] H^v$$

To find optimal H or H* we solve the first order condition $\partial TS/\partial H = 0$. Doing so produces the following quadratic equation which may be solved for H*.[4]

$$(14) \qquad (1 + v/e)H^2 - (2v + ev^2 - ev)H + ev(1+ev) = 0$$

Finally given H* we can construct a measure of the welfare loss created by the difference between H and H*. Accordingly as a proportion of industry revenue we define the welfare loss (WL):

$$(15) \qquad WL = [TS(H^*) - TS(H_0)]/(P_0 Q_0)$$

In the following sections we shall report WL and H* for 107 industries for a variety of circumstances defined by published levels for H, the estimates of economies of scale parameter v and different assumptions about demand elasticities and equilibrium conditions.

The Data and Some Conceptual Problems

In this section we introduce the data and discuss some conceptual problems presented by the economies of scale estimates. This will lead to some adjustments in the model and to the definition of an additional welfare loss measure based on control of pricing conduct.

Our calculations apply to 107 three and four digit S.I.C. industries for 1980. The data for H are from Statistics Canada (1983). For the demand elasticity e, since reliable estimates for a broad sample of manufacturing are rarely available, we use as our base case unitary elasticity but also consider e = .5 and e = 2. This range should encompass the elasticity values likely to characterize manufacturing industries.[5]

The data for the scale parameter, which is central given our interest in modelling the market power - efficiency trade-off, is taken from an important work by Baldwin and Gorecki (1986). For the year 1979 they report regression estimates of the labour and capital exponents in Cobb-Douglas production functions for 107 three and four digit manufacturing industries. The regressions for each industry are based on unpublished establishment data made available to them by Statistics Canada.[6] The estimates of the labour exponent (α) and capital exponent (β) determine our scale parameter w in the associated cost function as $w = (1/(\alpha+\beta))$. For constant returns to scale $w = 1$, $v = 0$ and for increasing returns $0 < w < 1$, $-1 < v < 0$ with lower values indicating greater economies.

Baldwin and Gorecki's results show that economies of scale typify Canadian manufacturing. Of the 107 industries 95 exhibit increasing returns to scale,

$(\alpha+\beta) > 1$, while the remaining 12 show decreasing returns. These decreasing returns to scale industries are problematical since they imply, rather improbably, that larger more successful firms have higher costs than smaller less successful firms and consequently that larger concentration does not force a trade-off between efficiency and market power. We regard these results as anomalous, perhaps an example of how even four digit SIC data are sometimes too coarse to represent meaningful economic markets, and impose for these 12 industries constant returns to scale.[7] The effect for these 12 industries is to define the value for optimal H as effectively zero since, with constant returns to scale, reductions in H produce market power benefits without adverse cost consequences.

A second problem is that Baldwin and Gorecki's scale elasticity results considerably exceed those obtained by Fuss and Gupta (1981) and Robidoux and Lester (1988) both of whom also use Canadian cross-section data. Both studies report weighted average scale elasticities of 1.06 (the proportional change in output for proportional change in inputs or costs) in contrast to the considerably higher 1.17 of Baldwin and Gorecki.[8] As pointed out by Robidoux and Lester the main reason for the discrepancy is Baldwin and Gorecki's use of value added rather than gross output as the output measure. Robidoux and Lester contend, since the evidence is that no economies of scale exist in general for intermediate inputs, that the exclusion of intermediate inputs from output and costs biases upward the estimates of scale elasticities.[9] Accordingly because we will be comparing our results to traditional deadweight loss studies that use value of shipments as the output and because we want our scale elasticity parameters to be broadly consistent with other studies, we adjusted our scale economy parameters. We did this by taking a weighted average assuming

constant returns to scale in intermediate inputs and using the estimated scale parameter for the value added inputs. The weights used were the ratio of intermediate inputs to shipments for intermediate inputs and value added to shipments for value added inputs.[10] The result is that the adjusted weighted average scale elasticity for our sample is 1.06.

The last problem concerns interpreting the scale economy parameters for firm outputs beyond the range of observed firm sizes. The problem can be illustrated using Figure 1 in which market demand and industry average cost curves are plotted against industry output. The output Q_0 represents the current Cournot equilibrium output. Two average cost curves are plotted, one for current H, H_0, and the other for the single firm case, H=1. As shown in Figure 1, given $AC = [Q^v/(v+1)100^v]H^v$, a higher H causes the industry average cost curve to shift down for a given industry output when there are economies of scale, i.e., $v < 0$.

A problem with this formulation is that whenever the estimated Cobb-Douglas exponents indicate economies of scale, as they usually do, then the first best solution (involving control of conduct) is a natural monopoly where firm output and industry output are the same. The conclusion that Canadian manufacturing is principally a sample of natural monopolies is intuitively unappealing. This conclusion results from extrapolating economies beyond the observed range of firm sizes. The Cobb-Douglas scale parameters indicate only that average costs fall throughout the range of observed firm output not throughout the range of all possible firm output. At some level of firm output we might expect economies of scale to be exhausted. In fact, for the control - of - structure case this issue is only an interpretive problem when the calculated H* implies an average firm size beyond the largest observed firm size. In contrast this issue is

always a concern when defining welfare loss based on control - of - conduct since when H = 1 optimal firm size is always beyond largest observed firm size.

To address the problem we assume that economies might reasonably be expected not to extend beyond maximum observed firm size. Maximum observed firm size q_m implies an associated maximum Herfindahl index H_m equal to q_m's share of industry output. If the calculated $H^* > H_m$ we will adjust H^* to equal H_m. This follows since if scale economies are exhausted at H_m there can be no cost benefits from increased concentration and only welfare-reducing market power effects. If $H^* < H_m$ no adjustment is required.[11]

The implications of this adjustment are shown in Figure 2 which presents three aggregate industry average cost curves each drawn given that scale economies are exhausted at largest firm size, q_m. The lowest curve represents the monopoly case. In monopoly when economies are exhausted at largest observed firm size q_m, industry average cost will first reach its lowest level at $Q = q_m$. In contrast, at existing concentration H_0, the relevant industry curve is the highest one shown in Figure 2 and economies are exhausted at industry output $Q' = q_m/H_0$. It now requires a larger industry output to exploit scale economies because there are more firms than when H = 1. The middle curve applies when concentration is H_m and all firms are scale efficient size q_m. Here scale economies are exhausted at industry output $q_m/H_m = Q_m$.

With the help of the industry demand curve a representative Cournot equilibrium is also shown in figure 2 by industry output Q_0 and price P_0 . If $H^* < H_m$ then the optimum structural policy will be a new Cournot equilibrium where the industry average cost curve will still be negatively sloped in the neighbourhood of the associated Cournot output. If $H^* \geq H_m$ then the structurally best Cournot equilibrium will be at industry output Q_m shown in

figure 2.[12]

Figure 2 also clearly shows the appropriate reference point for measurement of welfare loss when control of pricing conduct is possible. The reference point occurs when demand cuts lowest possible marginal cost which implies output Q_c in Figure 2. This reference point is considerably different from the one that would occur if scale economies were assumed, as in figure 1, to continue beyond largest observed firm size q_m.

We should also point out that figure 2 does not apply to the 12 industries with constant returns to scale. For these the average cost curve is always horizontal and the loss is the familiar deadweight loss triangle first measured by Harberger (1954). In this more traditional setting the distinction between regulation of conduct and structure dissolves since H* approaches zero. At H* = 0 the structural optimum is the same as the conduct optimum.

The last issue concerns the calculation of maximum observable plant size q_m. Since individual firm data is not available we use a proxy constructed from published data on concentration ratios (Statistics Canada 1983). The proxy is from Schmalensee (1977) and is calculated as $2a_1-a_2$ where a_1 is the average share of the four largest firms and a_2 is the average share of the fifth through eight largest firms. This proxy is used by Schmalensee to help construct a surrogate Herfindahl index (called MINL) in which the largest four firms have shares that decline linearly with the fourth firm having shares equal to the average shares of the fifth to eighth firms. Schmalensee concludes (p. 192) that the surrogate for H based on this estimate of q_m is "probably the most attractive for most circumstances".

The Results

The main results are summarized in Tables 1 to 3. Table 1 reports simple averages for existing H, optimal H and maximum H (the H occurring when all firms are at maximum observed firm size). Also given are average welfare losses for an optimal structural policy and an optimal conduct policy. For the former $H = H^*$ and conduct is not controlled. For the latter $H \geq H_m$ and conduct is controlled to ensure equality of price with average and marginal cost.

These results are presented for four cases. The first three assume Cournot equilibrium but consider alternative price elasticities of demand of .5, 1 and 2. The last case assumes unitary elasticity but with a Cournot-plus equilibrium wherein conjectural variation elasticities instead of being zero, as in a Cournot equilibrium, are set equal in each industry to existing H.[12] The Cournot-plus equilibrium upon which this formulation is based is derived in Clarke and Davies (1982) and is given by $(P-MC)/P = [H + \alpha(1-H)]/e$ where α is the conjectural variation elasticity which we set equal to H.[13]

Tables 2 and 3 report corresponding results for two sub-samples. Table 2 reports on those industries where existing concentration is too low because it is less than optimal concentration, while Table 3 reports on industries where existing concentration is higher than optimal concentration.

First examining the results for H^*, inspection of Table 1 indicates that in Cournot equilibrium higher elasticities lead to higher optimal concentrations, ceteris paribus. In fact equation (14) which determines H^* indicates that only the demand elasticity and scale economy parameter influence H^*. Concerning the scale economy parameter, the results (not shown in Table 1) are that greater scale economies lead to higher H^*, ceteris paribus. (The simple correlations

between H* and w are respectively -.88, -.79 and -.62 for e respectively .5, 1 and 2.) With respect to the industry equilibrium, Table 1 establishes that Cournot-plus behaviour decreases H*.[14]

These results make sense given H* is produced by opposing market power (welfare is increased when H is lower) and cost-reducing (welfare is increased when H is higher) forces. In particular higher demand elasticities imply lower price-cost margins, ceteris paribus, thereby weakening the market power effect and producing higher optimal H. Similarly when scale economies are greater the cost effect is strengthened with consequent increases in H*. In addition, since the Cournot-plus equilibrium has, compared to Cournot equilibrium, a more pronounced market power effect H* should be lower than in Cournot equilibrium.[15]

Interpreting the deadweight loss estimates in Table 1 is not as straightforward. First we will look at the size of the losses in Table 1 and then look at factors responsible for differences in welfare losses. The structural welfare losses generally hover around three percent of shipments. These results are higher than those reported in the more traditional welfare loss literature. For example Jones and Laudadio (1978) for the basic Harberger model (e = 1, constant returns to scale, measured profit margins) estimate losses in Canadian manufacturing in the late 1960's equal to .145 percent of GNP and Dickson and Yu (1989) estimate (e =1, constant returns and Cournot behaviour) losses equal to .88 percent of GNP for 1980 Canadian manufacturing. By comparison, if we assume our sample can be projected to represent all of Canadian manufacturing, these manufacturing welfare losses total 1.86 percent of GNP when e=1.[16] A more apposite comparison may be the conduct welfare loss estimates since these, like the traditional literature, determine losses by defining competitive

equilibrium as the relevant reference point or the opportunity foregone. For the e = 1 case welfare losses are 2.87 percent of GNP. Clearly recognizing scale economy effects makes a difference in the assessment of the welfare consequences of less than optimal structure or conduct. The losses are considerably higher than those suggested in the deadweight loss literature.[17]

The next question is what explains the differences in losses observed in Table 1? We expect welfare losses to depend on demand elasticities, scale economies, the difference between actual H and optimal H and the equilibrium concept. Looking first at demand elasticities, from Table 1 there is no discernible pattern between elasticities and structural losses but higher e's do produce lower conduct losses, ceteris paribus. This last pattern is reasonable since for a given H and market equilibrium the adverse market power consequences of H are alleviated by the higher elasticities which reduce mark-ups, ceteris paribus. Since the market power effect is blunted, the welfare loss from failure to be at the first best optimum is reduced.

This pattern is not seen in the structural losses reported in Table 1 because e also influences optimal H and structural welfare losses also depend on the difference between optimal H and actual H. With this in mind, an elasticity pattern does emerge for structural losses if we examine Tables 2 and 3. Table 2 gives results for those industries where actual H is less than H* and Table 3 gives results for those industries where actual H exceeds H*. The industries in Table 2 may be thought of as "von Weizsacker" industries while those in Table 3 are the more traditional "Bain" industries for which barriers-to-entry are too high. In Table 2 higher e's indicate higher structural losses while in Table 3 higher e's indicate lower losses. The explanation is that in Table 2 for any given actual H a higher e means higher H*, ceteris paribus, and a greater

difference between H and H* thus increasing the welfare loss. In contrast in Table 3 a higher e producing a higher H* produces a smaller difference between H and H*.

The influence of the equilibrium specification on the structural and conduct welfare losses can be determined by comparing the Cournot-plus equilibrium in Table 1 with the Cournot, e = 1 equilibrium. We see first that the conduct losses are larger for the Cournot-plus case. This occurs because, as compared to the Cournot case, more "collusive" behaviour reduces output (thus reducing output per firm and increasing unit cost) and produces a larger distortion from the competitive equilibrium or first best optimum.

The structural losses are also slightly higher for the Cournot-plus case. Justifying this is complicated because the reference point (the net surplus at H*) is different between the two equilibria because H* is different. Although average H* is considerably lower in the Cournot-plus case, the average difference between H and H* are about the same, .079 for the Cournot case and .067 for the Cournot-plus case. Since the structural loss is higher for the Cournot-plus equilibrium, it must be the case that as behaviour becomes more collusive the difference in the surplus is greater for a given difference between H and H*. In fact we know that for a perfectly collusive equilibrium a given difference between H and H* will produce the largest possible difference in surplus since there can be no adverse market power impact on net surplus (the fully collusive equilibrium being already in effect) and only beneficial cost effects.[18]

Finally, before examining the von Weizsacker and Bain industries in Tables 2 and 3, we briefly comment on the role of the scale economy parameter w in determining welfare loss. From our results it was clear that greater scale

economies, or lower w's, produced greater structural and conduct losses, ceteris paribus. This is reasonable since for a given difference between H and H* (or H_m) the gain, and hence the welfare loss from failure to be at H* or H_m, should be larger the more important are scale economies.[19]

The results for the two sub-samples given in Tables 2 and 3 indicate that for each of the four cases most industries fall into the Table 2 category for which actual concentration is less than optimal. What are the characteristics of these industries as opposed to those industries where H is too high? From Table 2 von Weizsacker industries have relatively low concentration and relatively high economies of scale (low w's), while Bain industries in Table 3 have higher H and lower scale economies. The industries in Table 2 suffer from what Mankiw and Whinston term a "business-stealing effect" wherein entry imposes an externality on other firms by reducing equilibrium output per firm and hence increasing unit costs. The result is a number of firms that is too high from a social planning perspective.

In contrast with the industries in Table 3 firm numbers are smaller but the cost benefits, given relatively minor scale economies, are too trivial to offset the adverse market power consequences from higher H. In fact this combination of low scale economies and high actual H seems to imply significant barriers to entry. However another possibility is our usual assumption that all industries share the same Cournot equilibrium is too sweeping, especially for those sub-samples with high H and almost constant returns to scale. With this kind of sub-sample a more appropriate equilibrium concept, one more consistent with the coexistence of high H and almost constant cost, may be one where there is limit pricing that yields close to a competitive outcome. If so the structural welfare losses for the Bain sub-sample would be smaller than indicated ~~and would~~

be closer to those in the traditional deadweight loss literature.

Finally we note that our calculations have examined how the surplus available jointly to producers and consumers is affected by moving to optimal concentration. A quick way to assess how consumers fare is to compare prices at existing H and optimal H. For the four cases (Cournot equilibrium with e = .5, 1 and 2 and Cournot-plus equilibrium with e = 1) the average percentage change in prices was respectively -5.5, 4.4, 5.4 and -1.9. Further inspection shows that prices increased when moving to H* for the von Weizsacker industries but decreased for the Bain industries. Thus the distribution of gains is different between the two categories. For von Weizsacker industries consumer surplus is reduced by the move to H* but this is more than offset by the increase in producers surplus. If policy makers wished to assign a greater weight to consumer welfare this would mean redefining H* to maximize an appropriately weighted sum of consumers and producers surplus. The greater the weight assigned to consumer welfare the greater the number of industries for which we would conclude that existing H is too high.

Conclusion

This paper has asked what are the optimal concentration levels in Canadian manufacturing and what are the welfare consequences of sub-optimal levels of concentration? It has answered these questions by modelling the trade-off between economies of scale and market power. This has also led to the identification of two welfare losses, one based on an optimal structural policy with conduct exogenous and the other based on the ability to control conduct and achieve a first-best optimum. In the traditional welfare loss literature no

such distinction is required since constant returns to scale are assumed and the optimal structure is always atomistic.

Our results indicate that welfare losses are more important than implied by the traditional literature whether measured using a structural or conduct reference. Furthermore in the small Canadian market the problem is generally a failure to capture scale economy benefits because H is too small, i.e., there are too many von Weizsacker industries.

Of course we need to be aware of the speculative nature of the calculations. To generate results we had to make assumptions about demand elasticities and market equilibrium and our results are conditioned by these assumptions. However at the very least the results should alert us to the real possibility that in small markets with significant scale economies there can be a problem of too many firms and too low concentration. These results also echo the Williamson (1968) injunction to pay more attention to the beneficial cost effects from merger. Finally, an implicit premise has been that controlling conduct is inherently more difficult than controlling structure. However one way to finesse the problem of controlling conduct is through the pursuit of freer trade as represented by the recently signed North American Free Trade Agreement. Through freer trade many domestic industries may face price and cost discipline that mandates both movement towards higher concentration levels and more competitive pricing.[20]

Endnotes

1. See for example Eastman and Stykolt (1967), Scherer et al (1975), Gorecki (1976), Gupta (1979), Caves et al (1980) and Cox and Harris (1984).

2. For more details on this approximation see Dickson (1994). One way to view this substitution is to interpret the estimates of optimal H as nothing more than the inverse of the optimal number of firms of equal size. Although our starting point are existing H's where firms are clearly not the same size, this initial imprecision does not affect our estimate of optimal H.

3. See, for example, Clarke and Davies (1982) for a derivation.

4. The larger root of the quadratic is used since it satisfies the requirement that H* be greater than zero but not exceed one. We should also note that a stable solution that helps ensure the second order conditions are satisfied requires the slope of the marginal revenue curve be less than the slope of the marginal cost curve. This collapses to the condition that $[1 - e(1-w)] > 0$. Also note that if $e < 1$ the definite integral does not converge. To deal with this we can think of using one unit instead of zero as the lower limit for output. This does not affect calculation of H*.

5. The rare empirical studies suggest demand elasticities generally less than one for manufacturing industries, e.g., Koutsoyiannis (1984).

6. As pointed out by Baldwin and Gorecki (p. 62) estimates built on a cross-section of individual establishment data should give more reliable results than traditional studies which use time series industry-level data. The authors also experimented with the translog form and concluded (p. 53) that for estimating scale elasticities little was to be gained from using the translog form.

7. Since Baldwin and Gorecki do not publish standard errors for their results, it may be that these 12 industries do not have returns to scale significantly different from $\alpha+\beta=1$. In fact using a probability level of .10, the authors report that two percent of their sample of industries exhibited decreasing returns.

8. These figures are as reported in Robidoux and Lester. Both studies estimate cost functions whose most general form is $lac = a + bq_i + c/q_i + f(w)$ where lac is the logarithm of average cost and $f(w)$ represents input prices. Fuss and Gupta use census size-class data for the late 1960's while Robidoux and Lester use the special data set also available to Baldwin and Gorecki.

9. Baldwin and Gorecki (p.276) did estimate a gross output Cobb-Douglas cost function for 8 industries and found considerably lower scale elasticities that were on average very close to the Robidoux and Lester results for the same industries.

10. Thus the adjusted w would be computed as 1(M/S) + w(V/S) where M designates intermediate inputs, V valued added, S Shipments and M/S + V/S = 1.

11. Without this adjustment for a Cournot equilibrium and unitary elasticity H* averages .165 for the sample, with the adjustment H* is .147. The average for H_m is .216.

12. Output Q_m, which identifies the kink in the industry average cost curve, is found using equation (12) with H set equal to H_m.

13. In generating results H, w and e are used to obtain industry price, output and cost figures. Occasionally industry price was less than industry average cost (AC). Since in Cournot equilibrium P = [we/(e-H)]AC, this meant we/(e-H) < 1 and tended to occur in industries with low H and significant scale economies (low w). Therefore to ensure a long-run equilibrium with non-negative profit we adjusted e downward in these industries to ensure we/(e-H)= 1. This problem occurred most for the e=2 case. After this adjustment the simple average elasticities for the four cases in Table 1 are respectively .48, .90, 1.49 and .97.

14. The optimal concentration equation implicitly defined by equation (14) will also be slightly changed since equation (10) will be replaced by the Cournot-plus equilibrium condition.

15. With Cournot-plus behaviour the quadratic implicitly defining H* is $B^2[(Bv(1-e)+e(1+v))/e]H^2 - B^2[2v+ev^2-ev\}H + Bev(1+ev) = 0$. The coefficient B indexes firm interaction. If B = 1 the equilibrium is Cournot and equation (14) results; if B > 1 there is Cournot-plus behaviour.

16. We note, however, that in the extreme case of a perfectly collusive or monopoly equilibrium $H^* \geq H_m$. This follows because when firms are already fully colluding there can be no additional market power effects when concentration increases, but there can be beneficial cost effects. It follows therefore that optimal concentration must occur when these cost effects are exhausted, i.e., at maximum H or larger.

17. In 1980 shipments for all manufacturing were 168 billion while our sample shipments were 99.3 billion. Hence welfare losses were increased by 168/99.3 before calculating the loss as a percentage of GNP.

18. For a review see Hay and Morris (1991). These losses are considerably less than one percent of GNP as long as unitary demand elasticity is assumed. Cowling and Mueller(1978) generate larger numbers (4 to 7 percent) by assuming monopoly pricing in industries and hence e = P/(P-MC). We also get large structural losses if we assume monopoly pricing. This is because increasing H is harmless because market power is already fully exploited and there are gains from the reduced production cost. In fact for e=2 we found an average structural welfare loss equal to 4.61 percent of GNP and a conduct welfare loss equal to 29 percent of GNP.

19. In fact if industries in our sample were perfectly collusive the structural welfare loss would be 7.74 percent of shipments when e = 2 as compared to 3.06 in Table 1. With perfect collusion H* is always identical to maximum H.

20. To illustrate for the Cournot e = 1 case we estimated with least squares a specification relating welfare loss to scale economies and the absolute difference between H* and H_0, HDIF, i.e., $WL = a_0 + a_1 HDIF + a_2 w$. For the 107 industries we found (t-statistics in brackets) WL = .32 + .26HDIF (6.4) - .34w (7.9).

21. A study by Cox and Harris (1985) estimates the cost of protection to Canada using a small general equilibrium model that specifies scale economies in the manufacturing sector. It is interesting to note that although they obviously employ a much different general equilibrium model, as opposed to our partial equilibrium approach, their estimate (4.13 percent of GNP) of the welfare gain from unilateral free trade (which would produce more efficient scale and more competitive pricing) is reasonably close to our conduct loss estimates.

References

Baldwin, J.R. and P.K. Gorecki (1986) The Role Of Scale in Canada-U.S. Productivity Differences in the Manufacturing Sector 1970-1979 (Toronto: University of Toronto Press)

Caves, R.E., M.E. Porter, A.M. Spence and J.T. Scott (1980) Competition in the Open Economy: A Model Applied to Canada (Cambridge: Harvard University Press)

Clarke, R. and S.W. Davies (1982) 'Market structure and price-cost margins.' Economica 49, 277-87.

Cowling, K. and D.C. Mueller (1978) 'The social cost of market power' Economic Journal 88, 727-48.

Cox, D. and R. Harris (1985) 'Trade liberalization and industrial organization: some estimates for Canada' Journal of Political Economy 93, 115-145.

Dickson, V.A. and W. Yu (1989) 'Welfare losses in Canadian manufacturing under alternative oligopoly regimes' International Journal of Industrial Organization 7, 257-67.

Dickson, V.A. (1994) 'Aggregate industry cost functions and the Herfindahl index' Southern Economic Journal 60, 445-52.

Eastman, H.C. and S. Stykolt (1967) The Tariff and Competition in Canada (Toronto: Macmillan)

Fuss, M. and V.K. Gupta (1981) 'A cost function approach to the estimation of minimum efficient scale, returns to scale and suboptimal scale' European Economic Review 15, 123-35.

Gorecki, P.K. (1976) Economies of Scale and Efficient Plant Size in Canadian Manufacturing Industries (Ottawa: Department of Consumer and Corporate Affairs, Bureau of Competition Policy Research Monograph No. 1)

Gupta, V.K. (1979) 'Sub-optimal capacity and its determinants in Canadian manufacturing industries' Review of Economics and Statistics 61, 506-12.

Harberger, A. C. (1954) 'Monopoly and resource allocation' American Economic Review Papers and Proceedings 44, 77-87.

Hay, D.A. and D.J. Morris (1991) Industrial Economics and Organization (New York: Oxford University Press)

Jones, J.C.H. and L. Laudadio (1978) 'The empirical basis of Canadian anti-trusr policy: resource allocation and welfare losses in Canadian industry' Industrial Organization Review 6, 49-59.

Koutsoyiannis, A. (1984) 'Goals of oligopolistic firms: an empirical test of competing hypotheses' Southern Economic Journal 50,540-67.

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

Perry, M.K. (1984) 'Scale economies, imperfect competition and public policy' Journal of Industrial Economics 32, 313-33.

Robidoux, B. and J. Lester (1988) 'Econometric estimates of scale economies in Canadian manufacturing' Working paper 88-4, Department of Finance, Ottawa

Scherer, F.M., A. Beckenstein, E. Kaufer, R. Bougeon-Maassen (1975) The Economics of Multi-plant Operation: An International Comparisons Study (Cambridge: Harvard University Press)

Schmalensee, R.C. (1977) 'Using the H-index of concentration with published data' Review of Economics and Statistics 59, 186-93.

Statistics Canada (1983) Industrial Organization and Concentration in the Manufacturing, Mining and Logging Industriues (Ottawa)

von Weizsacker, C.C. (1980) 'A welfare analysis of barriers to entry' Bell Journal of Economics 11,399-420.

Williamson, O.E. (1968) 'Economies as an anti-trust defence: the welfare trade-offs' American Economic Review 58, 18-36.

FIGURE 1

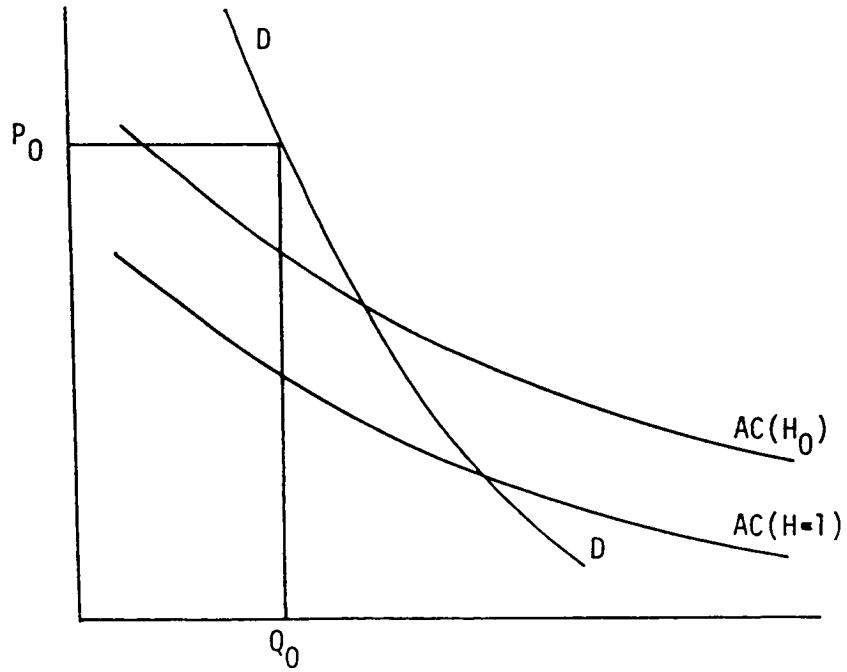


FIGURE 2

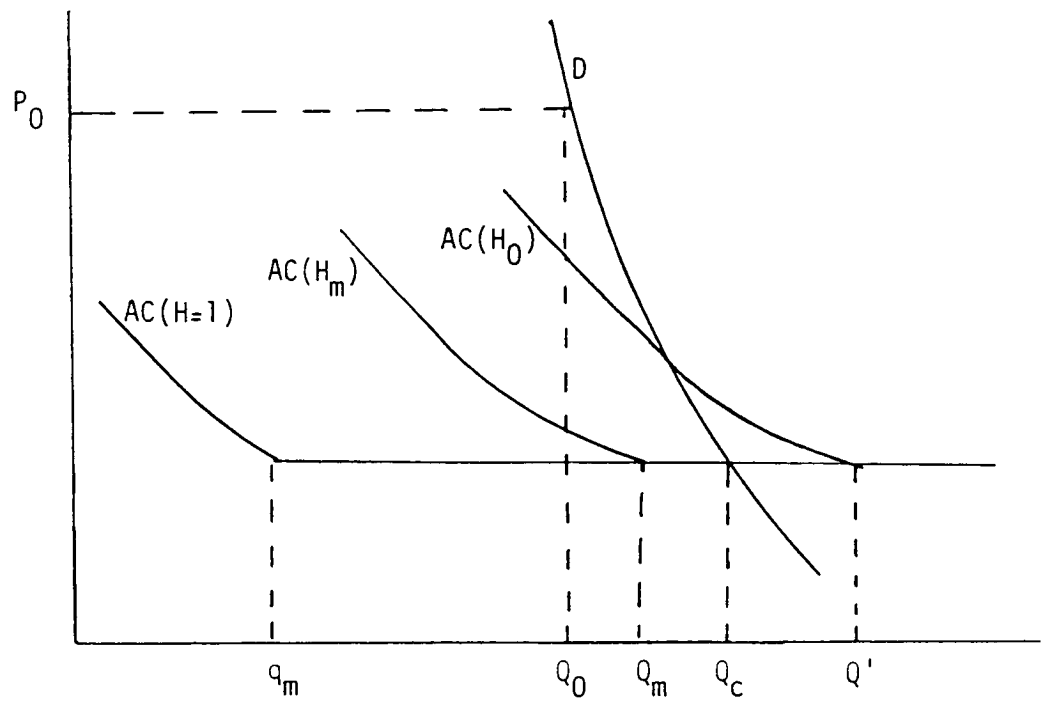


TABLE 1

OPTIMAL CONCENTRATION AND DEADWEIGHT LOSSES: ALL INDUSTRIES

Elasticity/ Equilibrium	H	Optimal H	Maximum H	WL (Structure)	WL (Conduct)
.5/Cournot	.117	.109	.216	2.82	5.67
1/Cournot	.117	.147	.216	2.75	4.57
2/Cournot	.117	.170	.216	3.06	4.19
1/Cournot-plus	.117	.118	.216	3.30	7.22

TABLE 2

OPTIMAL CONCENTRATION AND DEADWEIGHT LOSSES:

INDUSTRIES WITH H ≤ OPTIMAL H

Elasticity/ Equilibrium	Number of Industries	H	Optimal H	W	WL (Structure)	WL (Conduct)
.5/Cournot	61	.074	.126	.927	2.88	5.81
1/Cournot	79	.099	.173	.930	3.35	5.35
2/Cournot	85	.107	.199	.926	3.74	5.01
1/Cournot-plus	68	.084	.137	.925	3.83	6.77

TABLE 3

OPTIMAL CONCENTRATION AND DEADWEIGHT LOSSES;

INDUSTRIES WITH H > OPTIMAL H

Elasticity/ Equilibrium	Number of Industries	H	Optimal H	W	WL (Structure)	WL (Conduct)
.5/Cournot	46	.173	.087	.958	2.75	5.49
1/Cournot	28	.167	.072	.980	1.06	2.37
2/Cournot	22	.155	.060	.995	.47	.98
1/Cournot-plus	39	.175	.084	.967	2.39	8.00

APPENDIX A Optimal Concentration

We derive optimal concentration when e>1 and e<1.

1. Optimal concentration when e>1

Optimal concentration gives maximum total welfare. First we rewrite the expression for total welfare from equation (13):

$$TS = \int_0^Q P(Q)\, dQ - C(Q,H) \qquad (a\text{-}1)$$

$$\text{where } Q = 100\,[(e-H)/e]^{e/(1+ev)} / H^{ev/(1+ev)} \qquad (a\text{-}2)$$

$$P = (100 / Q)^{1/e} \qquad (a\text{-}3)$$

$$C = (1/(1+v))\ 100^{-v}\ Q^{1+v}\ H^{v} \qquad (a\text{-}4)$$

Substituting (a-3) and (a-4) into (a-1), we have:

$$TS = 100^{1/e}\ Q^{1-(1/e)} /[1-(1/e)] - [1/(1+v)]\ 100^{-v}\ Q^{1+v}\ H^{v} \qquad (a\text{-}5)$$

We substitute Q from (a-2) into (a-5), and have:

$$TS = 100^{1/e}\ \{\ 100\ [(e-H)/e]^{e/(1+ev)} / H^{ev/(1+ev)}\ \}^{1-(1/e)} /[1-(1/e)]$$
$$- (1/(1+v))\ 100^{-v}\ \{\ 100\ [(e-H)/e]^{e/(1+ev)} / H^{ev/(1+ev)}\ \}^{1+v}\ H^{v}$$

We simplify the above as:

$$TS = 100\ H^{(v-ev)/(1+ev)}\ \{\ [(e-H)/e]^{(e-1)/(1+ev)}/(1-1/e)$$
$$- [(e-H)/e]^{(e+ev)/(1+ev)} /(1+v)\ \} \qquad (a\text{-}6)$$

We then take partial differentiation on (a-6) with respect to H:

$$\partial TS/\partial H = 100\,[(v-ev)/(1+ev)]\,H^{[(v-ev)/(1+ev)]-1}\{\,[1/(1-1/e)]$$
$$[(e-H)/e]^{(e-1)/(1+ev)} - [1/(1-v)]\,[(e-H)/e]^{(e+ev)/(1+ev)}\,\}$$
$$+\ 100\,H^{(v-ev)/(1+ev)}\{\,[1/(1-1/e)]\,[(e-1)/(1+ev)]$$
$$[(e-H)/e]^{[(e-1)/(1+ev)]-1}\,(-1/e)$$
$$-\ [1/(1+v)]\,[(e+ev)/(1+ev)]\,[(e-H)/e]^{[(e+ev)/(1+ev)]-1}\,(-1/e)\,\} = 0$$

We simplify the above as:

$$100\,[(v-ev)/(1+ev)]\,H^{(v-ev)/(1+ev)}\,H^{-1}\{[1/(1-1/e)][(e-H)/e]^{(e-1)/(1+ev)}$$
$$-\ [1/(1-v)]\,[(e-H)/e]^{[(e-1)/(1+ev)]+1}\,\}$$
$$+\ 100\,H^{(v-ev)/(1+ev)}\{\,[1/(1-1/e)]\,[(e-1)/(1+ev)]$$
$$[(e-H)/e]^{[(e-1)/(1+ev)]-1}\,(-1/e)$$
$$-\ [1/(1+v)]\,[(e+ev)/(1+ev)]\,[(e-H)/e]^{(e-1)/(1+ev)}\,(-1/e)\,\} = 0$$

We then re-arrange the above as:

$$H^{(v-ev)/(1+ev)}\,[(e-H)/e]^{(e-1)/(1+ev)}\,\{\,[(v-ev)/(1+ev)]\,H^{-1}\,[1/(1-1/e)]$$
$$-\ [(v-ev)/(1+ev)]\,H^{-1}\,[1/(1-v)]\,[(e-H)/e]$$
$$+\ \ [1/(1-1/e)]\,[(e-1)/(1+ev)]\,[(e-H)/e]^{-1}\,(-1/e)$$
$$-\ [1/(1+v)]\,[(e+ev)/(1+ev)]\,(-1/e)\,\} = 0$$

Then we have:

$$[(v-ev)/(1+ev)]\,H^{-1}\,[1/(1-1/e)]$$
$$-\ [(v-ev)/(1+ev)]\,H^{-1}\,[1/(1+v)]\,[(e-H)/e]$$
$$+\ \ [1/(1-1/e)]\,[(e-1)/(1+ev)]\,[(e-H)/e]^{-1}\,(-1/e)$$
$$-\ [1/(1+v)]\,[(e+ev)/(1+ev)]\,(-1/e)\,\} = 0$$

We simplify the above as:

$$(1 + v/e)\ H^2 - (2v + ev^2 - ev)\ H + ev(1+ev) = 0 \qquad (a\text{-}7)$$

Since $1 \geq H \geq 0$, we find the optimal H as:

$$H^* = \frac{(2v+ev^2-ev) + \sqrt{(2v+ev^2-ev)^2 - 4(1+v/e)[ev(1+ev)]}}{2\ (1+v/e)}$$

2. Optimal Concentration when e<1

When e<1, the integral in the expression (a-1):

$$TS = \int_0^Q P(Q)\ dQ - C(Q,H)$$

does not converge if we integrate from zero. To deal with the problem, we integrate from Q=1 instead of zero and have:

$$TS = \int_1^Q P(Q)\ dQ - C(Q,H)$$

Assuming TS is TS_0 when $H=H_0$, we formulate the following formula to represent the difference between TS at any H and TS_0:

$$\begin{aligned} DS &= \int_1^Q P(Q)\ dQ - C(Q,H) - \int_1^{Q_0} P(Q)\ dQ - C(Q_0,H_0) \\ &= \int_{Q_0}^Q P(Q)\ dQ - [\ C(Q,H) - C(Q_0,H_0)\ ] \\ &= 100^{1/e}\ Q^{1-(1/e)} / (1-(1/e)) - (1/(1+v))\ 100^{-v}\ Q^{1+v}\ H^v \\ &\quad - [100^{1/e}\ Q_0^{1-(1/e)}/(1-(1/e)) - (1/(1+v))\ 100^{-v}\ Q_0^{1+v}\ H_0^v\ ] \end{aligned}$$

If there is a H^* which maximizes TS, then this H^* also

maximizes DS. We take partial differentiation on DS with respect to H, and have:

$$\partial DS/\partial H = \partial \{ 100^{1/e} Q^{1-(1/e)} /[1-(1/e)] - [1/(1+v)] 100^{-v} Q^{1+v} H^{v} - [100^{1/e} Q_0^{1-(1/e)} /[1-(1/e)] - [1/(1+v)] 100^{-v} Q_0^{1+v} H_0^{v}]/\partial H$$

Since

$$\partial [100^{1/e} Q_0^{1-(1/e)} /[1-(1/e)] - [1/(1+v)] 100^{-v} Q_0^{1+v} H_0^{v}]/\partial H = 0$$

We have:

$$\partial DS/\partial H = \partial\{ 100^{1/e} Q^{1-(1/e)}/[1-(1/e)] - [1/(1+v)] 100^{-v} Q^{1+v} H^{v} \}/\partial H \qquad (a-8)$$

We note that the expression in the bracket in the above equation is exactly the same as the expression for TS in equation (a-5). Therefore the first order condition from (a-8) to solve for H^* is the same as (a-7) and the expression to solve for optimal concentration is unchanged when e>1 and e<1.