A theory of markets with return-seeking firms

Murray, Cameron

University of Queensland

30 September 2013

A theory of markets with return-seeking firms*

Cameron K Murray †‡

September 30, 2013

Abstract

Neoclassical theory erroneously makes the assumption that firms maximise profits on a fixed endowment of physical capital leading to the pervasive rule of thumb that firms produce at a level of output where marginal revenues equal marginal costs. However this is merely a special case of the general goal of firms maximising returns on all costs. Firms adopting a return-seeking strategy make decisions that are consistent with fundamental assumptions of financial analysis and outperform profit maximising firms.

Introducing time and a measure of incremental capital unit into the model overcomes many limitations with mainstream analysis, particularly in relation to capital investment decisions. This new framework provides a more general model with which to consider market interactions and allows for observable pricing mechanisms, such as mark-up pricing, downward sloping cost curves at the firm level, and ignorance of marginal costs by firm managers. It also reveals that the leap between the positive descriptive model and the normative welfare implications of markets outcomes cannot be bridged by the fundamental welfare theorems.

JEL Classifications: D01, D21, D40

Keywords: Firm behaviour, pricing, return, profit, capital investment

1 Introduction

The neoclassical model of markets is considered the core theory of the economics discipline. So revered is the model that its worship is often parodied by insiders (eg. Figure 1). This defining core supports a discipline tasked with understanding the nature of resource allocation decisions, particularly through markets, and analyse patterns of growth, trade, and monetary and macro-economic variables. Few students of economics progress past this model to more advanced models which can capture more realistic behaviour of markets and most economic analysis of government policy decisions are fundamentally based on this model.

Despite widespread criticism[1] some fundamental problems of the core are routinely overlooked.

1. The market origins of the normal rate of return
2. No scope to understand economic rents (merely the abstract notion of surplus)
3. Firms minimise returns to normal when they maximise profits
4. Economies of scale are typically absent at the optimal output levels

*Accompanying interactive demonstrations of the new model are available at the research website http://buddyexperiment.blogspot.com.au/p/return-seeking-firms.html

†University of Queensland, School of Economics

‡Contact email: ckmurray@gmail.com

[1]For example, Google Scholar produces 54,600 results for the phrase "criticism of neoclassical market theory"

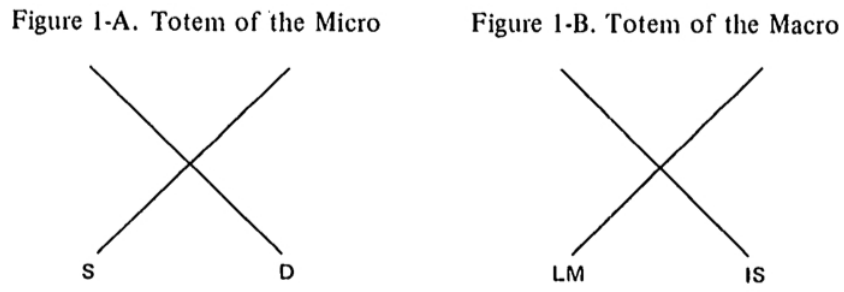

Figure 1: *A parody of the worshipped core models of econ tribes (Leijonhufvud 1973)*

5. Competition is merely a static choice of output level (capital investment decisions are excluded from the model)
6. The normative leap implied by the fundamental welfare theorems usually does not hold

A modification of the firm's decision problem from maximising profit under conditions of fixed capital, to one of maximising returns on all costs, enables a new model to emerge that offers explanations of these identified problems. This modification is compatible with many existing conceptual structures from neoclassical models, such as the market demand curve, cost curves under economies and diseconomies of scale, but shows that supply curves are an emergent phenomena. This new market model of return-seeking firms is detailed in Section 3 following a brief recap of the core market model in Section 2. Some implications for understanding markets, competition, and policy-making are discussed in Section 4, and directions for expansion of this new framework are identified in the concluding remarks.

2 The neoclassical market model

2.1 Defining features

The reader is assumed to be familiar with the basics of the family of core neoclassical market models, and only a brief recap of key features is discussed with reference to the market behaviours they seek to capture.[2]

Output, q, is a convex function of inputs such as labour and capital, $q_S = f(L, K)$, which are assumed to be substitutes at the margin. Demand is an inverse function of price, $q_d = f(p)$, markets clear, $q_d = q_s$, and profits are maximised to 'normal' under competitive conditions, $0 = \Pi = pq - wL - rK$. Labour and capital are paid at their marginal product, $\frac{\partial q_s}{\partial K}$ and $\frac{\partial q_s}{\partial L}$. The model is static in the sense that the even in its *dynamic* versions, output levels and prices are 'pre-solved' by the Walrasian auctioneer for all future periods. This is a defining feature of linear programming methods that removes all noting of determinism or path dependency when used for comparative statics.

In the short-run competitive model firms expand production until the marginal cost is increasing and cease expanding production when the supply curve intersects the marginal revenue curve, which is a horizontal demand curve under price-taking assumptions[3][4]. In the market aggregate the supply curve is the horizontal summation of the individual firm supply curves, and the demand curve takes on its usual downward slope, despite repeated criticisms of the necessary assumptions

[2]The defining reference for the neoclassical theory of markets is Mas-Colell et al. (1995)
[3]Marginal revenue is the derivative of revenue, $qP(q)$, with respect to q.
[4]Apart from the special cases of natural monopoly.

required for this aggregation to be valid (Keen & Standish 2012). The scope of these curves as described by (Mas-Colell et al. 1995) is shown in Figure 2.

While the price-taking assumption could be reasonable for some small scale production in large markets and other special circumstances, it is unlikely to be generally true in major commodity markets that are typically dominated by few large firms - agriculture, metals, fossil fuels, supermarkets, vehicles, construction and so forth are just some examples. More realistically firms face a downward sloping demand curve for their own output, and the Cournot model deals with optimal production decisions in these circumstances. As the number of firms increases, the Cournot model converges to the perfectly competitive outcome.[5] In the neoclassical world, competition is construed as a failure of firms to coordinate to maximise their joint profit - a failure that becomes more prevalent as the number of firms in a market increases.

The core model is typically used to analyse output and price responses to exogenous 'demand shocks' or 'supply shocks' in the 'short-run'. Yet under the general equilibrium assumptions in the model, if a market is not in equilibrium, then no other market can be in equilibrium. If there exists a source of a demand shock, then by definition all markets are not in equilibrium and the firm's cost curves during this 'demand shock' need not be fixed, as relative price levels must be changing. The grand assumption of equilibrium and the potential for demand shocks are merely additional simplifying assumptions made when analysing partial equilibrium of markets within this framework.

In the case of a positive demand shock the new profit-maximising level of output is not at all intuitive for most students learning the model - why increase output and price when you could simply increase price to a higher level and not adjust output? The story used to persuade first year economics students is that if the cost of production is below the price, or marginal revenue in more sophisticated versions, why wouldn't a firm keep producing to capture that little bit of extra profit? At first glance the logic can appear sound. Indeed students often respond with how profound that is, when they naively assumed that firms would cease expanding production when average total cost was at a minimum.

As will be later shown, the student's intuition was correct.

In the long run, since capital is no longer assumed to be fixed, all output is produced at the minimal cost meaning that all firms produce at the minimum average total cost, which is also the price. How one conceptually deals with the fact that any point is time is always a short-run and the long-run outcome of previous periods, is never fully considered. Time does not enter the model in a meaningful way.

2.2 Fundamental welfare theorems

The transition between the positive nature of the neoclassical descriptive model and its normative interpretations is bridged by the *fundamental welfare theorems* which are derived in full in Mas-Colell et al. (1995) and are briefly described as

1. Every Walrasian equilibrium is Pareto-efficient
2. Every Pareto-efficient allocation can be supported as a Walrasian equilibrium

A critical requirement of the second theorem is the convexity of preferences and production sets, or to be clear, the absence of economics of scale. Moreover, all the other assumptions for competitive markets are also required, such as perfect information, no externalities, free entry and disposal, complete markets including perfect credit markets and more. These theorems are the lever allowing economic analysis to rise from a mere description of markets, to a morally justifiable normative approach to policy assessment, using the weight of the utilitarian foundations of consumer theory.

[5] Cournot is just one of a family of models of imperfect competition

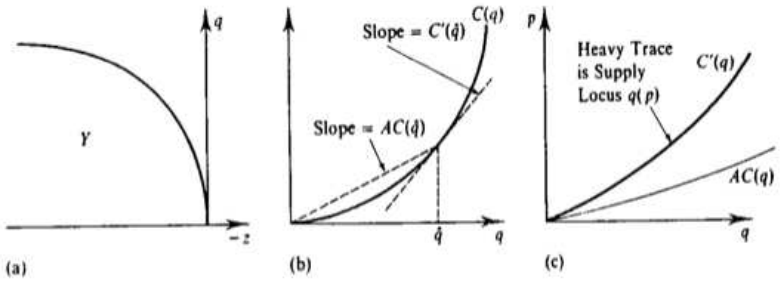

Figure 5.D.1
A strictly convex technology (strictly decreasing returns to scale).
(a) Production set.
(b) Cost function.
(c) Average cost, marginal cost, and supply.

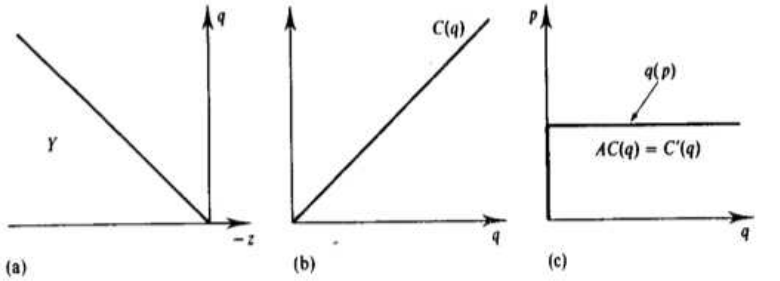

Figure 5.D.2
A constant returns to scale technology.
(a) Production set.
(b) Cost function.
(c) Average cost, marginal cost, and supply.

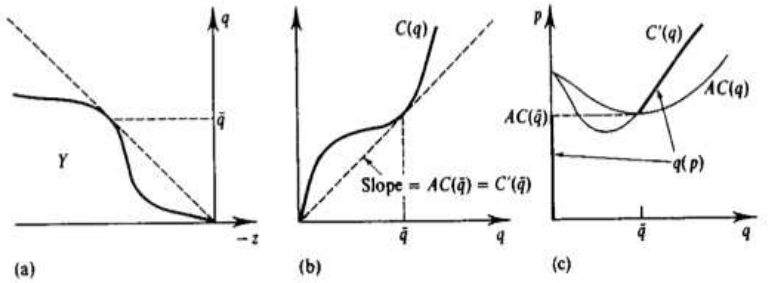

Figure 5.D.3
A nonconvex technology.
(a) Production set.
(b) Cost function.
(c) Average cost, marginal cost, and supply.

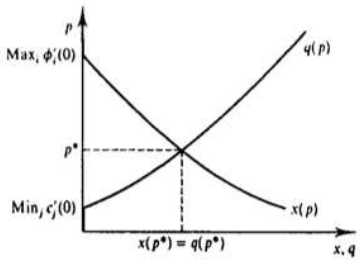

Figure 10.C.3
The equilibrium price equates demand and supply.

Figure 2: *Cost curves and market outcomes for simple model of the firm and market (Mas-Colell et al. 1995)*

Yet if markets operate in ways the conflict with any or all of the assumptions required for these welfare theorems to hold, the lever is broken and these moral foundations collapse.

2.3 Circular arguments over the source of 'normal' returns

Costs in the core neoclassical model include the opportunity cost of alternative investment - a 'normal' rate of return, or the 'risk-free' rate of return. In the neoclassical world the natural risk-free rate is a mere reflection of the time-preferences of consumers and producers. These deep parameters are can potentially be reflected in government policy either as a reaction to markets, or as a correction to some perceived imperfection of money and credit markets.

One persistent criticism is as follows.

> The rate of interest is normally positive for a quite different reason. Present purchasing power is valuable partly because, under the capitalist rules of the game, it permits its owner . . . to employ labour and undertake production which will yield a surplus of receipts over costs. In an economy in which the rate of profit is expected to be positive, the rate of interest is positive . . . [and so] the present value of purchasing power exceeds its future value to the corresponding extent. . . This is nothing whatever to do with the subjective rate of discount of the future of the individual concerned. . . (Robinson 1952)

If opportunity costs are embedded in costs, yet all markets are in equilibrium reflecting the opportunity costs, there is no other source of this normal rate of return.

> We have Marshall's theory that the rate of interest is the "reward of waiting" but "waiting" only means owning wealth. A man "may have obtained the defacto possession of property by inheritance or by any other means, moral or immoral, legal or illegal. But if, having the power to consume that property in immediate gratifications, he chooses to put it in such a form as to afford him deferred gratifications, then any superiority there may be in deferred gratifications over those immediate ones is the reward of his waiting" (1890, pp. 613-14).
>
> In short, a man who refrains from blowing his capital in orgies and feasts can continue to get interest on it. This seems to be perfectly correct, but as a theory of distribution it is only a circular argument. (Robinson 1972)

The same circular argument applied to rental returns to real property. The cost of creating land is zero. Therefore the opportunity cost must be the price at which one could sell the land. But that price is determined by the market and is merely the rent capitalised at a normal rate of return. The absence of time in the core model leads to the necessary assumption of some non-market exogenous source of normal returns.

Indeed, the real problem for this equilibrium notion of 'normal returns' is that if a market is perfectly competitive and no one firm can make returns higher than this rate, where is the incentive to invest in real productive activities? Why not put your funds in the bank rather than purchase capital goods, hire a labour force, and produce real products? If real firms do make returns above the 'normal rate' than we have automatically introduced a new dimension to our notion of competition - competition based on pricing that seeks to maximise this rate of return above the normal rate, rather than maximise profits taking this rate as given.

3 A market model with return-seeking firms

The premise of this model is that firms do not simultaneously solve their pricing and output decisions, but react to their environment in ways that can increase their rate of return over the

long run. Capital is fixed only for the single period for which output level decisions are made, while capital investment decisions occur across periods, allowing the next period cost curve to change as a result of capital investment. Costs exclude rents, but include include amortised capital investment from previous periods, labour and material inputs. Returns are nominal.

The core of the model reflects the routine financial decision-making process of firms whereby returns are maximised subject to constrains such as competition, demand, and legal and regulatory conditions. Given the uncertainty surrounding these constraint, the model is described as return-seeking, since any long term maximum is ultimately unknowable. Equation (1) expresses the firm's return-seeking objective of increasing their discounted revenue per dollar of cost over the foreseeable future, denoted h for the planning horizon.

$$\begin{aligned} \text{returns} &= \max \int_0^h \frac{p_t(q_t)q_t(K_t, L_t, P_t) - (w_t L_t + c_t K_t + r_t P_t)}{(w_t L_t + c_t K_t + r_t P_t)} e^{-\rho t} dt \\ &\Rightarrow \max \int_0^h \frac{p_t(q_t)q_t(K_t, L_t, P_t)}{(w_t L_t + c_t K_t + r_t P_t)} e^{-\rho t} dt \end{aligned} \tag{1}$$

This calculation is equivalent to maximising the internal rate of return for a capital investment decision. It needs no understanding of marginal cost or marginal revenue to be optimised. Before proceeding it is worth specifying that maximising *returns* can analytically determined by differentiating with respect to q at any point in time resulting in

$$\max returns \quad \Rightarrow \quad \frac{\partial r}{\partial q} = \frac{p'(q)ATC(q) - p(q)ATC'(q)}{ATC(q)^2} = 0. \tag{2}$$

The output condition for this to be true is when the gradient of the demand curve, time the average cost curve, is equal to the gradient of the cost curve times the demand curve, or

$$p'(q)ATC(q) = p(q)ATC'(q) \quad \text{and } ATC(q) \neq 0. \tag{3}$$

Of course this short-run maximising option need not be an option available to most firms in competitive markets, and later it will be shown how competition acts as a constraint on returns, allowing production beyond this point in markets where firms face downward-sloping cost curves. In light of both real constraints of firm activities, and the uncertainty inherent in the inter-temporal nature of their investment and output decisions, firm behaviour in this model is described as return-seeking, rather than return-maximisation. Firms make decisions seeking to improve their returns over the long run, but may not be able to maximise them in the short run because of the threat from existing and potential competitors.

Figure 3 shows a selection of cost and demand curves and the difference in optimal output and price under return-maximising, (Q_R, P_R), and profit-maximising, (Q_P, P_P), behaviour. Only in the case where variable/marginal costs are zero do the two models find that same optimising output level[6].

Those familiar with the family of neoclassical market models, including Cournot and Bertrand competition and other variations of semi-competitive models, might now be asking the question of how a market optimising solution is found under this proposed new framework. The answer is there is no hard analytical such solution. Without knowing about inherent risks in a particular market (the uncertainty over future demand), the limitations to competition from both property markets (either from access to physical locations, or access to intellectual property), and the economies of

[6]It should also be noted that the optimal output in a return-maximising situation can occur on the downward sloping section of the marginal cost curve.

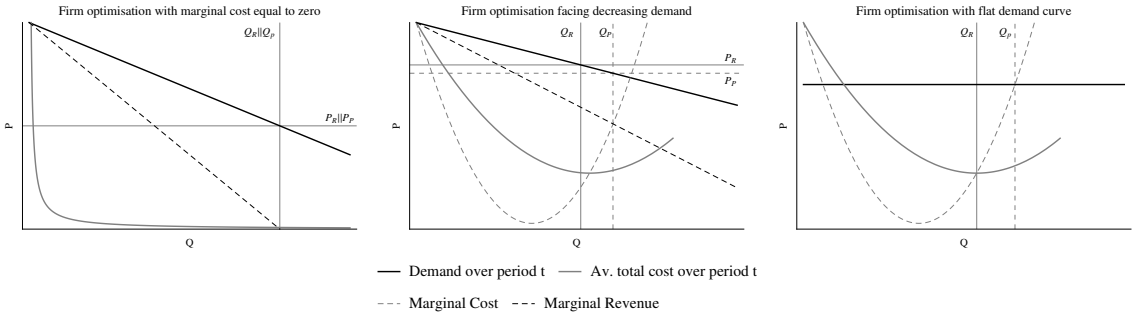

Figure 3: *Return-maximising and profit maximising output compared*

scale from the range of potential production techniques, little can be said about the likely market price, output level, or number of firms.[7]

4 Implications of the new model

Reframing the firm's decision problem leads to many new insights about market allocation and investment mechanisms, most of which are consistent with the business literature In particular, it demonstrates that the timing of new capital investment is the critical decision facing firms, with output decisions a secondary consideration. In fact the driver of output growth under a stable demand curve is new capital investment which capture economies of scale. Changing technology enter the model through the new cost function associated with the capital investment based on this technology.

It reveals that the shape of the supply curve is not reflective of the horizontal summation of firm cost curves, but of the market power of firms due to an inability for competitive entry into the market. It shows that even under the assumptions of the neoclassical market allocation model with fixed capital, the optimal output level for a firm is much lower, meaning that the capital per unit of output is generally higher than what is expected under the neoclassical models. And it shows that competing for market share is one of many effective return-seeking activities.

4.1 Capital investment decisions

The short story about capital investment decisions is that if you are a firm that faces diminishing average total cost (declining unit costs), when these economies of scale are exhausted it signals the need for new capital investment. There are no circumstances by which returns can be improved by producing output on the upward sloping section of the average total cost curve, even in the most extreme case of a firm facing a perfectly horizontal demand curve. The common student intuition about the importance of the minimum point in the average cost curve is correct.

In a market where any or all firms are operating at this point following an unexpected shift to the right of the demand curve, returns will be temporarily increased. These high returns will attract new capital (from existing or new firms) into the market, knowing that they can receive acceptable returns even once their new output is added to the existing market output.

How high do these returns have to be to attract new capital investment?

[7]Unlike the neoclassical market models, the exact number of firms in a market under this model is of little consequence to firm behaviour. It is the threat of alternative sources of production that limit returns

This depends on two important factors - the shape of the demand curve, and what I call the 'marginal capital unit output', which is a measure of the output level that justifies investment in a new increment of capital at the current price. The technicalities of this increment of capital are important. In many cases, existing production techniques can be augmented and incrementally improved. However, at other times a particular location may have exhausted its potential and a whole new production facility may be required to increase output.

Where incremental capital augmentation is no longer a viable investment option new market entrants are likely to appear. In the first case, the existing firms have a large advantage in being able to invest in smaller increments of capital, keeping returns low enough to protect them from new entrants who would require a very large increment of capital to begin any production.

The marginal capital unit output attracts investment when its cost can be justified by maintaining the level or return at the new level of output. This condition is represented in Equation (4), where the p function represents the demand curve as a function of market output q, C_u is the the marginal capital unit output, $c(C_u)$ is the cost of the marginal capital unit and c_t is the total cost of producing the previous output level prior to the capital investment in C_u. At the expect time k the new capital investment is justified. In some cases, new production technology may justify a k of zero as soon as it is viable. In other cases where production technology changes very slowly, k will mostly depend on the age of current capital.

$$\frac{p(q+C_u)_{t+k}(q+C_u)}{c_t + c(C_u)_{t+k}} = \frac{p(q)_t q}{c_t} \tag{4}$$

To be clear, the optimal timing of new investment in a simple case of a market with four firms[8] utilising the same capital and production techniques is presented graphically in Figure 4. With an arbitrary starting point, at time t all firms are producing at their return-seeking level facing the market demand curve D_t. The next marginal capital unit is only available at the scale of an existing firm production level. When the demand curve begins to shift, and firms expect this to be a permanent shift[9] that will at least move beyond D_{t+k}, they will undertake this new capital investment, allowing output to increase without an increase in price.

While this investment is being undertaken there may be short periods where one or other firm is operating on an upward sloping portion of their average total cost curve. What will be clear from this analysis however is that market power is likely to exist in markets where the marginal capital unit is a large share of the market. In this case, since there are no intermediate steps between output levels from say one production plant, and two plants, the period of time in which the demand schedule is such that prices do not justify investment in a second plant, the owner of the first plant will enjoy a temporary monopoly, despite free entry - due simply to the technicalities and scale of existing production techniques. Thus, market power is driven more by opportunities investment in marginal capital units, than the number of firms in an industry.

Of course nothing said here contradict long run neoclassical models, which espouse a flat supply curve. But without understanding the short-run in the context of these long run decisions, strange results such as upward sloping costs curves will emerge, greatly reducing the descriptive power of the model.

4.2 Firms minimise returns when they maximise profits

Section 3 showed that the return-seeking firm produces a lower quantity of output for a given level of capital investment than predicted by neoclassical market models under all scenarios except when marginal costs are zero. Figure 5 shows the situation from centre panel of Figure 3, where

[8] Or a market of 1-3 firms utilising 4 'units' of capital

[9] A temporary shift in the demand curve leads to alternative decisions discussed in a alter section

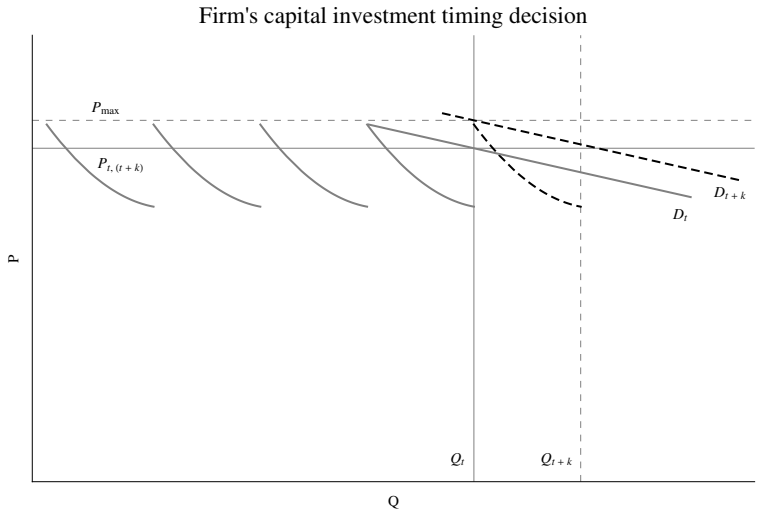

Figure 4: *Timing capital investment from a permanent shift in the demand curve*

the firm faces a downward sloping demand curve. The optimal output under a profit maximising assumption, Q_P, is much higher that output of a return maximising firm, Q_R.

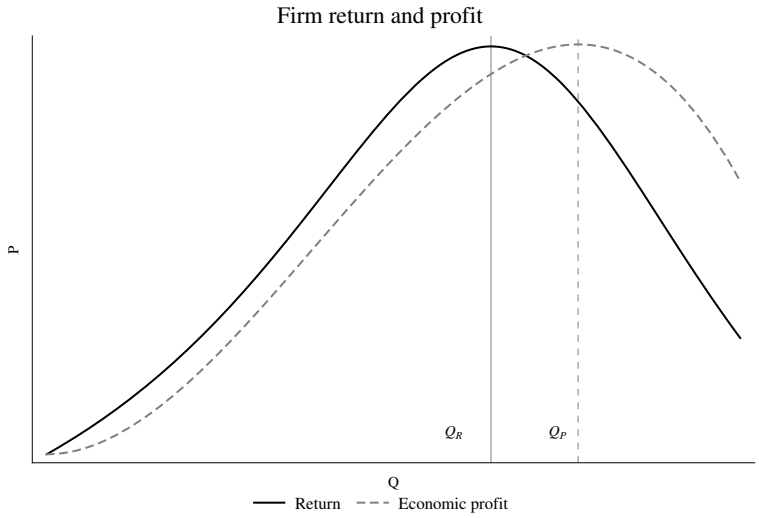

Figure 5: *Profit and returns against firm output*

The question to consider here is what assumption needs to hold for the profit-maximising output be financially superior decision to the return-seeking output level. And that assumption is that the marginal cost of new unit of output is zero over the time period under consideration. Any time marginal costs are non-zero, the return on these costs must also be factored into firm decisions.

If any firm in a market decided to operate at the profit-maximising level, they would be reducing their own returns with very little consequence for their competitors.

4.3 Supply curves can be any shape

Blinder's (1998) detailed research showed that firms typically operate on the downward sloping part of their average cost curve, and often on the downward sloping part of their marginal cost

curve. Indeed, 89% of interviewed firms had flat or downward sloping cost curves, despite economic theory suggesting that this should be a special case of a general rule of supply. By introducing returns on all costs as the firms goal, we are able to distinguish between the return-optimising behaviour that can lead to the appearance of an upwards sloping supply curve despite each firm facing downward sloping supply curves.

As a rule of thumb, the return-seeking model of firms finds that an upward sloping supply curve can arise in two situations. First, when a temporary monopoly situation arises from unanticipated short-term demand shock. Second, a market with increasing 'marginal capital unit' costs. This means, for example, that investing in the next unit of capital is more expensive than the last unit. The common example is in mining and natural resources, where expanding production requires capital investment in more expensive to extract resources. Both of these situations are described in this section and compared with markets where firms face decreasing 'marginal capital unit' costs and declining average cost curves.

In the first case we have a market with few firms who face little to no short term threat of losing market share to their competitors - at least within the period of time the demand shift is expected to last. In Figure 6 an apparent upward sloping supply curve can be derived when firms choose increase prices, output and returns in response to a known short-term shock, a shock which will not trigger any new capital investment. Since no firm is likely to lose their market share to another (a process described in more detail in Section 5), each is able to increase returns over this short period without risking future returns due to competition from other firms.

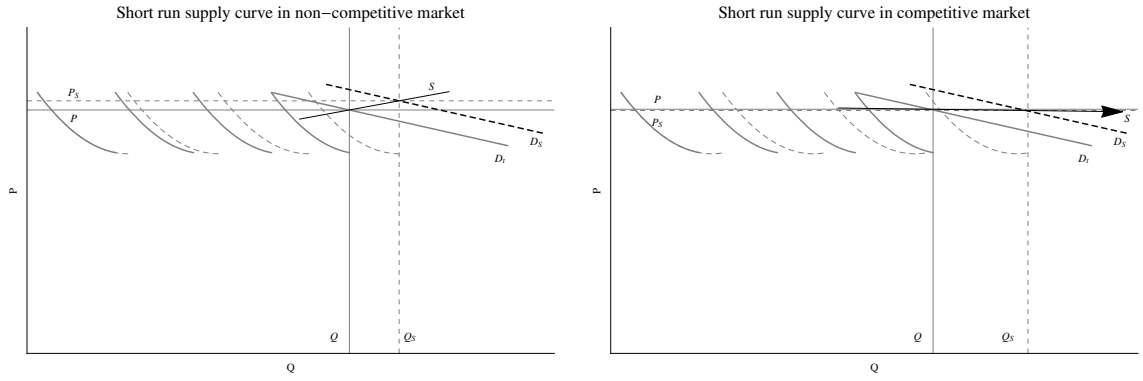

Figure 6: *Apparent supply curve from demand shock (non-competitive and competitive short-run)*

However, on the right panel of Figure 6 it can be seen that in a highly competitive environment, where market share could be lost during the period of the demand shock, firms may choose to maintain their existing level of returns, allowing them to compete on price win market share. The net result of this behaviour is an increase in output with lower prices during the demand shock.

This type of analysis lead to an alternative interpretation of estimated slope of supply curves in the short run. Namely, that the less short run competitive exists, the steeper the upward slope of the short run supply curve. Estimating the shape of the supply curve provides a corresponding estimate of the competitiveness of the market.

In the second case, of markets with increasing 'marginal capital unit' costs, it is best to have in mind resource markets such as minerals, where new deposits can only be mined at higher cost than currently mined deposits. Capital is invested in periods prior to when the product is delivered, although contracts may also be negotiated prior to the capital investment itself. Essentially average unit costs falls rapidly once the capital is invested, and beyond the optimal capacity there may be some small scope to use the capital more intensely before new investment in the marginal capital unit is required.

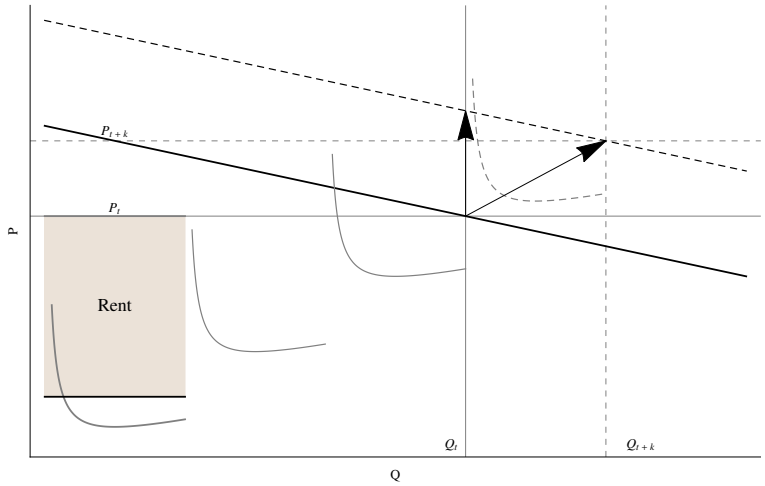

Figure 7: *Supply curve range from demand shift in resource market (and of rents)*

Figure 7 shows a descriptive model of such a market. Three units of capital, potentially three mine sites, possibly three firms, are operating at time t. The mines are at different sites, with unit costs reflecting the remoteness and richness of deposits. The ordering of the firm cost curves is such that the most costly mine is the marginal mine. When demand shifts there is a period in which output cannot be increased noticeable prior to the point when the next unit of capital investment become viable. Thus, depending on how well this shift is anticipated, the apparent supply curve will fall somewhere between the two arrows. For example, if the demand shift is anticipated then miners will contract for future deliveries from the new site (the dashed cost curve) prior to their investment, and the existing mines will only be able to charge at this new price, P_{t+k}.

The final important feature of such markets is that they attract economic rents. Those who own property rights to more advantageous locations - be it mineral deposits, or simply land itself - will earn rents over time as demand shifts and marginal units of these property rights attract capital investment. The rent is the revenue above the price that would have attracted the capital investment, this price being defined as the current returns that attract investment in this market on capital costs (excluding the cost of the property right itself, since this cost equals the capitalised value of the rent). By not combining money, capital (the physical produced commodities that contribute to the production of commodities in future periods), and property rights into a single fungible concept of capital, the model allows for potential extensions that consider the interaction of debt creation, physical capital investment, and property rights values.

5 Competition is a process and a constraint

"*Innovation and heterogeneity are the true hallmarks of competition, yet these concepts are effectively excluded by the neoclassical model*" - Keen and Standish (2011)

By ignoring time, and invoking the Walrasian auctioneer, the neoclassical model can ironically be said to model competition by assuming it already happened before any trade took place. For most analysis of real markets at real points in time this is not a useful assumption to make. Even the comparative statics that arise from neoclassical models fail to describe the way one equilibrium could transition to another. For example, Keen and Standish (2010) simulate an almost perfect market, where firm choice of output arises in response to the change in profit from the previous period. If increasing output increased profits in the previous period, they will continue to increase in the next. Without the ability of a central planner to prearrange the output levels of all firms,

there is no way to generate any signal for firms to stop their behaviour when they reach the Cornout output level. A signal only arises when the market as a whole is producing at the monopoly level of output.

Problems of competition are typically framed by neoclassical economists as problems of coordination of output levels given fixed capital costs. A market with many firms will be unable to collude, and therefore they will increase output at the expense of others, and the group as a whole. Competitive markets will give up producer surpluses to consumers, and increase total surplus in the process.

Apart from the notion of producer surplus being ill-conceived[10], competition is not a static decision to increase output over a brief period of time. Competition is a process of maximising returns on costs by the entrepreneur that are generated by their innovative and coordinative efforts.

Equation (1) is a useful guide to this process, recalling that h is the planning horizon of a firm, c is the cost of capital and material inputs at time t, r is the rent for real property, and w is the wage cost of labour.

$$returns = \max \int_0^h \frac{p_t(q_t)q_t(K_t, L_t, P_t)}{(w_t L_t + c_t K_t + r_t P_t)} e^{-\rho t} dt$$

Each variable in this equation provides a clue as to how firms can pursue their return-seeking objectives. These include (but are not limited to):

1. via a reduction in input quantities per unit of output, be it labour, L, or capital equipment K through new production techniques and updated capital equipment,
2. via a reduction in unit costs, be it labour costs, w, or the cost of capital equipment c,
3. via product differentiation, which generates a steeper firm demand curve due to less substitutability,
4. via increasing market share and the pursuit of customer loyalty to shift the firm demand curve (both higher and steeper),
5. via pricing strategies that reduce the attractiveness of your market to potential competitors (preserving p and q over future time periods),
6. via regulatory capture and regulation to increase the demand curve and steepen its slope, or
7. via mergers, which increase market share and reduce unit costs.

These actions and more define the process that is competition, and are typical of firm behaviours in real markets. Some are features of the economics textbooks, while others rarely feature in economic analysis, but are commonly described in business schools. For example, increasing market share in a competitive market of where firms face increasing marginal costs reduces profits. So why is it such a common competitive strategy?

One critical item on the above list is the use of pricing strategies. As noted earlier, the return-maximising price and output level for a firm may be possible in the short-run, during periods of unanticipated demand. However, if this short-run return is likely to create future costs from loss of market share, an alternative pricing strategy that results in lower short-run returns will instead be pursued. Remember the model maximises returns over the foreseeable planning horizon rather than a single period. Thus the threat of new entrants into the market, or the threat of losing market share to existing competitors, is the key constraint on prices that competition provides.

Competition strategies that arise from the model can be clustered into three main areas - pricing, reducing unit costs and augmenting the firm demand curve.

[10]Producers are rewarded by a real return on their costs

5.1 Competition through pricing

Probably the most direct element of firm competition the is pricing. In the neoclassical market models, prices are set only as an allocation mechanism that clears markets in the short-run. However, in the return-seeking model, prices are set to in regards to dynamic considerations such as future capital investment, changing market share, and defending the market from new competitors.

The last item is crucial. Even when the number of firms in a market is low, if the marginal unit of capital is small relative to the market, and technology changes mean new capital can reduce costs, firms will have an incentive to price in a way that their returns will not attract competitors to invest in the marginal capital unit before they do. Mark-up pricing, for example, is a rule of thumb that firms can use to ensure their returns are broadly consistent with their current competitors. Thus, pricing strategy depends on the size of the marginal unit of capital relative to the market demand, and the technical and legal abilities of competitors or new entrants to invest. In terms of Equation (1) pricing strategies aim to preserve revenue in the over the planning horizon.

The wide existence of this strategy provides the possibility that duopolies can still appear very competitive, and have reasonably low and stable returns, even though traditional Cournot theory would say that they restrict output and generate 'abnormal returns'. This can occur even with high sunk costs, which is not the usual case in contestable market theory. Where duopolies and oligopolies may be able to increase returns is through erecting legal, contractual and regulatory barriers for potential new competitors.

The opposite of defensive pricing would be understood as price gouging. When a business has a monopoly and the ability to change prices regularly, it will exploit seasonal shifts in the demand curve. For example, petrol retailing is a local monopoly facing a weekly demand cycle. As such, the rational pricing behaviour is to change prices daily to maximise returns. When retailers have less local monopoly power, these price fluctuations are lower[11]

To reiterate an earlier point, the message here is that price increases in response to demand changes are hallmarks of markets with low levels of competition and high levels of firm pricing power.

5.2 Competition through unit cost reduction

5.2.1 Capital investment

Consider a firm in a competitive market. Each firm faces a particular cost profile based on the vintage of their capital. The firm that invests in new technology capital first enables them to capture higher returns for a period until other firms invest in similar, or improved capital. The optimal point of investment in new capital obviously depends on the age, cost, depreciated value and productivity of current capital, compared to new capital.

This process occurs in all markets where existing or new competitors are a threat. The process of new competitors entering the market decreases prices because it expands the productive base and reduces the market-wide average total cost curve. While Baumol et al. (1982) provide the basis for mainstream theories of contestable markets, there is little reason to require the fundamental assumptions of Baumol's theories, such as no entry barriers, no sunk costs, and access to the same technology. Existing and new firms can compete, based on new technology capital and in the face of sunk costs, if returns are high enough to attract them.

Take the example market in Figure 8. The three grey cost curves, included the dashed curve, show three firms using capital of different vintage, with prices are set by each firm such that returns are

[11] See Hogg et al. 2012 for spatial and temporal analysis of petrol retail pricing. http://www.ncer.edu.au/papers/documents/WP86.pdf

satisfactory for all capital to be utilised (price P). However these returns are also high enough to attract a new competitor with access to new technology capital. The black dashed curve is the cost curve of the new competitor. If they invest in capital and enter the market, the new price of this good when the new supply is added will be P_C.

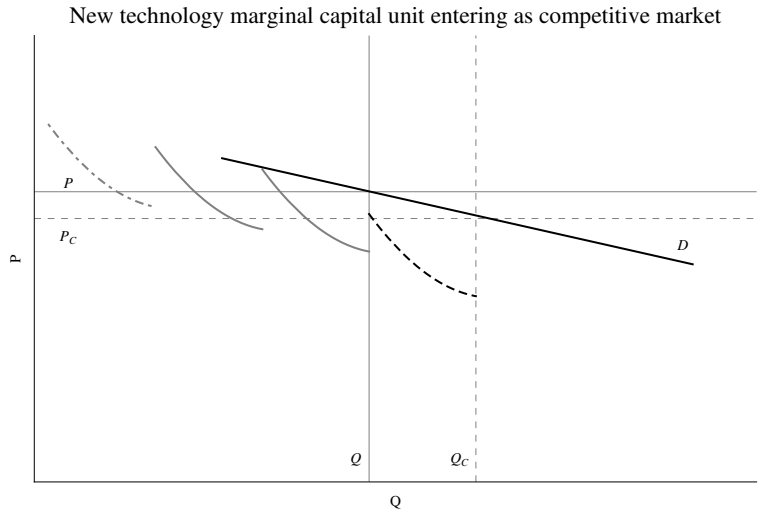

Figure 8: *Market supply and firm output with new competitor*

However, since the new price is below the cost of the highest cost incumbent firm, that firm (or that part of a larger firm with multiple capital units) will retire their capital until they can invest in the low cost technology capital being utilised by the new entrant. When the incumbent leaves the market the new output is determined by the three lower cost firms. Figure 9 shows this outcome, and that each remaining firm improves their return by increasing their output. Whether the change in output and price stops at this point depends on the competitive dynamics involved. If the displaced incumbent was a single firm, it might be viable for these remaining firms to increase production beyond this point, accepting temporarily lower returns, in order to defend agains the displaced firm reinvesting in the market.

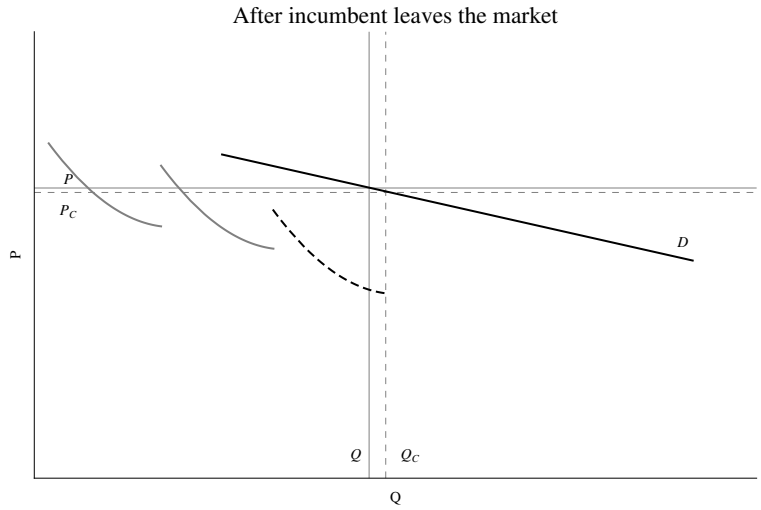

Figure 9: *Market supply, optimal firm output, and new competitor with incumbent displaced*

We expect therefore, that a period where new product and production techniques are being implemented, a competitive environment will emerge as firms seek to capture high returns from being

the first to invest in new capital and methods of production. Indeed, if economies of scale are present we would expect the number of firms surviving this period of competition to relatively low, which then shields them from Baumol-type short term competitive threats.

It is no surprise that one of the core competitive actions in a capitalist economy is to invest in better capital.

How exactly do these investments transmit through the return function? New capital typically allows reduced labour input L, per unit of output, which comes at a price cheaper that the increase in capital cost cK. Thus the term $wL + cK$ is reduced, and returns increased.

5.3 Wage costs

The timeless political battle between wages and profits must have its ultimate source in markets or the policy framework that supports private market production. By the definition in Equation (1) a reduction in w always results in an increase in returns, and if it occurs across the whole economy more broadly, results in an increase in rents.

The simply way to envisage this mechanism is to think of a shopping centre owner. Retail rents are determined by the willingness to pay by the marginal shop tenant, who determines this by subtracting costs and their hurdle rate of return from their expected revenue. This rate of return is usually quite market specific. If just one tenant in the shopping manages to reduce wage costs they simply increase their own returns. It is highly unlikely that they will be the marginal rent setter for the centre as a whole. Yet if every shop tenant benefits from the same reduction in wages costs, through a policy change or some kind of fundamental reorganisation of retail businesses in general, then the shopping centre owner can now capture these reduced wages as rents by holding out for a rent price where the marginal tenant will still meet their hurdle rate of return.

How this trade-off aggregates at a macro-level is unclear. For example, wages determine household incomes, which determine the price able to be paid for housing. If wages are falling, owners of residential land face falling rents, even if owners of non-residential land and other assets face increases in rents. Thus, the competition between rent-seeking owners of real property can result in a reasonable bargaining position for wage earners, being supported by those whose rents directly rely on wage levels.

Additionally, reduction in general business costs, typically referred to as 'red tape', is a common tactic of rentiers. In the same vein as reducing economy-wide wages, reducing c will simply produce higher rents. The outcomes of the policy and technology changes are very much conditional on the patterns of ownership of real property and commercial business. One fundamentally requires a market theory that supports disaggregation of rents and returns to properly analyse the output and distributional effects of wage-related policy and technology.

5.4 Competition through demand augmentation

5.5 Product differentiation

Firms can increase returns is to be the first to innovate and bring a new product to market. Creating a new market niche allows firms to face a much higher and steeper firm demand curve, as products are less substitutable and competition is yet to take place in that new market. The first mover advantage in new markets is a clear driver of competition.

5.6 Marketing and customer loyalty

Manipulating the shape of the firm demand curve can be a beneficial a return-seeking strategy. Marketing attaches an emotional signal to a product that is hard for customers to replace, making the product of now firm less substitutable with another. In the return-seeking model a steeper demand curve increases returns.

A complementary method for creating less price sensitive customers is through the promotion of firm, or brand, loyalty. Through loyalty rewards schemes, or simply through the trust generated by returns policies, warranties and other initiatives, loyalty can be a tool for firms for create less price sensitive customers and increase returns.

5.7 Market share

If cost curves are downward sloping it makes sense to capture market share from competitors to increase returns that arise from gaining the largest economies of scale, and decrease the returns of competitors who will be unable to harness as significant economies of scale. Attracting market share from competitors is usually the goal, rather than a method of competition, but it needs to be stated here why this strategy is consistent with the return-seeking model.

Simply put, market share means the share of total market output a firm can produce while maintaining the price level. For a firm facing a downward sloping cost curve and few new capital investment opportunities, increasing market share is one of the few options to increase returns by facilitating lower unit costs of production.

An alternative method of gaining market share is through mergers. Mergers are generally attractive in markets where production occurs with increasing returns to scale, allowing the merged entity to lower their cost curve. The difference between the neoclassical models of restricted competition and the return-seeking model reveals that threat of merger to competition is dependent on the cost profile and size of the marginal capital unit.

5.8 Lobbying for regulatory change

Since all markets operate within a legal and regulatory framework, the ability of changes to the rules of the game to shift the curves facing individual firms or industries is a particular avenue for increasing returns of a particular firm. This motive is well developed by public choice economists, although integrating this idea into the neoclassical market model is conceptually difficult. In the return-seeking model we can interpret policy changes as having price impacts through the potential removal for future competitors, which were being considered in the pricing decision of firms.

We can imagine that firms may wish to increase 'red-tape' to defend their market from new entrants, especially if their are economies of scale to the administrative burden. A firm may find that a new law would cost them 0.1% of the price per unit of output at their current level of production, but for a new firm produce just a tenth of their output, the administrative cost might be 2% of the price.

6 Normal returns and welfare considerations

Two concepts need to be cleared up before concluding overview. First is that the only rates of return considered here are nominal, not real. Only nominal variables are observed in the market place, along with expectations of the growth in these variables over the planning horizon. Rates of return are therefore emergent from subjective risk-weightings of firms in any particular market.

Second, the welfare theorems that support the normative position of economic theory have no basis. The very concept of surpluses fails to enter any calculation. All that exist are returns, rents and consumption, distributed in various ways between individuals and organisations in the economy. Those seeking a measure of 'success' need to choose their own moral foundations.

7 Final remarks

The neoclassical market model forms the core of the economics discipline. However when seeking to understand real markets economists are forced to drop many of the basic features of the core model. A model of return-seeking firms is consistent with empirical evidence on the shape of demand curves, the true costs structure faced by firms, and is suggestive of the mechanisms by which the process of competition operates. A key ingredient is the inclusion of economic rents, and the revelation of the trade-offs between production costs and rents, taxes, wages and the costs of regulation.

The model also preempts some critical concepts that may be relevant for improving our knowledge of the economy. For instance, return-seeking firms are likely to be attracted to debt in order to leverage their return. Competitive pressures may generate an appetite for debt amongst firms that permeates through the competitive pressures of markets - if one firm leverages their returns, others must follow. Such a cascading nature of debt/credit issuance could generate business cycles with 'micro-foundations' consistent with this model.

Identifying the nature of rents in the model leads opens new avenues of analysis of capital investment. Particularly, we can identity that lowering rents will increase returns from capital investment, which opens up new scope for policy to both incentivise capital investment in general, and to respond to recessionary conditions. For example, when a new location is required for a manufacturing plant, existing land owners may not sell at a price that makes the new investment attractive, withholding their land from development consistently with their own long-run return-seeking strategy.

An important element that has been presently overlooked for the purpose of introducing the model is whether there are any normative interpretations to be made in a similar vein to the fundamental welfare theorems. Is the proper measure of welfare simply total macroeconomic output, adjusted for legitimate notions of social equity?

It needs to be noted that this model of a market with return-seeking firms is intended to be a trigger for a large scale departure from neoclassical market theories. Marginalist approaches are not required to understand markets, nor does equating marginal revenue with marginal cost maximise returns for a firm even in the short-run Moreover, this departure means that wages and rents are not primarily determined by markets, but are in fact social decisions, or institutional constructs, that emerge from the legal and political environment to which markets respond.

I will conclude by preempting reactions to this exposition from the economics profession. Reactions will fall into discrete categories. First, there will be those who roughly agree with most of what I have written and will say that there is nothing new in here; that there are economic models that allow for positive profits and natural monopolies which handle the cases described above. To them I say please consider the value of a core model that requires such extensive ad hoc modification to be consistent with empirical regularities, and still fails to be predictive.

The second category will be those who think I have misrepresented the neoclassical market models. That I am criticising only what is covered in introductory economics courses and not the more sophisticated modern research. To them I say that I have in fact based the description of neoclassical models on the most popular post-graduate micro-economics textbook, since it is far more honest about its assumptions than any undergraduate text. Further, that even the most

modern models are typically variations of the same core, pre-solved and ignorant of any method for reaching their solutions.

The final category is those who believe I have left out something of crucial importance in this short exposition. I admit that in all likelihood I have. I cannot perfectly rewrite the core model of a whole discipline. The idea is to introduce a framework, a scaffolding, to bridge the gap between economic analysis of markets and the actual self-reported behaviour of firms. When the truly fundamental feature of the model, the process of firms seeking returns over profits, cascades through a consistent aggregation method into market, the interactions will be dependent on additional assumptions, most of which I hope are drawn from observations, such as legal constraints and limits to knowledge.

References

Baumol, W. J., Panzar, J. C. & willig, R. D. (1982), *Contestable Markets and the Theory of Industry Structure*, Harcourt Brace Jovanovich.

Blinder, A. S., Canetti, E. & Lebow, D. (1998), *Asking about Prices: A New Approach to Understanding Price Stickiness*, Russell Sage Foundation Publications.

Keen, S. & Standish, R. (2012), 'Debunking the theory of the firm - a chronology', *Real World Economics Review* .

Leijonhufvud, A. (1973), 'Life among the econ', *Western Economic Journal* **11**(3), 327.

Mas-Colell, A., Whinston, M. & Green, J. R. (1995), *Microeconomic Theory*, Oxford University Press.

Robinson, J. (1952), *The accumulation of capital*, MacMillan.

Robinson, J. (1972), 'The second crisis of economic theory', *The American Economic Review* **62**(1/2), 1–10.
URL: *http://www.jstor.org/stable/1821517*