


Spring 2020

## Non-Increasing Cost functions: Making the Connection between Keynesian and Neoclassical Macroeconomics

Max von Holtendorff

Follow this and additional works at: [https://digitalcommons.bard.edu/levy\\_ms](https://digitalcommons.bard.edu/levy_ms)

 Part of the [Economic History Commons](#), [Economic Theory Commons](#), and the [Macroeconomics Commons](#)

# Non-Increasing Cost functions: Making the Connection between Keynesian and Neoclassical Macroeconomics

Max von Holtendorff  
Levy Economics Institute MA Thesis

May 26, 2020

## **Part I: Real and Financial Variables in Economics**

In physics a great deal of attention is paid to the units in which quantities are denominated. Care is taken to ensure that the units in every equation are consistent. Dimensional analysis, i.e. checking to ensure that the units in equations properly balance is a useful check on any theory and can, in fact, sometimes lead to physical insights. For example take Newton's second law, which states that force equals mass times acceleration. Since acceleration has units of meters per second squared, and mass has units of kilograms the units of force must be kilogram meters per second squared. If force was not measured in those units the equation would be incorrect.

In contrast, dimensional analysis is not used in economics. Most quantities in economics are denominated in money terms, with a few exceptions, such as labour time. Thus, dimensional analysis is typically not required. There are however two very different types of variables studied by economists: real variables and financial variables. Despite the fact that both are often measured in terms of dollars the two variables actually share very different properties. The real economy is composed of thousands of dissimilar objects; aggregation must somehow be done. In contrast, financial variables are all measured in directly comparable units. The way in which real objects are combined to facilitate analysis done is through prices. Thus, when real elements of the economy are measured they are measured in value terms, and denominated in terms of dollars. Despite both being denominated in the

same units, real and financial variables are very different and should not be directly compared.

Real variables are variables related to real physical objects. The amount of real objects we have can increase; real objects can be created and they can be destroyed, and all the properties we expect to find in physical objects hold. Unfortunately, analysis based on real objects is limited since real objects of different types cannot be compared without a common unit of measure. Therefore, aggregation is difficult and many of the types of comparisons economists wish to make are impossible. Did an economy that produced 10 shirts and 5 pairs of shoes produce more than an economy which produced 10 pairs of shoes and 5 shirts? Without some common unit of measure, comparisons of that sort become impossible. The difficulty in making comparisons of dissimilar bundles of goods is not just theoretical. At some points, in centrally planned economies such as China's, production was measured in terms of actual output. As a result China produced far more of certain goods than were required in the economy(Zhao, 2009).

### **Behaviour of Real Variables**

To facilitate the comparison of dissimilar goods and their aggregation into total quantities economists use prices. Often the prices used are the prices that are observed empirically. Each good is attributed a value based on what the good is exchanged for and then that value can be used in the calculation of economic aggregates. However, there can be a certain level of arbitrariness in the prices of goods, and often certain goods are worth more or less than they exchange for on the market. Ideas about the fallibility of market prices led classical economists to attempt to define a set of prices that were natural and would in some sense capture the real value of the goods and not be affected by the vagaries of market prices. Marx in particular formulated a labour theory of value, where objects are valued according to the labour time that contributed to their creation. However, Marx's attempts to create such a system of value led to certain theoretical difficulties. Despite some interest in continuing his approach, attempts to define an objective set of prices at which to evaluate goods have generally been abandoned. Economists today tend to value goods at the prices for which they are exchanged when creating economic aggregates.

Defining value in terms of prices leads to an immediate analytical problem when making comparisons across time. Economists must determine what set

of prices ought to be used, the prices at an initial period or the current price. Sometimes the prices for new goods do not exist in all time periods, and the types of goods sold changes over time, so defining a particular reference year can be difficult. In order to manage intertemporal comparisons, economists define a price level, which is an aggregate of the prices of all consumer goods weighted by the percentage each good makes of total consumption. The change in the price level is called the inflation rate, and is used to facilitate the comparison of prices in different years.

It is important to remember that price level is an arbitrary concept. When inflation is measured, a certain set of goods is chosen as the basis on which price comparisons over time are to be made. Some economists argue that inflation rates can be misleading, as they do not account for improvements to the products that are being sold each year. However, inflation rates are not an objective measure of total value; they are just a somewhat subjective mechanism to allow intertemporal comparisons of prices. Price level is also an aggregate, and while it might be tempting to think that all prices change at the same rate, the overall inflation rate can include some goods that decrease in price while some goods increase in price drastically.

The consensus about how to compare different physical goods and about how to account for the changes in prices over time can occasionally lead to incorrect conclusions. Goods can only really be valued when there is a well developed market for them, and when that market is large enough to not simply reflect market power.

In trade between two people, there are often a large number of prices at which both the buyer and the seller would be willing to make the exchange, with the final value of the object being determined by social factors or social conventions. It would be incorrect to take such a measurement of the value of a physical object as an objective market value. Yet occasionally, economic analyses attribute numerical value to goods for which there can not possibly be a well-developed market. For example, attributing monetary value to clean air or pricing a healthy environment cannot possibly account for the true value of interconnected natural systems. Postulating a market for ecosystems or the environment is impossible, yet often, economic analyses attempt to measure the value of such goods.

Thus, a fundamental problem that arises when aggregating real goods in monetary terms, is the tendency, in many branches of economics, to forget that real and financial variables are very different objects. The two share few properties and the rules that govern their behaviour are very different.

## Behaviour of Financial Variables

In the years leading up to 2008 Wynne Godley wrote several papers predicting that, contrary to the beliefs of most economists of the time, there were fundamental problems in the world economy. Godley predicted that a recession was inevitable. In fact, he predicted in 2000 that if the world were to go 10 years without a recession, the day of reckoning would be sensational (Godley, 2000). His predictions were proven true when global financial markets crashed in 2008.

Wynne Godley made his prediction based on careful analysis of the financial side of the economy. Since 2008 there has been a surge of attention to models similar to the ones Godley used to make his predictions: stock-flow consistent (SFC) models. SFC models are characterised by their careful accounting, and the attention to the financial side of the economy. In SFC models the accounting structure of the economy is paramount: auxiliary equations are made only when necessary to close the model.

SFC models recognize several accounting principles that financial variables must follow. Unlike real assets every financial asset is someone else's financial liability. Cash is a financial asset of the holder and a financial liability of the government. Mortgages are financial liabilities of the borrower and financial assets for whoever holds them and bonds are financial liabilities for the seller and financial assets for whoever holds them. A corollary of the above is that net financial assets must sum to zero for the economy as a whole. The financial wealth of any economy, in contrast to the real wealth, is precisely zero. Similarly if any party is to increase their holdings of financial assets (run a surplus) some other party in the economy must decrease their holdings of financial assets (run a deficit). The sum of all increases and decreases in financial assets in a given period is precisely zero. Ensuring the principles of financial assets are respected ensures that SFC models can model the financial side of the economy with great accuracy.

The principles of financial consistency can be summarized by borrowing a concept from physics. Financial assets are a conserved quantity, that is they cannot be created or destroyed. Similar to how charge works in physics, if a positive financial balance is created, a negative financial balance must also be created. In physics, positive charge and negative charge can be created from nothing; similarly, in financial economics, cash balances in accounts and government liabilities are created simultaneously when no assets or liabilities existed previously.

The implications of the properties of financial assets can be quite surprising. Many of the most paradoxical results in macroeconomics can be understood as implication of the properties of financial assets. Since we are used to thinking in terms of real goods the implications of financial balances are often quite surprising. For example the paradox of thrift, where an increase in aggregate saving leads to a decline in the total value of trade is paradoxical precisely because saving in terms of real goods obviously does not lead to a decrease in the amount anyone else is able to hold. Similarly the Kalecki profit equation, derives the paradoxical result that under certain assumptions capitalists earn what they spend, is deeply counterintuitive because usual intuitions about real goods do not apply to financial assets.

### **Domar's Growth Model**

The properties of financial assets are thus very different from those of real assets. It is very possible for accumulation of real assets to occur, in fact every time a real asset is created net real wealth increases. The difference between real and financial assets has been the subject of economic research in the past. In fact Evsey Domar (Domar, 1946) derived his famous growth model by making an assumption about the relationship between the desired amount of physical capital and the financial side effect of an increase in capital to derive under what conditions the two could be balanced.

Domar assumed that the rate of increase of productivity was given by a parameter times the amount of investment. The previous assumption formed the real side of his model. The financial side of his model is derived from accounting, and is given by the assumption that the rate of growth of GDP depends on the rate of growth of investment. In order for the real and financial side of the economy to be in constant proportions Domar derives that investment must grow at an exponential rate.

The Domar model is, of course, extremely simple. In particular, the Domar model neglects the role of government. The central focus of Domar's paper, an investigation of the conditions under which the real and the financial side of the economy will grow at the same rate, has often been misunderstood. In fact, the Domar model fell out of favour among economists because of the instability of the solutions. Domar's fundamental conclusion was that the real and the financial side of the economy do not converge. Without government investment needs to take a very particular form for the real and financial side of the economy to remain balanced, and, in general,

nothing ensures that will be the case.

### **Adjustment between Real and Nominal in SFC and DSGE models**

Neoclassical models of the economy, such as Real Business Cycle (RBC) models, or the New Keynesian Dynamic Stochastic General Equilibrium (DSGE) models, describe the economy very differently than SFC models. In principle, these models are consistent with the SFC accounting framework that models the financial side of the economy but, in contrast to SFC models, the real side of the economy is their focus. It is, however, possible to understand the difference between RBC models and SFC models in terms of how the real side of the economy and the financial side of the economy are brought into balance. Nominal output is equal to the price level times real GDP by identity.

If nominal output changes, then either the price level or the level of real output must change. In SFC models, typically the price level is assumed to be roughly constant, and therefore changes in nominal GDP will increase real GDP proportionally. In contrast, neoclassical RBC and DSGE models typically assume that real output is fixed. Thus, adjustment necessarily occurs through price level. The non-impact of financial variables on the real side of the economy in neoclassical models is referred to as neutrality of money: since real variables are unaffected by financial variables adjustment occurs only through inflation.

There is an additional method that real GDP and financial GDP could move together: financial GDP could respond to changes in real GDP. However it is difficult to create realistic descriptions of the financial sector that have this property, and so, while some neoclassical models incorporate elements of the connection between financial and real GDP, most models primarily use either inflation or output adjustment. When nominal GDP changes, some combination of output and inflation must change to ensure that real and financial variables remain balanced.

### **Part II: Increasing vs Non-Increasing Cost Functions**

The realisation that a model must incorporate either a quantity adjustment or a price adjustment mechanism to ensure that real and nominal variables remain balanced immediately leads to the question of which adjustment mechanism occurs in practice. In order to understand that question it is necessary

to look into the details of how goods are produced in the economy.

If the costs of goods rises as the quantity produced increases then inflation ought to be the primary adjustment mechanism. If the cost of goods remains constant then output will be the adjustment mechanism.(graphic?) While SFC models do not typically incorporate production functions, the assumption that prices do not change with output, which is embedded in many SFC models, can be understood as the assumption that companies face constant or decreasing costs of production.

The most common reason for making a constant cost assumption is that returns to scale must be constant and that goods markets, in particular the labour market, are willing to supply more at the current price. For the labour market, that condition is equivalent to the belief that we are not at full employment. Constant costs can also be obtained with rising input prices, if a firm or the economy faces increasing returns to scale. In that case, increasing factor prices at higher production levels can be compensated for by improved efficiency of production leading to overall constant per unit prices.

If costs per unit are constant, then prices will be roughly constant as long as the markup over per unit costs is constant. Many SFC models do not explicitly incorporate inflation into their description of the economy. However, those SFC models that do, base their modelling on labour market power and markups. Such models have a mix of quantity and price adjustments. Models without changing markups have prices, and thus price levels, that do not change with output. Therefore, as nominal GDP rises real GDP must rise. Adjustment occurs through output and so the financial side of the economy determines aggregate production.

The adjustment mechanism in SFC models is also implicitly present in much of post-Keynesian economic thought. Modern Monetary Theory (MMT) proponents have claimed that, in practice, most inflation in first world countries is not caused by excessive monetary spending. The MMT claim can be interpreted in terms of cost functions in the economy; cost functions are understood to be constant or decreasing in output. Barring an increase in the markup over costs(which MMT proponents believe is usually the cause of inflation), real output must adjust as nominal output increases.



## Cost functions in mainstream economic thought

Neoclassical economics is often criticised for premising theory upon assumptions which are not valid, such as infinitely lived representative agents, the assumption of equilibrium, and the idea that consumers calculate the optimal consumption bundle out of an infinite number of potential choices. Neoclassical economic models also typically include production functions in which marginal costs increase with quantity. However, it is less commonly known how foundational the assumption of increasing marginal costs is to mainstream economic theory. Many of the points of contention between post-Keynesian and neoclassical economic thought are nullified if neoclassical models are modified to include cost functions that are not increasing.

The reliance on increasing cost functions in neoclassical economics begins with economics 101. The profit maximizing problem of a representative firm does not have a solution at a given price in the case of decreasing marginal costs, and in the case of constant marginal costs the solution is degenerate<sup>1</sup>. This means that the entire framework of perfect competition does not apply without increasing marginal costs. The fact that the price maximizing quantity of a firm is infinite without increasing marginal costs also implies that it is not possible to draw supply curves for a firm or industry in the case of non-increasing costs of production. Since much of econ 101 relies upon the supposedly foundational concepts of supply and demand without integrating increasing costs of production, most of the lessons of economics 101 cease to apply. Standard economic ideas about the effect of price controls, the efficiency of markets, and the efficiency of taxation are not discussed except in the special case of increasing costs of production.

Similarly, in more advanced economics such as Walrassian general equilibrium theory, on which most modern macroeconomic models are based, increasing costs of production are assumed. Constant returns to scale can sometimes be assumed but the costs of inputs are always assumed to be increasing. Standard economic assumptions about the efficiency of markets do not apply without the assumption of increasing costs.

---

<sup>1</sup>Firm profits are given by  $\pi = PQ - C(Q)$ , where  $\pi$  is profit,  $P$  is price, and  $Q$  is quantity.  $\frac{d\pi}{dQ} = 0$  implies  $P = \frac{dC}{dQ}$  at the critical points. To be a maximum  $\frac{d^2\pi}{d^2Q} = -\frac{d^2C}{d^2Q} < 0$ , which implies marginal cost must be increasing. When  $\frac{d^2\pi}{d^2Q} = 0$  inspection shows that any quantity is a maximum when  $P = \frac{dC}{dQ}$ , zero is a maximum when  $P < \frac{dC}{dQ}$ , and a firm wants to sell as much as possible when  $P > \frac{dC}{dQ}$ . Thus firms only want to sell a finite amount when marginal costs are increasing

The difficulties increasing returns (and therefore non-increasing marginal costs in some cases) creates for economic modelling, and the very concept of supply and demand, have been known since increasing returns was first discussed. Young(1928) discusses increasing returns as part of the reason for economic growth but due to the difficulties increasing returns creates for the competitive framework he is careful to emphasize increasing returns that are external to the firm. If returns are not increasing at the firm level but only on the industry level then efficiency can endogenously increase as the economy grows while competitive markets are maintained. However, much of Young's discussion was qualitative and as economics became increasingly mathematized it became difficult to discuss a topic such as the economics of non-increasing costs that did not lend itself to easy mathematical analysis. Kaldor (1966) also referred to increasing returns as part of his analysis of the slow growth of the United Kingdom, but his work was presented in the formal mathematical style that was, at that time, becoming fashionable in economics.

In 1977 Avinash Dixit and Joseph Stiglitz developed a model of imperfect competition. This model of competition allowed elements of increasing returns to finally be incorporated into some economic models. Dixit-Stiglitz competition is far from a perfect method to analyse perfect competition; its relevance is that it allowed increasing returns to be incorporated into formal neoclassical economic models without entirely ruining the competitive structure that neoclassical models tend to be based on.

Dixit Stiglitz model of imperfect competition was used to incorporate increasing returns into research in a variety of economic fields. Paul Krugman used it for his work in trade theory [Footnote], and economists such as Rommer used it to develop mathematical models of the ideas that Young outlined in the 1920s. It is interesting to read Romer's discussion of the implications of increasing returns on traditional economic results. He outlines the dramatic changes which the study of increasing returns has had on many economic fields and expresses hope that that impact will be extended. The 1980s were a time when macroeconomics was ready for a revolution. The postwar Keynesian synthesis was faced with increasing theoretical and empirical problems, and the new work allowing economic modelling to incorporate increasing returns ought to have had a large impact.

In 1980 Yew Kwang Ng (Ng 1980) published a paper that could have been the start of a revolution in macroeconomics, similar to the revolution that occurred in neoclassical growth theory and trade theory. In his paper, Ng

presented, in a simple model, the reason for neutrality of money in most neo-classical economic models. Firms' supply curves slope upward, and demand curves slope downward. If there is an increase in demand due to an increase in nominal values then initially firms will increase their prices, which will increase the price level. However, since firms' costs are based on the price level their costs will then increase. The increase in costs will compensate for the increase in demand and the end result is a change in the price level with no change in output.

Ng extended the above analysis to the case of imperfect competition. He showed that if firms set prices at a markup over costs, then if certain conditions were met, conditions which amount to the claim that costs of production are not increasing in quantity<sup>2</sup> then money is non-neutral. However, with imperfect competition, money remains neutral if costs are increasing. For similar reasons, money is neutral in the case of perfect competition. The argument for neutrality in the case of perfect competition still holds if costs are proportional to the price level and not equal to it.

Ng's work did not receive much attention. An article was written in response by Brian, Lambert and Turner(1982), in which they argued that the technical condition, not imperfect competition, was what led to non-neutrality of money. Their article is of course correct, as Ng acknowledged. Unfortunately further attention to the paper was limited. A natural response to Ng's paper ought to have been empirical work demonstrating that marginal costs are increasing, or including a discussion of the regime where it held in macroeconomic modelling, however few macro papers since Ng published his paper have done either.

By the middle of the 1990s the the consensus view within macroeconomics was that imperfect competition did not lead to the non-neutrality of money. Dixon and Rankin (1994) embody the consensus view when they argue that only with other market imperfections does imperfect competition lead to non-neutrality. They cite Blanchard and Kiyotaki (1987) as the clearest statement of the argument. However, Blanchard and Kiyotaki make assumptions about the production function and labour supply func-

---

<sup>2</sup>Since there are a large number of firms in Ng's model  $\frac{dC}{dq} = W/MPL$ , with C=cost, W=wage and MPL=marginal product of labour and q=quantity. taking the derivative we obtain  $\frac{d^2C}{dq^2} = -\frac{dMPL}{dN} * \frac{dN}{dq} \frac{1}{MPL} + \frac{dW}{dN} \frac{dN}{dq} \frac{1}{W}$ . Since we are only interested in the sign we can multiply by  $N$  and divide by  $\frac{dN}{dq}$  since both are always positive, giving us a condition for marginal cost not increasing which is the same as Ng's condition for money non-neutrality

tion in their paper that imply that Ng's technical condition does not hold, or equivalently that firms' marginal costs are increasing in output. No attempt is made in the paper to ascertain whether or not that assumption is justified empirically: the paper simply states that increasing returns to scale are ruled out to guarantee the existence of an equilibrium. Blanchard and Kiyokaki's views on the topic have become the mainstream view on the use of perfect competition in macroeconomic models despite this neglect. Most macroeconomic models since then follow them in simply assuming, with no justification, that Ng's technical condition does not hold.

For example in Smets and Wouters (2007), one of the most cited macroeconomics papers of the last twenty years, the authors assume a constant return to scale production function. When combined with the assumption that the elasticity of the labour supply and the elasticity of the demand for capital are positive, this guarantees that firms do not face constant or decreasing costs. Smets and Wouters paper is typical of much of modern macroeconomics: many modern papers use a very similar methodology. There has been a decreased interest in formal modelling within macroeconomics since the 2008 crisis, with a turn to more empirical work. Thus even now the state-of-the-art models used in macroeconomics are based on an unjustified assumption without which they would give dramatically different results <sup>3</sup>.

Since, with increasing costs, money is neutral, in order to explain financial crises, which are difficult to understand as being caused by changes in real variables, modern New Keynesian models add various short term frictions to their models. Typically, it is assumed that firms are not able to adjust prices instantaneously, but that in the long run prices are flexible. This assumption ensures that the behaviour of the models is able to superficially incorporate some level of money non-neutrality, while the fundamental behaviour of the model is unchanged. Short term non-neutrality of money is added to models in which money is fundamentally neutral in an ad hoc way in order to match

---

<sup>3</sup>Note that in some cases results that are similar to those of Ng can be obtained by assuming constant costs of production per unit and a fixed cost. The logic is similar to that used in papers discussing New Trade Theory. If firms are constrained to have zero profits, for example, by free entry then an increase in demand can sometimes result in constant or decreasing average costs, which can lead to firm behaviour similar that seen with non-increasing marginal costs. However, in order for fixed costs to give the nontraditional results, the number of firms in the economy must be allowed to vary. Assuming standard Dixit-Stiglitz competition with a continuum of firms rules out much of the interesting non-traditional behaviour.

empirical data. In contrast, if costs were assumed to be non-increasing then non-neutrality of money arises naturally from the analysis.

### **Duplication argument for Increasing Returns**

There is a theoretical argument in economics that is often used to argue that firms cannot face decreasing returns to scale. The argument is one from duplication: if a firm does face decreasing returns to scale then it could simply duplicate production arrangements that exist at a smaller scale and thus achieve at worst constant returns to scale [Footnote]. The duplication argument was known as early as Koopmans (1957) and today is well-known enough to be referenced in graduate economics textbooks, such as Mas-Colell and Whinston(1995). Mas-Colell and Whinston state that many economists use the duplication argument to justify assuming constant returns to scale. However, strictly speaking, the argument only sets a lower bound on the efficiency of production, it does not provide a reason to prefer constant returns to increasing returns. Assuming constant returns given the duplication argument would require that no elements of production benefit from being done at a larger scale, an assumption which seems somewhat dubious.

If constant returns to scale are assumed, the only way for the economy and for individual firms to face increasing costs of production is if input prices are rising [Footnote]. Generally the input prices driving the cost increases must be inputs that are not produced, or else the duplication argument would apply to their own prices. Therefore, things like labour, land, and natural resources are the inputs which must be driving cost increases. Increasing costs of financing could also drive cost increases in some instances depending on the nature of the financial system.

If, on the other hand, increasing returns exist to some degree, then in general, marginal costs will not be increasing, even with some level of increase in input prices. The relationship between input price increases and returns to scale that must exist in order for money to be non-neutral is given by Ng(1980)'s technical condition, which is that the elasticity of the labour supply is less than the elasticity of the marginal product curve

In input factor prices are not rising then the assumption of increasing costs of production which is so central to modern macroeconomics cannot hold. In particular, since labour is such a large part of the costs of most businesses, in a less than full employment situation mainstream macroeconomics cannot apply. Since more workers are willing to work at the current wage in a

less than full employment situation the elasticity of the labour supply is zero, and with constant returns Ng's condition for non-neutrality of money is guaranteed to be satisfied.

The idea that increasing costs must be driven by rising input costs of unproduced goods and labour opens up exciting opportunities for policymakers and economists. The prices of resources sold on the open market, and the prices of labour are largely set at the economy wide level, and so, unlike proprietary industry information, they can at least theoretically be measured. That means that, if inflation is becoming high, policymakers can potentially find more specifically what inputs of production are driving the increase and the target specific supply side policies to reduce it. For example, if there is a shortage of oil and rising oil prices are driving cost increases and therefore inflation in industry, macroeconomic inflation fighting tools might not be the ideal solution to the inflation. Reducing aggregate demand, to the point at which demand for oil drops, will reduce many forms of economic activity that are not oil dependent unnecessarily.

Instead, supply side policies that directly reduce the demand or increase the supply of oil are likely to be far superior inflation fighting tools. By measuring the price of unproduced inputs that are a substantial percentage of cost increases more precise inflation fighting mechanisms become possible. While western countries currently seem far from inflation constraints, this inflation fighting tool could be extremely useful if inflation begins to become a concern.. Less developed countries could also benefit from the improved understanding of inflation that can be reached by finding what input of production is causing the price increases.

### **Evidence against increasing cost functions**

Some of the implications of a perfectly competitive framework, if taken seriously, imply very counterintuitive results about individual firm behaviour. If firms face increasing costs of production, and the quantities they are normally selling imply that they are selling when price equals marginal cost, then they ought to not be willing to sell more than they usually sell at the current price. However it is extraordinarily unusual for firms to be turn down customers who make large orders and if such behaviour occurs it is typically a short term inability to produce within a short amount of time. Given enough time firms, far from raising prices, seem willing to offer discounts to consumers who produce bulk orders, behaviour which is incompatible with rising prices

of production.

Similarly, most firms seem to devote considerable effort to attempting to expand operations. Such behaviour is perfectly sensible in a world without increasing costs, however it seems strange that a firm would advertise to sell products when any increase in the amount of product they sell will cost them more than the current price to produce. However, in a world with decreasing costs, such behaviour by firms makes much more sense. If advertising increases sales, the greater expenditure on advertising is made up for by reduced costs of production, and overall profits increase.

Of course many of the above problems arise only in a perfectly competitive framework. Imperfect competition with increasing costs is not necessarily inconsistent with the above anecdotal outcomes, although perhaps the explanation is less intuitive than the explanation based on increasing costs. It is therefore useful to consider empirical evidence in order to answer the question about what cost functions prevail for most firms.

Empirical evidence that directly investigates the cost functions of firms, reveals that, as the duplication argument would indicate, few firms face increasing costs of production. Binder et al (1998) conducted a survey of over 200 industry executives and found that only 11 percent reported increasing marginal costs. Their work was initially designed to test the prevalence of the sticky prices used to arrive at short-run money neutrality by New Keynesian economic models; however, they instead found that Ng's technical condition for non-neutrality held. Their estimates are actually higher than previous research conducted in the same way, such as Eiteman and Guthrie (1952) who find that roughly five percent of businesses face increasing marginal costs.

An approach to measuring costs that has received more attention recently is to attempt to measure the overall level of returns to scale in industries. Research in this area tends to find a range of values for the aggregate level of returns to scale. Junius (1997) surveys many of the studies in this area and concludes that most of them find evidence of some degree of returns to scale. Lee and Ohanian (1999) discuss difficulties with measuring the degree of returns to scale from aggregate data, and conclude that such estimates suffer from several statistical problems that cannot be easily overcome, indicating that perhaps other sources of information on the presence of increasing returns are more reliable.

## Cost functions in modern macroeconomic research

While mainstream macroeconomic thought has largely focussed on the regime in which costs are increasing, some authors have explored the implications of decreasing or constant cost functions. Generally such work finds results that are similar to many claims made by post-Keynesian economists. Citation? The economy does not naturally reach an unique efficient equilibrium with non-increasing costs, which means that government intervention can potentially improve overall welfare. Some authors also find that expectations can be self-fulfilling. If participants in an economy believe that growth will be high, growth is high; while, if they believe growth will be low, growth is low. Such ideas provide justification for many of Keynes statements in Chapter 12 of the *General Theory*, in which he states that the fundamental business of the stock market is not determining real factors within the economy but guessing how others will act. Finally, some authors find that government spending increases can boost growth of the economy.

One strand of work that considers the implications of decreasing or constant costs is the work of Roger Farmer on indeterminacy, which is surveyed in Benhabib and Farmer (1999). The indeterminacy school of economics uses models similar to the DSGE models used by new Keynesians, but modifies them to incorporate increasing returns. These models can exhibit a multiplicity of equilibria. The indeterminacy school resolves the indeterminacy of their model by allowing agents self-fulfilling beliefs to determine which set of economic outcomes are realised. In this case, the fundamentals of the economy no longer determine the values of economic variables, which creates the possibility of the government improving economic outcomes.

Devereaux et al (1996) explore the effect of tax-funded government fiscal stimulus in an economy with increasing returns to scale. They find that, in the presence of increasing returns, an increase in entirely wasteful government spending can lead to increases in consumption for workers. The condition under which their nontraditional results hold is equivalent to that derived by Ng(1980).

Yew Kwang Ng has published additional work on returns to scale, much of which is collected in Ng(2009). Ng shows that equilibria can exist with increasing returns to scale in a general equilibrium framework. He also restates the argument of his earlier paper, that an increase in spending can increase output, and shows that it can be efficient for governments to subsidize industries that face a high degree of returns to scale.



The overall lesson that can be learned from neoclassical treatments of non-increasing costs regimes is that the results are dramatically different from those of standard theory. The exploration of non-increasing costs is not well explored within neoclassical macroeconomics, and so it is difficult to make any statements about potential results with certainty. However, the work that has been done seems to indicate results that seem in line with much of modern post-Keynesian thought.

An additional point that can be made from work incorporating increasing returns into standard economics models is that analysis becomes much more conceptually and mathematically difficult. Ng (1980), for example, notes that to deal adequately with the case of decreasing costs of production the economics of equilibrium is no longer adequate: analysis of system dynamics becomes required. In addition, the economics of competition in the presence of non-increasing costs becomes less straightforward. The standard prediction of neoclassical economic theory is, if production of a good faces increasing returns, then an industry is a natural monopoly. Some departures from the standard economic assumptions are therefore required.

Dixit-Stiglitz competition assumes that consumers prefer variety. However, it is not clear why the same firm cannot simply produce multiple varieties of the same product; in fact, few modern firms produce a single product. Dixit-Stiglitz competition is not a fully general theory of the competitive environment. There are other potential explanations for competitive behaviour in the presence of increasing returns; however, many of them are difficult to model with mathematical certainty.

Studying increasing returns in macroeconomics might necessitate abstracting away from competitive issues and simply assuming mark-up pricing. Making this assumption might at first seem ad hoc and less elegant than the methodology of the standard theories, but theories such as New Keynesian DSGE models are forced to make many ad hoc assumptions, such as sticky prices and unexplained exogenous shocks. in order to achieve sensible results. Studying increasing returns might necessitate more appeals to simplifying assumptions or stylized facts initially, but such research has the potential to explain non-neutrality of money without relying on ad hoc assumptions, potentially leading to fewer assumptions overall.

## General Implications of Non-Increasing Costs

Within mainstream economics there is widespread dissatisfaction with the New Keynesian DSGE macroeconomics models that are the most commonly used models in the field (Stiglitz, 2018; David and Wills, 2018). At the moment there is no consensus about an alternative, and no agreed upon way to improve such models. Moving towards a model which includes non-increasing costs of production is an avenue which ought to be further explored. Given the 2008 financial crisis, non-neutrality of money would seem to be an important element of macroeconomic analysis. Incorporating the non-neutrality of money naturally into the foundations of modern models would seem to provide a natural explanation for the importance of the financial side of the economy. In addition, the fact that in models without increasing costs the inflation constraint is far less important would indicate that such models are suitable for studying the recent macroeconomic situation, in which general inflation has been persistently low.

Incorporating non-increasing costs will likely involve substantial departures from both the mathematical techniques that are used in macroeconomics today and from some of the underlying economic theory. Competition with non-increasing costs is far more complicated, and there isn't a consensus explanation for competitive behaviour in such a situation. Appealing to insights from industrial organisation will likely be important as work in this area proceeds.

Economic education also has a strong tendency to neglect situations where non-increasing costs are prevalent. Economics students develop a set of intuitions and learn a set of tools that only apply in the presence of increasing costs of production. For example some graduate level economics texts on public economics treat perfect competition in a few chapters of a twenty chapter book, which implies that the vast majority of the analysis rules out non-increasing costs of production. If macroeconomics is to incorporate more realistic cost functions into the analysis, such cost functions need to be focussed on more in economics education.

For example, many students come away from introductory economics with the idea that supply and demand is the fundamental idea of economic analysis. Fewer realise that supply and demand analysis is only applicable to firms that operate in a perfectly competitive structure. When analysis of price ceilings, welfare economics and market efficiency is done in introductory economics, students understand that the models they are learning are

generally true. However, given evidence on the prevalence of non-increasing costs, many of the conclusions learned may only be true for a small minority of markets. The attention given within economics teaching to what amounts to a small percentage of the market leads to disproportionate research attention towards research based on those ideas, even if the empirical evidence to support that focus is weak.

There are also political consequences of the fact that economics teaching focuses on the economics of non-increasing costs. Students learn that taxes are not welfare improving, that unregulated markets are generally efficient, and many other statements that have political relevance, without understanding that those statements are only true under very limited sets of circumstances. This leads to the public perception that pro free market policies have the backing of economic research. Moving away from supply and demand in introductory economics education could fix elements of that perception. The focus on perfect competition also extends to the graduate level. For example, Public Economics by Myles Garrett only discusses imperfectly competitive markets in a single chapter (Garrett 1996).

### **Implications for post-Keynesian Research**

Post-Keynesian economics has received increased attention following the economic crash of 2008. In particular Modern Monetary Theory (MMT) has come to the attention of politicians, the media, and increasingly even mainstream economists. Mankiw, Krugman, Larry Summers and many other titans of mainstream economic thought have expressed interest in MMT and more generally post-Keynesian thought. However post-Keynesians and mainstream economists come from very different theoretical paradigms and so communication between the two is sometimes difficult. In addition, different elements of post-Keynesian thought are at different levels of development. Most of the attention mainstream theorists have given to post-Keynesian thought have been to post Keynesian macroeconomics, and in particular post-Keynesian financial macroeconomics. post-Keynesian microeconomics is less well developed.

In a sense, some of the issues facing post-Keynesian economics are the same as those that faced the work of Keynes after he wrote the General Theory. Keynesian macroeconomic ideas were revolutionary at the time they were published and received widespread attention due to their obvious relevance to the macroeconomic situation at the time. However, the ideas of the

General Theory were difficult to reconcile with the microeconomics of the time. Various attempts at reconciling the two were put forth, but most of the postwar Keynesian synthesis was both deficient on theoretical grounds (ie the stock flow inconsistency of Hicks IS-LM model) and not really consistent with the writing of Keynes. The theoretical weakness of the so-called Keynesian Synthesis became apparent in the 1970s and Keynesian ideas were largely abandoned within mainstream macroeconomics.

History threatens to repeat itself in modern economics. Interest in real Keynesian ideas has surged after the failure of ideas based on neoclassical microeconomics to explain macroeconomic phenomena. However without a theory of firm behaviour that is consistent with Keynesian macroeconomics the integration of it into mainstream economics cannot possibly proceed. Understanding Keynesian macroeconomics as the macroeconomics of non-increasing costs connects Keynesian ideas to microeconomics and allows them to be fully integrated into the mainstream economic curriculum.

In addition, understanding Keynesian economics as the economics of non-increasing costs also enables a better understanding of the situations where Keynesian economics does not hold. Since, as this paper has demonstrated, most firms do not face increasing costs, most of the time Keynesian economics is more relevant to understanding economic behaviour. However when firms do face increasing costs Keynesian ideas do not apply. Understanding this will enable post-Keynesians to be aware of situations in which inflation might become a concern. In particular, developing countries might well face increasing costs of production, which would explain the greater relevance of the inflation constraint in such cases.

Making the connection between Keynesian macroeconomics and non-increasing costs of production could be done quite easily. SFC models that model primarily the financial side of the economy could simply mention that in the case of constant costs output becomes the adjustment mechanism between the real and the financial side of the economy. This would enable mainstream economists to more clearly understand the point of departure of post-Keynesian and neoclassical models.

Similarly when critiquing mainstream macroeconomic models the assumption of increasing costs of production is an excellent focus point. Often too many critiques of neoclassical theory are made, obscuring the differences between neoclassical and Keynesian thought that are most relevant. If, when MMT proponents argue that inflation is not a relevant concern, they stated that the reason for the lack of concern about inflation was that firms were

far from the point at which their costs were rising, communication between mainstream economics and MMT would be improved. Putting some of the claims of MMT proponents into such neoclassical terminology could enable better communication between MMT proponents and neoclassical macroeconomists, facilitating a more productive research dialogue between the two schools.

## References

- [1] Garrett, Myles. *Public Economics*, Cambridge University Press, 2001
- [2] Stiglitz, Joseph E. "Where modern macroeconomics went wrong." *Oxford Review of Economic Policy* 34.1-2 (2018): 70-106.
- [3] Young, Allyn A. "Increasing returns and economic progress." *The economic journal* 38.152 (1928): 527-542.
- [4] Vines, David, and Samuel Wills. "The rebuilding macroeconomic theory project: an analytical assessment." *Oxford Review of Economic Policy* 34.1-2 (2018): 1-42.
- [5] Romer, Paul M. "The origins of endogenous growth." *Journal of Economic perspectives* 8.1 (1994): 3-22.
- [6] iyang, Zhao. *Prisoner of the State: The secret journal of premier Zhao Ziyang*. Simon and Schuster, 2009.
- [7] Domar, Evsey D. "Capital expansion, rate of growth, and employment." *Econometrica, Journal of the Econometric Society* (1946): 137-147.
- [8] Ng, Yew-Kwang. "Macroeconomics with non-perfect competition." *The Economic Journal* 90.359 (1980): 598-610.
- [9] Blanchard, Olivier Jean, and Nobuhiro Kiyotaki. "Monopolistic competition and the effects of aggregate demand." *The American Economic Review* (1987): 647-666.
- [10] Smets, Frank, and Raf Wouters. "An estimated dynamic stochastic general equilibrium model of the euro area." *Journal of the European economic association* 1.5 (2003): 1123-1175.

- [11] Hillier, Brian, Peter Lambert, and Roy Turner. "Macroeconomics with Non-Perfect Competition: A Comment." *The Economic Journal* 92.367 (1982): 701-705
- [12] Dixon, Huw, and Neil Rankin. "Imperfect competition and macroeconomics: A survey." *Oxford Economic Papers* (1994): 171-199.
- [13] Cole, Harold L., and Lee E. Ohanian. "Aggregate returns to scale: why measurement is imprecise." *Federal Reserve Bank of Minneapolis Quarterly Review* 23.3 (1999): 19-28.
- [14] Blinder, Alan, et al. *Asking about prices: a new approach to understanding price stickiness*. Russell Sage Foundation, 1998.
- [15] Eiteman, W.J. Guthrie, G.E., (1952). "The shape of the average cost curve", *American Economic Review* 42: 832-838.
- [16] Junius, Karsten. "Economies of scale: A survey of the empirical literature." Available at SSRN 8713 (1997).
- [17] Ng, Yew. *Increasing returns and economic efficiency*. Springer, 2009.
- [18] Benhabib, Jess, and Roger E. Farmer. "Indeterminacy and sunspots in macroeconomics." *HANDBOOKS IN ECONOMICS* 15.1A (1999): 387-448.