Debunking the theory of the firm—a chronology

Steve Keen and Russell Standish [Universities of Western Sydney and New South Wales, Australia]

Copyright: Steve Keen and Russell Standish, 2010

A personal introduction by Steve Keen

I have been an economics renegade for almost 40 years, and for most of that time I have had to tolerate neoclassical economists either dismissing my work or ignoring it. Since the Global Financial Collapse began, that has ceased to be the case. Neoclassical economists, although hardly likely to become born-again Post Keynesians, are chastened and to some degree humbled by their inability to anticipate this crisis. Moreover, it’s hard for them to hide behind the “Black Swans” defence when the speaker is someone credited with having predicted the crisis before it happened.

So that’s all and good; one might even think that the seeds have been laid to finally achieve a true reform of economics, to make it empirical rather than a priori.

However, one also notices a worrying trend among neoclassicals today. True, after having had their macroeconomics lambasted and having been shown how Minsky’s analysis makes sense of what they can’t understand, they are often willing to admit that neoclassical macroeconomics is a shambles. No-one, they will say, defends rational expectations anymore. But then they assert: at least neoclassical microeconomics is sound.

This is patently absurd – especially since neoclassical macro was built largely by subverting its “Keynesian” predecessor on the grounds that it "did not have good microfoundations", and then casting macroeconomics as applied neoclassical microeconomics. However, despite this, neoclassicals still cling to the purity of their vision of the utility maximising consumer on one side of the market, and the profit-maximising firm on the other.

But absurd though it may be, it is an echo of what happened when Keynes tried to bring economics into the real world eighty years ago. Led by Hicks, Samuelson and the like, the neoclassicals dragged the profession back into fantasy via their vision of a beautiful microeconomics.

For society’s sake we need to prevent this intellectually reactionary behavior from being repeated after this crisis. A dispassionate analysis of neoclassical microeconomics shows that it is absurd on its own grounds – that, to coin a phrase "neoclassical microeconomics lacks good microeconomic foundations". But the superficial elegance of the theory remains seductive, and when nations have got beyond this crisis, the same seductive superficiality could give rise to a neoclassical resurgence.

It therefore seems extremely important to emphasize and demonstrate again that their microeconomics is irreparably flawed. This was the task of my book Debunking Economics: to point out that, for example "downward-sloping market demand curves" don’t slope downwards unless there is just one consumer and one commodity (the Sonnenschein-Mantel-Debreu conditions, treated properly as a proof by contradiction that individual demand curves can’t be aggregated), that a marginal product of capital can’t be defined (Sraffa’s critique), and so on.

In most of Debunking Economics, I simply repackaged established critiques by previous authors – I stood on the shoulders of the giants that neoclassical economics ignored. But there was one part of the theory that, when I began the book, simply appeared irrelevant rather than analytically false: the theory of the firm. From my own experience as a young
believer in mainstream economics, I realised how powerful that vision of intersecting supply and demand curves in a competitive, efficient market is, but all I could do, I thought, was parody its irrelevance rather than analytically dismiss it.

Then I spotted a flaw that, to my knowledge at the time, had not been noted before: that the assumptions of a downward sloping market demand curve and a horizontal firm demand curve in the perfectly competitive model were mutually incompatible.

I since have discovered that I wasn't the first to point this out – Stigler, of all people, had done it in 1957. But I added a proof that what neoclassical economics calls profit-maximising behaviour – equating marginal cost and marginal revenue – provably does not maximise profits.

Further analysis found many other flaws in the superficially seductive "logic" of the neoclassical theory of the firm. In many ways, the flaws in this crucial part of neoclassical microeconomics are worse, and more easily proven, than those in consumer theory or capital theory or the like.

Of course, attempts to get this analysis published in mainstream economic journals prior to 2008 predictably failed; my various papers appeared in non-mainstream books and journals, and even the journal of interdisciplinary physics, *Physica A*.

I was content to leave them there and focus on my main interest today, of extending Minsky's work to understand the financial crisis. But then Australia's Treasury drafted a new tax whose foundation was the intellectually flawed vision of a competitive market, with firms facing a horizontal demand curve, and profit-maximising by equating marginal cost and marginal revenue.

So I decided to put together this compendium of all the reasons why this widely believed theory is nonsense. I start with the simplest disproofs that I have developed, and progress right through to a critique of Cournot-Nash theory. It's heavy going, but I fear that unless we really drive a stake through the heart of this vampire logic called neoclassical economics, it will rise again and lead us back to ignorance once we forget the Global Financial Collapse, as our predecessors forgot the Great Depression.

**Stigler 1957**

The proposition that Keen had thought original in *Debunking Economics*—that, under conditions of "atomism", the slope of the demand curve facing the individual competitive firm was the same as the slope of the market demand curve—had in fact been made in 1957 by that arch defender of neoclassicism, George Stigler, and in a leading neoclassical journal: *The Journal of Political Economy* (Stigler 1957—see Figure 1).
Figure 1: Stigler 1957

Stigler’s simple application of the chain rule showed that the underlying assumption of the Marshallian model—atomism, that firms in a competitive industry do not react strategically to the hypothetical actions of other firms—is incompatible with each firm facing a horizontal demand curve. In an $n$-firm industry where the output of the $i^{th}$ firm is $q_i$, this assumption, means that $\frac{\partial q_i}{\partial q_j} = 0 \quad \forall i \neq j$. As a result, $\frac{dQ}{dq_i} = 1$, $\frac{dP}{dq_i} = \frac{dP}{dQ}$, and hence $\frac{dQ}{dq_i} = \frac{dP}{dq_i}$.
\[
\frac{dP}{dq_i} = \frac{dP}{dQ} \frac{dQ}{dq_i} \\
= \frac{dP}{dQ} \left( \sum_{j=1}^{n} \frac{\partial q_j}{\partial q_i} \right) \\
= \frac{dP}{dQ} \left( \frac{\partial q_i}{\partial q_i} + \sum_{j \neq i}^{n} \frac{\partial q_j}{\partial q_i} \right) \\
= \frac{dP}{dQ} \left( 1 + \sum_{j \neq i}^{n} 0 \right) \\
= \frac{dP}{dQ}
\]

(0.1)

It is thus impossible for the market demand function \( P(Q) \) (where \( Q = \sum_{i=1}^{n} q_i \)) to have the dual properties that \( P'(Q) < 0 \) and \( P'(q_i) = 0 \)—and Stigler had shown this in 1957! Yet the claim that the market demand curve is negatively sloped, while the individual perfectly competitive firm faces a horizontal demand curve, has graced the opening chapters of every economics textbook published in the last half century.

Mendacity in education—another personal observation

One of my motivations for writing *Debunking Economics* was my belief that an education in economics was mendacious. I had in mind the failure to note the Cambridge Controversy arguments when discussing the concept of an aggregate production function (see Chapter 6 of Debunking Economics), or the avoidance of the Sonnenschein-Mantel-Debreu conditions when deriving a market demand curve from the aggregation of individual ones (Chapter 2).

When I discussed these issues with any of the minority of neoclassical economists who were themselves aware of those critiques, the even smaller minority who did not dismiss them outright would raise the pedagogic defense of difficulty. These topics are complex, and require an advanced knowledge, not only of economics, but of mathematics. Better to give new students a simple introduction—well behaved aggregate production functions, nice downward sloping market demand curves, and so on—and cover the nuances when they have more knowledge.

No such defense applies here: the only mathematical knowledge needed to comprehend that Marshallian atomism is incompatible with a horizontal demand curve for the firm is elementary calculus.

The responses I have received on this point from neoclassical economists to date have been disingenuous. At best, they have referred to Stigler’s attempt to recast perfect competition as the limiting case as the number of firms in an industry increases (discussed in
the next section). At worst, they have claimed that the laws of mathematics do not apply to economics.

The latter claim is of course nonsense for an approach to economics which, from its founding father to today’s leading exponents, exalted itself over its rivals because it was mathematical:

those economists who do not know any mathematics … can never prevent the theory of the determination of prices under free competition from becoming a mathematical theory. Hence, they will always have to face the alternative either of steering clear of this discipline … or of tackling the problems of pure economics without the necessary equipment, thus producing not only very bad pure economics but also very bad mathematics. (Walras 1900 [1954]: 47)

This raises the question of why neoclassical economists defend commencing an education in economics with such bad mathematics? We expect it is because the fantasy of perfect competition is essential to fulfilling the vision of rational self-interested behavior being compatible with welfare maximization. If one admits that the individual firm faces a downward-sloping demand curve, then the elimination of deadweight loss that is the hallmark of perfect competition can’t possibly be compatible with individual profit-maximization.

This is easily illustrated using another standard mathematical technique, the Taylor series expansion.3

Perfect competitors aren’t profit maximizers

Consider a competitive industry where all firms are producing at the “perfect competition” level where price equals marginal cost. In general, profit for the \( i \)th firm is:

\[
\pi_i(q_i) = P(Q) \cdot q_i - TC(q_i)
\]

What happens to the \( i \)th firm’s profits if it changes its output by a small amount \( \delta q_i \)?

Under the Marshallian condition of atomism, industry output also changes by the same amount. The change in profit \( \delta \pi(\delta q_i) \) is thus

\[
\pi_i(q_i + \delta q_i) - \pi_i(q_i) = \left( P(Q + \delta q_i) \cdot (q_i + \delta q_i) - TC(q_i + \delta q_i) \right) - \left( P(Q) \cdot q_i - TC(q_i) \right)
\]

This can be approximated by applying the first order Taylor series expansion, and by making the substitution that, at this output level, price equals marginal cost: \( P(Q) = TC'(q_i) \). The symbolic mathematics engine in Mathcad makes fast work of this approximation:4

1 A referee for the Economic Journal wrote that “we always consider the perfect competition case as a polar case which represents an extreme scenario, and is largely a benchmark. I would prefer to see the equation: \((AR - MR)/AR = 1/(nE)\), so that for \( E \) at a normal value of say 2, and \( n \) at 1000, then the divergence of \( AR \) and \( MR \) is 1/20t of 1%. Then price equals MR seems a pretty good approximation!”

2 A referee for the Journal of Economic Education commented that “Stigler’s many attempts to save neoclassical theory have always caused more problems than they have solved. His version of the chain rule is contrary to the partial equilibrium method and thus is irrelevant”.

3 This proof was first developed by John Legge, of La Trobe University.
Therefore \( \delta \pi (\delta q_i) \approx q_i \cdot \delta q_i \cdot \frac{dP}{dQ} \). Since \( \frac{dP}{dQ} P < 0 \), if \( \delta q_i < 0 \) —if, in words, the firm reduces its output—it \s in profit will rise. Thus the output level at which price equals marginal cost is not a profit maximum for the individual competitive firm, and if such a firm is indeed a profit maximizer, it will reduce its output below this level.

Some neoclassical economists have thrown the “perfect knowledge” assumption at us at this point: perfectly informed consumers will instantly stop buying from the firm that has reduced its output and increased its price, and switch to those that are still setting price equal to marginal cost. But this argument is still based on the “horizontal demand curve” assumption, which itself is a furphy,\(^5\) and the market price in the model has already risen because of the change in output by one firm—there is no “cheaper supplier” to whom omniscient consumers can turn.

“Price equals marginal cost” is, therefore, not an equilibrium under the Marshallian assumption of atomism. As a result, the coincidence of collective welfare and the pursuit of individual profit is impossible: if neoclassical economists want to pull that particular rabbit out of a hat, they need another hat. Stigler attempted to provide one.

\*\* Stigler’s limiting case \*\*

Stigler, of course, was not trying to bury perfect competition when he showed that
\[
\frac{dP}{dq_i} = \frac{dP}{dQ} : \text{he was one of the pre-eminent defenders of the neoclassical model against empirically-oriented researchers like Eiteman and Means (see Freedman 1998). He therefore devised an alternative explanation of perfect competition, as the limiting case of competition as the number of firms in an industry increased. His analysis, shown in Figure 1, footnoted the derivation of the expression shown in Figure 3:}
\]

\[
\text{Marginal revenue = Price + Number of sellers \times Market elasticity}
\]

\*\* Figure 3: Stigler’s expression for marginal revenue (Stigler 1957: 8) \*\*

---

\(^4\) We are using a symbolic mathematics program both to reduce the need for some tedious manual calculations, and because on several occasions, neoclassical economists have disputed the results of manual calculations—by in effect distorting the definition of a derivative.

\(^5\) “furphy” is a delightful Australian word meaning “an irrelevant or minor issue raised to specifically divert attention away from the real issue”. It deserves wider currency—especially amongst economists!
Stigler then asserted that “this last term goes to zero as the number of sellers increases indefinitely” (Stigler 1957: 8). Marginal revenue for the $i$th firm thus converges to market price. Perfect competition thus appeared to be saved, despite a downward-sloping firm’s demand curve: profit-maximizers would set marginal cost equal to marginal revenue, and this would converge to price as more firms entered a market.

Stigler’s convergence argument is technically correct, but in conflict with the proof shown above that “price equals marginal cost” is not a profit maximum for the individual firm. The resolution of this conflict led to Keen’s first truly original contribution to this literature: the proposition that equating marginal revenue and marginal cost maximizes profit is also a furphy.

Equating MC and MR doesn’t maximize profits

Generations of economists have been taught the simple mantra that “profit is maximized by equating marginal cost and marginal revenue”. The proof simply differentiates (0.2) with respect to $q_i$. However, the individual firm’s profit is a function, not only of its own output, but of that of all other firms in the industry. This is true regardless of whether the firm reacts strategically to what other firms do, and regardless of whether it can control what other firms do. The objectively true profit maximum is therefore given by the zero of the total differential: the differential of the firm’s profit with respect to total industry output.

We stress that this issue is independent of whether the individual firm can or cannot work out this maximum for itself, whether the firm does or does not interact with its competitors, and whether the firm does or does not control the variables that determine the profit maximum. Given a mathematically specified market inverse demand function that is a function of the aggregate quantity supplied to the market, and a mathematically specified total cost function for the individual firm that is a function of its output, the question “what is the level of the firm’s output that maximizes its profit?” is completely independent of the question of “will the firm, in any given environment, or following any given behavioral rule, actually determine or achieve this level?”. That objective, profit-maximizing level is given by the zero of the total differential of profit:

$$\frac{d}{dQ} \pi(q_i) = \frac{d}{dQ} \left( P(Q)q_i - TC(q_i) \right) = 0 \quad (0.4)$$

This total derivative is the sum of $n$ partial derivatives in an $n$-firm industry:

$$\frac{d}{dQ} \pi(q_i) = \sum_{j=1}^{n} \left\{ \left( \frac{\partial}{\partial q_i} \pi(q_i) \right) \cdot \frac{d}{dQ} q_j \right\} \quad (0.5)$$

In the Marshallian case, atomism lets us set $\frac{d}{dQ} q_j = 1 \forall j$ (we address the Cournot case in section 0). Expanding (0.5) yields

$$\frac{d}{dQ} \pi(q_i) = \sum_{j=1}^{n} \left( \frac{\partial}{\partial q_i} \left( P(Q)q_i - TC(q_i) \right) \right) \quad (0.6)$$

Continuing with the product rule, (0.6) can be expanded to:
\[
\frac{d}{dQ} \pi(q_i) = \sum_{j=1}^{2} \left( P(Q) \frac{\partial}{\partial q_j} q_j + q_i \frac{\partial}{\partial q_j} P(Q) - \frac{\partial}{\partial q_j} TC(q_j) \right)
\]  

(0.7)

Under the Marshallian assumption of atomism, the first term in the summation in (0.7) is zero where \( j \neq i \), and \( P(Q) \) where \( j = i \). The second term is equal to \( q_i \frac{d}{dQ} P(Q) \forall j \); the third is zero where \( j \neq i \), and equal to \( \frac{d}{dq_i} TC(q_i) \) (or marginal cost \( MC(q_i) \)) where \( j = i \). Equation (0.7) thus reduces to

\[
\frac{d}{dQ} \pi(q_i) = P(Q) + n \cdot q_i \frac{d}{dQ} P(Q) - MC(q_i)
\]  

(0.8)

The true profit maximum—under the Marshallian condition of atomism—is thus given by equation (0.9):

\[
\pi(q_i)_{\text{max}} : MC(q_i) = P + n \cdot q_i \frac{dP}{dQ}
\]  

(0.9)

The error in the standard “Marshallian” formula is now obvious: it omits the number of firms in the industry from the expression for the individual firm’s marginal revenue. With this error corrected, the correct profit-maximizing rule for a competitive firm is very similar to that for a monopoly: set marginal cost equal to *industry level marginal revenue*.6

---

**Monopoly, competition, profit and hyper-rationality**

Neoclassical economics assumes that, given revenue and cost functions, there is some output level that will maximize profits, and another that will maximize social welfare (by eliminating deadweight loss).7 The argument that the two coincide under perfect competition has been shown to be nonsense. So too is the argument that a single rational firm could work out the profit maximum, but a bunch of rational *non-interacting* firms couldn’t, as the calculus in the previous section shows.

Of course, an objection can be made to the above mathematical logic that solving equation (0.9) requires knowledge of the number of firms in the industry, which the individual competitive firm can’t be assumed to have.8 Here, we can turn Milton Friedman’s methodological defense of the theory of the firm against itself. Friedman, as is well known, argued that while firms didn’t in fact do calculus to work out their profit-maximizing output

---

6 Though not necessarily identical, since \( n \cdot q_i \) only equals \( Q \) if \( q_i = \frac{Q}{n} \forall i \). This impact of dispersal in firm size may explain some of the simulation results shown later.

7 We use standard undergraduate terms here because the analysis we are challenging is, up to this point, that served up to undergraduates. We address game theoretic concepts later.

8 Equation (1.9) can be put in another form which partly addresses this criticism, and also emphasizes the error in the conventional formula. The profit-maximizing level of output is not to equate firm-level marginal revenue and marginal cost, but to make the gap between them a fraction of the gap between price and marginal cost: \( MR(q_i) - MC(q_i) = \frac{n-1}{n} (P - MC) \). The fraction tends to 1 as \( n \to \infty \), so the more “competitive” an industry is, the easier it is to apply this formula.
levels, we could model their behavior “as if” they did, because unless the behavior of businessmen in some way or other approximated behavior consistent with the maximization of returns, it seems unlikely that they would remain in business for long. (Friedman 1953: 22)

We are not arguing that firms do the calculus to work out this profit-maximizing level either. Instead, we are simply showing that the calculus can be done, and the profit-maximizing level is not the one asserted by neoclassical economists. However, it is possible now—in a way that wasn’t possible in 1953—to actually carry out Friedman’s “billiard players” experiment. Citing him again:

Now, of course, businessmen do not actually and literally solve the system of simultaneous equations in terms of which the mathematical economist finds it convenient to express this hypothesis, any more than leaves or billiard players explicitly go through complicated mathematical calculations or falling bodies decide to create a vacuum. The billiard player, if asked how he decides where to hit the ball, may say that he “just figures it out” but then also rubs a rabbit’s foot just to make sure; and the businessman may well say that he prices at average cost, with of course some minor deviations when the market makes it necessary. The one statement is about as helpful as the other, and neither is a relevant test of the associated hypothesis. (Friedman 1953: 22)

A “relevant test of the associated hypothesis” is to set up a virtual market that conforms to neoclassical assumptions—with a static downward sloping market demand curve, and given cost functions subject to diminishing marginal productivity, so that there is indeed a profit-maximizing level of output for each firm—and see what happens. Figure 4 shows a Mathcad program that implements this.10

---

9 In fact, we argue later that the assumption that there is some profit-maximizing level for a firm is a furphy. The profit-maximizing strategy for actual firms is simply sell as much as possible, at the expense where possible.

10 The behavior modeled was deliberately made as simple as possible, to avoid the rejoinder that the results were the product of our algorithm rather than raw profit-motivated behavior. It could only have been simpler by having each firm vary its output by one unit at each time step—a modification which, as it happens, results in a much slower but absolute convergence to the Keen equilibrium.
Figure 4: Simulation of instrumental profit maximizers

Working through the program line by line:

1. A random number generator is seeded
2. A for loop iterates from a minimum number to a maximum number of firms
3. If there is more than one firm in the industry, each firm is randomly allocated an initial output level. The amounts are uniformly distributed from a minimum of the Keen prediction for a profit-maximizing firm, \( q_K \), to a maximum of the neoclassical prediction \( q_C \).
4. If there is only one firm in the industry, its output starts at the level predicted by the neoclassical model—which coincides with \( q_K \).
5. An initial market price is set, based on the sum of initial outputs.
6. Line 6 sets the market price in the case of a monopoly.
7. Each firm is randomly allocated an amount by which it varies output. The distribution has a mean of zero and a standard deviation of 1% of the neoclassical prediction for a profit-maximizing firm’s output (this is the last aspect of the program that involves probability).
8. Line 8 allocates a change amount of 1% of the predicted output for a monopoly.
9. A for loop iterates over a number of runs where each firm varies its output trying to increase its profit from the initial level.

10. Firstly each firm adds its change amount to its initial output. This is a vector operation: if there are 100 firms in the industry $Q_0$ is a vector with 100 initial output amounts, and $dq$ is a vector with 100 (positive or negative) output changes.

11. A new market price is calculated on the basis of the new aggregate output level.

12. Line 12 again allows for a monopoly.

13. Each firm then calculates whether its profit has risen or fallen as a result of its change in output, and the collective impact of all the changes in output on the market price. If a firm finds that its profit has risen, it continues to change output in the same direction; if its profit has fall, it changes its output by the same amount but in the opposite direction.

14. Each step in the iteration is stored in a multi-dimensional array.$^{11}$

15. The multidimensional array is returned by the program.

The program was run with identical cost functions for each firm, set up so that the market aggregate marginal cost curve was independent of the number of firms in the industry (we return to this issue in the Appendix). The number of firms was varied from 1 to 100. The eventual aggregate output at the end of 1000 iterations is shown in Figure 5, and the corresponding market price is shown in Figure 6, against the predictions of the Neoclassical and the Keen approach respectively.

---

$^{11}$ In effect, $F$ is a matrix where the $j^{th}$ and $i^{th}$ column contains the vector of outputs by an $i$ firms industry at the $j^{th}$ iteration.
As is obvious, the number of firms in an industry had no impact on the eventual market output level or price: the Neoclassical prediction that price would converge to the level at which price equals marginal cost clearly was not fulfilled.

Some neoclassical referees thought that the results occurred because, though all firms were acting independently, they were all doing the same thing (reducing output from the initial level), and thus acting in a semi-collusive way. In fact, as Figure 7 and Figure 8 show, though the average outcome conformed to Keen’s predictions, the individual firms all pursued very different strategies. The aggregate outcome, which contradicted the neoclassical prediction and confirmed Keen’s, was the result of quite diverse individual firm behavior—despite all firms having identical cost functions.

Figure 7 shows the output levels of 3 randomly chosen firms from the 100 firm industry, the average for all firms, and the predictions of the Keen and neoclassical formulae. Firm 1 began near the neoclassical output level, rapidly reduced output towards the “Keen” level, but then reversed direction; Firm 2 began halfway between the neoclassical and Keen predictions, then reduced output below the Keen level and stayed there; Firm 3 began closer to the neoclassical level and meandered closer to the Keen level.

---

12 A referee for the Economic Journal commented that “if firms act the same way, they will all get higher profits if and only if they reduce outputs. Then the algorithm will continue to lead them to the monopoly outcome since there is no chance any firm can realize the true impact of its own output change. Thus the result is not surprising.”
The sole source of the volatility of each firm’s behavior is the complex impact of interactions between firms, in the context of a very simply defined market—there is no random number generator causing this volatility. As Figure 8 shows, each firm made its changes in response to the impact of both its changes in output, and the collective changes in output, on its profit. Some firms made larger profits than others—notably the firms with the larger output made the larger profits. However, the average profit was much higher than predicted by the neoclassical model.
This model indicates that, in this game of competitive profit maximization, the virtual equivalents of Friedman’s “billiard players” follow the laws of mathematics in their search for profits, as Friedman argued. However, these laws of mathematics are incompatible with the beliefs of neoclassical economists.

Since hyper-rational profit-maximizers cannot be relied upon to save neoclassical belief, there are only two avenues left: irrational behavior, and Cournot-Nash game theory.

**Price-taking behavior is irrational**

A regular neoclassical rejoinder to our analysis has been that we are “cheating” by not assuming rational, price-taking behavior. Our standard reply that the assumption of “price-taking” behavior is itself cheating, with regard to the laws of mathematics: as shown in Section 0, the assumption that $P'(q_i) = 0$ is incompatible with the assumption of a downward-sloping market demand curve ($P'(Q) < 0$). However, it is also easily shown that “price-taking behavior” is irrational.

The assumption of price-taking behavior appears regularly in neoclassical economics, from the level of Marshallian analysis through to the foundations of general equilibrium analysis (see for example Mas-Colell et al 1995: 314, 383). Neoclassical economists do not seem to realize that this is a classic “rabbit in the hat” assumption: if it is assumed, then the “perfectly competitive” result of price equaling marginal cost follows, regardless of how many firms there are in the industry.
The essence of price-taking is the belief that a firm’s change in its output doesn’t affect market price: this amounts to setting $\frac{\partial P(Q)}{\partial q_i} = 0$ in equation (0.7). This results in the “profit-maximizing strategy” of setting price equal to marginal cost, independently of the number of firms—that is, once this assumption is made, even a monopoly produces where price equals marginal cost. This behavior is clearly irrational for a monopoly, and it is only the “fog of large numbers”—the confusion of infinitesimals with zero, as Keen noted in *Debunking Economics*—that led neoclassical economists to regard price-taking as rational behavior for competitive firms.

Figure 9 illustrates that price-taking behavior is *irrational*: an agent who behaves this way is necessarily making a logical error. If the market demand curve slopes downwards, then the *a priori* rational belief is that any increase in output by the firm will depress market price.

Figure 9: Irrationality of "price-taking" behavior

The desired neoclassical of price equal to marginal cost is thus dependent on irrational behavior (in the context of Marshallian competition—we address Cournot competition later). We quantify the degree of irrationality needed by modifying the program shown in Figure 4, so that a proportion of firms actually do behave irrationally: if a strategy has caused an increase in profit, a fraction of firms respond by reversing that strategy.

The modified program is shown in Figure 10. The outer loop (line 2) now iterates the counter \( i \) from 0 to 50, with the value representing the fraction of firms who behave irrationally...
at each iteration. The only change to the inner loop is that the change in output by each firm is now reversed for \( i \%) of firms at each iteration.\(^\text{13}\)

\[
\text{Firms:=} \quad \text{Seed(rand)} \\
\text{for } i \in 0..50 \\
\quad Q_0 \leftarrow \text{round}\left(\text{runif}\left(\text{firms},0,\text{q}_C(\text{firms})\right)\right) \\
\quad p_0 \leftarrow P\left(\sum Q_0\cdot a, b\right) \\
\quad dq \leftarrow \text{round}\left(\text{rnorm}\left(\text{firms},0,\text{q}_C(\text{firms})\right)\right) \\
\text{for } j \in 0..\text{runs} - 1 \\
\quad Q_{j+1} \leftarrow Q_j + dq \\
\quad p_{j+1} \leftarrow P\left(\sum Q_{j+1}\cdot a, b\right) \\
\quad dq \leftarrow \left[\text{sign}\left(\text{runif}\left(\text{firms}, -i/100, -i/100 + 1\right)\right)\right] \left[\left(p_{j+1}Q_{j+1} - p_jQ_j\right) - \left(tc(Q_{j+1}\text{ firms}) - tc(Q_j\text{ firms})\right)\right] \\
\quad F_{j+1} \leftarrow Q_{j+1} \\
\]

\text{Figure 10: Analyzing the impact of irrationality}

Figure 11 shows the aggregate outcome for a 100 firm industry. With no irrationality, the industry produces the amount predicted by the Keen formula. Output then increases almost monotonically as the degree of irrationality rises—until, when 20 per cent of firms are behaving irrationally at each iteration, market output converges to near the neoclassical output level.

For a degree of irrationality between 20% and 45%, the neoclassical outcome continues to dominate the simulation results. Then as irrationality rises above this level, the market effectively follows a random walk—where, curiously, profits in general tend to be \textit{higher} than what would apply if each firm equated marginal revenue and marginal cost.

\(^{13}\) The function call \( \text{runif}(\text{firms}, -i/100, -i/100 + 1) \) generates a vector of numbers between \(-i/100\) and \(1 - i/100\); when \(i=0\), all these numbers will be positive and thus not affect the value of the \textit{sign()} function; when \(i>0\), \(i\%) of these numbers will be negative and thus the sign of the \textit{sign()} function will be reversed. The firms that have this randomly assigned negative number against their output change will increase output at the next step if profit rose when the decreased output on the previous step (and vice versa). This is instrumentally irrational behavior.
**Market outcome and model predictions**

Figure 11: Market output as a function of the degree of irrationality

Figure 12 shows the behavior of three randomly chosen firms, and the average behavior, at a 20% level of irrationality—i.e., when one firm in five reverses any strategy that benefited it on the previous iteration.

Figure 12: Sample outputs at 20% irrationality
Figure 13 shows the impact that a degree of irrationality of 20% has on firms’ profits. Profit falls throughout the run, until by the end, it is almost (but not quite) as low as that caused by equating marginal revenue and marginal cost.

Ironically, higher profits apply if firms simply follow a random walk than if they apply the neoclassical formula (see Figure 14).
A degree of irrational behavior thus saves the neoclassical preferred outcome of price equal to marginal cost—though with some collateral damage, since it is now clearly neither profit-maximizing, nor rational. The question remains, what might help ensure this level of irrationality? Cournot-Nash game theory appears to provide an answer in strategic interactions between firms—though this answer is only unequivocal at a very superficial level of analysis.

**Strategic interaction and competition**

Unlike the strictly false Marshallian model of competition, Cournot-Nash game theory provides a prima facie sound basis for “perfect competition” as the outcome of strategic interactions between competitors. In Cournot-Nash game theoretic analysis, firms decide their own actions on the basis of the expected reactions of other firms, in such a way that each firm's best response is set \( MR(q_i) = MC(q_i) \). This is profit-maximizing for the firm, in the context of the expected response of competitors to its actions, though it results in a lower level of profit than if firms “collude” to share the monopoly level of output between them.

Up to this point, our contribution has been to show that what neoclassicals call “collusive behavior” can actually result from firms not reacting strategically to what other firms do—in the notation of the early part of this paper, when firms set \( \frac{\partial q_i}{\partial q_j} = 0 \ \forall i \neq j \).

This paradox—that what neoclassical theory labels “collusion” actually occurs when firms do not react to each other—inspired us to attempt to integrate (corrected) Marshallian and Cournot-Nash theory, by making the level of strategic interaction between firms a variable. Defining the response of the \( i \)th firm to an output change by the \( j \)th firm as \( \theta_{ij} = \frac{\partial q_i}{\partial q_j} \),
we then had to rework the expression for profit into one that depended entirely upon the level of strategic interaction. The result of this research was a second original contribution, a generalized formula for profits in terms of the level of strategic interaction—and the discovery that the optimal level of interaction, in the context of identical firms, was zero. The derivations involved are quite complex, and they are reproduced below in their entirety.

We start from the same position as equation (0.4). For profit-maximization, we require

$$\frac{d}{dQ} \pi(q_i)$$

the zero of which we then expand this as per equation (0.5), but rather than then setting

$$\frac{d}{dQ} q_j = 1 \forall j$$

we work out what

$$\frac{d}{dQ} q_i$$

is in terms of the strategic reaction coefficient \(\theta_{i,j}\):

$$\frac{d}{dQ} q_i = \sum_{j=1}^{n} \frac{\partial}{\partial q_j} q_i$$

$$= \sum_{j=1}^{n} \theta_{i,j}$$

(0.10)

As a result, our next equation differs from equation (0.6):

$$\frac{d}{dQ} \pi(q_i) = \sum_{j=1}^{n} \left( \frac{\partial}{\partial q_j} \left( P(Q) \cdot q_i - TC(q_i) \right) \cdot \frac{d}{dQ} q_j \right)$$

$$= \sum_{j=1}^{n} \left( \frac{\partial}{\partial q_j} \left( P(Q) \cdot q_i \right) \cdot \frac{d}{dQ} q_j \right) - \sum_{j=1}^{n} \left( \frac{\partial}{\partial q_j} TC(q_i) \cdot \frac{d}{dQ} q_j \right)$$

(0.11)

Working firstly with the total cost component,

$$\frac{\partial}{\partial q_j} TC(q_i) = 0 \forall i \neq j$$

and

$$\frac{\partial}{\partial q_j} TC(q_i) = MC(q_i) \forall i = j$$

Thus the cost component of the profit formula reduces to:

$$\sum_{j=1}^{n} \left( \frac{\partial}{\partial q_j} TC(q_i) \cdot \frac{d}{dQ} q_j \right) = MC(q_i) \cdot \frac{d}{dQ} q_j$$

$$= MC(q_i) \cdot \sum_{j=1}^{n} \theta_{i,j}$$

(0.12)

The revenue component involves some more intricate steps:

$$\sum_{j=1}^{n} \left( \frac{\partial}{\partial q_j} \left( P(Q) \cdot q_i \right) \cdot \frac{d}{dQ} q_j \right) = \sum_{j=1}^{n} \left( P(Q) \cdot \frac{\partial}{\partial q_j} (q_i) \cdot \frac{d}{dQ} q_j \right) + \sum_{j=1}^{n} \left( q_i \cdot \frac{\partial}{\partial q_j} (P(Q)) \cdot \frac{d}{dQ} q_j \right)$$

(0.13)

14 This work was first published in Keen & Standish 2006.
\[
\frac{\partial}{\partial q_j} (P(Q)) \quad \text{ reduces to } \quad \frac{d}{dQ} (P(Q))
\]
as before, though the logic is slightly more complicated:

\[
\frac{\partial}{\partial q_j} (P(Q)) = \frac{\partial}{\partial Q} (P(Q)) \cdot \frac{\partial Q}{\partial q_j} \\
= \frac{\partial}{\partial Q} (P(Q)) \cdot 1 \\
= \frac{d}{dQ} P(Q)
\]

(0.14)

Making this substitution into (0.13), and using \( P \) rather than \( P(Q) \) for the sake of clarity, yields:

\[
\sum_{j=1}^{n} \left( P \cdot \frac{\partial}{\partial q_j} (q_j) \cdot \frac{d}{dQ} q_j \right) + \sum_{j=1}^{n} \left( q_j \cdot \frac{dP}{dQ} \cdot \frac{d}{dQ} q_j \right) \\
= P \cdot \sum_{j=1}^{n} \left( \theta_{i,j} \cdot \frac{d}{dQ} q_j \right) + q_i \cdot \frac{dP}{dQ} \cdot \sum_{j=1}^{n} \left( \frac{d}{dQ} q_j \right)
\]

(0.15)

Care has to be taken with expanding the expression \( \frac{d}{dQ} q_j \) in (0.15), since

\[
\frac{d}{dQ} q_j = \sum_{k=1}^{n} \theta_{j,k}
\]
but the \( i \) suffix here is just a placeholder for iterating over the \( n \) firms in the industry. We therefore make the substitution of \( k \) for \( i \) in this subscript so that we define

\[
\frac{d}{dQ} q_j = \sum_{k=1}^{n} \frac{\partial}{\partial q_k} q_j = \sum_{k=1}^{n} \theta_{j,k}
\]

\[
P \cdot \sum_{j=1}^{n} \left( \theta_{i,j} \cdot \frac{d}{dQ} q_j \right) + q_i \cdot \frac{dP}{dQ} \cdot \sum_{j=1}^{n} \left( \frac{d}{dQ} q_j \right)
\]

(0.16)

Equation (0.11) finally reduces to:

\[
\frac{d}{dQ} \pi(q_i) = P \cdot \sum_{j=1}^{n} \left( \sum_{k=1}^{n} \theta_{i,j} \cdot \theta_{j,k} \right) + q_i \cdot \frac{dP}{dQ} \cdot \sum_{j=1}^{n} \left( \sum_{k=1}^{n} \theta_{j,k} \right) - MC(q_i) \cdot \sum_{j=1}^{n} \theta_{i,j}
\]

(0.17)

The zero of this equation determines the profit maximum for any given level of strategic interaction between firms. We can now rephrase the corrected Marshallian and the Cournot-Nash profit maxima in terms of their “conjunctural variation” levels.
The Marshallian substitution is rather easy. Given \( \frac{\partial q_i}{\partial q_j} = 0 \forall i \neq j \) and

\[
\sum_{j=1}^{n} \theta_{i,j} = 1 \quad \sum_{k=1}^{n} \sum_{j=1}^{n} \theta_{j,k} = n
\]
is the trace of an identity matrix so that

\[
\sum_{j=1}^{n} \left( \sum_{k=1}^{n} \theta_{j,k} \right) = n \quad \text{and} \quad \theta_{i,j} \cdot \theta_{j,k} = 1 \forall i = j = k \quad \text{and zero otherwise},
\]

therefore in the case of atomism, the maximum of (0.17) reduces to

\[
\frac{d}{dQ} \pi(q_i) = P + q_i \cdot \frac{dP}{dQ} \cdot n - MC(q_i) = 0
\] (0.18)

This reproduces the formula derived in equation (0.9).

For the Cournot case, we start from the general situation where \( \theta_{i,j} = \theta \forall i \neq j \) and

\[
\sum_{j=1}^{n} \theta_{i,j} = 1 \quad \sum_{j=1}^{n} \sum_{k=1}^{n} \theta_{j,k} = n
\]

Then

\[
\sum_{j=1}^{n} \left( \sum_{k=1}^{n} \theta_{j,k} \right) = n + (n^2 - n) \cdot \theta = n \left( 1 + (n - 1) \cdot \theta \right)
\]
is the sum of each column of the matrix—which is \( (n-1) \cdot \theta + 1 \) multiplied by each element of one of its columns, so that we have \( (n-1) \cdot \theta + 1 \) copies of \( (n-1) \cdot \theta + 1 \). Thus

\[
\sum_{j=1}^{n} \left( \sum_{k=1}^{n} \theta_{i,j} \cdot \theta_{j,k} \right) = \left( (n-1) \cdot \theta + 1 \right)^2
\]

Making these preliminary substitutions and factoring the common element \( (1 + (n-1) \cdot \theta)^2 \),

we derive:

\[
\frac{d}{dQ} \pi(q_i) = \left( 1 + (n-1) \cdot \theta \right) \cdot \left( P + q_i \cdot \frac{dP}{dQ} \cdot n - MC(q_i) \right)
\] (0.19)

Given that the Cournot-Nash "best response" results in each firm setting

conventionally defined marginal revenue \( \frac{dP}{dQ} \) equal to marginal cost, we can now

---

15 The alleged neoclassical equilibrium occurs where \( P = MC(q_i) \); for long-run equilibrium, only the most efficient scale of output applies so that marginal cost is identical for all firms, therefore all firms must produce at the same level of output \( q_i = q = Q/n \) . For this to be stable, all firms must have the same level of strategic interaction with each other, \( \theta = \theta \).

16 Since \( \theta \) lies in the range \( [0, 1/n \cdot E] \), \( 1 + (n-1) \cdot \theta \neq 0 \); it can therefore be factored out.
work out the corresponding value for $\theta$. This is $\theta = \frac{1}{n \cdot E}$, where $n$ is the number of firms in the industry and $E$ the market elasticity of demand ($E = \frac{P}{Q} \frac{dQ}{dP}$).

It is now also possible to work out the optimum value for $\theta$, from the view of a profit-maximizing individual firm: what level of strategic response should a firm have to its rivals, given that its objective is to maximize its profit?

In this generalized case of identical firms, the answer is obvious: the optimal value of $\theta$ is zero. As shown by equation (0.18), the profit maximum is where $\frac{d}{dQ} \pi(q_i) = P + q_i \cdot \frac{dP}{dQ} \cdot n - MC(q_i) = 0$. Given equation (0.19), this is only possible for $\theta = 0$. Cournot-Nash game theory is thus “A curious game. The only winning strategy is not to play”\textsuperscript{17}. It is therefore, on closer examination, a very poor defense of the concept of perfect competition.\textsuperscript{18}

This interpretation is given additional weight by the observation that, though the standard “Prisoners’ Dilemma” presentation implies that the Cournot strategy is stable and the Keen strategy is unstable (both in a Nash equilibrium sense), the Cournot strategy is locally unstable, while the Keen strategy is locally stable.

**Local Stability and Instability**

In the Cournot-Nash game-theoretic analysis of duopoly, if firms “cooperate” and split the monopoly-level output, they make equally high profits. However, each firm has an incentive to “defect” and produce a larger amount where its marginal revenue equals its marginal cost, because it will make a higher profit still—if the other firm continues to produce its share according to the monopoly formula. This gives both firms an incentive to defect, resulting in both producing where marginal revenue equals marginal cost. This results in a lower profit for each firm than when they split the monopoly output between them, but it is a globally stable strategy, whereas all other strategy combinations are unstable.

As a result, output is higher and price lower under duopoly than monopoly, and the limit of this process as the number of firms increases is “perfect competition”. This is illustrated with the example of a duopoly facing a linear market demand curve $P(Q) = a - b \cdot Q$, with identical quadratic total cost functions $tc(q) = k + c \cdot q + \frac{1}{2} d \cdot q^2$.

Figure 9 shows the output combinations produced by two firms producing at either the Cournot or Keen predicted level, in terms of the demand and cost arguments.

\textsuperscript{17} For those who do not know, this is a line from the 1980s movie *War Games.*

\textsuperscript{18} It may be thought that this result is an artifact of an accidental aggregation effect from using the same reaction coefficient for all firms; we refute this by generalizing the analysis to allow for each firm to have a different reaction coefficient to the market. This research will be published in a subsequent paper.
Figure 15: Output levels (symbolic) under Cournot & Keen strategy combinations

Figure 16 shows the numeric outcomes with parameter values of \( a=800, \ b=10^{-7}, \ c=100 \) & \( d=10^{-8} \). Clearly, the Keen/Keen combination results in the lowest aggregate output, and the highest price; Cournot/Cournot gives the highest aggregate output and lowest price; while the mixed strategy results in the highest output for one firm and the lowest for the other, with an intermediate aggregate output.

Figure 17 shows why firms are tempted to "defect"—or in our terms, to move from not interacting to behaving strategically at this level of analysis. The firm that reacts to its competitor and prices where marginal revenue equals marginal cost will produce a greater quantity, which is only partly offset by a lower market price—so long as its competitor does not change its strategy. It unambiguously increases its profit, while that of its competitor falls. However, the same temptation also applies to the competitor, so both are likely to switch to interacting strategically. This is the temptation that makes the Cournot/Cournot combination a Nash Equilibrium, even though it involves an unambiguously lower profit for both firms.
Figure 17: Change in profit (symbolic) from Keen/Keen combination

Figure 18: Change in profit (numeric) from Keen/Keen combination

So far, the argument looks conclusive for the Cournot-Nash Equilibrium as the outcome of strategic interaction, and competition thus works to cause higher aggregate output and lower prices than would apply with fewer firms in the industry. Add more firms, and ultimately you converge to where price equals marginal cost—the Holy Grail of perfect competition.

The acknowledged wrinkle in this argument is that, with infinitely repeated games, the Nash equilibrium is the Keen strategy—called “collusion” or “cooperate” in the literature because, before our critique, it was believed that the only way firms could work this amount out was by acting as a cartel.\(^{19}\) It’s possible to “rescue” perfect competition by assuming finitely repeated games, showing that “defect” (or Keen) dominates the final play, reverse-iterating back to the second last, and by finite backwards regression arrive at “defect” as the ideal strategy for all iterations. But this is obviously weak as an analogy to actual competition, where the infinitely repeated game is closer to the reality of an indefinite future of competition—even if some competitors do exit an industry, their rivals can never know when this might happen.

\(^{19}\) Of course, neoclassical economists still believe this today, and will doubtless continue believing it, given how dogma has almost always overruled logic in the development of economic theory.
Most game theorists express puzzlement with this dilemma: a strategy is dominant in a one shot, but not in a repeated game. So “collusion” (or more correctly, “non-interaction”) appears dominant, and it appears that firms will tend not to compete over time.20

There is an additional wrinkle that possibly explains this known dilemma (and the simulation results shown in Figure 8):21 while the Cournot strategy is a Nash Equilibrium, it is locally unstable, and while the Keen strategy is not a Nash Equilibrium, it is locally stable. This occurs not because of collusion, but because strategic interactions—which we might describe as a “Meta-Nash Dynamic”—lead from the Cournot equilibrium to the Keen.

One firm may benefit from a strategic change—say, Firm 1 increasing its output from that in the Keen/Keen output pair, while Firm 2 reduces it. The strategy pair would then be “increase, decrease” (or “+1/-1”) and the profit outcomes “increase, decrease”. In an iterative search for the profit-maximizing level, this would encourage Firm 1 to continue increasing its output, which would take it in the direction of the Cournot equilibrium. However Firm 2, having lost from that strategic combination, will change its strategy—and rather than decreasing its output further, it will increase its output. Then the strategy pair will be “increase, increase” and the profit outcomes are “decrease, decrease”. As a result, both firms will change their strategy to “decrease”, and head back to the Keen equilibrium.

Figure 19 illustrates this using the example parameters above.22 The top half shows the outcome for Firm 1; the bottom half, for Firm 2; strategy settings by Firm 1 are shown by column 1, and settings by Firm 2 by row one. A strategy pair of “+1/-1” results in Firm 1 increasing profit by 333, which clearly encourages Firm 1 to continue increasing production. However, that combination causes a drop in profits of 333 for Firm 2, which will cause Firm 2 to swap strategies—say from “-1” to “+1”. That will then switch the market situation to the “+1/+1” combination, where both firms suffer a fall in profits (and the fall in profits gets larger for larger output increases). Both firms are then likely to switch to reducing output. The Keen equilibrium is thus locally stable because of strategic interactions.

20 We suspect that this dilemma explains the paradox that neoclassical economists, who are normally so opposed to government intervention, support “competition policy”, which in effect forces firms to compete with each other.

21 One curious feature of this simulation is that the convergence result is dependent, not on the number of firms—as neoclassical theory falsely asserts—but on the dispersal of output changes by each firm. The higher the size, relative to output, of the randomly allocated output changes, the higher the likelihood that the end result will be convergence to the Cournot equilibrium rather than the Keen. This result is reported in Keen & Standish 2006.

22 The outcome applies so long as a>c, b<a and d<c; all these are fundamental conditions for a market to exist in the first instance. a<c, for example, would set the equilibrium market output at less than zero.
The Cournot equilibrium, on the other hand, is locally unstable. Figure 20 shows the outcomes for changes of one unit for each firm. The strategy pair "+1/-1" results in increase in profits for Firm 1 and a fall in profits for Firm 2, as it did in the Keen equilibrium. Firm 1 will then be encouraged to continue increasing production, while Firm 2 will be encouraged to switch from reducing output to increasing output. The next strategy pair is thus likely to be "+1/+1" (or some higher combination). This also causes a loss for both firms, so another switch in strategy is likely—to reducing output.

Unlike the Keen/Keen situation, the strategy pair "-1/-1" from the Cournot equilibrium results in an increase in profits for both firms—and larger reductions in output cause larger increases in profit. Further movement away from the Cournot equilibrium is rewarded, so that both firms are likely to adopt the strategy of reducing output, until they reach the Keen equilibrium—with absolutely no "collusion" taking place. The Cournot equilibrium is thus locally unstable, not because of collusion, but because of strategic interactions.

Figure 21 and Figure 22 put the impact of strategic interactions graphically: in each case the predicted output pair (Keen/Keen and Cournot/Cournot respectively) is in the middle of the box. While firms are not behaving collusively, the only strategy pairs that have a chance...
to be self-sustaining are those that have a positive impact on the profit of both parties—since as explained above, any strategy that has a negative impact will necessarily mean a change in behavior by one or both firms. Therefore, the shape of the aggregate profit “hill” indicates whether any sustaining strategic interactions exist.

**Figure 21** confirms that there are no such interactions in the vicinity of the Keen equilibrium: all strategic pairs involve a fall in aggregate profits relative to the starting point. The Keen equilibrium is thus locally stable.

*Figure 21: Impact of strategic interactions on profit near Keen equilibrium*

The Cournot equilibrium, on the other hand, is locally unstable, because aggregate profit will rise if both firms reduce output (see **Figure 22**).
Thus, though the Cournot-Nash defense of perfect competition is not strictly false, in practice it is fragile. It appears that, if a profit-maximizing level of output per firm can be identified, then rational profit-maximizing firms will identify it, regardless of how many of them there are in an industry.\textsuperscript{23} The Holy Grail of "perfect competition", though theoretically attainable via strategic interaction, \textsuperscript{,} is a will o’ the wisp.

So too, ironically, is the argument that there is a profit-maximizing level of output per firm.

The empirical reality of competition

A plethora of empirical studies have established that at least 89 per cent of output—and perhaps as much as 95 per cent—is produced under conditions of constant or falling marginal cost, and rising economies of scale.\textsuperscript{24} Given such circumstances, there is no profit-maximizing level of output for the individual firm: so long as the sale price exceeds average costs, the firm will profit from additional sales. The key presumption of the neoclassical model—that there is a profit-maximizing level of sales—is thus not fulfilled in the real world.

\textsuperscript{23} Subject to the one caveat mentioned in Footnote 21.

\textsuperscript{24} See Lee 1998 for a comprehensive survey of the 20\textsuperscript{th} century studies.
The most recent such survey was carried out by Alan Blinder and a team of PhD students in 1998. Blinder’s results are also arguably the most authoritative, given the scale of his study, and Blinder’s prestige as an economist.

Blinder et al. surveyed a representative weighted sample of US non-agricultural corporations with annual sales of more than US$10 million; a 61% response rate resulted in a study of 200 corporations whose combined output represented 7.6% of the USA’s GDP. The interviews were face to face, with Blinder and a team of Economics PhD students conducting the interviews; the interviewees were top executives of the firms, with 25% being the President or CEO, and 45% a Vice President.

Blinder summarized the results in the following way:

“First, about 85 percent of all the goods and services in the U.S. nonfarm business sector are sold to “regular customers” with whom sellers have an ongoing relationship … And about 70 percent of sales are business to business rather than from businesses to consumers…

Second, and related, contractual rigidities … are extremely common … about one-quarter of output is sold under contracts that fix nominal prices for a nontrivial period of time. And it appears that discounts from contract prices are rare. Roughly another 60 percent of output is covered by Okun-style implicit contracts which slow down price adjustments.

Third, firms typically report fixed costs that are quite high relative to variable costs. And they rarely report the upward-sloping marginal cost curves that are ubiquitous in economic theory. Indeed, downward-sloping marginal cost curves are more common… If these answers are to be believed … then [a good deal of microeconomic theory] is called into question… For example, price cannot approximate marginal cost in a competitive market if fixed costs are very high.” (p. 302)

The key final point about falling marginal cost deserves elaboration. Given that, as they discovered, “marginal cost is not a natural mental construct for most executives,” they translated marginal cost into ‘variable costs of producing additional units,” and posed the following question:

B7(a). Some companies find that their variable costs per unit are roughly constant when production rises. Others incur either higher or lower variable costs of producing additional units when they raise production.

How would you characterize the behavior of your own variable costs of producing additional units as production rises? (Blinder 1998: 102)

The survey team collated the responses into five groups, as summarized in Table 1:

<table>
<thead>
<tr>
<th>Structure of Marginal Costs</th>
<th>Percentage of firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decreasing</td>
<td>32.6</td>
</tr>
<tr>
<td>Decreasing with discrete jumps</td>
<td>7.9</td>
</tr>
<tr>
<td>Constant</td>
<td>40</td>
</tr>
<tr>
<td>Constant with discrete jumps</td>
<td>7.9</td>
</tr>
<tr>
<td>Increasing</td>
<td>11.1</td>
</tr>
</tbody>
</table>

Table 1: Marginal cost structure of American corporations (Blinder et al. 1998: 102-103)
Blinder et al. pithily observed that:

“The overwhelmingly bad news here (for economic theory) is that, apparently, only 11 percent of GDP is produced under conditions of rising marginal cost.” (102)

The overall results of Blinder’s survey are summarized in Table 2. Given the empirically common circumstances detailed here, the pre-requisites for being able to identify a profit-maximizing level of output do not exist for at least 89 per cent of US firms.\(^{25}\) Instead, for these firms, the only profit-maximizing strategy is to sell as much as they can—and at the expense, where possible, of competitors’ sales.

<table>
<thead>
<tr>
<th>Summary of Selected Factual Results Price Policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median number of price changes in a year</td>
</tr>
<tr>
<td>Mean lag before adjusting price months following Demand Increase</td>
</tr>
<tr>
<td>Demand Decrease</td>
</tr>
<tr>
<td>Cost Increase</td>
</tr>
<tr>
<td>Cost Decrease</td>
</tr>
<tr>
<td>Percent of firms which Report annual price reviews</td>
</tr>
<tr>
<td>Change prices all at once</td>
</tr>
<tr>
<td>Change prices in small steps</td>
</tr>
<tr>
<td>Have nontrivial costs of adjusting prices of which related primarily to the frequency of price changes</td>
</tr>
<tr>
<td>the size of price changes</td>
</tr>
<tr>
<td>Sales</td>
</tr>
<tr>
<td>Estimated percent of GDP sold under contracts which fix prices</td>
</tr>
<tr>
<td>Percent of firms which report implicit contracts</td>
</tr>
<tr>
<td>Percent of sales which are made to Consumers</td>
</tr>
<tr>
<td>Businesses</td>
</tr>
<tr>
<td>Other (principally government)</td>
</tr>
<tr>
<td>Regular customers</td>
</tr>
<tr>
<td>Percent of firms whose sales are Relatively sensitive to the state of the economy</td>
</tr>
<tr>
<td>Relatively Insensitive to the state of the economy</td>
</tr>
<tr>
<td>Costs</td>
</tr>
<tr>
<td>Percent of firms which can estimate costs at least moderately well</td>
</tr>
<tr>
<td>Mean percentage of costs which are fixed</td>
</tr>
<tr>
<td>Percentage of firms for which marginal costs are Increasing</td>
</tr>
<tr>
<td>Constant</td>
</tr>
<tr>
<td>Decreasing</td>
</tr>
</tbody>
</table>

\(^{25}\) We say at least because all previous surveys have reported a lower proportion of products that are produced under conditions of diminishing marginal productivity—typically 5 per cent of output (Eiteman & Guthrie 1952).
The only practical way that this can be done is via product differentiation, and that indeed is the obvious form that real competition actually takes. Innovation and heterogeneity are the true hallmarks of competition, yet these concepts are effectively excluded by the neoclassical model.

A model of how this actual form of competition works would be extremely useful to economic theory—and perhaps even to economic policy, if we could scientifically identify those industry structures that truly promote innovation. The continued teaching of the neoclassical model, and the continued development of a research tradition in which rising marginal cost plays a key role, are a hindrance to developing an adequate model of real world competition.

Our closing observation on this theory is perhaps the most important. A theory is more than a scholastic exercise: a good theory is also an attempt to understand reality, and, where possible, to alter it for the better. There are, therefore, few things more dangerous than an applied bad theory. Unfortunately, neoclassical competition theory is applied throughout the world, in the guise of policies intended to promote competition.

The anti-capitalist nature of neoclassical competition policy

The neoclassical vision of competition has been enshrined in competition policies adopted by governments and applied to key industries such as telecommunications, power, sanitation, and water supply. The major practical implications of accepted theory are that more firms equates to increased competition, increased competition means higher output at lower prices, and market price should ideally be equal to marginal cost.

Since the theory is flawed, these implications are at best unproven, and at worst false. There are now numerous instances around the world where competition policies have resulted in deleterious outcomes; a special issue of *Utilities Policy* in 2004 details several of these for the USA (and Australia). Loube, for example, examined the US Telecom Act of 1996, and found that “this policy has actually raised prices for residential customers” (Trebing & Miller 2004: 106).

Proponents of competition policy normally ascribe such outcomes to poor implementation of policy, poor regulatory oversight, or unanticipated circumstances. However, if the theory is flawed as we argue, then these outcomes are not accidents, but the systemic results of imposing a false theory on actual markets. Some predictable negative consequences are rising costs due to reduced economies of scale, underinvestment caused by imposed prices that lie below average cost, and reduced rates of innovation in related industries, caused by the inadequate “competitive” provision of infrastructure.

That these policies were imposed in a well-meaning attempt to improve social welfare cannot detract from the fact that, if the theory guiding these policies was false, then the policies are likely to cause more harm than good. Real world markets would function far better if competition policy, as it stands, were abolished.

Conclusion

A careful examination of the neoclassical theory of competition thus finds that it has little, if any, true content.
The Marshallian argument, which forms the backbone of neoclassical pedagogy, is strictly false in its belief that a downward-sloping market demand curve is compatible with horizontal individual firm demand curves. Once this error is corrected, the model’s major conclusion, that competitive industries are better than concentrated ones, is overturned. Given identical demand and cost conditions, competitive industries will produce the same output as monopolies, and sell at the same price—and there are good grounds for expecting that monopolies would have lower costs (see Appendix One).

The Cournot analysis is mathematically correct, but subject to a problem of local instability as well as the known dilemma of repeated games. If it is interpreted as an “as if” explanation for what happens between competing firms in an industry—i.e., it proposes that firms do not actually solve the mathematics to find their Nash equilibrium output levels, but instead undertake an iterative search of the output-profit space—then this iterative search will lead to the Keen equilibrium, not the Cournot-Nash one, because the former is locally stable under strategic interactions, while the latter is not.

Given this intrinsic barrenness of the theory, its empirical irrelevance is even more important. Neoclassical economists have ignored a multitude of empirical papers that show that marginal cost does not rise, that firms do not compete on price, and so on, on the basis of Friedman’s belief that asking businessmen what they do is not “a relevant test of the associated hypothesis.” But if the “associated hypothesis” is in fact false, or irrelevant, then “asking businessmen what they do” is at least a good place from which to derive stylized facts that a relevant hypothesis would have to explain. It is high time that economists abandoned what superficially appears to be “high theory”, and got their hands dirty with real empirical research into actual firms and actual competition.

Here the picture that emerges from even a cursory examination of the data is very different to neoclassical belief. Table 3 shows the aggregate distribution of firm sizes in the USA in 2002: large firms make up well under 0.3 per cent of the total number of firms, but are responsible for over 60 per cent of sales.

<table>
<thead>
<tr>
<th>Industry</th>
<th>Total</th>
<th>0-499</th>
<th>500+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>5,697,759</td>
<td>5,680,914</td>
<td>16,845</td>
</tr>
<tr>
<td>Estab.</td>
<td>7,200,770</td>
<td>6,172,809</td>
<td>1,027,961</td>
</tr>
<tr>
<td>Emp.</td>
<td>112,400,654</td>
<td>56,366,292</td>
<td>56,034,362</td>
</tr>
<tr>
<td>Ann. pay.($000)</td>
<td>3,943,179,606</td>
<td>1,777,049,574</td>
<td>2,166,130,032</td>
</tr>
<tr>
<td>Receipts($000)</td>
<td>22,062,528,196</td>
<td>8,558,731,333</td>
<td>13,503,796,863</td>
</tr>
</tbody>
</table>

Table 3: US firm size data (US Office of Small Business Advocacy)

At the same time, small firms are not negligible: all industries are characterized by a wide distribution of firm sizes, from sole trader through to large conglomerates (see Table 4). Perhaps the real story of competition is the survival of such diversity.
Table 4: Distribution of firm sizes in manufacturing (US SBA)

In the light of both its theoretical weaknesses and its irrelevance to the empirical data, Sraffa’s advice in 1930 about what to do with Marshall’s theory bear repeating today, not only in relation to Marshall’s theory, but even to the Cournot-Nash approach:

the theory cannot be interpreted in a way which makes it logically sell-consistent and, at the same time, reconciles it with the facts it sets out to explain. Mr. Robertson’s remedy is to discard mathematics, and he suggests that my remedy is to discard the facts; perhaps I ought to have explained that, in the circumstances, I think it is Marshall’s theory that should be discarded. (Sraffa 1930: 93)

The neoclassical theory of competition is a hindrance to understanding real markets and real competition, and it should be abandoned.

Appendices

Appendix One: Conditions for comparability of cost structures

Economists blithely draw diagrams like Figure 23 below to compare monopoly with perfect competition. As shown above, the basis of the comparison is false: given Marshallian assumptions, an industry with many “perfectly competitive” firms will produce the same amount as a monopoly facing identical demand and cost conditions—and both industry structures will lead to a “deadweight loss”. However, in general, small competitive firms would have different cost conditions to a single firm—not only because of economies of scale spread result in lower per unit fixed costs, but also because of the impact of economies of scale on marginal costs.
Rosput (1993) gives a good illustration of this latter point in relation to gas utilities. One of the fixed costs of gas supply is the pipe; one of the variable costs is the compression needed to move the gas along the pipe. A larger diameter pipe allows a larger volume of gas to be passed with lower compression losses, so that the larger scale of output results in lower marginal costs:

Simply stated, the necessary first investment in infrastructure is the construction of the pipeline itself. Thereafter, additional units of throughput can be economically added through the use of horsepower to compress the gas up to a certain point where the losses associated with the compression make the installation of additional pipe more economical than the use of additional horsepower of compression. The loss of energy is, of course, a function of, among other things, the diameter of the pipe. Thus, at the outset, the selection of pipe diameter is a critical ingredient in determining the economics of future expansions of the installed pipe: the larger the diameter, the more efficient are the future additions of capacity and hence the lower the marginal costs of future units of output (Rosput 1993: 288).

Thus a single large supplier is likely to have lower costs—in which case, the marginal cost curve for the monopoly should be drawn below that for the competitive industry. Given the same demand curve and the same profit-maximizing behavior, a monopoly is thus likely to produce a higher output than a competitive industry, and at a lower cost.

The cost examples in this paper were artificially constructed to ensure that the assumption of identical costs embodied in Figure 23 were fulfilled—something that we doubt has been done by neoclassical authors in comparable papers. The cost functions were:
Obviously, it is very arbitrary to have the number of firms in an industry as an argument in the marginal cost function of a single firm—and also highly unlikely. Yet without that “heroic” assumption, the aggregate of marginal costs curves for a competitive industry will necessarily differ from the marginal cost curve for a monopoly. If a monopoly has greater access to economies of scale than smaller competitive firms, as in Rosput’s example of gas transmission, then on conventional profit-maximizing grounds, a monopoly would produce a higher output for a lower price.

It is also easily shown that the neoclassical pedagogic assumption that the same marginal cost curve can be drawn for a competitive industry and a monopoly is true in only two circumstances: either the monopoly simply changes the ownership of plants in the industry—so that there is no technical difference between one industry structure and the other—or both industry structures face identical constant marginal costs.\(^\text{26}\)

Marginal cost is the inverse of marginal product, which in turn is the derivative of total product. The condition of identical marginal costs—that is, that the marginal cost curve for a monopoly is identically equal to the sum of the marginal cost curves of an industry with many competitive firms, for all relevant levels of output—therefore requires that the total products of two different industry structures differ only by a constant. This constant can be set to zero, since output is zero with zero variable inputs.

Consider a competitive industry with \(n\) firms, each employing \(x\) workers, and a monopoly with \(m\) plants, each employing \(y\) workers, where \(n > m\). Graphically this condition can be shown as in Figure 24.

\[ \text{Monopoly: } MC(Q) = C + D \cdot Q + E \cdot Q^2 \]
\[ \text{Competitive: } mc(q,n) = C + D \cdot n \cdot q + E \cdot n^2 \cdot q^2 \]

\((0.20)\)

\(^\text{26}\) This argument was first published in Keen 2004a.
Using $f$ for the production function of the competitive firms, and $g$ for the production function of the monopoly, the equality of total products condition can be put in the following form:

$$ n \cdot f(x) = m \cdot g(y) \quad (0.21) $$

Substitute $y = \frac{n \cdot x}{m}$ into (1) and differentiate both sides of (0.21) by $n$:

$$ f(x) = \frac{x}{m} \cdot g\left(\frac{n \cdot x}{m}\right) \quad (0.22) $$

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

$$ \frac{g\left(\frac{n \cdot x}{m}\right)}{n} = \frac{x}{m} \cdot g\left(\frac{n \cdot x}{m}\right) \begin{cases} \text{or} \\
\frac{g'\left(\frac{n \cdot x}{m}\right)}{g\left(\frac{n \cdot x}{m}\right)} = \frac{m}{n \cdot x} \end{cases} \quad (0.23) $$

The substitution of $y = \frac{n \cdot x}{m}$ yields an expression involving the differential of the log of $g$: 
\[
\frac{g'(y)}{g(y)} = \frac{1}{y}
\]

(0.24)

Integrating both sides yields:

\[
\ln(g(y)) = \ln(y) + c
\]

(0.25)

Thus \( g \) is a constant returns production function:

\[
\begin{aligned}
g(y) &= C \cdot y
\end{aligned}
\]

(0.26)

It follows that \( f \) is the same constant returns production function:

\[
\begin{aligned}
f(x) &= \frac{w}{n} \cdot C \cdot \frac{x}{m}
\end{aligned}
\]

(0.27)

With both \( f \) and \( g \) being identical constant returns production functions, the marginal products and hence the marginal costs of the competitive industry and monopoly are constant and identical. The general rule, therefore, is that welfare comparisons of perfect competition and monopoly are only definitive when the competitive firms and the monopoly operate under conditions of constant identical marginal cost.

References


SUGGESTED CITATION: