Aggregation problems in the non-interactive equilibrium theory of markets

Steve Keen, University of Western Sydney,
John Legge, La Trobe University,
Geoffrey Fishburn,University of New South Wales, &
Russell Standish, University of New South Wales[1]

Abstract: The standard (non-game-theory) arguments used to derive the model of perfect competition are erroneous. Stigler disproved the proposition that the individual firm's demand curve is horizontal in 1957, and his number-of-firms convergence argument is invalid given a minimum firm size; a firm in a multi-firm industry maximises profit by setting its marginal revenue above marginal cost; and definitive output comparisons cannot be made between different industry structures under conditions of diminishing marginal productivity. We conclude that the theory of firms and markets should be made consonant with empirical research on the output and pricing behaviour of actual corporations.

Keywords: Microeconomics, competition, monopoly, perfect competition, oligopoly, profit maximization, Industrial Organization

JEL Classifications: A20, C71, D20, D21, D41, D42, D43, D46, D50, D60, L11, L13, L21

Aggregation problems in the non-interactive equilibrium theory of markets

Abstract: The standard (non-game-theory) arguments used to derive the model of perfect competition are erroneous. Stigler disproved the proposition that the individual firm's demand curve is horizontal in 1957, and his number-of-firms convergence argument is invalid given a minimum firm size; a firm in a multi-firm industry maximises profit by setting its marginal revenue above marginal cost; and definitive output comparisons cannot be made between different industry structures under conditions of diminishing marginal productivity. We conclude that the theory of firms and markets should be made consonant with empirical research on the output and pricing behaviour of actual corporations.

Keywords: Microeconomics, competition, monopoly, perfect competition, oligopoly, profit maximization, Industrial Organization

JEL Classifications: A20, C71, D20, D21, D41, D42, D43, D46, D50, D60, L11, L13, L21

Aggregation problems in the non-interactive equilibrium theory of markets

Undergraduate instruction in economics normally puts as follows the proposition that, in competitive markets with profit-maximising atomistic firms, equilibrium price equals marginal cost.

Firstly, demand (P), profit (Π), total revenue ($TR = P \cdot q$) and total cost (TC) are defined as functions of quantity (q for the individual firm, $Q = \Sigma q$ for the industry). For an individual firm, profit is therefore:

A. $\Pi(q) = P(Q)q - TC(q)$ (1)

Next, given the assumed properties of $P(Q)$ and $TC(q)$, it is asserted that profit is maximised where the first derivative of profit is zero, and as profit maximisers, all firms are assumed to produce this level of output in equilibrium:

$$\Pi_{\max}(x) : P(Q) + q \cdot P'(q) - TC'(q) = 0 \quad (2)$$

Finally, the assumption is made that the individual competitive firm is so small relative to the entire industry that the derivative of market price with respect to its output q is zero:

$$P'(q) = \frac{dP(Q)}{dq} = 0 \quad (3)$$

Given this assumption, marginal revenue equals market price for the individual competitive firm:

$$MR(q) = TR'(q) = P(Q) \quad (4)$$

Profit for the competitive firm is therefore maximised when its marginal cost ($MC(q) = TC'(q)$) equals the market price:

$$MC(q) = P(Q) = MR(q) \quad (5)$$

However, as Stigler showed in 1957, the slope of the demand curve for the individual firm cannot be zero, but is instead identical to that for the market (Stigler 1957: 8). His expression of this was succinct:

$$\frac{dP}{dq_i} = \frac{dP}{dQ}\frac{dQ}{dq_i} = \frac{dP}{dQ} \quad (6)$$

Thus the demand curve for the individual firm can only be horizontal if the market demand curve itself is horizontal. However, Stigler proposed a reformulation of marginal revenue for the individual firm which he alleged showed that the individual firm's marginal revenue converged to price as the number of firms in an industry increased. He began with $\frac{dP}{dq_i} = \frac{dP}{dQ}$ and then worked with the simplification of identical output levels ($Q = n \cdot q$) to derive:

$$\begin{aligned} \frac{d}{dq_i}(P \cdot q_i) &= P + q\frac{dP}{dQ} \\ &= P + \frac{Q}{n}\frac{P}{P}\frac{dP}{dQ} \\ &= P + \frac{P}{n \cdot E} \end{aligned} \quad (7)$$

where $E = \frac{P}{Q}\frac{dQ}{dP}$ is the market elasticity of demand. Curiously writing this out in full English as

$$MarginalRevenue = Price + \frac{Price}{NumberofSellers \times MarketElasticity} \quad (8)$$

he argued that "this last term goes to zero as the number of sellers increases indefinitely" (Stigler 1957: 8). If this were true, then for the very large number of firms that are necessary for the model of perfect competition, "marginal revenue equals price" would be strictly false, but a reasonable approximation to the truth.

However, it is obvious from (7) that $\frac{P}{n \cdot E} \equiv q\frac{dP}{dQ}$, and therefore the only conditions under which Stigler's supposition could be true are that $q \to 0$, or that $\frac{dP}{dQ} \to 0$.[1] If instead q is strictly greater than zero and $\frac{dP}{dQ} < 0$, then $\frac{P}{n \cdot E}$ will not go to zero as $n \to \infty$, because changes in E and P will exactly counterbalance changes in n. This can be made obvious by substituting $Q/q = n$ into Stigler's final line:

$$
\begin{aligned}
P + \frac{P}{n \times E} &= P + \frac{P}{\frac{Q}{q} \cdot \frac{P}{Q}\frac{dQ}{dP}} \\
&= P + \frac{P}{\frac{P}{q} \cdot \frac{dQ}{dP}}
\end{aligned}
\tag{9}
$$

The final line of (9) is of course the same as RHS of the first line of (8). If there is a minimum firm size$q = q_{min} > 0$, then the gap between market price and the firm's marginal revenue is completely independent of the term n. As an illustration, consider a linear demand curve, with n identical firms each producing q_{min} units of output:

$$
\begin{aligned}
P &= a - bQ \\
n &= \frac{Q}{q_{min}} \\
E &= \frac{P}{Q}\frac{dQ}{dP} = \frac{a-bQ}{Q} \cdot \frac{-1}{b} \\
\frac{P}{n \times E} &= \frac{a-bQ}{\frac{Q}{q_{min}} \cdot \frac{a-bQ}{Q} \cdot \frac{-1}{b}} \\
&= -q_{min} b
\end{aligned}
\tag{10}
$$

The difference between market price and own-output marginal revenue is thus a constant, regardless of the number of firms in the industry, and is equal to the minimum feasible firm size times b — the linear equivalent to Stigler's more general term in (7). Given that b>0 (and in general $\frac{dP}{dQ} < 0$), the only way that price can equal marginal revenue is if the representative firm in the industry produces an infinitesimal output.

This is a commonplace assumption — derived from the proposition that an infinite number of firms produce a finite output — in the literature on the convergence of Cournot/Nash equilibrium to the competitive equilibrium (Novshek & Sonnenschein 1983; Mas-Colell 1983; Novshek 1985; and others). However, this assumption raises insurmountable logical problems, both for the theory of exchange and for the concept of diminishing marginal productivity.

With respect to the theory of exchange, if each firm produces an infinitesimal output, then a consumer has to purchase all the output of a (lesser) infinity of producers in order to acquire a single unit of a commodity. The consumer must then integrate this infinity of infinitesimals to produce the single unit, thus making the consumer into a producer.

The consumer is also clearly in a monopsony position with respect to the infinity of firms from which it purchases, which raises further logical conundrums. Any attempt to counter this by presuming that each consumer is in turn an infinitesimal part of demand for each firm rapidly reduces the theory to the level of farce.

With respect to the theory of production, an essential aspect is diminishing marginal productivity due to the existence of fixed costs in the short run. If there were no fixed costs then all inputs would be variable, diminishing marginal productivity could not occur, and marginal cost would therefore be constant. Yet for the output of the representative firm to be infinitesimal, fixed costs must be zero. The assumption of a limit of zero output for the individual firm is therefore inconsistent with fixed costs and diminishing marginal productivity.

We therefore contend that any propositions derived from the assumption of infinitesimal output levels should be inadmissible in economics. There must be some minimum size for the firm $q_{min} \gg 0$ so that fixed costs exist, and so that the firm's customers are not in a near-monopsony position with respect to it.

Given the existence of a minimum firm size, $\frac{P}{n \cdot E}$ does not converge to zero as $n \to \infty$, but becomes a constant once this minimum size is reached. For a numerical example, consider an industry with a linear demand curve given by $P = 100 - \frac{1}{100000000}Q$ and technology such that the output of minimum size firm is $q_{min} = 1000$. $\frac{P}{n \cdot E}$ will fall as n rises, until such time as $q = \frac{Q}{n}$ reaches this minimum scale: from that point on, $\frac{P}{n \cdot E} = q_{min}\frac{dP}{dQ} = -\frac{1}{100000}$. While $-\frac{1}{100000}$ is clearly small, it is not zero, and therefore own-output marginal revenue for the individual firm must be less than the market price.

Stigler's argument therefore does not justify the propositions that marginal revenue converges to price for the individual competitive firm, or that equating marginal cost to price is in the limit a profit maximising equilibrium. However, Stigler's analysis is strongly related to a proposition first put by Cournot, that the market equilibrium converges to the competitive equilibrium as the number of firms rises, if each firm equates its own-output marginal revenue and marginal cost.[2]

This proposition is true, but there are two fundamental problems with using it either to model non-interactive profit-maximising behaviour, or to underwrite the relevance of perfect competition. Firstly, equating own-output marginal cost to own-output marginal revenue is not profit maximising in a multi-firm industry. Secondly, a firm cannot be presumed to know its own-output marginal revenue without requiring each firm to possess complete information about the behaviour of other firms.

1 Profit maximisation in a multi-firm industry

Leaving aside the second problem for the moment, if every firm equates its own-output marginal revenue to its marginal cost, then industry output will converge to the perfectly competitive level as the number of firms increases. However, it is easily shown that this is profit maximising behaviour only for a monopoly, because only then is the firm's marginal revenue exclusively determined by its own output. In a multi-firm industry, total revenue for the *i*th firm is a function of its own output q_i plus that of the rest of the firms in the industry Q_R, so that a change in the firm's revenue can be described as:

$$dTR_i(Q_R, q_i) = \left(\frac{\partial}{\partial Q_R} P(Q) \cdot q_i\right) dQ_R + \left(\frac{\partial}{\partial q_i} P(Q) \cdot q_i\right) dq_i \quad (11)$$

Without knowledge of the output variation decisions of all other firms, a single firm in a multi-firm industry is therefore not able to identify changes in its total revenue that are exclusively the result of variations in its own output. However it is possible to work out a general profit-maximisation formula for the single firm by establishing the aggregate industry output level that would result if each firm in the industry did equate its marginal cost to its own-output marginal revenue. In this derivation we continue with $\frac{dP}{dq_i} = \frac{dP}{dQ}$, and assume constant identical marginal costs ($MC(q_i) = MC$). This assumption is justified later in the paper, and the results generalised to differing rising marginal costs.

$$\begin{aligned}\sum_{i=1}^{n}\left(\frac{d}{dq_i}(P(Q)\times q_i - TC_i(q_i))\right) &= \sum_{i=1}^{n}\left(P(Q)+q_i\frac{d}{dq_i}P(Q)\right)-\sum_{i=1}^{n}\left(\frac{d}{dq_i}TC_i(q_i)\right)\\ &= nP(Q)+\sum_{i=1}^{n}\left(q_i\frac{d}{dQ}P(Q)\right)-\sum_{i=1}^{n}MC\\ &= nP(Q)+\frac{d}{dQ}P(Q)\sum_{i=1}^{n}q_i - n\cdot MC\\ &= nP(Q)+Q\frac{d}{dQ}P(Q)-n\cdot MC\\ &= (n-1)P(Q)+\left(P(Q)+Q\frac{d}{dQ}P\right)-n\cdot MC\\ &= (n-1)P(Q)+MR(Q)-n\cdot MC\\ &= 0\end{aligned} \tag{12}$$

This can be rearranged to yield

$$MR(Q)-MC=-(n-1)(P(Q)-MC) \tag{13}$$

Since price exceeds marginal cost, the second term on the RHS of (16) is positive, while -(*n-1*) is negative for *n>1*. The RHS is thus negative, indicating that industry marginal cost *exceeds* marginal revenue if each firm equates its own-output marginal revenue to marginal cost. Equating own-output marginal cost to marginal revenue is therefore profit maximising only for a single firm industry (where $n=1$). In multi-firm industries, setting own-output marginal revenue equal to marginal cost ignores the fact that the firm's true marginal revenue is affected not only by its own output, but also by the outputs of all other firms.

A profit maximising firm facing constant marginal costs should therefore be able to deduce that to maximise its profits, it should produce where its own-output marginal revenue *exceeds* its marginal cost. This conclusion, which generalises to any firm facing "well behaved" demand and cost functions, of course contradicts a long-held belief in economics.[3]

2 The equilibrium profit maximisation rule for a multi-firm industry

We use the aggregate relation under Cournot output levels to work out the profit-maximising output level for the individual firm. Since producing where own-output marginal revenue equals marginal cost results in industry marginal cost exceeding marginal revenue by $(n-1)(P(Q)-MC)$, a reasonable deduction is that to maximise profits, each firm should produce an output where the gap between its own-output marginal revenue and marginal cost equals $1/n$ times this aggregate excess of marginal cost over its marginal revenue (we now write $MC(q_i)$ in place of MC since the result will be generalised to differing quantity-dependent marginal cost functions):

$$MR_i(q_i)-MC_i(q_i)=\tfrac{n-1}{n}(P(Q)-MC_i(q_i)) \tag{14}$$

This formula obviously works for a monopoly, where it confirms that a monopoly maximises profit equating it marginal cost and marginal revenue. For a multi-firm industry, it indicates that industry output will be the same as for a monopoly regardless of the number of firms. We know from (12) that the LHS of (14) sums to $(n-1)P(Q)+MR(Q)-n\cdot MC$. Summation of the RHS yields:

$$\sum_{i=1}^{n}\Big[\frac{n-1}{n}(P(Q)-MC_i(q_i))\Big]=\frac{n-1}{n}\sum_{i=1}^{n}[P(Q)]-\frac{n-1}{n}\sum_{i=1}^{n}[MC_i(q_i)]$$
$$=\frac{n-1}{n}n\cdot P(Q)-\frac{n-1}{n}\sum_{i=1}^{n}MC \qquad (15)$$
$$=(n-1)\cdot P(Q)-(n-1)\cdot MC$$

The only output level under which these summations are consistent is where $MR(Q)=MC$: the "monopoly" output level where aggregate marginal revenue equals aggregate marginal cost.

We illustrate this outcome and generalise it to differing marginal cost functions by considering firstly the n-firm case with identical constant marginal costs and a linear demand curve $P(Q)=a-bQ$, secondly a duopoly with linear total cost functions $(k_i+c_iq_i)$, and finally the general formula for firms with a well-behaved market demand function and differing marginal cost functions.

Consider n identical firms each producing q units at a constant marginal cost $MC_i=c$. If every firm produce the Cournot output level, then

$$MR_i-MC_i=P-bq-c=0$$
$$a-bQ-bq-c=0$$
$$bnq+bq=a-c \qquad (16)$$
$$q=\frac{a-c}{b(n+1)}$$
$$Q=nq=\frac{n}{n+1}\frac{a-c}{b}$$

If, on the other hand, every firm produces the profit-maximising amount given by (14), then we have that in equilibrium:

$$MR_i-MC_i=P-bq-c=\frac{n-1}{n}(P-c)$$
$$q=\frac{P-c}{bn}$$
$$q=\frac{a-bnq-c}{bn} \qquad (17)$$
$$q=\frac{1}{2}\frac{a-c}{nb};and$$
$$Q=nq=\frac{1}{2}\frac{a-c}{b}$$

The monopoly level of output thus applies, regardless of the number of firms in the industry.

In the duopoly case with differing marginal costs, the profit functions for the two firms are:

$$\Pi_1(q_1)=(a-b(q_1+q_2))q_1-(k_1+c_1q_1)$$
$$\Pi_2(q2)=(a-b(q_1+q_2))q_2-(k_2+c_2q_2) \qquad (18)$$

As is well known, the Cournot equilibrium levels of output in this situation are:

$$\begin{bmatrix} q_1 \\ q_2 \end{bmatrix}=\begin{bmatrix} \frac{1}{3}\frac{a-2c_1+c_2}{b} \\ \frac{1}{3}\frac{a+c_1-2c_2}{b} \end{bmatrix} \qquad (19)$$

Total market output is $Q=q_1+q_2=\frac{2a-(c_1+c_2)}{3b}$, and if $c_1=c_2=c$, this is $\frac{2}{3}\frac{a-c}{b}$, which is larger than the monopoly level of $\frac{1}{2}\frac{a-c}{b}$. According to our guidelines however, profit maximising firms should set the derivatives of these profit functions to:

$$\frac{d}{dq_1}(\Pi_1(q_1)) = \frac{1}{2}(P(q_1+q_2)-c_1)$$
$$\frac{d}{dq_2}(\Pi_2(q_2)) = \frac{1}{2}(P(q_1+q_2)-c_2) \qquad (20)$$

The solutions to these equations are:

$$\begin{bmatrix} q_1 \\ q_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{8}\frac{2a-3c_1+c_2}{b} \\ \frac{1}{8}\frac{2a+c_1-3c_2}{b} \end{bmatrix} \qquad (21)$$

and the sum of these two amounts is

$$Q = q_1 + q_2 = \frac{a-\frac{c_1+c_2}{2}}{2b} \qquad (22)$$

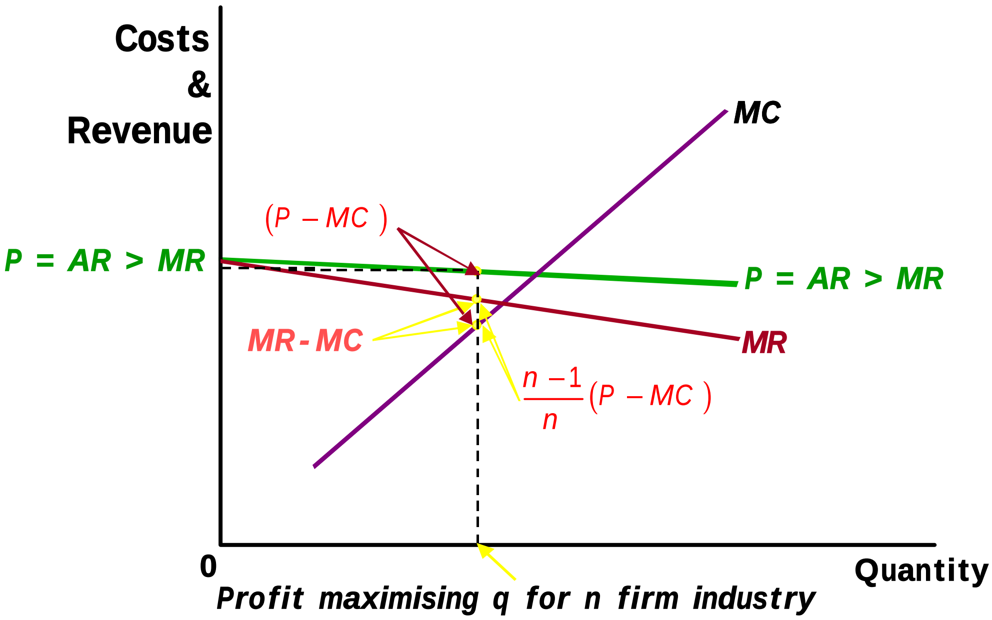
This quantity is evidently less than the Cournot level, and equivalent to the output of a monopolist if $c_1 = c_2 = c$, with the attendant maximum level of profit. This result scales to the n-firm case with differing constant marginal cost functions.

The formulae for an *n*-firm industry with any "well-behaved" demand and marginal cost functions are:

$$q_i = \frac{1}{n}\frac{P-MC_i(q_i)}{-\frac{dP}{dQ}}$$
$$Q = \frac{\left(P-\frac{1}{n}\sum_{i=1}^{n} MC_i(q_i)\right)}{-\frac{dP}{dQ}} \qquad (23)$$

We emphasise that these results are achieved not through collusion, but through rational non-interactive profit maximising behaviour. The "monopoly" level of output is reached by non-interacting firms independently working out the production level which yields them the highest equilibrium level of profit — assuming that they know their own-output marginal revenue. In graphical terms, this means that the point of profit maximisation for an individual firm in a multi-firm industry is not given by the point of intersection of the marginal cost and marginal revenue curve, but by the location at which the *gap* between marginal revenue and marginal cost equals $\frac{n-1}{n}$ times the gap between price and marginal cost. Figure 1 illustrates this:

Figure 1. Graphical illustration of true profit maximisation point.



Though derived with respect to a single market without entry, this profit-maximisation formula also applies to multi-market models with entry, since it is obvious that the profit maximising condition $MR_i(q_i) - MC_i(q_i) = \frac{n-1}{n}(P(Q) - MC_i(q_i))$ will apply in all industries. Free entry (Novshek 1980) will therefore not change the outcome that each industry will converge to the monopoly solution, regardless of the number of firms, if the condition $\sum_{i=1}^{n} MC_i(q_i) = n \cdot MC$ holds. We return to the interpretation of this shortly.

3 The information problem

The problem with the above analysis (and with much of game theory) is that it presumes a firm in a multi-firm industry can identify the change in its total revenue that is exclusively the product of changes in its own output. This assumption is only sustainable if either (a) own-output marginal revenue is a known constant independent of changes in own output or (b) the firm has knowledge of the output decisions of all other firms. Condition (a) has been decided in the negative above. Condition (b) would enable the firm to differentiate changes in its revenues that emanate from its own actions from those that emanate from the actions of others. However, it requires "competitive" firms to act semi-collusively, by sharing information on what their output decisions are, yet not to act collusively otherwise. This "schizophrenic-monopoly" condition is hardly compatible with the modelling of competitive behaviour, whether this presumes non-interactive or interactive (game theory) postures.

If this information condition is rejected, then it is impossible for firms to identify their own-output marginal revenue curves (though they can be assumed to know their marginal cost functions). All they can do is "grope" in output/revenue space amongst all the other "groping" firms in the industry, varying output in some algorithmic way in an attempt to maximise profits. The question as to the output/price combination this process converges to, and whether it is dependent on the number of firms in the industry, can be answered by computer modelling.

We take our lead here from Friedman 1953, who argued that economic agents do not consciously solve the optimising formulae taught to students — whether that of the conventional belief or the correct formula provided in (14) — but that any agent who did not behave "as if" it was doing this would fail. Giving the example of expert billiard players, Friedman argued that:

> "Excellent predictions would be yielded by the hypothesis that the billiard player made his shots *as if* he knew the complicated mathematical formulas …, and could then make the balls travel in the direction indicated by the formulas. Our confidence in this hypothesis ... derives rather from the belief that, unless in some way or other they were capable of reaching essentially the same result, they would not in fact be *expert* billiard players." (Friedman 1953: 21)

We now pose the question: if a firm simply "groped" for a level of output on the basis of whether its profit rose or fell, would its output level converge to where its own-output marginal revenue equals its marginal cost, and therefore industry output converge to where price equals marginal cost as the number of firms increases? Or would each firm produce where its own-output marginal revenue exceeds its marginal cost as in equation (14), so that industry output converges to the monopoly level where aggregate marginal revenue equals marginal cost?

We consider this question using three programs: a simple "blind groping" algorithm where all firms commence from the Cournot output level (Figure 2); a second where firms begin with randomly allocated initial outputs and the experiment is repeated multiple times (Figure 3); and a third, multi-agent implementation.[4]

In Program One, firms begin with a specified level of output, and a specified amount by which they all alter output in a search for the level that returns the maximum profit. The search procedure followed independently by each firm is to change output in an initially randomly determined positive or negative direction. If this change in output increases profit, then the firm continues altering its output in this direction; if it reduces profit, then the firm alters direction. The algorithm is repeated a large number of times by each firm.

Figure 2. Program One: fixed step size from specified initial output level

```
Mkt(n, initial, stepsize) :=  | F ← matrix(n, 1, M)·initial
                              | dq ← sign(rnorm(n, 0, 1))·stepsize
                              | for  i ∈ 0.. 2·round( q_CO(n) / stepsize )
                              |     | Quantity_i ← ΣF
                              |     | Profit_A ← P(Quantity_i)·F − tc(F)
                              |     | F ← F + dq
                              |     | Profit_B ← P(ΣF)·F − tc(F)
                              |     | dq ← vectorize( sign(Profit_B − Profit_A)·dq )
                              | Quantity
```

Our hypothetical industry has the following linear demand curve:

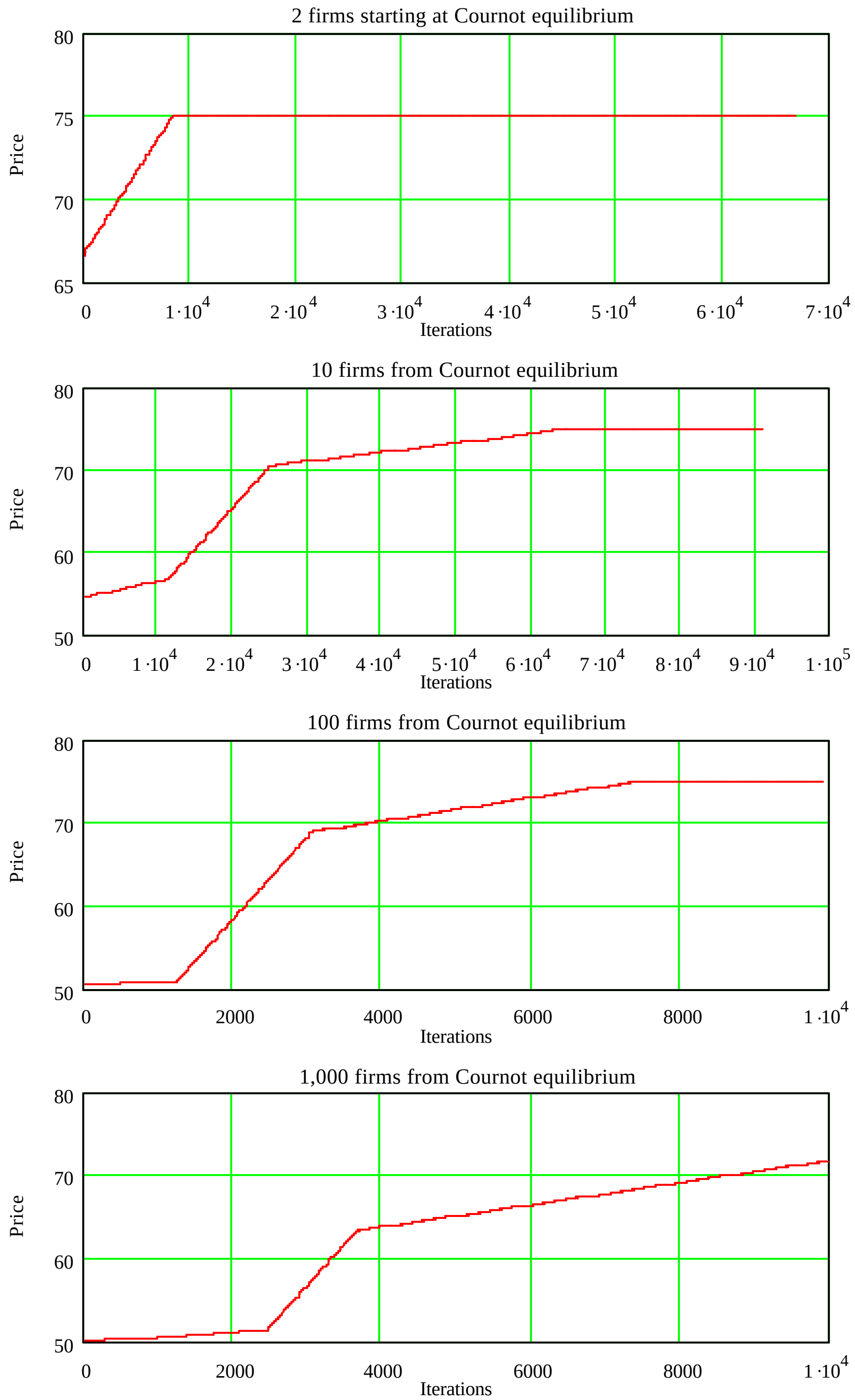
$$P(Q) = 100 - \frac{1}{100000000}Q \qquad (24)$$

Our example uses a constant cost production function, with of course the same function for all firms:

$$TC(q) = 50q \qquad (25)$$

With these parameters, the profit maximising level of output for a monopoly is *2,500* million units at a price of *75*, while the quantity at which price equals marginal cost is *5,000* million units at a price of 50. In the simulations shown in Figure 3, all firms begin with the output level predicted for the specified number of firms by the accepted formula (equation (16)). Each firm then varies its output by *stepsize* (50,000 for the duopoly, 10,000 for the ten and hundred firm case, and 1,000 for the thousand firm simulation). As Figure 3 shows, in all cases the aggregate output and price converge to the monopoly solution from the Cournot.

Figure 3. Convergence to Monopoly equilibrium from Cournot "equilibrium"



Program Two (Figure 4) assigns a random initial production level to each firm (bounds are set to avoid aggregate output levels that could result in an initially negative market price). Each firm works out its profit, and makes a random change in its output level. If this change in output decreases its profit, the firm changes its output in the opposite direction, but by a smaller amount. The same process is repeated on subsequent iterations. If accepted theory is correct, the point of convergence should be a function of the number of firms, with a single firm converging to the monopoly values and a many-firm industry converging towards perfect competition. 50 runs were undertaken, and the initial and final aggregate quantities and prices calculated.

Figure 4. Program Two: multiple runs, random initial outputs, diminishing stepsize

$$
\begin{aligned}
&\mathrm{Mkt}(n,m,l) := \ \text{for } j \in 0..m \\
&\qquad F \leftarrow \mathrm{round}\left(\mathrm{rnorm}\left(n, \frac{q_{pc}+Q_M}{2\cdot n}, \frac{Q_M}{2\cdot n}\right)\right) \\
&\qquad \mathrm{Quantity}_{ini_j} \leftarrow \sum F \\
&\qquad dq \leftarrow \mathrm{round}\left(\mathrm{rnorm}\left(n, 0, \frac{Q_M}{2\cdot n}\right)\right) \\
&\qquad \text{for } i \in 0..l \\
&\qquad\qquad \mathrm{Profit}_{pre} \leftarrow \overrightarrow{\left(P\left(\sum F\right)\cdot F - tc(F)\right)} \\
&\qquad\qquad F \leftarrow F + dq \\
&\qquad\qquad \mathrm{Profit}_{post} \leftarrow \overrightarrow{\left(P\left(\sum F\right)\cdot F - tc(F)\right)} \\
&\qquad\qquad dq \leftarrow \overrightarrow{\left(\mathrm{sign}\left(\mathrm{Profit}_{post} - \mathrm{Profit}_{pre}\right)\cdot dq\right)}\cdot\frac{l-i}{l} \\
&\qquad \mathrm{Quantity}_j \leftarrow \sum F \\
&\quad \mathrm{augment}\left(\mathrm{Quantity}_{ini}, \mathrm{Quantity}\right)
\end{aligned}
$$

.

As Figure 5 shows, output and price converge to the monopoly level no matter how many firms are in the industry.

Figure 5. Initial and final output for different numbers of firms with constant marginal cost

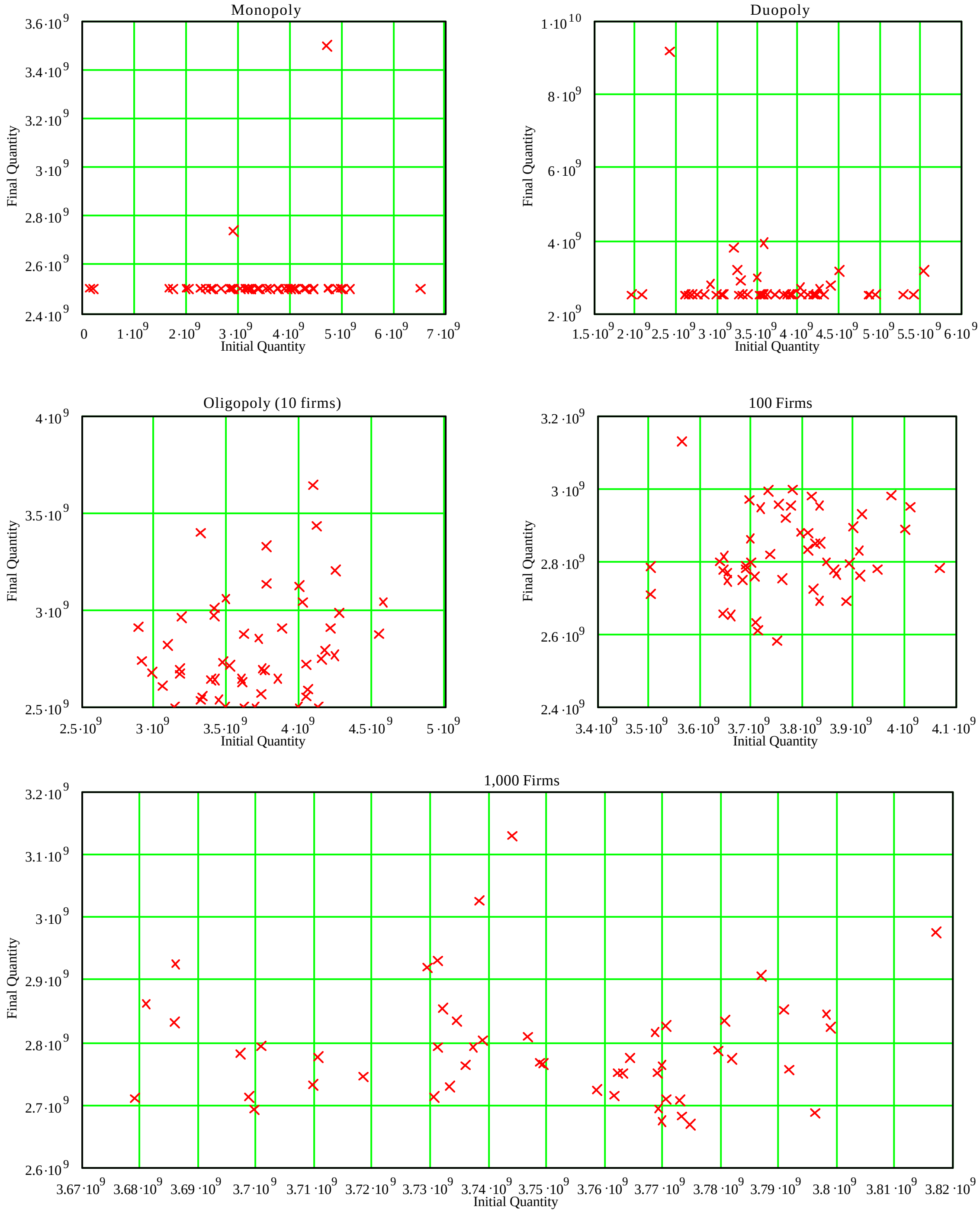


Table 1 shows the mean and standard deviations for final prices and quantities. While the munti-firm output levels were higher than those for the monopoly, there was no tendency for the number of firms to be positively correlated with the market level of output, and in all cases the price and output levels were much closer to those predicted for a monopoly than even those predicted for a duopoly — let alone the perfectly competitive levels. This indicates that, in line with Friedman's billiard players, our non-interactive firms have behaved as if they were following equation (14), rather than the conventionally believed equating of own-output marginal revenue to marginal cost.

Table 1: Mean and standard deviations of price and quantity in Program Two

	Price		Quantity	
	Mean	Stdev	Mean	Stdev
1 firm	75	1.42	2,524,265,631	141,788,991
2 firms	72.4	9.56	2,755,783,129	956,017,286
10 firms	71.9	2.68	2,810,743,977	267,519,760
100 firms	71.8	1.12	2,823,338,331	112,474,569
1000 firms	72.1	0.91	2,793,353,392	91,467,180

Program Three (see Figure 6 for the firm behaviour program fragment), implemented in the agent-based environment EcoLab,[5] creates a market containing a number of firms initialised to the Cournot production level. At each time step, the market provides the price to each firm and asks it what its production is for the next time step. After summing up the production levels for all agents, it calculates a new price from the demand function *P(Q)*. Each firm performs a hill-climbing algorithm, changing production levels in the direction of increasing profit, but with a probability parameter *irrationality* that it will in fact change production in the opposite direction.

With *irrationality=0*, eventually all firms start decreasing production synchronously, with all firms experiencing rising profits until the aggregate market production and price equal the monopoly values. With *irrationality>0*, then occasionally some firms will break ranks with the others and increase production. They thus gain market share (and increase profit) at the expense of the other firms. However, this parameter has little effect on the overall behaviour of the market until it is greater than around 40%. This issue will be explored further in a later paper on the game-theoretic implications of our analysis.

Figure 6. Program Three sample code

```
double firm::production(double price)
{
  double last_profit=prev_production*price - cost(prev_production);
  if (last_profit-prev_profit==0) delta=0;
  delta = (last_profit-prev_profit)>0? delta: -delta;
  if (ratrand && irrationality>0 && irrationality>ratrand->rand())
    delta=-delta;
  prev_production+=delta;
  if (prev_production<0) prev_production=0;
  prev_profit=last_profit;
  return prev_production;
}
```

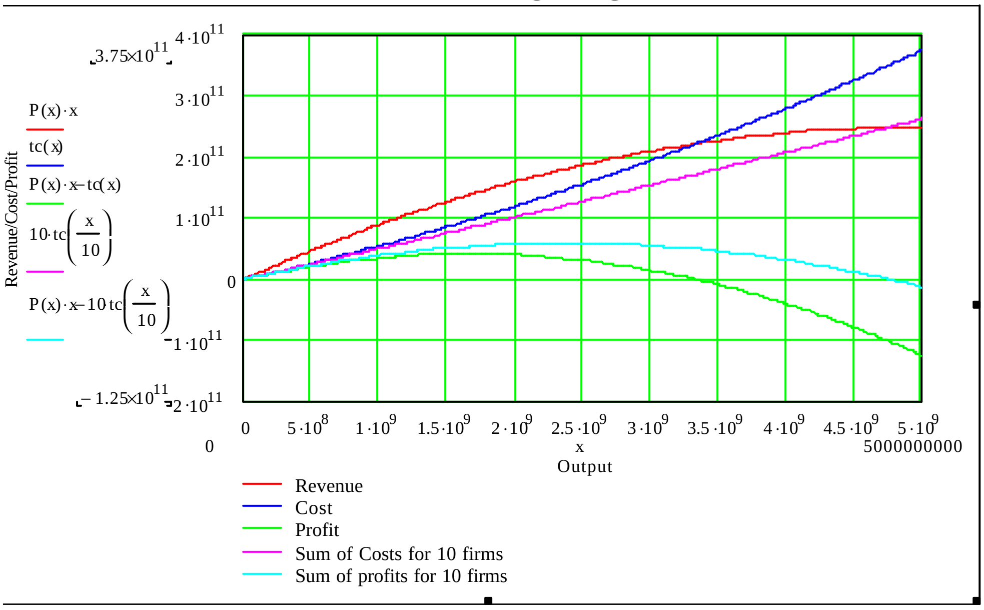
All three programs thus support our conjecture that rational profit maximisers will produce where their own-output marginal revenue exceeds marginal cost as per equation (14), and that market output will converge to the monopoly level regardless of the number of firms.

4 Rising marginal cost and aggregation

These results may be thought to be an artifact of the assumption of constant costs. Simulations with rising marginal costs do indeed give results where competitive industries produce a greater output at a lower cost, but this apparent confirmation of the theory is due to a condition that must be fulfilled for marginal cost curves to be aggregated. The results differ, not because industry output converges to the perfect competition ideal of price equal to marginal cost as the number of firms rises, but because with diminishing marginal productivity, the cost function for a monopoly necessarily differs from the sum of cost curves in an industry with more than one producer.

Figure 7 compares the total cost and profit curves for a single firm industry and a ten firm industry where each firm has the same rising total cost function $TC(q) = 10,000,000 + 50q + 5 \times 10^{-9} q^2$. It is apparent that, given the same functional form for the total cost curve, the sum of the total cost curves for the ten firm industry is *lower* than the total cost curve for the single firm industry. The higher level of output of the multi-firm industry occurs because of *lower* aggregate costs for the many-firm instance than for the monopoly, given the same nonlinear cost function. This returns us to the interpretation of the aggregation condition used earlier to derive equation (12), that $\sum_{i=1}^{n} MC_i(q_i) = n \times MC$.

Figure 7. Total revenue and aggregate total cost functions for different numbers of firms with rising marginal cost



In order to make a definitive comparison between monopoly and a competitive market, the marginal cost curve for the monopoly must be identical to the sum of the marginal cost curves for the competitive market. If cost curves differ, definitive price/output comparisons cannot be made because it is quite possible for a monopoly to produce a larger quantity at a lower price than a competitive industry because of lower costs (Schumpeter 1942: 81-83). It is easily shown that this condition of identical marginal cost functions is only possible where all firms have identical constant marginal costs.

The condition that the marginal cost curve of a single plant[6] and the sum of marginal cost curves for two or more plants must be identical is simultaneously the condition that their marginal products must be identical — since differences in marginal product are the only allowable source of differences in per unit cost. If the marginal product curves are identical, then total product curves can only differ by a constant. If we take labour as our variable factor of production, then output is zero at zero input, and the constant can be set to zero.

We can put this condition into the following form: given the same number of variable inputs, the output of the single plant must be identical to the sum of the outputs of the other plants. If we consider (without loss of generality) a multiple firm industry with n firms, each having one plant employing x workers, then for the aggregate cost curve to be identical to that of the monopoly, the output of these n firms must be equal to the output of the single plant firm employing nx workers.

Using f for the production function of the competitive firms, and g for the production function of the monopoly, this condition is Euler's Equation with a constant of zero:

$$n \cdot f(x) = g(n \cdot x) \quad or \quad f(x) = \frac{g(n \cdot x)}{n} \tag{26}$$

The only production functions that satisfy this condition are ones that yield identical constant marginal product, and therefore identical constant marginal cost. Differentiating both sides of (26) by n yields:

$$f(x) = x \cdot g'(n \cdot x) \tag{27}$$

This gives us a second expression for f. Equating these two definitions yields:

$$\frac{g(n \cdot x)}{n} = x \cdot g'(n \cdot x) \tag{28}$$

Making the substitution of $u = n \cdot x$, this yields an expression involving the log of the differential of g:

$$\frac{g'(u)}{g(u)} = \frac{d}{du}\ln(u) = \frac{1}{u} \tag{29}$$

Integrating both sides yields:

$$\ln(g(u)) = \ln(u) + c \tag{30}$$

Taking exponentials and substituting, it is obvious that g must be a constant returns production function:

$$g(n \cdot x) = C \cdot n \cdot x \tag{31}$$

It follows from (25) that f is the *same* constant returns production function:

$$f(x) = \frac{g(n \cdot x)}{n} = C \cdot x \tag{32}$$

Thus only in the case of constant identical marginal products will the marginal cost curve of a single plant be identical to the sum of the marginal cost curves of two or more plants. If diminishing marginal productivity applies, then the cost functions *must* differ, and no definitive output, price or welfare comparisons can be made between one industry structure and any other.

Furthermore, if economies of scale exist and apply to variable as well as fixed costs (larger plants can afford more specialised machinery allowing less wastage of variable inputs, more joint production, etc.), then more concentrated markets with smaller numbers of plants will have lower average fixed and marginal cost structures than more competitive markets with a larger number of plants. Given our result that the equilibrium output for the industry will be where $MR(Q) = MC(Q)$ regardless of the number of firms, these lower marginal costs will translate into lower prices and higher output levels, with obvious if unquantifiable implications for consumer welfare. In contrast to conventional belief, the neoclassical equilibrium theory of the firm prima facie implies that more concentrated industry structures should be preferred over more competitive ones on welfare grounds.

Of course, whether this theoretical conjecture carries over to the real world depends on whether the standard assumptions of the theory apply in reality.

5 The real world

Numerous researchers (Eiteman 1947 et seq., Haines 1948, Means 1972, Blinder et al. 1998 — see Lee 1998 and Downward & Lee 2001 for surveys) have shown that the vast majority of actual firms have production functions characterised by enormous fixed costs and constant or falling marginal costs. This cost structure, which is described as “natural monopoly” in academic literature and portrayed as an exception to the rule of rising marginal cost, actually appears to be the empirical reality for 95 per cent or more of real firms and products.

In general, these empirical researchers concluded that firms determine their prices by a markup on variable costs, with the size of the markup reflecting partly by the need to cover fixed costs at a levels of output well within the current production capacity of the firm, the degree of competition (so that empirical research gives some grounds by which a more competitive industry can be preferred to a less competitive one), the desire to finance investment and/or repay debt with retained earnings, and the impact of the trade cycle. Price is set by the firm prior to the market, and the firm attempts to sell as much of its output as it can at this price. Firms produce competing but heterogeneous products, and the main form of competition between firms is not price but product differentiation (by both marketing and R&D). Once a minimum sales level has been achieved, each new unit sold adds significantly to profit, and this continues out to the last unit sold — so that marginal revenue is always significantly above marginal cost.

Means coined the term "the administered price thesis" for this perspective on the pricing behaviour of the firm (Means 1972), while the competitive behaviour is consonant with the dynamic theory of competition and creative destruction proposed by Schumpeter (1936).

6 Conclusion

Our results call into question much of supply and demand analysis, to the extent that it is based upon non-interactive profit-maximising behaviour (whether game theory provides an independent foundation for supply and demand analysis is a separate issue). The proposition that the individual firm maximises profit by equating its marginal revenue and marginal cost is true only for a monopoly. For multi-firm industries, setting own-output marginal revenue to be greater than marginal cost is the proper rule for profit maximisation, and the generally believed formula is a special case relevant only when $n = 1$. Firm output is not determined by the intersection point of marginal revenue and marginal cost, but by the locus at which the gap between these two curves equals $\frac{n-1}{n}$ times the gap between price and marginal cost (where n is the number of firms).

At the partial equilibrium level, the well-known result that a supply curve cannot be derived for a monopoly generalises to any industry structure. It can no longer be argued that price is set by the intersection of supply and demand, but at best that output is determined by the intersection of industry marginal revenue and marginal cost, and the market price is set by the demand curve at this quantity.

At the general equilibrium level, the incompatibility of the assumption that price equals marginal cost with profit maximising behaviour means that any model that makes this assumption must now justify it on the basis of interactions between agents which results in them producing *more* than the profit maximising level. Welfare comparisons between different industry structures cannot be made in the context of diminishing marginal productivity, and welfare theorems that rely upon both profit maximising behaviour and the equality of marginal cost and price are also prima facie invalid — as are those that rely upon individual firms equating marginal cost and marginal revenue in a multi-firm industry. The conventional theory of the firm is therefore, to borrow a phrase from Kirman, an "empty citadel". What then should replace it?

Marshall once famously remarked that an economist should "(1) Use mathematics as a shorthand language, rather than as an engine of inquiry. (2) Keep to them till you have done. (3) Translate into English. (4) Then illustrate by examples that are important to real life. (5) Burn the mathematics. (6) If you can't succeed in 4, burn 3. This last I did often" (Marshall, cited in Pigou 1925: 427). Modern economics has not to date followed this advice, and in fact the profession has a history of ignoring empirical research — something akin to Marshall's step 4 — which contradicts its model of the firm (the proposition that this research *could* be ignored was an important aspect of Friedman 1953). We surmise that this failure to recognise empirical data that did not conform to the accepted model was in part due to the belief that the mathematics behind the theory of the firm was incontrovertible (but see also Moss 1984). We hope we have demonstrated that this is not so. In this situation, the only sensible approach is to develop a theory of the firm which conforms to the substantial but neglected literature on the pricing, output and competitive behaviour of actual corporations.

References

Arrow, K.J., (1959). "Towards a theory of price adjustments", in Abramovitz, M. et al., (eds.), *The Allocation of Economic Resources*, Stanford University Press, Stanford.

Aslambeigui, N, & Naples, M.I., (1997). "Scissors or horizon: Neoclassical debates about returns to scale, costs, and long run supply, 1926-1942", *Southern Economic Journal*, vol. 64: 517-530.

Aumann, R.J., (1964). "Markets with a continuum of traders", *Econometrica*, vol. 32: 39-50.

Aumann, R.J., (1966). "Existence of competitive equilibria in markets with a continuum of traders", *Econometrica*, vol. 34: 1-17.

Bishop, R.L., (1948). "Cost discontinuities, declining costs and marginal analysis", *American Economic Review*, vol. 38: 607-617.

Blinder, A.S., Canetti, E., Lebow, D., & Rudd, J., (1998). *Asking About Prices: a New Approach to Understanding Price Stickiness*, Russell Sage Foundation, New York.

Conkling, R.L., (1999). "Marginal cost pricing for utilities: a digest of the California experience", *Contemporary Economic Policy*, vol. 17: 20-32.

Cournot, A. (1838[1988]) "Of the competition of producers", Chapter 7 of *Researches into the Mathematical Principles of the Theory of Wealth*, L. Hachette, Paris, reprinted in Daughety, A.F., (ed.), 1988. *Cournot Oligopoly*, Cambridge University Press, Cambridge.

Downward, P. & Lee, F., (2001). "Post Keynesian pricing theory reconfirmed? A critical review of Asking About Prices", *Journal of Post Keynesian Economics*, vol. 23: 465-483.

Eiteman, W.J., (1947). "Factors determining the location of the least cost point", *American Economic Review*, vol. 37: 910-918.

Eiteman, W.J., (1948). "The least cost point, capacity and marginal analysis: a rejoinder", *American Economic Review*, vol. 38: 899-904.

Eiteman, W.J. & Guthrie, G.E., (1952). "The shape of the average cost curve", *American Economic Review*, vol. 42: 832-838.

Friedman, J., (1983). *Oligopoly Theory*, Cambridge University Press, Cambridge.

Friedman, M., (1953). "The methodology of positive economics", in *Essays in Positive Economics*, University of Chicago Press, Chicago.

Groenewegen, P., (1995). *A soaring eagle: Alfred Marshall 1842-1924*, Edward Elgar, Cheltenham.

Haines, W.W., (1948). "Capacity production and the least cost point", *American Economic Review*, vol. 38: 617-624.

Harrod, R.F., (1934). "Doctrines of Imperfect Competition", *The Quarterly Journal of Economics*, vol. 48: 442-470

Keen, S., (2001). *Debunking Economics*, Pluto Press & Zed Books, Sydney.

Keen, S., (2003). "Standing on the toes of pygmies: why econophysics must be careful of the economic foundations on which it builds", Physica A (forthcoming).

Kuenne, R.E., (1998). *Price and Nonprice Rivalry in Oligopoly*, Macmillan, London.

Langlois, C., (1989). "Markup pricing versus marginalism: a controversy revisited", *Journal of Post Keynesian Economics*, vol. 12: 127-151.

Lee, F., (1996). "Pricing and the business enterprise", in C.W. Whalen (ed.), *Political Economy for the 21st century*, ME Sharpe, Armonk.

Lee, F., (1998). *Post Keynesian Price Theory*, Cambridge University Press, Cambridge.

Lipsey, R. G., & Chrystal, K. A., (1999). *Principles of Economics: 9e*, Oxford University Press, Oxford.

Mankiw, N.G., (2001). *Principles of Microeconomics*, Harcourt, New York.

Marshall, A., (1907 [1961]). *Principles of Economics*, Macmillan, London.

Marshall, A. "Correspondence with Professor A. L. Bowley", in Pigou, A.C., (ed.), (1925). *Memorials of Alfred Marshall*, Macmillan Publishing Company, London.

Martin, S., (2001). *Advanced Industrial Economics*, Blackwell, Cambridge.

Mas-Colell,, A. (1983). "Walrasian Equilibria as Limits of Noncooperative Equilibria. Part I: Mixed Strategies", *Journal of Economic Theory*, vol. 30: 153-70.

Means, G.C, (1972). "The administered-price thesis reconsidered", *American Economic Review*, vol. 62: 292-306.

Moss, S.,(1984). Thehistoryof the theory of the firm from Marshall to Robinson to Chamberlin: the source of positivism in economics", *Economica*, 51: 307-318.

Novshek, W. (1980). "Cournot Equilibrium with Free Entry", *Review of Economic Studies*, vol. 47: 473-86.

Novshek, W. (1985). "Perfectly Competitive Markets as the Limits of Cournot Markets", *Journal of Economic Theory*, vol. 35: 72-82.

Novshek, W. & Sonnenschein, H., (1983). "Walrasian Equilibria as Limits of Noncooperative Equilibria. Part II: Pure Strategies", *Journal of Economic Theory*, vol. 30: 171-87.

Pigou, A.C., (1922). "Empty economic boxes - a reply", *Economic Journal*, vol. 36: 459-465.

Pigou, A.C., (1927). "The law of diminishing and increasing cost", *Economic Journal*, vol. 41: 188-197

Pigou, A.C., (1928). "An analysis of supply", *Economic Journal*, vol. 42: 238-257.

Robertson D.H., (1924). "Those empty boxes", *Economic Journal*, vol. 34: 16-31.

Robertson D.H., (1930). "The trees of the forest", *Economic Journal*, vol. 44: 80-89.

Robbins, L. (1928). "The representative firm", *Economic Journal*, vol. 42: 387-404.

Sato, K., (1979). "A note on capital and output aggregation in a general equilibrium model of production", *Econometrica*, vol. 47: 1559-1568.

Schumpeter, J. A., (1936). *The Theory of Economic Development*, Harvard University Press, Cambridge MA.

Schumpeter, J. A., (1942 [1976]). *Capitalism, Socialism and Democracy*, George Allen and Unwin, London.

Shepherd, W.G., (1984). "'Contestability' vs. competition", *American Economic Review*, vol. 74: 572-587.

Shove, G.F., (1930). "The representative firm and increasing returns", *Economic Journal*, vol. 44: 93-116.

Simon, J.L., Puig, C.M., & Aschoff, J., (1973). "A duopoly simulation and richer theory: an end to Cournot", *Review of Economic Studies*, vol. 40: 353-366.

Sraffa, P. (1926). "The Law of returns under competitive conditions", *Economic Journal*, vol. 40: 538-550.

Sraffa, P. (1930). "The trees of the forest—a criticism", *Economic Journal*, vol. 44: 89-92.

Stigler, G.J. (1957). "Perfect competition, historically considered", *Journal of Political Economy*, vol. 65: 1-17.

[1] The authors would like to thank Mike Honeychurch (University of Melbourne), Trond Andresen (Norwegian Institute of Technology), Joanna Goard (University of Wollongong), Michael Kelly (Illinois Institute of Technology), & Greg Ransom (UCLA Riverside) and members of the HAYEK-L internet discussion list for useful insights and comments on earlier drafts.

[1] The latter condition is feasible if industry output tends to infinity, but this gives rise to obvious logical conundrums for a theory of prices.

[2] This proposition is independent of the game-theoretic/Nash-equilibrium interpretation of Cournot's argument.

[3] Whether game-theoretic interactions might induce a firm to produce at a level greater than that which maximises profits is a separate issue: our primary concern in this paper is the classic "non-interactive" analysis of monopoly versus competitive markets.

[4] These programs are available at http://parallel.hpc.unsw.edu.au/getaegisdist.cgi/getbranch/firmmodel.

[5] Ecolab and supporting documentation is available at http://parallel.hpc.unsw.edu.au/ecolab.

[6] We consider plants here rather than firms since the only interesting comparison is where physical production systems differ rather than merely ownership. If a single firm were able to convert a competitive industry into a monopoly by buying all plants in an industry and centralising their management, then it could easily apply equation (14) and set each plant's marginal revenue above its marginal cost.