Available online at www.sciencedirect.com
SCIENCE DIRECT
Utilities Policy 12 (2004) 109–125
UTILITIES POLICY
www.elsevier.com/locate/jup

Deregulator: Judgment Day for microeconomics

Steve Keen *

School of Economics and Finance, University of Western Sydney, Locked Bag 1797 Penrith 1797, NSW, Australia

Received 24 March 2004; accepted 4 April 2004

Abstract

The economic theory that motivated the deregulation and privatization of the US electricity industry is seriously flawed in three crucial ways. First, the Marshallian theory of the firm is based on two mathematical errors which, when amended, reverse the accepted welfare rankings of competitive and monopoly industry structures: on the grounds of corrected neoclassical theory, monopoly should be preferred to competition. Second, while proponents of deregulation expected market-clearing equilibrium prices to apply, it is well known that the equilibrium of a system of spot market prices is unstable. This implies that imposing spot market pricing on as basic an industry as electricity is likely to lead to the kind of volatility observed under the deregulation. Third, extensive empirical research has established that on the order of 95% of firms do not produce under conditions of rising marginal cost. Requiring electricity firms to price at marginal cost was therefore likely to lead to bankruptcies, as indeed occurred. The economic preference for marginal cost spot market pricing is therefore theoretically unsound, and it is no wonder that the actual deregulatory experience was as bad as it was.

JEL classification: D0; D4; D5; D6; K2; L; L2; L5; L9

Keywords: Microeconomics; General equilibrium; Deregulation

1. Introduction

Deregulation of the US electricity market was driven by the belief that a free market would result in a more efficient outcome than either regulated competition or the public provision of electric power. The story told to the economic layman was that both the cost of production and final consumer prices would fall. The story told to the economic cognoscenti was that the elimination of regulation would enable a closer matching of the marginal benefits and marginal costs of electricity production. Both explanations anticipated a substantial rise in social welfare.

The promotions for "Deregulator" thus promised a Disneyland future, but in a metamorphosis worthy of Kakfa, the experience was closer to *Terminator*'s "Judgment Day." Electricity prices rose dramatically, and with unheard of volatility—a typical instance being the increase in the mid-Atlantic wholesale market price from $5 per MW h to $177 per MW h during the first quarter of 2001 (Trebing, 2003: p. 298). Supply was curtailed, leading to widespread power outages and power rationing at exorbitant prices.

Many of these problems were blamed on the behaviour of unscrupulous and now largely bankrupt businesses (whose managers have, however, generally avoided personal bankruptcy). Some even blamed the regulations that ushered in deregulation—the group responsible for the "Manifesto on the California Electricity Crisis" (2003 and 2001) alleged that these regulations forced excessive reliance upon spot markets, and that "economic losses due to the crisis would have been greatly reduced if the utilities had not been required by regulation to rely on the spot markets for over 50% of their supplies..."

These arguments have some merit, but they leave untold the profound story: the chaos of deregulation

* Tel.: +61-2-4620-3016.
E-mail address: s.keen@uws.edu.au (S. Keen).

0957-1787/$ - see front matter
doi:10.1016/j.jup.2004.04.010

was an entirely predictable outcome of applying conventional economic theory to this crucial real-world market, because this theory is flawed in three fundamental ways:

- New research shows that the theory of markets contains mathematical and economic fallacies which, when corrected, reverse the accepted welfare predictions. The assertions that prices are lower, output higher and welfare is maximized by competitive markets are theoretically false. Corrected theory proves that deadweight welfare losses are unavoidable even in competitive markets, and that (according to neoclassical theory) monopolies are likely to result in lower prices, higher output and greater consumer welfare than competitive markets;
- The presumption that free market spot prices will converge to a market-clearing equilibrium set of prices is mathematically false. Though a market-clearing equilibrium set of prices and outputs can be defined, that set is unstable, so that the market will never reach equilibrium. It is hardly remarkable that pricing chaos followed the imposition of a spot market-clearing price system in so basic an industry as electricity; and
- Real-world firms, including electricity producers, do not have the cost structures assumed by economic theory, with the result that setting price equal to marginal cost would cause the vast majority of firms to go bankrupt. There are good practical and theoretical reasons why most products are not produced under conditions of diminishing marginal productivity, so that in practice marginal costs are constant or falling and well below average costs. The spate of bankruptcies that followed the imposition of marginal cost pricing, though not intended by the regulatory authorities, was nonetheless no accident.

Amending these three flaws leads to three very different policy recommendations:

- Pricing policy should return to the historical emphasis upon cost recovery and adequate profitability
- Given the crucial and fundamental role of electricity in production, the primary focus of regulation should be on the maintenance of reliable supply with low price volatility; and
- The economic fetish for the so-called competitive firms without market power should be abandoned, and the industry allowed to evolve towards the situation that characterizes real-world competitive industries, of a Zipf/Power-law distribution of both firm size and effective market power (Axtell, 2001: pp. 1818–1820).[1]

2. Fallacies in the theory of markets[2]

Though economics has developed some new modes of analysis in recent decades, the Marshallian vision of the firm remains a core belief, and is certainly dominant in the "applied" microeconomics under which the California Public Utilities Commission (CPUC) ushered in marginal cost pricing. A few economists are aware of at least one problematic component of this theory (Stigler, 1957), but most economists believe it is incontrovertible.

This belief is false: the theory contains two crucial fallacies which, when corrected, not only destroy the theory but also invert the conventional economic ranking of competitive industries and monopolies. I will outline first the orthodox understanding of the theory, and then the fallacies (full proofs are given in Appendix A).

2.1. The belief

The conventional Marshallian/neoclassical theory of markets argues that market price and quantity are set by the intersection of supply and demand. The supply curve represents the marginal cost of production of a commodity, while the demand curve represents the marginal benefit of its consumption. Where the two marginals are equal, the gap between total benefits and total costs is greatest.

However, ever since Harrod developed the concept of marginal revenue (Besomi, 1999: pp. 16–18) it has been known that this picture of social harmony via market equilibrium only applies in competitive markets. In the other extreme of a monopoly, the monopolist sets price where marginal cost equals not price, but marginal revenue. A monopoly therefore sells a smaller output at a higher price than a competitive market.

[1] Zipf/Power-law distributions are statistical spreads of some key characteristic (in this case the size of firms in terms of employees, dollar turnover per annum, etc.) that are proportional to this size measure raised to a power. When the relationship between the size measure (number of employees per firm) and the frequency of this size measure in the data (percentage of firms with this many employees) is plotted on a log–log plot, the relationship is a straight line. Axtell (2001: p. 1819) found that this relationship fitted all US firms to a high degree of accuracy (the relationship had an R^2 of 0.992). The same relationship is likely to apply within industries—though with less accuracy—so that industries are characterized by very few large firms and very many small ones, rather than the neoclassical taxonomy of "monopoly" or "perfect competition".

[2] This section (and Appendix A) summarize new research that is contained in a more technical paper currently being refereed for another journal. The paper (Keen et al., 2004) can be download from http://www.debunking-economics.com

A crucial part of this analysis is the relationship between price and marginal revenue for the individual firm. Marginal revenue is the rate of change of total revenue, where total revenue equals price times output. For the monopolist, quantity equals total market output, and the demand curve faced by the firm is the market demand curve, which is assumed to be negatively sloped. Mathematically, this means that marginal revenue is less than price:

$$\mathrm{MR_M} = \frac{\mathrm{d}}{\mathrm{d}Q}(P \cdot Q) = P \cdot \frac{\mathrm{d}Q}{\mathrm{d}Q} + Q \cdot \frac{\mathrm{d}P}{\mathrm{d}Q} = P + Q \cdot \frac{\mathrm{d}P}{\mathrm{d}Q} < P \quad (1)$$

Since the monopolist maximizes profit by producing the quantity at which marginal cost equals marginal revenue, the monopoly price will exceed marginal cost.

It is also not possible to derive a "supply curve" for a monopoly, since there is a different marginal revenue curve for every demand curve, and the monopolist produces well above its marginal cost curve. Supply and demand analysis, that mainstay of conventional economic logic, is impossible if industries are characterized by monopolies or similar uncompetitive structures.

On the other hand, competitive firms individually produce a much smaller amount than the monopoly, and are "price-takers" who cannot influence the market price. The demand curve experienced by each individual firm is therefore a horizontal line at the market price, because while market price falls if the aggregate market quantity produced rises $(\mathrm{d}P/\mathrm{d}Q) < 0$, the market price is unaffected by changes in the output of a single firm $((\mathrm{d}P/\mathrm{d}q_i) = 0)$:

$$\mathrm{MR}_i = \frac{\mathrm{d}}{\mathrm{d}q_i}(P \cdot q_i) = P \cdot \frac{\mathrm{d}q_i}{\mathrm{d}q_i} + q_i \cdot \frac{\mathrm{d}P}{\mathrm{d}q_i} = P + q_i \cdot 0 = P \quad (2)$$

Thus, while competitive firms follow precisely the same profit-maximizing guideline of equating marginal revenue and marginal cost, the proposition that the firm's marginal revenue equals the market price means that each firm produces where marginal cost equals the market price. Therefore, the rising portion of the marginal cost curve of the firm becomes its supply curve, and the sum of all firms' supply curves equals the supply curve for the industry. At the aggregate market level, the intersection of this supply curve with the market demand curve determines the equilibrium price. With the area above the price and below the demand curve representing consumer welfare, and the area below price and above the supply curve representing producer welfare, overall social welfare is maximized by perfect competition. On the other hand, with a monopoly supplier, there is a transfer of surplus from consumers to the producer, and a deadweight loss of welfare due to the monopoly.

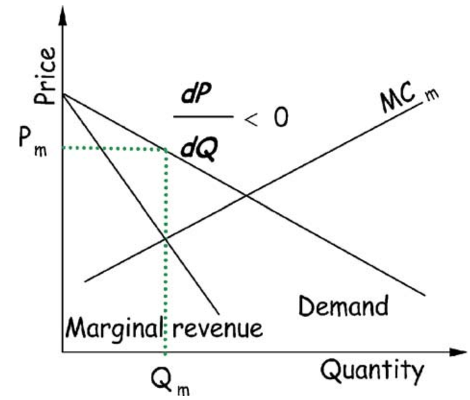

Fig. 1. Monopoly prices where marginal cost equals marginal revenue.

These arguments are summarised in Figs. 1–3. Fig. 1 shows a monopolist with a rising marginal cost curve and a downward sloping market demand curve. The firm produces the quantity Q_m at the price P_m.

Fig. 2 shows the situation for the market and the individual firm in a competitive industry. Market price is set by the intersection of demand and supply, while each individual firm takes this market price as given and produces where market price equals its marginal cost. The sum of the marginal cost curves of all firms in the industry then determines the industry supply curve. Each individual firm produces the output q_e at the price P_e, while the aggregate industry output is Q_e.

The output of the competitive industry exceeds that of the monopoly, while the competitive price is lower, resulting in the welfare comparison shown in Fig. 3. There is no deadweight loss of surplus with the competitive market (Fig. 3a), and price is lower and output higher. The socially optimum price level occurs where price equals marginal cost, and therefore the competitive market is welfare superior to the monopoly.

2.2. *The application of microeconomic theory to electricity*

The welfare propositions of standard economic theory are evident in the policy discussions of economists on electricity pricing. Hogan, for example, argued that:

> The standard determinant of competitive market pricing is system marginal cost. This is the simple definition of the market-clearing price where supply equals demand. This production level just balances the marginal benefit of additional consumption with the marginal cost of production. Under the usual competitive assumptions, this

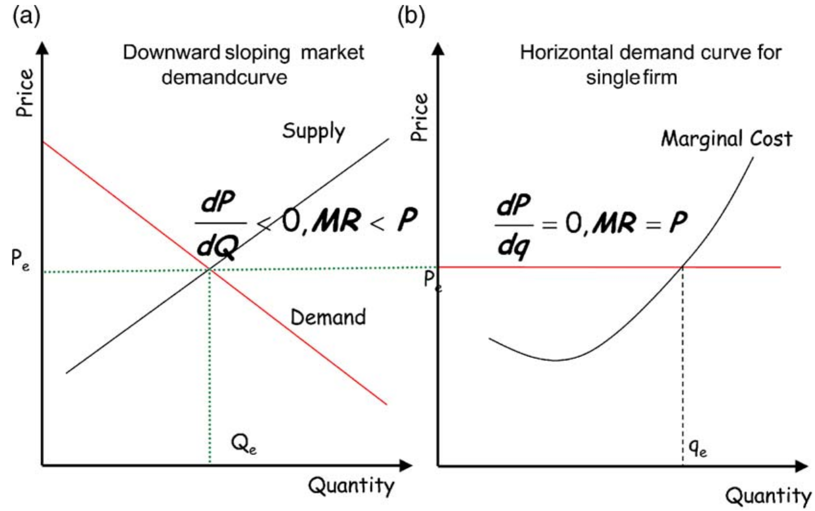

Fig. 2. Competitive firms price where marginal cost equals price.

textbook market equilibrium condition also provides the welfare maximizing economic outcome, which is the definition of economic efficiency (Hogan, 2001: p. 13).

For most industries, this theoretical ideal would remain of academic interest only. But as the regulatory authorities for utilities became more enamoured of economic theory, the theoretical ideal became a regulatory objective. Conkling (1999) dates the shift from theory to practice to a 1974 Act of the California legislature (ACR 192, August 31 1974) "that directed the [California Public Utilities Commission] (CPUC) to investigate marginal cost pricing as one of six alternatives to existing rate structures" (Conkling, 1999: p. 23). The consequent Decision 85559 of the CPUC in 1976 adopted marginal cost pricing, inaugurating the shift away from cost recovery and adequate profitability in regulatory oversight. Conkling observed that:

> In what easily could be mistaken for a high level debate within the economics profession, the testimony presented argued the pros and cons of using marginal costs versus average costs in rate-making... The Commission concluded that "efficient resource allocation requires that all prices be set equal to their 'incremental' costs"... The Commission adopted a policy "to make conservation in the sense of efficient allocation of elec-

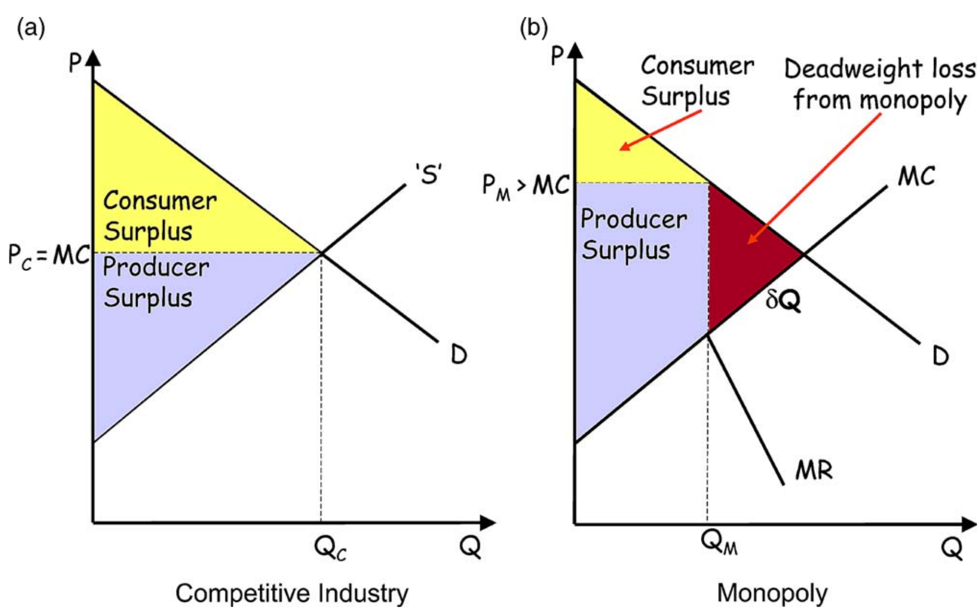

Fig. 3. Welfare maximization from competition, deadweight loss from monopoly.

tricity the keystone of the rate structure'' (p. 7) (Conkling, 1999: p. 23).

Subsequent decisions of the CPUC put this policy shift into effect. Decision 91107 in 1979 attempted to define company-specific marginal costs, but had to grapple with the unexpected outcome that the resulting price recommendation caused a US$ 1.49 billion deficit for PG&E (Conkling, 1999: p. 25). Decision 92549 in 1981 was, according to the CPUC's own ''Standard Practices'' Handbook, the first decision ''recognizing the desirability of marginal cost pricing applied to electricity ratemaking,'' because in it ''[t]he Commission concluded that '[m]arginal costs provide the acceptable approach to allocating cost recovery between customer groups''' (Conkling, 1999: p. 25).

That statement was number 12 in the ''Findings of Fact'' of Decision 91107 (p. 230). However, though the vast majority of economists believe it to be fact, it is in reality a fallacy, even under the ''usual competitive assumptions'' referred to by Hogan. These assumptions are in turn based on two theoretical fallacies, one affecting the shape of the market supply curve, the other the slope of the demand curve faced by the individual firm.

2.3. The fallacies

The supply curve fallacy emanates from the superficially innocuous assumption that the marginal cost curve of a monopoly supplier is identical to the ''supply curve'' of a competitive industry. This assumption is true in only two circumstances; in general, these two curves *must* differ, thus making any definitive comparison of the welfare effects of different market structures impossible.

The demand curve fallacy arises from a mathematical error that Stigler first identified in 1957. Stigler provided an alleged alternative solution, but more careful analysis shows that this was a ''red herring'' (see Appendix A).

When these two fallacies are corrected, it is easily established that, in circumstances where competitive and monopoly industry organisation can be compared, there is no welfare difference between them: in both industry structures, market output is determined by the intersection of industry-level marginal cost and marginal revenue. Where the two cannot be definitively compared, it is likely that, contrary to conventional theory, monopoly will result in higher consumer welfare than competition.[3]

[3] As discussed in Section 3, reality—rather than flawed economic theory—gives some reasons to restore the traditional economic preference for competition over monopoly; but this involves a very different picture of what competition is.

2.3.1. Identical marginal cost curves

The conventional welfare comparison of competition and monopoly (see Fig. 3) makes the assumption that the marginal cost curve for a monopoly is identical to the sum of the marginal cost curves for the competitive industry, and that both these curves slope upward because of diminishing marginal productivity.

These assumptions are mutually incompatible: if the cost curves are identical, then diminishing marginal productivity cannot apply and the marginal cost curves are identical horizontal lines. If diminishing marginal productivity does apply, then the marginal cost curve of a single producer cannot be the same as the sum of marginal cost curves for several producers. Since economies of scale will give a monopoly an advantage at high levels of output, the welfare comparison of competitive supply versus monopoly will be akin to that shown in Fig. 4: the monopoly will generate *greater* consumer surplus than the competitive market, even if the monopoly prices where marginal cost equals marginal revenue while the competitive industry prices where marginal cost equals price.

Appendix A gives the formal proof of this conundrum. The intuition behind it is that the equivalence of marginal cost curves imposes a condition not merely on these curves themselves, but also on the total cost and total product curves from which they are derived. If the marginal cost curves are identical, then so too are the marginal product curves (since marginal productivity determines marginal cost). If the marginal product curves are identical, then the total product curves can only differ by a constant. If we consider labor as the variable input and capital as the fixed, then output is zero with zero variable input. The total product curve for the monopoly must therefore be

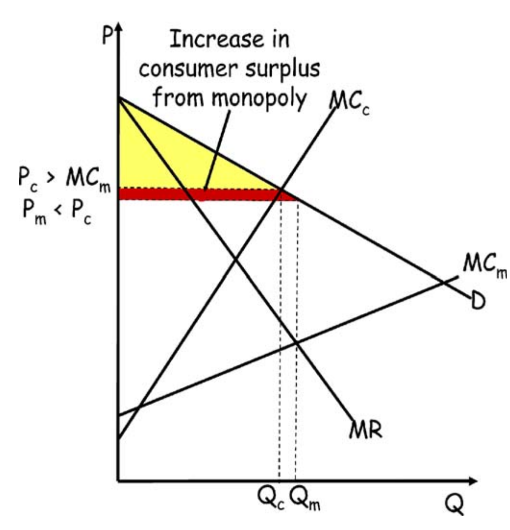

Fig. 4. Higher consumer surplus from monopoly with lower marginal costs.

identical to the sum of the total product curves for the competitive firms.

Mathematical analysis shows that there are only two ways this can occur: either the monopoly simply takes over all the competitive firms and operates them at exactly the same scale as before (in which case its marginal cost curve is the linear sum of the marginal cost curves of the competitive firms), or both the monopoly and the competitive firms have *identical and constant* marginal costs (see Appendix A, or Keen 2001 Chapter 4 for a verbal exposition).

In general, monopolies do not come about simply by one firm taking over the management of many and continuing to operate them exactly as before: when WalMart "monopolises" a previously competitive retail market, it does so by building a supermarket and wiping out (most of) the corner stores—not by making all the Ma and Pa Kettles and their shops into its retail outlets. The general situation is that monopolies use larger-scale production facilities while competitive firms use smaller scale ones.

The cost structures of large firms will therefore differ from those of smaller ones, and these will result in large firms having *lower* marginal costs than small firms (as well as lower average costs). Rosput (1993) gives a good illustration of this 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: p. 288).

Returning to the theory, correcting this fallacy means that the only circumstance under which we can (a) incorporate the real-world phenomenon that monopolies operate a smaller number of plants at a higher level of output than competitive firms and (b) make a definitive comparison of the welfare effects of monopoly and competitive markets is where marginal costs are both identical and constant.[4] This of course raises a conundrum for the conventional model of perfect competition where each firm faces a horizontal demand curve—which brings us to the second fallacy.

2.3.2. *Horizontal firm demand curves*

Though possibly millions of economists have been taught that the individual competitive firm faces a horizontal demand curve, this proposition has been known to be mathematically false since 1957. Writing in the prestigious *Journal of Political Economy*, the leading neoclassical economist George Stigler showed with a single line of calculus that the slope of the supply curve facing the individual competitive firm was identical to the slope of the market demand curve:[5]

$$\frac{\mathrm{d}P}{\mathrm{d}q_i} = \frac{\mathrm{d}P}{\mathrm{d}Q}\cdot\frac{\mathrm{d}Q}{\mathrm{d}q_i} = \frac{\mathrm{d}P}{\mathrm{d}Q} \quad (3)$$

The English rendition of this mathematics is that the slope of the individual firm's demand curve is equivalent to the product of the slope of the market demand curve, and the amount by which total industry output changes given a change in the output of one firm. If a single firm increases its output, industry output will rise by that same amount: therefore, the ratio of the change in industry output to the change in output by a single firm is one. Hence, the slope of the demand curve facing the individual firm is identical to the slope of the market demand curve.

The graphical intuition is shown in Fig. 5, which shows a market demand curve for an industry with a large number of firms. The overall movement from Q_1 to Q_2 involves a change of ΔQ in output and ΔP in price, consisting of changes in the output of each firm of δq that cause a corresponding change of price by δP. The slope of any tiny line segment $\delta P/\delta q$ is equivalent to the slope of the overall section $\Delta P/\Delta Q$.

Economists have presumably accepted the mathematically mutually exclusive propositions that the slope of the market demand curve is negative $(\mathrm{d}P/\mathrm{d}Q) < 0$ while at the same time the demand curve for a single firm is horizontal $(\mathrm{d}P/\mathrm{d}q) = 0$, because it appears similar to saying that the elasticity of demand at the market level $E = (P/Q)\cdot(\mathrm{d}Q/\mathrm{d}P)$ is very small compared to the elasticity of demand with respect to changes in the output of a single firm $e = (P/q)\cdot(\mathrm{d}q/\mathrm{d}P)$. However, since $(\mathrm{d}Q/\mathrm{d}P) = (\mathrm{d}q/\mathrm{d}P)$, this truism is determined simply by the relative size of Q and q, and

[4] This is of course unrealistic—but the point of this paper is that the entire theory is both unrealistic and internally inconsistent.

[5] Appendix A elaborates upon this one line analysis.

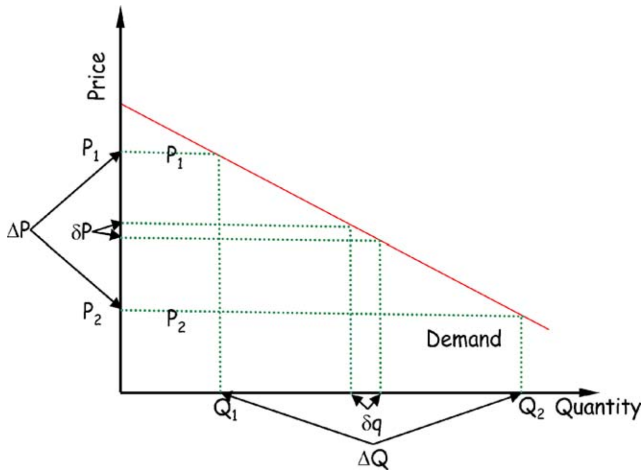

Fig. 5. Slope of firm's demand curve identical to market demand curve.

these terms do not appear at all in the expression for marginal revenue.

Economists also cling to the "small actor" argument that even if the individual firm's demand curve does slope downwards, the firm knows that its impact on market price is minuscule, so it behaves "as if" the market price is given. There is some merit in this argument, but it begs the question: what is the market price? Economists assume that it is the price given by the intersection of the market demand and supply curves;[6] but the latter only exists if $(\mathrm{d}P/\mathrm{d}q) = 0$, and this is mathematically incompatible with a downward-sloping market demand curve. The market price therefore *cannot* be that set by the intersection of supply and demand: it must be something else.

Accurate mathematics (see Appendix A and Keen et al., 2004) shows that the price set by the intersection of market marginal revenue and marginal cost will rule, and this is borne out by computer simulations. The price that the myriad small firms will take as "given"—once it is found via an iterative process—is the "monopoly" price.

2.3.3. The consequences

A number of significant consequences follow from the fact that the slope of the demand curve for the individual competitive firm is the same as the industry demand curve. The most intuitive is that all industries price above marginal cost; the most surprising is that the standard mantra that "a profit-maximizing firm maximizes its profit by equating marginal revenue and marginal cost" is false.

[6] A market supply curve only exists if the demand curve faced by each individual firm is strictly horizontal, so that each firm produces strictly on its marginal cost curve.

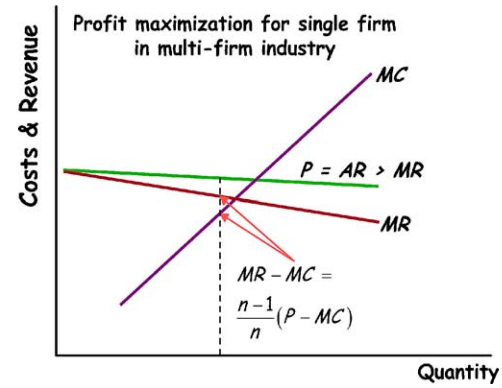

Fig. 6. True profit-maximization level of output.

There is a simple but deceptive aggregation error in the conventional belief. Only for a monopoly is "marginal revenue" entirely the result of the output changes of a single firm. In a multi-firm industry, changes in a firm's revenue are caused not only by changes in its own output, but also by changes in the output of all other firms. In a multi-firm industry, a firm profit maximizes not by equating its marginal revenue and marginal cost, but by choosing an output level at which its own-output marginal revenue *exceeds* its marginal cost.[7]

The true profit-maximizing formula is (see Appendix A):

$$\mathrm{MR}_i(q_i) - \mathrm{MC}_i(q_i) = \frac{n-1}{n}(P(Q) - \mathrm{MC}_i(q_i)) \quad (4)$$

where q_i is the output of the ith firm and n is the number of firms in the industry. Aggregation of this true profit-maximizing position results in an industry-level output that (depending on the nature of marginal cost curves) is identical to the monopoly level of output and independent of the number of firms in the industry.[8] The aggregate output position for an industry is thus as shown in Fig. 1,[9] while Fig. 6 shows the profit-maximizing position for each firm in a multi-firm industry. Curiously, for an academic discipline that has been obsessed with the intersection of curves, this profit-maximizing level of output occurs not where the curves intersect, but where a gap exists between them.

This puts into sharp relief the absurdity of forcing a firm to sell its output at a price equal to its marginal cost. Even under standard neoclassical assumptions about the shape of demand and marginal cost functions, marginal revenue is well below marginal cost at

[7] The full proof of this proposition is given in Appendix A (and more fully in Keen et al., 2004).

[8] Stigler attempted to evade the implications of his proof that $(\mathrm{d}P/\mathrm{d}q) = (\mathrm{d}P/\mathrm{d}Q)$ by reworking the expression for marginal revenue in terms of the market elasticity of demand and the number of firms. Appendix A shows that the number of firms is irrelevant to the profit-maximizing position.

[9] With some nuances that are discussed next.

this point, so the firm is being forced to sell part of its output at a loss. The loss is even larger than economists might anticipate, because losses (incremental output being sold for less than its cost of production) begin not where own-output marginal revenue and marginal cost intersect, but substantially prior to this point. The US$ 1.49 billion loss that the unamended Decision 91107 would have forced upon PG&E was obviously unintentional, but it was nonetheless no accident.

The most significant consequence for the so-called competition policy is that a competitive market is not inherently superior to a monopoly. Since the profit-maximizing behaviour of competitive markets is in the aggregate identical to that of a monopoly—in that the profit-maximizing equilibrium occurs where aggregate marginal cost equals marginal revenue—then there is no welfare difference between competition and monopoly. Both will behave like the right hand side of Fig. 3. The ‘‘deadweight loss’’ of consumer and producer surplus that has in the past been ascribed simply to monopoly is instead the deadweight loss from profit-maximizing behaviour.

However, this loss is likely to be *greater* for a more competitive market than for a less competitive one, because there are theoretical and practical reasons why larger firms will have lower marginal costs than smaller firms. With lower marginal costs and exactly the same aggregate price setting behaviour, monopoly is welfare-superior to competition.[10]

The conventional theory of the firm is thus a shambles. Its ‘‘ideal’’ price level of marginal cost, which regulators like the CPUC imposed upon regulated industries, forced real-world corporations to produce a substantial proportion of their output at a loss. Its ideal market structure of many competitive unregulated firms results not in the ‘‘welfare-maximizing’’ equivalence of marginal cost and price, but to marginal cost significantly exceeding marginal revenue. The marginal cost fetish thus bankrupted real firms, driving prices higher and consumer welfare lower.

This was not the end of the damage. This critique has so far accepted that the deregulated market price would be an equilibrium one, but a substantial body of research indicates that competitive spot market prices are unstable.

[10] At least in terms of the neoclassical theory of markets. Other reasons to distinguish competitive markets from monopolies can be preferred—the impact of competition upon markups over cost as an argument for competition; the impact of economies of scale on cost as an argument for monopoly. But these lack the definitive bias of uncorrected neoclassical theory in favour of competitive markets: whether one situation is superior is now a question of empirical analysis rather than definitive theory.

3. The instability of spot markets

Ever since Walras, the economist’s Nirvana has been a world of free markets: a place of complete tranquility and harmony where supply and demand alone determine prices, and prices in all markets are in complete equilibrium. But how do you get from here to Nirvana? How do you get from an initial set of prices that involve disequilibrium, to a set that simultaneously clears all markets?

Walras himself believed that the journey was feasible, though he failed to prove it, and he was conscious of the practical difficulties that both production and out-of-equilibrium trades would cause. He instead imagined a pure exchange economy where prices were coordinated by an ‘‘auctioneer’’ who did not permit trading until such time as all markets cleared. The auctioneer declared an initial set of prices (which could be the equilibrium set only by a miracle) and then orchestrated a process of ‘‘tatonnement’’, or ‘‘groping’’, adjusting up the price of commodities where demand exceeded supply (and vice-versa).

Walras presumed this process would converge to equilibrium because the ‘‘direct’’ effects of reducing the price of a commodity whose supply exceeded demand would necessarily push the market towards equilibrium, while the indirect effects of this market on all others would cause some to move closer to equilibrium and others further away (Walras, 1874). On balance, Walras expected this tatonnement process to iterate prices towards a general equilibrium.

Utility reformers clearly shared Walras’ faith in the dynamic stability of free spot markets. Unfortunately, Walras’ faith was misplaced—but he was not in a position to know any better. One cannot be so charitable about modern-day ‘‘reformers’’ whose naive faith in the stability of spot markets finds little support in the literature. The one defence that spot markets advocates could mount against a charge of gross negligence on this issue is that the literature on the stability of general equilibrium is neither conclusive nor erudite.[11]

That said, even a modicum of an understanding of the theoretical literature should have caused policy economists to be extremely wary of spot market pricing for electricity. Theorists who have considered the stability of Walras’ tatonnement process have found that it

[11] The key reference in general equilibrium theory—Debreu’s (1959) *Theory of Value*—completely ignores dynamic stability with such absurd contrivances as the proposition that “a production plan (made now for the whole future) is a specification of the quantities of all his [a producer’s] inputs and all his outputs… The certainty assumption implies that he knows now what input–output combinations will be possible in the future (although he may not know the details of technical processes which will make them possible)…” Even Walras’ highly stylised market was far more realistic than this.

is unstable under quite plausible conditions upon endowments and tastes. Hurwicz observes that:

> From static analysis (going back to Walras and Marshall), it is known that, even under very plausible circumstances [Walrasian tatonnement] systems ... have multiple equilibria... Hence, it is not to be expected that, in a reasonably broad class of economic environments (i.e., here, aggregate excess demand functions) every equilibrium point of a Walrasian tatonnement process will be stable (Hurwicz, 1986: pp. 46–47).

The obvious implication is that, if Walras' highly abstract tatonnement process is unstable under plausible conditions, then real-world spot markets where both production and out-of-equilibrium trades do occur must also be unstable.[12] Real-world prices therefore cannot be equilibrium, spot market-clearing prices.

This implication can be made far more concrete by considering the dynamics of a hypothetical production system with spot prices.[13] Imagine a simple world in which there are just two commodities—corn and iron—and where there is no final demand, so that each year's total outputs of corn and iron become the inputs into the next year's production process.[14] This pure "supply side" economy removes any complications that might arise from variations in demand (or anything else for that matter), so it should be the simplest world of all to manage. If a spot market is going to work smoothly anywhere, it should work smoothly here.

Each commodity is a necessary input in the production of both commodities (iron is needed to make agricultural implements, while corn is needed to fire furnaces, feed workers, etc.). In this hypothetical world, it takes 9/10 of a quarter of corn and 1/20th of a tonne of iron to produce 1 quarter of corn, and 3/5 quarters of corn plus 1/5th of a tonne iron to produce 1 tonne of iron. This economy can grow stably at a growth rate of 6.325% per annum if the ratio of corn output (in quarters) to iron output (in tonnes) is 14.81[15] (see Fig. 7).

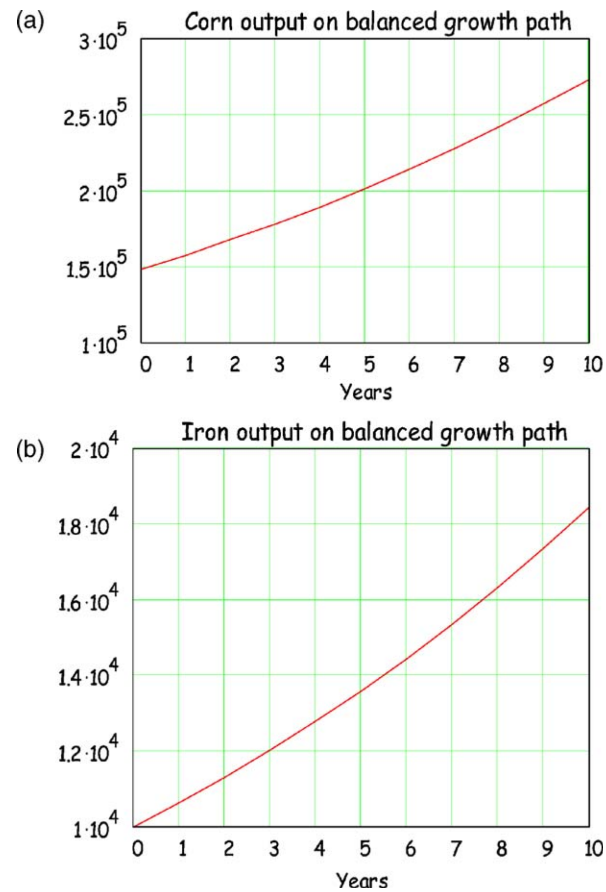

Fig. 7. Equilibrium growth path.

If spot market prices are going to support this stable growth path, then the price times quantity of corn sold by corn producers to iron producers must equal the price times quantity of iron sold by iron producers to corn producers (Walras's law). If the economy starts off with the ratio of corn output to iron output of 14.81, then this is indeed what happens: the price ratio of corn to iron remains stable at 1.23 (see Fig. 8a): using iron as the numeraire, one quarter of corn can be purchased with 1.23 tonnes of iron.[16] This may seem perverse—a quarter of corn costs more than a tonne of iron even though many more quarters of corn are produced than tonnes of iron. But this is because much more corn is used up in the production process than

[12] Unfortunately, in a tendency that is all too rife in theoretical economics, Hurwicz instead concluded that since tatonnement as Walras envisaged it was likely to be unstable, Walras' auctioneer should be advised to use another adjustment process! "From a normative and computational point of view it is natural to conclude that the possible absence of global stability calls for replacing the Walrasian tatonnement by another dynamic process" (Hurwicz, 1986: p. 47). He proposed one based on adjusting the excess demand functions directly rather than simply adjusting quantities. He lamented that "Clearly the informational burden of this system is greater than that of [Walras]," but stoically concluded that "one must, in general, be prepared to require a bigger message space when stability is demanded" (Hurwicz, 1986: p. 48). Similar unreal deductions can be found in related literature (such as Hands, 1983: pp. 399–411).

[13] This section paraphrases the arguments provided in Blatt (1983: pp. 111–146). Appendix B supplements the numerical example given here with a general dynamic argument.

[14] This does not rule out consumption, since you can imagine that the consumption needs of workers and capitalists have been collapsed into the input–output matrix.

[15] To two decimal places; the ratio to 15 decimal places is 14.8102496759067.

[16] To two decimal places; to 15 it is 1.23418747299222.

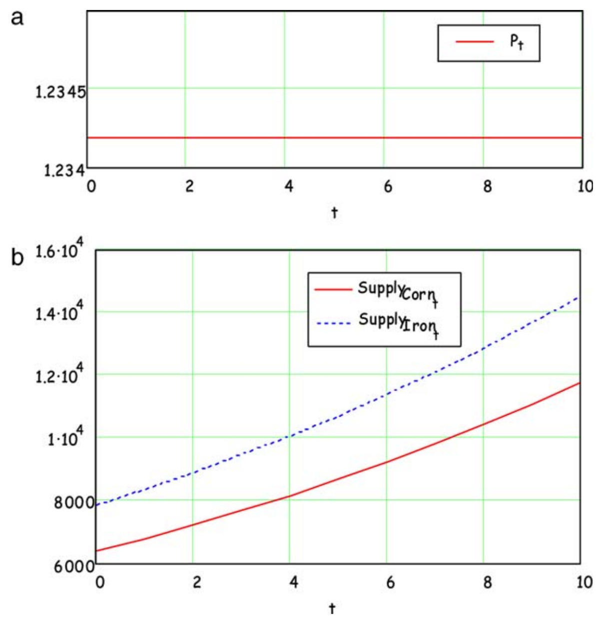

Fig. 8. Equilibrium price and supply to spot markets.

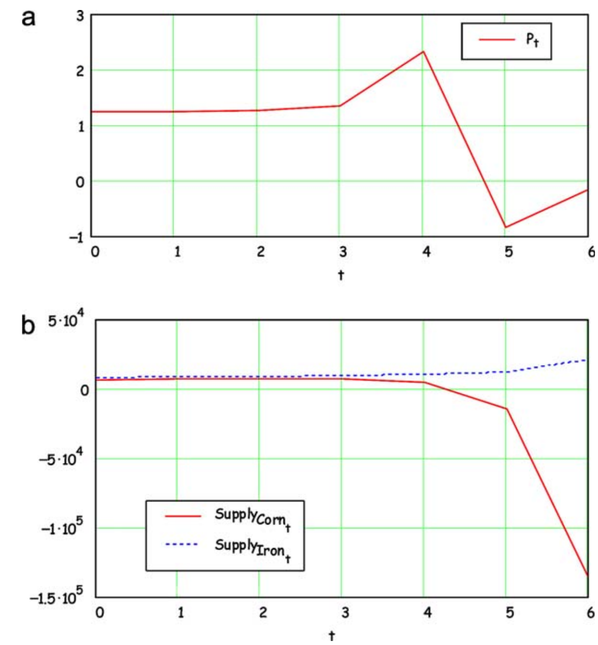

Fig. 9. Price and supply to spot market dynamics away from equilibrium.

iron: thus even though the output of corn exceeds the output of iron, the amount of iron supplied to the market exceeds the amount of corn (see Fig. 8b).

So if this toy economy starts in equilibrium, there it remains forever. But what if it starts a slight distance away from equilibrium? Imagine that corn output starts just 0.1% above the ideal ratio: will the spot market price system reduce the price of corn and result in the output ratio returning to the equilibrium level?

Fig. 9 says no: even though corn output exceeds the equilibrium level, the spot market dynamics result in the price of corn *rising*, not falling. The output of corn continues to rise and iron output falls until, bizarrely, iron output becomes negative! What is going on? Appendix B provides the full technical explanation, but colloquially, the vagaries of the input–output system mean that an excess production of corn results in a *fall* in the amount of corn supplied to the market. This fall in supply to market then sets off a price rise in the spot market for corn, which drives the economy away from equilibrium, not towards it.

This may seem perhaps to be the product of the particular example used, but as Blatt (1983: pp. 117–146) explains (also see Appendix B to this paper), the instability of the equilibrium growth path is a general property that must apply to any production system that allows growth to occur. Nor can it be blamed either on the input–output model itself, as opposed to a system with flexible production ratios. It is elementary mathematics that an input–output model is both a component of any more flexible production system, and the component that determines the stability or otherwise of the equilibrium. The other, non-linear components of a more general production function may constrain the instability away from equilibrium, but the equilibrium itself remains unstable (see Appendix B).

Therefore, since we live in an economy that can and (most of the time) does grow, this is probably a property of the actual production system in which we live, and the additional aspects of the real world that this model omits are unlikely to correct this fundamental instability.

A system of spot markets could therefore be expected to display continuous disequilibrium. But the real world is *not* a system of spot markets: despite economists' proclivity to model the world as if supply and demand are always in equilibrium (and therefore spot prices rule), in the vast majority of real-world markets (everything from ice creams to automobiles), firms maintain a substantial stock of unsold commodities which they adjust according to demand.

What happens, however, if economic policy forces one market to use spot prices? If this is a non-essential commodity (say, ice cream), then probably little of import: ice cream prices would undoubtedly fluctuate much more than they currently do, but there would be little impact on the rest of the economy.

But if the commodity is one that is essential for the production of almost everything—as is electricity—then quite conceivably, the inherent instability of a production process with spot prices could manifest itself in the price behaviour of this one commodity, with potentially calamitous consequences, not just for this one market, but for all others into which the commodity is an input.

Arguably this is part of what happened with electricity. Ignoring all other sources of price instability (such as the incredible amount of re-selling of electricity in the spot market, the outright price rigging of Enron etc.), the input–output nature of production alone, combined with the essential nature of electricity, means that price instability should have been expected.

Instead, economists seduced by a belief in equilibrium expected price stability, and even after the California experience, they call for further reliance upon competitive market prices (Ad-hoc group, 2003). Informed economic analysis recommends precisely the opposite.

The theoretical implication of this research is that economics should develop the ability to analyze disequilibrium systems; the practical implication is that, if limits on instability are desired, then spot prices should be discouraged in favour of buffer pricing, stock adjustment pricing, administered pricing, and so on. Fortunately, as extensive research has established, these are the pricing mechanisms that real-world businessmen actually implement.

4. Costs in the real world

One essential assumption in the conventional theory is that firms experience rising marginal costs because of diminishing marginal productivity. The first section of the paper shows that, even given this assumption, the theory is an empty shambles: all industry structures set price well above marginal cost, the profit-maximizing level of output occurs where own-output marginal revenue exceeds marginal cost, and, since a supply curve cannot be derived, demand and supply analysis is untenable.

However, the theory is even emptier than this, because extensive empirical research[17] has established that the vast majority of firms do not produce under conditions of diminishing marginal productivity. Instead, the typical modern firm experiences constant or falling average variable costs (and of course falling average fixed costs): for at least 95% of firms, the "U-shaped cost curve" that dominates economic thinking about costs is false. Instead, firms experience falling average costs of production as output rises, and, after a "breakeven" volume of production is reached, each additional sale adds to profit.

One of the most interesting empirical papers (Eiteman and Guthrie, 1952: pp. 832–838) asked firms

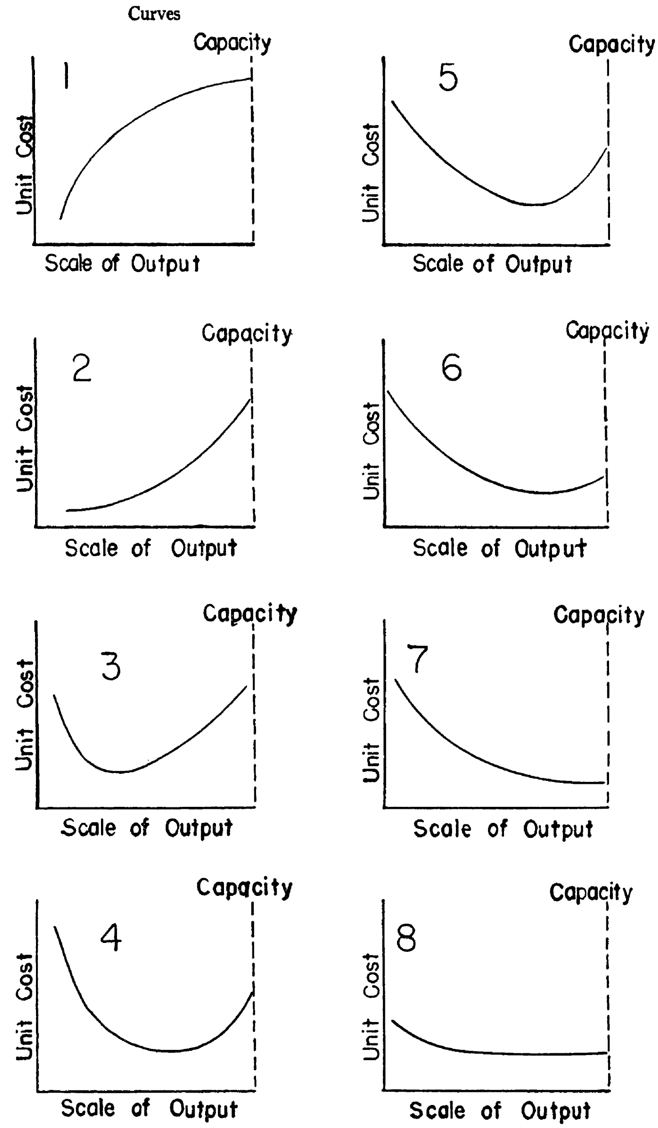

Fig. 10. Eiteman and Guthrie (1952): cost functions shown to managers.

to choose which of a set of eight graphs (see Fig. 10) most closely resembled their cost functions.

Eiteman and Guthrie classified diagrams 3–5 as consistent with neoclassical theory, and diagrams 6–8 as inconsistent with it one thousand questionnaires were sent to firms with between 500 and 5000 employees. The 334 responses they received where the respondents made no distinction between the different products they manufactured[18] are shown in Table 1. Applying their classification, just under 95% of firms chose a curve

[17] Downward (1994: pp. 23–43) cites over 120 surveys; Lee (1998) gives a historical overview of research into actual corporate pricing; Blinder et al. (1998) is the most recent such study.

[18] 32 other respondents occasionally chose different diagrams for different products in their catalog, but the aggregate picture remained the same: 95% of products are not produced under conditions of rising marginal cost.

Table 1
Eiteman and Guthrie (1952): cost functions as seen by managers

Curve indicated	Number of companies
1	0
2	0
3	1
4	3
5	14
6	113
7	203
8	2
Total	336

that contradicted neoclassical expectations of diminishing marginal productivity.[19]

Eiteman and Guthrie's analysis has recently been reconfirmed by Blinder et al. (1998). Downward and Lee (2001) provide a useful summary of the key results: "over 89% of respondents indicated that 'marginal' costs either declined or stayed constant with changes in output (sometimes involving discrete jumps)" (Downward and Lee, 2001: p. 469).

This extensive research has been ignored by most economists because it contradicted what appeared to be a watertight theory, and seemed counter-intuitive—"If firms do not experience rising costs, then what stops them from producing an infinite amount?"

In fact, what appears counter-intuitive to economists is a product of belief in an invalid theory. The assumptions of conventional theory omit the very aspects of the real world that constrain the output levels of real firms.

The neoclassical model assumes homogeneous products where producers compete only on price, and where price is the sole issue of relevance to consumers. But in the real world, both products and consumers are heterogeneous, and price is only one of a range of indicators that consumers consider when deciding between one firm's product and another.

As Sraffa (1926: pp. 548–560) argued, the heterogeneous nature of output and consumption puts limits on the capacity that an individual firm will build. Kornai's theory of "demand constrained" output (Kornai, 1990) explains that firms in a market economy will operate within their installed capacity, because without spare capacity they will be unable to

[19] Diagram 6 may appear consistent with neoclassicism, but Eiteman and Guthrie's argument is that for real firms, costs fall until very near capacity output is reached, and this is what curve 6 indicates. Their textual explanation of diagrams 5–7 in their survey form indicates this: "5. If you choose this curve you believe that unit costs are high at minimum output, that they decline gradually to a least-cost point near capacity, after which they rise sharply. 6. If you choose this curve you believe that unit costs are high at minimum output, that they decline gradually to a least-cost point near capacity, after which they rise slightly; 7. If you choose this curve you believe that unit costs are high at minimum output, that they decline gradually to capacity at which point they are lowest." (Eiteman and Guthrie, 1952: p. 835)

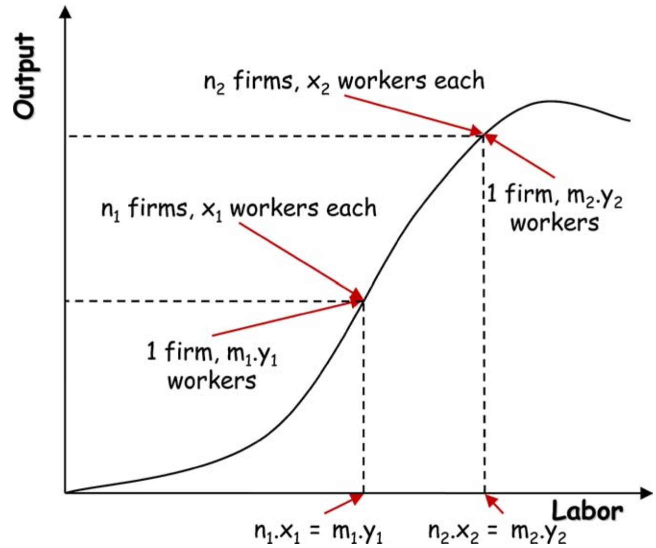

Fig. 11. Production functions required for identical marginal cost curves.

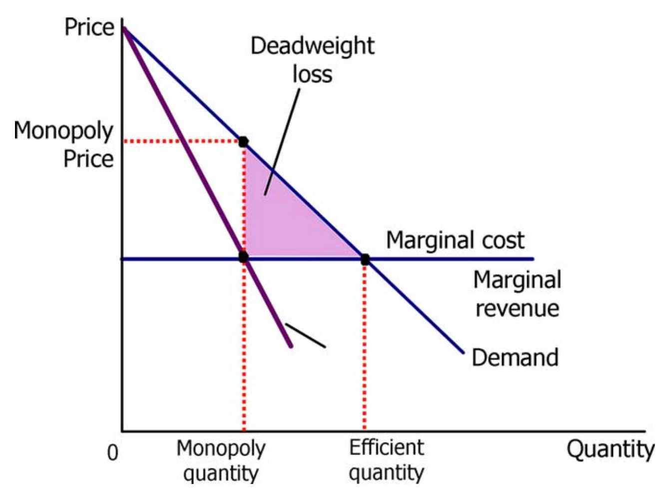

Fig. 12. Only marginal cost curve for definitive welfare comparison with $n > m$.

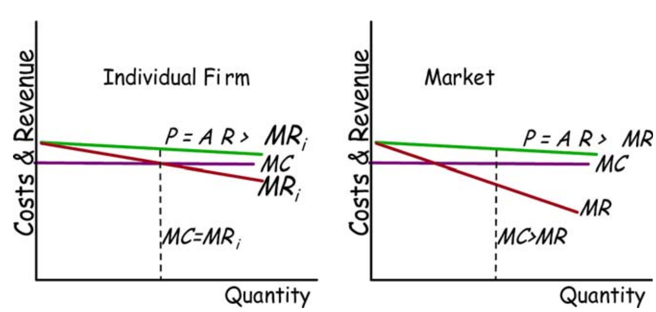

Fig. 13. Aggregation of MR = MC leads to market output where MC > MR.

take account of opportunities that may arise in the market but cannot be anticipated (such as the safety recall of Firestone tyres in 2000 that proved such a boon to other tyre manufacturers). Eiteman (1947: pp. 910–918, 1948: pp. 899–904) explains that engineers design modern factories to produce at near maximum efficiency right up to the point of full capacity—and argues that the neoclassical concept of diminishing marginal productivity was derived from observation of the now largely defunct farming practices of the 19th century.

These real-world cost structures mean that firms will suffer substantial losses if they are forced to set price equal to marginal cost. With high fixed costs and constant or falling marginal costs, average costs lie well above marginal cost, and a marginal cost regime will force firms to absorb unrecoverable losses on their fixed costs. P&G's $ 1.5 billion catastrophe (Conkling, 1999: p. 25), though an unintended consequence of the imposition of marginal cost pricing, was an entirely predictable outcome to anyone acquainted with the real world.

5. Conclusion

Neoclassical microeconomics is conventionally regarded as a "pro-market" theory, whereas Marxist economics (for example) is "anti-market". But the pro-market perspective of neoclassical economics is ideological only: While it is a useful weapon with which to assert the superiority of the market over central planning, it is a dangerous notion to apply to actual economies and firms.

Firstly, even on its own terms, the theory mis-specifies the point of profit maximization for the individual firm. Secondly, it falsely presumes that competition can force profit-maximizing firms to produce where marginal cost equals price. Thirdly, it assumes equilibrium market-clearing spot prices will apply, when mathematically, spot prices are unstable in a multi-commodity model of production with growth. Fourthly, it believes that marginal cost pricing is compatible with profit maximization, when given real-world cost structures, marginal cost is well below average cost.

Neoclassical economics is thus a totally inappropriate tool to use to decide how real-world prices should be set, especially in so crucial a market as electricity. Pricing of electricity should return to the historic practices of the industry, while policy should return to the oversight on adequate profitability and cost recovery, and the maintenance of reliable supply with low price volatility.

Appendix A. Fallacies in the theory of firms[20]

A.1. Constant marginal cost

Marginal cost is the inverse of marginal product, which in turn is the derivative of total product. The condition of identical marginal costs[21] therefore requires that total products differ only by a constant, which can be set to zero if 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 Fig. 11.

Using f for the production function of the competitive firms, and g for the production function of the monopoly, the condition can be put in the form shown by equation (A.1):

$$n \cdot f(x) = m \cdot g(y) \qquad \text{(A.1)}$$

Substitute $y = (n \cdot x)/m$ into (A.1) and differentiate both sides of (A.1) by n:

$$f(x) = \frac{x}{m} \cdot g'\left(\frac{n \cdot x}{m}\right) \qquad \text{(A.2)}$$

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

$$\frac{g(n \cdot x/m)}{n} = \frac{x}{m} \cdot g'\left(\frac{n \cdot x}{m}\right) \text{ or } \frac{g'(n \cdot x/m)}{g(n \cdot x/m)} = \frac{m}{n \cdot x} \qquad \text{(A.3)}$$

The substitution of $y = (n \cdot x)/m$ yields an expression involving the differential of the log of g:

$$\frac{g'(y)}{g(y)} = \frac{1}{y} \qquad \text{(A.4)}$$

Integrating both sides yields:

$$\ln(g(y)) = \ln(y) + C \qquad \text{(A.5)}$$

Thus, g is a constant returns production function:

$$g(y) = C \cdot y \qquad \text{(A.6)}$$

It follows from (A.6) that f is the *same* constant returns production function:

$$f(x) = \frac{m}{n} \cdot C \cdot \frac{n \cdot x}{m} \qquad \text{(A.7)}$$

With both f and g being identical constant returns production functions, it follows that the marginal products and hence the marginal costs of the competitive industry and monopoly are constant and identical. The gen-

[20] Some of the argument in this paper is necessarily technical. These appendices contain the detail needed to support the generally verbal arguments of the paper.

[21] i.e., the marginal cost curve for the monopoly being identically equal to the sum of the marginal cost curves of the competitive firms for all relevant levels of output.

eral 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 (as illustrated by Fig. 12).

The only exception to this occurs where $n = m$ and therefore $x = y$, in which case the condition collapses to $f(x) = g(x)$, which can be fulfilled by any production function—including one displaying diminishing marginal productivity. However, in general, diminishing marginal productivity is incompatible with a definitive comparison of the welfare effects of monopoly and perfect competition.

We can now use this condition for welfare comparability to identify the next fallacy, that equating marginal cost and marginal revenue is not profit-maximizing behavior (except for a monopoly). First, we recap and elaborate upon Stigler's proof that the demand curve for a competitive firm has the same slope as the market demand curve.

A.2. Profit-maximizing behavior

The assumption that $(\mathrm{d}P/\mathrm{d}q_i) = 0$ while $(\mathrm{d}P/\mathrm{d}Q) < 0$ is easily invalidated using the Chain Rule, as Stigler did in 1957:

$$\frac{\mathrm{d}P}{\mathrm{d}q_i} = \frac{\mathrm{d}P}{\mathrm{d}Q}\frac{\mathrm{d}Q}{\mathrm{d}q_i} = \frac{\mathrm{d}P}{\mathrm{d}Q} \tag{A.8}$$

Elaborating upon this, the relation $(\mathrm{d}Q/\mathrm{d}q_i) = 1$ is a simple consequence of the proposition that firms are independent:

$$\begin{aligned}\frac{\mathrm{d}Q}{\mathrm{d}q_i} &= \frac{\mathrm{d}}{\mathrm{d}q_i}\left(\sum_{j=1}^{n} q_j\right)\\ &= \frac{\mathrm{d}}{\mathrm{d}q_i}\left(\sum_{j=1}^{n}(q_1 + q_2 + \ldots + q_i + q_n)\right)\\ &= \left(\sum_{j=1}^{n}\left(\frac{\mathrm{d}}{\mathrm{d}q_i}q_1 + \frac{\mathrm{d}}{\mathrm{d}q_i}q_2 + \ldots + \frac{\mathrm{d}}{\mathrm{d}q_i}q_i + \frac{\mathrm{d}}{\mathrm{d}q_i}q_n\right)\right)\\ &= 1\end{aligned} \tag{A.9}$$

Stigler's relation $(\mathrm{d}P/\mathrm{d}q_i) = (\mathrm{d}P/\mathrm{d}Q)$ can now be used to establish that equating own-output marginal cost to marginal revenue is *not* a profit-maximizing strategy in a multi-firm industry. The accepted formula is only true for a monopoly; in a multi-firm industry, a firm's total revenue is a function not only of its own behavior, but also the behavior of all the other firms in the industry:

$$\mathrm{TR}_i = \mathrm{TR}_i\left(\sum_{j\neq i}^{n} q_j, q_i\right) \tag{A.10}$$

Defining Q_R as the output of the rest of the industry ($Q_\mathrm{R} = \sum_{j\neq i}^{n} q_j$), a change in revenue for the ith firm is properly defined as:

$$\mathrm{dTR}_i(Q_\mathrm{R}, q_i) = \left(\frac{\partial}{\partial Q_\mathrm{R}}P(Q)\cdot q_i\right)\mathrm{d}Q_\mathrm{R} + \left(\frac{\partial}{\partial q_i}P(Q)\cdot q_i\right)\mathrm{d}q_i \tag{A.11}$$

The accepted formula ignores the effect of the first term on the firm's profit. However, this can easily be calculated by considering what would happen if all firms did in fact equate their own-output marginal cost and marginal revenue. Then, we would have equation (A.12):

$$\sum_{i=1}^{n}\left(\frac{\mathrm{d}}{\mathrm{d}q_i}(P(Q)\cdot q_i - \mathrm{TC}_i(q_i))\right) = 0 \tag{A.12}$$

where $\mathrm{TC}_i(q_i)$ is the total cost function for the ith firm. Expanding this relation using Stigler's relation and the condition that marginal costs (MC) are identical and constant yields the relation:

$$(n-1)P(Q) + \mathrm{MR}(Q) - n\cdot\mathrm{MC} = 0 \tag{A.13}$$

Equation (A.14) shows the workings in full:

$$\begin{aligned}&\sum_{i=1}^{n}\left(\frac{\mathrm{d}}{\mathrm{d}q_i}(P(Q)\times q_i - \mathrm{TC}_i(q_i))\right)\\ &= \sum_{i=1}^{n}\left(P(Q) + q_i\cdot\frac{\mathrm{d}}{\mathrm{d}q_i}P(Q)\right) - \sum_{i=1}^{n}\left(\frac{\mathrm{d}}{\mathrm{d}q_i}\mathrm{TC}_i(q_i)\right)\\ &= n\cdot P(Q) + \sum_{i=1}^{n}\left(q_i\cdot\frac{\mathrm{d}}{\mathrm{d}Q}P(Q)\right) - \sum_{i=1}^{n}(\mathrm{MC})\\ &= n\cdot P(Q) + \frac{\mathrm{d}}{\mathrm{d}Q}P(Q)\cdot\sum_{i=1}^{n}q_i - n\cdot\mathrm{MC}\\ &= (n-1)\cdot P(Q) + \left(P(Q) + Q\cdot\frac{\mathrm{d}}{\mathrm{d}Q}P\right) - n\cdot\mathrm{MC}\\ &= (n-1)P(Q) + \mathrm{MR}(Q) - n\cdot\mathrm{MC} = 0\end{aligned} \tag{A.14}$$

Equation (A.14) can be rearranged to yield

$$\mathrm{MR}(Q) - \mathrm{MC} = -(n-1)(P(Q) - \mathrm{MC}) \tag{A.15}$$

Since $n-1$ exceeds 1 in all industry structures except monopoly, and price exceeds marginal cost in all industry structures (except, allegedly, perfect competition), the RHS of (A.15) is negative. Therefore, if all firms equate their marginal cost to their own-output marginal revenue, aggregate marginal cost exceeds marginal revenue because aggregate marginal revenue is less than the individual firm's marginal revenue. Graphically, the situation is as shown in Fig. 13: the equating of own-output marginal revenue (MR_I for the ith firm) to marginal cost results in aggregate marginal cost exceeding aggregate marginal revenue.

The mathematics behind this aggregation problem is shown in equation (A.16) (where industry marginal revenue is MR and the firm's marginal revenue is $MR_i(q_i)$):

$$\begin{aligned}\sum_{i=1}^{n}(MR_i(q_i)) &= \sum_{i=1}^{n}\left(\frac{d}{dq_i}(P\cdot q_i)\right)\\ &= \sum_{i=1}^{n}\left(P\cdot\frac{d}{dq_i}(q_i)+q_i\cdot\frac{d}{dq_i}P\right)\\ &= \sum_{i=1}^{n}\left(P+q_i\cdot\frac{d}{dQ}P\right)=n\cdot P+Q\cdot\frac{d}{dQ}P\\ &= (n-1)\cdot P+MR > MR \text{ for } n>1 \end{aligned} \tag{A.16}$$

Thus, equating own-output marginal revenue to marginal cost is clearly not profit-maximizing behavior! What is? This can be derived from equation (A.15): if equating MR_i and MC results in the aggregate loss shown there for each of n identical firms, then each firm should produce where the *gap* between their own-output marginal revenue and marginal cost equals $1/n^{th}$ of this loss. The profit-maximizing output level is thus where the gap between own-output marginal revenue and marginal cost equals $(n-1)/n$ times the gap between market price and marginal cost:

$$MR_i(q_i)-MC=\frac{n-1}{n}\cdot(P-MC) \tag{A.17}$$

Equation (A.18) shows that this leads to a profit-maximizing level of output for the industry and hence for the firms in it:

$$\sum_{i=1}^{n}(MR_i(q_i)-MC)=\sum_{i=1}^{n}\left(\frac{n-1}{n}\cdot(P-MC)\right) \tag{A.18}$$

The LHS of (A.18) sums to $(n-1)\cdot P(Q)+MR(Q)-n\cdot MC$(see equation A.14). Summing the RHS yields:

$$\begin{aligned}\sum_{i=1}^{n}\left(\frac{n-1}{n}\cdot(P-MC)\right)&=\frac{n-1}{n}\cdot\sum_{i=1}^{n}(P-MC)\\ &=\frac{n-1}{n}\cdot(n\cdot P-n\cdot MC)=(n-1)\cdot(P-MC)\end{aligned} \tag{A.19}$$

Equating the LHS and RHS of (A.18) thus yields:

$$\begin{aligned}&(n-1)P(Q)+MR(Q)-n\cdot MC\\ &\quad=(n-1)\cdot(P-MC)\end{aligned} \tag{A.20}$$

Equation (A.20) can be simplified to yield:

$$MR(Q)=MC \tag{A.21}$$

This establishes that this output selection strategy by each of n identical firms leads the profit level of each of the n firms being maximized, in which case total industry profit is also maximized. Keen et al. (2004, 7–12) show that this result holds for an n-firm industry facing a linear demand curve, and that in all cases, industry output corresponds to the "monopoly" level regardless of the number of firms.

These results indicate why Stigler's reworking of the marginal revenue of the ith firm in an industry with n identical firms to:

$$MR_i=P+\frac{P}{n\cdot E} \tag{A.22}$$

(where E is industry elasticity of demand, $E=(dQ/dP)\cdot(P/Q)$) is correct but irrelevant. Though MR_i can be made to approach P by increasing n, this is exactly countered by the point of profit maximization being not where $MR_i=MC$, but where there is a gap between these functions that is a function of n (Equation A.17). When this profit-maximizing level is calculated, it results in an output level that is independent of n, so that all industry structures produce the so-called monopoly output level.

The general profit-maximizing formula in the case of differing non-constant marginal costs[22] is:

$$\begin{aligned}q_i&=\frac{1}{n}\frac{P-MC_i(q_i)}{-(dP/dQ)}\\ Q&=\frac{\left(P-1/n\sum_{i=1}^{n}MC_i(q_i)\right)}{-(dP/dQ)}\end{aligned} \tag{A.23}$$

Appendix B. The instability of spot market prices

The model in the paper reproduces the example from Blatt (1983, 114–119). Here, I will start with a simple expression for an input–output system with $\mathbf{x(t)}$ representing the vector of outputs and $\mathbf{A}$ the input–output matrix.[23] Then, the simplest possible linear discrete time model of multi-commodity production with no fixed capital, perfect thrift, and no technical change is:

$$\mathbf{x(t+1)}=\mathbf{A}\cdot\mathbf{x(t)} \tag{B.1}$$

For this system to grow stably over time, there has to be a stable rate of growth α at which all sectors grow:

$$\mathbf{x(t+1)}=(1+\alpha)\cdot\mathbf{x(t)} \tag{B.2}$$

These two equations yield (B.3):

$$\begin{aligned}&(1+\alpha)\cdot\mathbf{x(t)}=\mathbf{A}\cdot\mathbf{x(t)}\\ &(1+\alpha)\cdot\mathbf{x(t)}-\mathbf{A}\cdot\mathbf{x(t)}=0\\ &((1+\alpha)\cdot I-\mathbf{A})\cdot\mathbf{x(t)}=0\end{aligned} \tag{B.3}$$

This is only consistent with non-zero output levels if the determinant of $((1+\alpha)\cdot I-\mathbf{A})$ equals zero:

$$|(1+\alpha)\cdot I-\mathbf{A}|=0 \tag{B.4}$$

This is the crucial relationship that basically determines the results to follow, since the stability of a linear dif-

[22] In this case, the profit-maximizing position varies given the number of firms in the industry, but this reflects variations in the cost functions and not differences in profit-maximizing strategies given the number of firms.

[23] This avoids having to invert and transpose $\mathbf{A}$ immediately as in Blatt's example (Blatt used a numerical example which was easier for a non-mathematical reader to follow).

ference equation is determined by the dominant eigenvalue of the matrix (the largest root of the polynomial $|(1+\alpha)\cdot I - \mathbf{A}| = 0$). If this dominant eigenvalue exceeds zero (for a continuous time system) or one (for a discrete time system, such as this example), then the equilibrium of the system will be unstable.

This cannot be evaded by considering a non-linear system, such as (in the case of production) a Cobb–Douglas production function,[24] since any such function can be reduced to a polynomial expansion whose first variable term is an input–output matrix. Though the higher polynomial terms may stabilize an unstable system far from equilibrium, the input–output matrix alone determines the stability of the model close to equilibrium.

The matrix **A** consists of all non-negative entries, and by the Perron–Frobenius theorem, the real part of the dominant eigenvalue of such a matrix exceeds zero. The inverse of this matrix will thus also have a dominant eigenvalue also greater than zero, which is the inverse of **A**'s dominant eigenvalue. One of these will necessarily exceed one,[25] so either **A** or its inverse will have a dominant eigenvalue greater than one. Therefore, any dynamic system involving both **A** and its inverse will necessarily be unstable.

This is the rub for a system of spot-price markets with production: both **A** and its inverse are involved, one in production and the other in price setting; therefore, either quantities or relative prices must be unstable. This "dual [in]stability theorem" was first identified by Jorgenson in 1960, but in predictable neoclassical fashion, he subsequently considered how the system might be made stable by various adjustments—rather than accepting the conclusion that spot market prices must be unstable. His further considerations (Jorgenson, 1961, 1963) involved mathematical errors pointed out by McManus (1963) and confirmed by Blatt (1983).

Continuing with the analysis, the equilibrium relative price vector for this production system with a uniform rate of profit of π will be

$$\mathbf{p} = (1+\pi)\cdot\mathbf{p}\cdot\mathbf{A} \tag{B.5}$$

Manipulating this to get a compact expression for **p** yields (B.6):

$$\begin{aligned}
&\mathbf{p} = (1+\pi)\cdot\mathbf{p}\cdot\mathbf{A}\\
&\mathbf{p}\cdot\mathbf{A}^{-1} = (1+\pi)\cdot\mathbf{p}\cdot\mathbf{A}\cdot\mathbf{A}^{-1}\\
&\mathbf{p}\cdot\mathbf{A}^{-1} - (1+\pi)\cdot\mathbf{p}\cdot I = 0\\
&\mathbf{p}\cdot(\mathbf{A}^{-1} - (1+\pi)\cdot I) = 0
\end{aligned} \tag{B.6}$$

This is only consistent with non-zero prices if the determinant of $(\mathbf{A}^{-1} - (1+\pi)\cdot I)$ is zero:

$$\left|\mathbf{A}^{-1} - (1+\pi)\cdot I\right| = 0 \tag{B.7}$$

This is the "dual instability problem": the stability of output depends on the dominant eigenvalue of **A**, while stability of prices depends on the dominant eigenvalue of $\mathbf{A}^{-1}$. Since both of these must have positive real part, one of these must be greater than one—and hence the part of the system it describes (either relative prices or outputs, and possibly both)[26] must be unstable.

Neoclassical authors (Jorgenson included, in later papers) tried to fudge stability by bringing in all sorts of additional mechanisms. But because the stability of equilibrium itself is determined solely by the linear component of a function, these are irrelevancies: a system of spot markets cannot be in simultaneous equilibrium with demand equal to supply in all markets. Spot markets of the kind forced upon electricity will have unstable prices and/or quantities. Though the reforms did not and could not force all markets to be spot markets, the crucial role of electricity—in that it is an input to the production of all commodities including itself—implies that its price and/or output would be destabilized by spot market pricing.

Jorgenson himself made such a deduction in a convoluted way: since the equilibrium of a model of spot market prices is necessarily unstable, other non-market-clearing pricing mechanisms are necessary and probably what actually exists in the real world:

> The conclusion is that excess capacity (or positive profit levels or both) is necessary and not merely sufficient for the interpretation of the dynamic input–output system and its dual as a model of an actual economy (Jorgenson, 1960: p. 893).

Unfortunately, this objective contribution was lost amid later attempts to reclaim the ideological preference for spot markets and market-clearing prices.

References

Ad-hoc group of concerned professors, former public officials, and consultants. January 30, 2003; January 26, 2001. Manifesto on the

[24] The Cobb–Douglas production function, much beloved of neoclassical economists (see for example Mankiw, 1995), is another content-free pseudo-concept. As Shaikh (1974) and others have shown, the Cobb–Douglas "production function" is simply a transformation of the national income identity *Income* = *Wages*+*Profits* under conditions of relatively constant income shares. Its "impressive correlation" with economic growth data occurs because it is a correlation of x with approximately x.

[25] The eigenvalues of the example **A** are 0.941 and 0.159; its inverse $\mathbf{A}^{-1}$ has eigenvalues 1.063 and 6.27.

[26] A non-dominant eigenvalue of **A** or its inverse could also be greater than one.

California Electricity Crisis. Institute of Management, Innovation and Organization. Berkeley: University of California.
Axtell, R.L., 2001. Zipf distribution of U.S. firm sizes. Science 293, 1818–1820.
Besomi, D., 1999. The Making of Harrod's Dynamics. St Martin's Press, New York.
Blatt, J.M., 1983. Dynamic Economic Systems. ME Sharpe, Armonk, New York.
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 17, 20–32.
Debreu, G., 1959. Theory of Value: An Axiomatic Analysis of Economic Equilibrium. Yale University Press, New Haven.
Downward, P., 1994. A reappraisal of case study evidence on business pricing: Neoclassical and Post Keynesian perspectives. British Review of Economic Issues 16 (39), 23–43.
Downward, P., Lee, F.S., 2001. Post Keynesian pricing theory 'reconfirmed'? A critical review of asking about prices. Journal of Post Keynesian Economics 23, 465–483.
Eiteman, W.J., 1947. Factors determining the location of the least cost point. American Economic Review 37, 910–918.
Eiteman, W.J., 1948. The least cost point, capacity and marginal analysis: a rejoinder. American Economic Review 38, 899–904.
Eiteman, W.J., Guthrie, G.E., 1952. The shape of the average cost curve. American Economic Review 42, 832–838.
Hands, D.W., 1983. Stability in a discrete time model of the Walrasian tatonnement. Journal of Economic Dynamics and Control 6, 399–411.
Hogan, W.W., 2001. Electricity market restructuring: reforms of reforms. In: 20th Annual Conference, Center for Research in Regulated Industries, May 23–25, Rutger University.
Hurwicz, L., 1986. On the stability of the tatonnement approach to competitive equilibrium. In: Sonnenshein, H.F. (Ed.), Lecture Notes in Economics and Mathematical Systems. Springer-Verlag, Berlin.
Jorgenson, D.W., 1960. A dual stability theorem. Econometrica 28, 892–899.
Jorgenson, D.W., 1961. Stability of a dynamic input–output system. Review of Economic Studies 28, 105–116.
Jorgenson, D.W., 1963. Stability of a dynamic input–output system: a reply. Review of Economic Studies 30, 148–149.
Keen, S., 2001. Debunking Economics: The Naked Emperor of the Social Sciences. Pluto Press & Zed Books, Sydney, London.
Keen, S., Legge, J., Fishburn, G., Standish, R., 2004. Aggregation fallacies in the theory of the firm. University of Western Sydney Working Paper. Available from http://www.debunking-economics.com/Papers.
Kornai, J., 1990. Vision and Reality, Market and State. Routledge, New York.
Lee, F., 1998. Post Keynesian Price Theory. Cambridge University Press, Cambridge.
Mankiw, G., 1995. The growth of nations. Brookings Papers on Economic Activity 1, 275–326.
McManus, M., 1963. Notes on Jorgenson's model. Review of Economic Studies 30, 141–147.
Rosput, P.G., 1993. The limits to deregulation of entry and expansion of the US gas pipeline industry. Utilities Policy October, 287–294.
Shaikh, A.M., 1974. Laws of algebra and laws of production: the humbug production function. Review of Economics and Statistics 61, 115–120.
Sraffa, P., 1926. The law of returns under competitive conditions. Economic Journal 36, 535–550.
Stigler, G.J., 1957. Perfect competition, historically considered. Journal of Political Economy 65, 1–17.
Trebing, H.M., 2003. Market failure in public utility industries: an institutionalist critique of deregulation. In: Tool, M.R., Bush, P.D. (Eds.), Institutional Analysis and Economic Policy. Kluwer Academic Publishers, Boston, Dordrecht, London, pp. 287–314.
Walras, L., 1874. Elements of Pure Economics. George Allen & Unwin, London [Further editions: 1900, 1954].

Further reading

Lee, F.S., Downward, P., 1999. Retesting Gardiner Mean's evidence on administered prices. Journal of Economic Issues 33, 861–886.
Sraffa, P., Robertson, D.H., Shove, G.F., 1930. Increasing returns and the representative firm: a symposium. Economic Journal 40, 79–116.