MACROECONOMICS: SCIENCE OR FAITH BASED DISCIPLINE?

Bill Russell*

26 April 2016

ABSTRACT

Whether or not macroeconomics is a science depends on the scientific nature of macroeconomic theories and how the discipline responds when the empirical evidence fails to match the underlying assumptions and predictions of the theories. By way of an example, four conditions for macroeconomics to be a science are developed and used to examine the 'modern' theories of the Phillips curve. It is found that while the discipline in general maintains one condition it routinely violates the other three. This suggests the macroeconomics discipline has some way to go before it can call itself a 'pure science'.

Keywords: Methodology, Phillips curve, inflation, structural breaks, non-stationary data, macroeconomics
JEL Classification: B41, C20, E31.

---

* Bill Russell, Economic Studies, School of Business, University of Dundee, Dundee DD1 4HN, United Kingdom. +44 (0)1382 385165 (work phone), +44 (0)1382 384691 (fax), email: brussell@brolga.net. I would like to thank David Bloor, Geoff Harcourt, David Hendry, Matthias Klaes, Tord Krogh, Donald MacKenzie, Ragnar Nymoen, Dennis Petrie, Michael Pollitt, Victoria Sparrman and Simon Wren-Lewis for their very thoughtful comments and encouragement to pursue this topic. I particularly thank Adrian Pagan who has carried out a long and patient (on his part) correspondence with me concerning the implied statistical process of inflation considered in the appendix.

1. INTRODUCTION

This is an applied methodology paper that develops a framework to answer the question: where does macroeconomics lie in the spectrum between 'pure' faith and 'pure' science?[1] This question initially appears not to have an answer as it has a number of dimensions and is complicated by the spectrum not being uniquely defined. However, in the next section it is demonstrated by way of an example that for a discipline to be a 'pure' science it needs to meet four conditions. These conditions identify the types of theories that are relevant for a science and how a discipline uses empirical evidence to examine the theories.

In Sections 3 and 4 the conditions are applied to the Friedman (1968) and Phelps (1967) (F-P) expectations augmented and New Keynesian (NK) theories of the Phillips curve as two examples of dominant macroeconomic theories of the past five decades. A similar analysis can be applied to other macroeconomic theories. It is found that while the 'modern' theories of the Phillips curve are scientific in the sense of Popper (1959) it appears that the empirical validity of the assumptions underlying the model and their associated predictions are somewhat compromised.[2] In particular, important model defining underlying assumptions of the F-P and NK models appear to be empirically invalid. It is common to defend the use of empirically invalid assumptions in macroeconomics on the grounds that agents behave 'as if' the assumption is true. This defence is considered in Section 5 before returning to the question concerning the status of macroeconomics as a science in Section 6. Finally Section 7 discusses briefly how this framework relates to the philosophy of science and concludes.

2. FOUR CONDITIONS FOR A DISCIPLINE TO BE A SCIENCE

Imagine an island community that does not have the technology to understand how the island is part of the world, solar system and universe. On this island the dominant theory of the 'sun discipline' is that the passage of the sun through the sky is controlled by the god Tron who at

---

1 This paper grows out of a long literature that examines a range of methodological issues surrounding macroeconomics including the thought provoking papers of Sims (1980), Mäki (1994, 2009), Sugden (2000), Caballero (2010), Fair (2012), Howitt (2012), Kozicki (2012), Uhlig (2012) and the contributors to Medema and Samuels (1996).

2 Popper (1959) defines a theory is 'scientific' if it produces empirically testable predictions.

the end of the day extinguishes and dismantles the sun before next morning reassembling, lighting and throwing it back up into the air from the other side of the island. Tron theory predicts that there is day and night and the sun 'goes down' before 'rising' the next day. These predictions are empirically falsifiable and therefore Tron theory is scientific in the sense of Popper (1959). Furthermore, as the predictions cannot be empirically falsified, Tron theory is a valid description of the data. One can imagine that over time Tron theory becomes more 'rigorous' utilising high level mathematics, internally consistent micro-foundations and optimising behaviour of some form. Note that these are notional characteristics of the theory and not necessary for the theory to be scientific.[3]

After many years an alternative theory develops that assumes (i) the island is on a planet, and (ii) the planet is spinning on its own axis with respect to the sun. Consequently, it is the spinning of the planet that causes the sun to set and rise on a daily basis. The alternative theory has the same predictions as Tron theory, it is also scientific and it is also consistent with the data. Given that the predictions of both theories are empirically valid and the underlying assumptions cannot be examined due to the lack of technology the dominant theory is the one with the most desirable notional characteristics. In this case Tron theory dominates the alternative theory as the later does not incorporate the god Tron, has no rigorous micro-foundations and is not mathematical.[4]

Consider now an improvement in technology whereby it is observed that the island is indeed on a planet spinning on its axis with respect to the sun. The predictions of both theories are still consistent with the data *but* now the underlying assumptions of Tron theory are empirically observed to be perfectly incorrect in the sense that the assumptions are never correct.

In this example the two theories are both valid empirical descriptions of the data which allows us to focus on the role of the underlying assumptions on whether the 'sun discipline' is a science. The Tron theorists may reject the empirical evidence concerning the assumptions and argue that while the assumptions of the model, and in particular the role played by the god Tron,

---

[3] Notional is defined as an idea existing only in theory and/or only in the imagination.

[4] It is likely that the proponents of the alternative theory will develop the same notional modelling characteristics as Tron theory in an attempt to avoid the criticism of not being 'rigorous'.

are not literally true the passage of the sun through the sky can be modelled ‘as if’ Tron exists. This response may be motivated by a non-scientific attachment to the underlying assumptions possibly due to the theorists not wanting their human capital, influence and status to be extinguished by the proponents of the alternative theory or they politically support the policy recommendations that follow directly from the underlying assumptions.[5] Alternatively, Tron theorists may reject or change their theory not because the predictions are invalid but because the underlying assumptions of the theory are empirically invalid.

At such critical moments in the evolution of theories not everyone in a discipline will respond in the same way. However, those who remain Tron theorists in the above example might be classified as ‘faith’ based theorists in the sense that they are faithfully adhering (for whatever reason) to previously held views in the face of the empirical evidence. Those who reject Tron theory and accept the alternative theory as a better description (but not necessarily a perfect description) of the data are scientists in the sense that their ‘test’ of the two models is based on empirical evidence as emphasised by Popper. Note that it is not whether the theorist prior to the new empirical evidence believed in Tron theory that determines if they are guided by ‘faith’ or ‘science’ but how the theorist responds to the new empirical evidence concerning the assumptions.

Extending this concept from the individual theorist to the discipline is more complicated. If Tron theory is rejected and the alternative theory becomes the dominant view then the discipline is a science. Alternatively if Tron theory co-exists in a significant way with the alternative theory following the new evidence then the discipline is composed of faith based and science based sub-disciplines. Finally, if the alternative theory is either rejected or marginalised then we can conclude the discipline is based on faith alone.

---

[5] See Dasgupta and David (1994) and Hands (2001) from the economics of science literature and Bloor’s (1991) ‘strong programme’ within the sociology of scientific knowledge for explanations of why researchers faithfully adhere to empirically invalid assumptions and methodologies.

This example suggests the following four axiomatic conditions for a discipline to be a science:

C1. Only theories that conceptually provide empirically falsifiable predictions are relevant.

C2. Theories are rejected in their current form when the predictions are empirically falsified.

C3. If more than one theory produces empirically valid descriptions of the data then the dominant theory is the one where the assumptions better approximate in an empirical sense the focus of the assumption.

C4. When an assumption of the theory is perfectly incorrect then the assumption along with the theory in its present form are rejected.

C1 and C2 are a restatement of Popper (1959) in terms of what theories are scientific and the role of falsification in the evolution of theories. In contrast, C3 and C4 focus on the assumptions to ensure they evolve in a way that is consistent with the 'real' world.

The standard hypothesis testing procedure that underpins falsification in C2 focuses on whether or not the predictions (or forecasts) of a theory are correct. For example, we might state a null hypothesis as the coefficient value predicted by a theory and then choose; (i) the test statistic; (ii) the distribution of the test statistic; and (iii) the level of significance for the test. This allows the critical values and associated 'reject' and 'non-reject' regions of the decision rule to be identified. Finally the test statistic is calculated from the estimated coefficient and the decision is made to either reject or not reject the null hypothesis.

Unfortunately we cannot repeat this procedure when testing the null hypothesis that an assumption is correct. This is because assumptions are mostly abstract concepts of what they are describing and so may be only approximately correct. The test statistic that we need to calculate is one that measures how good the approximation is but this is undefined as it depends on the prior views of the theorist and is therefore in the 'eye of the theorist'. Therefore, as we cannot estimate the test statistic, its associated standard error, and we do not know the distribution of the test statistic, there is no objective test of whether an assumption is approximately correct.

C4 solves this problem by re-forming the null hypothesis as the assumption is *perfectly incorrect* where the latter is defined as *never* observed to be correct. The corresponding test statistic is the proportion of times the assumption is observed to be correct and the rejection region of the decision rule is any non-zero value of the proportion. Furthermore, as there is no measurement error around zero (either it is zero or it is not zero) there is no standard error of the estimate and so the distribution of the test statistic, its standard error and the appropriate level of significance are all irrelevant to the test. If the test statistic is zero then we cannot reject the null hypothesis that the assumption is perfectly incorrect and the assumption is rejected. Alternatively, if the test statistic is non-zero then the assumption is approximately correct and whether it is a good approximation is undefined as it depends on the prior views of the theorist.[6] Finally, if the test statistic is 1 (i.e. the proportion is 1) the approximation is defined again as the assumption is perfectly correct in the sense that the assumption is always observed to be correct. This framework for testing an assumption is shown in schematic Diagram 1.

Underpinning C4 is the idea that any two theorists may not agree on how good an approximation an assumption is of 'reality' but they can agree on whether an assumption is never observed to be correct. This means that a well-defined perfectly incorrect point on the decision rule is a valid and objective critical value for the rejection of an assumption. To not reject the assumption when it is perfectly incorrect implies the theorist is faithfully retaining the assumption despite the empirical evidence that it is always observed to be incorrect. In the same way that falsification in C2 leads to the rejection of a *theory*, the finding that an assumption is perfectly incorrect leads to the rejection of an *assumption* in C4.

We sometimes observe in a 'science' based discipline an 'old' theory that contains a perfectly incorrect assumption continuing to be used and some might argue that this means C4 does not apply. On the contrary, the continued use of an 'old' theory demonstrates that theories with perfectly incorrect assumptions are rejected. This is because the usefulness of the old theory as a simple approximation of the new theory can only be understood from the perspective of the new theory. If the old theory is found to be a poor approximation it will not continue to be used. However, if the old theory is a good approximation of the new theory possibly over a

[6] This implies there is no unique one-to-one mapping between the test statistic and the 'goodness' of the approximation for all theorists. For example, if the test statistic is a proportion of 0.4 this may be a very good approximation for one theorist but a very poor approximation to another.

restricted range of conditions then the computational and theoretical simplicity of the old theory may mean that it has some advantage over the more elaborate new theory. However, so that the new, and dominate, theory can be developed the assumption underpinning the old theory has first to be rejected.[7] That is, C4 applies. This issue is returned to briefly in Section 5.

C4 can be explained with the following example. Assume we wish to test an assumption that a box contains green apples. In this case we are going down the left hand side of Diagram 1 as we can directly measure the test statistic of when the assumption is correct. The null hypothesis is that the assumption is perfectly incorrect and that the box contains no green apples. The test statistic is the proportion of green apples in the box and the decision rule contains three regions: perfectly incorrect when the proportion is zero, perfectly correct when 1, and approximately correct between 0 and 1. The box is opened and the contents are observed (i.e. measured). If there are no green apples in the box then we cannot reject the null hypothesis that the assumption is perfectly incorrect. C4 then suggests that we should reject the assumption.[8]

It may be that the perfectly incorrect point of the test statistic is not well-defined because of the form of the assumption or we cannot observe or measure directly when the assumption is correct. In this case we may be able to transform the assumption into an associated prediction where the test statistic has a well-defined perfectly incorrect point and this allows an indirect 'test' of the assumption. This case is shown on the right hand side of Diagram 1. For example, a theory may assume individuals are motivated by 'X' which cannot be measured directly and so the test statistic is undefined everywhere. However, the motivation assumption may imply individuals behave in a way that can be observed and so the test statistic (i.e. the proportion of cases where the prediction is observed to be correct) is well defined allowing the null hypothesis that the prediction is perfectly incorrect to be tested. If the null hypothesis cannot be rejected then the implied behaviour and the assumption are rejected simultaneously. That is, if the implied behaviour is never observed we may conclude indirectly that the assumption of the

---

[7] For example, Newtonian physics, Galileo's law of falling objects, and Boyle-Charles's law of ideal gases are three 'old' physics theories based on perfectly incorrect assumptions that continue to be used. In each case the perfectly incorrect assumption has been rejected so that a new more elaborate theory is developed that allows us to understand when the old theory is a good approximation of the new theory.

[8] An alternative null hypothesis is the assumption is perfectly correct. In this case if the test statistic is 1 then we cannot reject the alternative null hypothesis.

motivation of individuals is perfectly incorrect. Note that the reverse does not follow. If the implied behaviour is observed we cannot conclude the assumption is correct as the implied behaviour may also be consistent with another assumption.[9] The 'testing' of an assumption by its transformation into the implied behaviour of agents is used in Section 4 to examine some of the behavioural assumptions of 'modern' theories of the Phillips curve.

Return now to C4 which states that when an assumption is observed to be perfectly incorrect it is rejected. It is C4 that is critical in the evolution of scientific theories as it allows the discipline to focus on what is incorrect with the existing theory and encourages improvement specifically in that area. If the assumptions are not rejected when identified as perfectly incorrect then time is wasted building theories based on the empirically invalid assumptions instead of developing and exploring alternative empirically valid assumptions. Note that it is the empirical analysis of C2 and C4 that provides a 'real' anchor for the evolution of theories both in terms of improved prediction of the data *and* stopping theories from evolving along the lines of notional assumptions that are only valid in the imagination of the theorist.

The conditions allow for the co-existence of rival theories.[10] This is due in part to the falsification in C2 being subject to the Duhem–Quine problem that any test of a hypothesis depends on a number of ancillary assumptions that may or may not be true.[11] Therefore, a rejection of the hypothesis may be due to either the hypothesis or the ancillary assumptions being false.[12] However, if we believe the ancillary assumptions of the test are logically correct and not themselves empirically invalid then falsification can occur conditional on the ancillary assumptions being 'agreed' to be correct.

The co-existence of theories is also allowed under C3 because there is no one-to-one mapping between the assumptions and the approximation as argued above. However, if the assumption

---

9 Note also that when rejecting the implied behaviour of assumption X we are simultaneously rejecting all assumptions that generate the same implied behaviour.

10 Two theories are co-existing if the theories are based on different methodologies and assumptions and seek to explain the same outcome and the assumptions of both theories are approximately correct. Co-existence does not apply if one theory is a simplification (i.e. a restricted version) of the other as in the examples in footnote 7.

11 It may be that the predictions cannot be tested given available technology. This allows the co-existence of rival theories until technology improves when C2, C3 and C4 apply.

12 See Gillies (1998).

(or the transformed prediction) is not rejected when observed in only 3 per cent of cases the theorist would look somewhat 'non-scientifically' attached to the assumption especially if the alternative assumption is observed in 80 per cent of cases.[13]

The hypothetical 'sun discipline' example above mirrors numerous examples in the physical, chemical, medical and biological sciences where the predictions of the competing theories are identical and so cannot be used to distinguish between the theories. In these cases the competition is often resolved with the advent of empirical evidence that a 'model defining' assumption underpinning one of the theories is found to be invalid or 'perfectly incorrect' in the terminology used above. For example, Miasma theory persisted for many centuries and explained the spread of diseases such as cholera and the plague by miasma or 'bad air'.[14] This theory is scientific in the sense of Popper and consistent with the data in that it explains the spread of disease. However, from 1546 when Italian physician Girolamo Fracastoro suggested that diseases were transferred between people by small spores an alternative theory of the spread of diseases existed which was also consistent with the data. The alternative theory was eventually accepted as valid in the mid to late 19th Century when improved technology allowed the identification of germs as the carriers of disease and Miasma theory was rejected. Of course the medical profession could have argued that disease spread 'as if' by miasma but this would have violated C4 indicating medicine was a faith based discipline and not a science.[15] Furthermore, if medicine had persisted with Miasma theory one may hypothesise how this would have hindered (or stopped) all the developments that followed the discovery of germs and the necessary rejection of Miasma theory.

---

[13] The philosophy of science when examining how science is organised, undertaken and progresses provides some understanding of why rival theories co-exist under C2 and C3. For example see Kuhn (1970), Lakatos (1978), Bloor (1991), Dasgupta and David (1994), Bird (1998) and Hands (2001).

[14] A more contemporary example that has not yet reached a conclusion is the alternative theories underpinning symmetry breaking in the standard model of particle physics. The Higgs field and associated boson is one of a number of theories that explains symmetry breaking and this theory received considerable support following the identification of a particle using the Large Hadron Collider on 4 July 2012 which was shown subsequently to have many characteristics in common with the Higgs boson.

[15] In this case Miasma theory from the perspective of 'germ theory' is a poor approximation of the latter and so does not continue to be used.

3. EMPIRICAL VALIDITY OF THE 'MODERN' THEORIES OF THE PHILLIPS CURVE

In this section we consider the empirical validity of the 'modern' theories of the Phillips curve to examine if macroeconomics conforms with C1 and C2. 'Modern' theories of the Phillips curve can be thought of in terms of restrictions to the hybrid Phillips curve:

$$\pi_t = \delta + \delta_f E_t(\pi_{t+1}) + \delta_b \pi_{t-1} + \delta_\mu \mu_t + \epsilon_t \quad (1)$$

where inflation, π_t, depends on expected inflation, $E_t(\pi_{t+1})$, conditioned on information available at time t, lagged inflation, π_{t-1}, and a 'forcing' variable, μ_t. In the F-P Phillips curve $\delta_f = 0$ and $\delta_b = 1$ and agents are purely backward looking with adaptive expectations. At the other extreme, the NK Phillips Curve of Clarida, Galí and Gertler (1999) and Svensson (2000) agents are purely forward-looking with rational expectations and $\delta_f = 1 - d$ and $\delta_b = 0$ where d is the discount rate. Finally, in the hybrid models of Galí and Gertler (1999) and Galí, Gertler and Lopez-Salido (2001) agents are both backward and forward looking and $\delta_f + \delta_b = 1 - d$. The F-P theory predicts the dynamic inflation terms sum to one while in the NK theory the coefficient on expected inflation is the discount rate. However, the empirical NK literature largely ignores this and considers the sum of the dynamic inflation terms to be one. For simplicity of exposition the sum of the coefficients on the dynamic inflation terms is assumed to be one in the NK and hybrid models unless otherwise stated.

'Modern' Phillips curve theories predict that inflation is an integrated or near integrated process as $\delta_f + \delta_b$ is either 1 or approximately 1. This assertion is straightforward and can be explained by the following proof-by-contradiction.[16] Assume (i) there are two mutually exclusive and exhaustive states of the world whereby inflation, π_t, is either an integrated I(1) process or a stationary I(0) process, and (ii) the forcing variable μ_t in equation (1) is a stationary I(0) process.[17] Consider now we estimate with ordinary least squares the hybrid Phillips curve of equation (1). When inflation, π_t, is an integrated process of order 1 then future inflation, π_{t+1},

---

[16] This argument is made in more detail in Russell (2016).

[17] Conceptually the forcing variable represents the business cycle which is mean reverting and therefore a stationary I(0) process. If the forcing variable appears to be integrated then it is most likely that breaks in the series have not be accounted for in the application of the unit root test as explained by Perron (1989).

and past inflation, π_{t-1}, are also integrated of the same order. Furthermore, we can ignore the stationary forcing variable as it will not enter the asymptotics of the estimation. Finally, standard cointegration theory indicates that when π_t, π_{t+1} and π_{t-1} are I(1) processes then $\delta_f + \delta_b = 1$. If this is not the case and $\delta_f + \delta_b < 1$ then $\Delta\pi_t$ would be I(1) and therefore π_t is I(2) which contradicts our original assumption that inflation is I(1). Conversely, if inflation is I(0) then $\delta_f + \delta_b < 1$. Again, if this is not the case and $\delta_f + \delta_b = 1$ then inflation is I(1) which contradicts the original assumption. These conclusions can be generalised with respect to the sum of the dynamic inflation terms in the Phillips curve and, by implication, if the sum of the dynamic terms is close but not equal to 1 then the statistical process of inflation is near integrated.

Therefore, we can conclude that on the assumption that the Phillips curve is a 'true' description of the data, 'modern' theories where $\delta_f + \delta_b = 1$ implies, or predicts, inflation is an integrated I(1) process otherwise the theory would be inconsistent with the data and we would conclude the theory is not a 'true' descriptions of the data.

This prediction has a number of important shortcomings. One, the statistical process of inflation in the 'modern' theories is due to the characteristics of agents. In the F-P model the dynamic terms sum to one because agents hold adaptive expectations while in the NK model the coefficient on expected inflation is the discount rate of agents. This is contrary to standard monetary theory where the long-run rate of inflation is determined by the behaviour of central banks partly in response to the shocks experienced by the economy.

Two, if inflation is an integrated process as in the F-P theory then central banks accommodate all shocks and hold no target rate of inflation. The implications for central banks from the NK theory are similarly unbelievable. If households and firms are risk neutral, face a symmetrical loss function in the region of the optimum price and their expectations about future prices are rational and unbiased then assuming a real annual interest rate of four per cent the quarterly value of δ_f is of the order of 0.99. This implies that the mean adjustment lag in the NK theory to a shock to inflation is around 25 years which on a practical level implies that central banks again do not hold a target rate of inflation.[18] Note that this conclusion does not apply to any one central bank but applies to central banks of every country at all times in history. To argue that central banks never hold a target rate of inflation is hard to sustain.[19] Furthermore, it is hard to accept that central bankers are entirely passive players in the inflationary process and by implication drawing salaries for not running monetary policy.[20]

Three, inflation cannot be a 'truly' integrated process as it appears to have an upper boundary at some moderate rate and a lower boundary around zero in the developed world.[21] The bounded nature of inflation is reinforced by what it implies for the price level. If inflation is an integrated process of order 1 then the price level is integrated of order 2 which cannot be correct because prices by definition are non-negative and therefore bounded at zero. The problem for both the F-P and NK theories is that they predict 'inflation is always and everywhere an *integrated* phenomenon'.[22] Consequently, both theories argue that the statistical processes of inflation is a singleton set containing only the element that inflation is an integrated process. This implies that when inflation is a stationary process (which it may have been since the early 1990s) then both the F-P and NK theories are invalid empirical descriptions of the data. The only way the 'modern' theories can be consistent with the data is if we believe that inflation cannot, is not, and can never be a stationary process at any time in history and in any country. If we are unwilling to accept this belief then we must be equally unwilling to accept the 'modern' theories of the Phillips curve as a general description of inflation.

---

18 Calculated as $\delta_f/(1-\delta_f) = 0.99/(1-0.99) = 99$ quarters or approximately 25 years.

19 Note that in this case the concept of the 'Volker deflation' at the start of the 1980s is misplaced as it implies the Governor of the Federal Reserve Bank of America targeted a lower rate of inflation which would be out of his control if inflation is an integrated process.

20 It is ironic that the NK theory is so popular in central banks given what it implies about the behaviour of central bankers.

21 While some economies have experienced periods of positive hyperinflation there has not been any recorded instance of an economy experiencing negative hyperinflation. It may be that 10,000 years of recorded history is insufficient time to conclude the probability of negative hyperinflation is zero but it appears a 'reasonable' conclusion given available information. In any case, given price adjustments are in discrete units, negative hyperinflation would conceptually disappear once the price level quickly hits the lower bound of zero.

22 This paraphrases Milton Friedman famous quotation from his Wincott Memorial Lecture delivered in London on 16 September 1970.

Four, some observers may argue that the F-P or NK theories are correct but there is a missing equation such as a Taylor Rule that describes the behaviour of the central bank.[23] If the Taylor Rule dominates then inflation is a stationary process around the central bank's target rate of inflation and $\delta_f + \delta_b$ must be less than 1 by the definition of stationarity. Alternatively, if either the F-P or NK theories are correct and agents dominate the inflationary process then the Taylor Rule is redundant and conceptually incorrect.[24] Given inflation is not an integrated process then we might conclude that the central bank dominants the F-P and NK agents in the inflationary process implying the 'modern' Phillips curve theories are conceptually incorrect.

And finally five, the behaviour of agents in the 'modern' theories is questionable. These agents are very sophisticated and well informed so as to undertake the highly complicated optimising processes in the theories. It is therefore illogical to believe that such sophisticated and well informed agents in these models would make the systematic error (more importantly the permanent error) of thinking that inflation is always and everywhere an integrated process given the bounded nature of inflation.

What then is the likely statistical process of inflation? 'Modern' theories of the Phillips curve argue that with no change in monetary policy and with mean zero inflationary shocks then inflation will vary around the long-run rate of inflation. A change in monetary policy results in inflation converging on and varying around a new long-run rate of inflation. We might therefore expect that inflation is a stationary process around a shifting mean. The latter is due to changes in monetary policy and allows for the numerous expected rates of inflation that are central to all the 'modern' theories of the Phillips curve since Friedman (1968) and Phelps (1967). If inflation is a stationary process around a shifting mean then the unit root commonly found in inflation data has no behavioural relevance and is simply due to not accounting for the structural breaks in mean inflation when testing for a unit root.[25]

---

23 See Taylor's (1993) seminal article and the extensive 'Taylor Rule' resources on his personal web site.

24 When $\delta_f + \delta_b = 1$ in the 'modern' Phillips curve then, as argued above, inflation must be an integrated process with no long-run rate of inflation. This fundamentally conflicts with the concept of the Taylor rule where inflation is a stationary process around a target rate of inflation due to the actions of the central bank.

25 Perron (1989) argues that stationary processes with breaks are easily mistaken for integrated processes.

In summary, the prediction of the 'modern' Phillips curve theories that inflation is an integrated process is hard to sustain. This prediction appears at the end of the 1960s with Friedman (1968) and Phelps (1967) and has not been rejected on empirical grounds even though (i) inflation appears bounded, and (ii) the rhetorical underpinning of the theories themselves suggest inflation is a stationary process around periodic shifts in mean due to changes in monetary policy. The failure to empirically falsify the F-P model and reject the basic stylised fact that inflation is an integrated process means that the objective of subsequent theories of the Phillips curve has been to explain the incorrect 'stylised fact' that inflation is integrated. Consequently, while we can say that the 'modern' theories of the Phillips curve are scientific and meet condition C1 the macroeconomics discipline has consistently violated condition C2 in not rejecting the invalid prediction of the 'modern' theories of the Phillips curve. In a real sense macroeconomic 'theory' has come to dominate and ignore empirical reality for an extended period of time.

4. EXAMINING THE MODEL DEFINING UNDERLYING ASSUMPTIONS

In this section we make use of the perfectly incorrect point of the test statistic to examine three model defining assumptions of the 'modern' Phillips curve theories and see if the macroeconomics discipline conforms with Condition 4.

4.1 Marginal Costs equal Marginal Revenues

The advent of marginal analysis in the mid to late 19th Century by Jevons (1871), Menger (1871) and Walras (1874-77) appeared to solve how firms set, or take, prices and choose the level of output.[26] Since then 'acceptable' macroeconomic models assume firms maximise profits requiring marginal costs to equal marginal revenues. The profit maximising solution is relatively straightforward in theoretical models given the underlying assumptions concerning the abilities of the agents, available information and the form of the production function.

[26] Discussing marginal costs and revenues makes it appear the argument is straying into microeconomics. However, the 'accepted methodology' of modern macroeconomics is to use optimising micro-foundations which assume profit maximising firms where marginal costs equal marginal revenues.

However in practice marginal costs and revenues are not just difficult to measure they are undefined.[27]

Marginal costs are undefined as all output is a joint product of the labour, non-labour and capital inputs.[28] That is, all outputs have at least one input that simultaneously produces more than one output.[29] Examples of how inputs produce joint products include (i) managers of a car plant producing more than one model of car; (ii) sales assistants in a supermarket selling more than one product; (iii) a lamb slaughtered to produce a number of cuts of meat; (iv) a lathe in a workshop producing a range of metal items; and (v) the cost of electricity and cleaning products in a warehouse storing many lines of chocolates and other sweets.

There are three forms of joint products. The first are the joint products of labour where there is no clear delineation between fixed and variable costs. This simple but insightful observation is the basis of much of Kalecki's work.[30] He recognises there is a category of labour that does not change with small changes in output which he terms 'overhead' labour and should be treated as a part of fixed costs. For example, the number of managers, cleaners and security guards do not change with small increases in output. Therefore there are as many measures of marginal costs as there are delineations between variable and overhead labour available to the firm.

The second form is the joint products of non-labour inputs. Consider, a lamb that produces a range of joint products including legs of lamb and lamb cutlets. What is the marginal cost of a

[27] Turvey (2000) argues marginal costs can only be calculated using forecasted changes in costs and for a substantial change in output. Consequently while the concept of marginal cost may be agreed between economists, in practice the calculation requires an agreed set of rules. Taking Turvey's argument to its logical conclusion this implies there are as many measures of marginal costs as there are infinite sets of rules and therefore marginal costs are undefined. See also Turvey (1969).

[28] There is a long literature on the concept of joint products. Important early work includes Marshall (1920, 1927) and Sraffa (1960), and more recently Baumol (1976, 1977), Panzar and Willig (1977) and Willig (1979).

[29] Some observers may feel that arguing *all* products are joint outputs is an exaggeration and cite single output firms such as water and electricity utilities as counter-examples. However, while these firms appear to have only one output, in practice the supply of the output differs by geographic location. Consider two end-users 'a' and 'b' and the supply network passes 'a' to service 'b'. In this case the marginal depreciation costs of the network before 'a' need to be apportioned to both 'a' and 'b' which is not uniquely defined implying marginal costs are undefined even with these apparently single output firms.

[30] For example see Kalecki (1954, 1971).

leg of lamb? If the lamb is slaughtered for a single leg then the marginal cost is very high. If a range of products is produced from the slaughtered lamb then the marginal cost is lower. The marginal cost depends on how the cost of the lamb (as well as the costs of the person who slaughters the lamb, depreciation and other input costs) is apportioned to the individual outputs.

The third form is the joint product of capital. Leaving aside how we delineate between capital and non-capital inputs, marginal costs should include a charge for the 'true' depreciation of fixed capital. Consider again the example of slaughtering the lamb. The marginal cost of the leg of lamb should include the actual cost of the accelerated depreciation of the equipment involved in the production process and this is not just a simple engineering problem to determine. A single cut with the knife may split the lamb into a leg and the remainder. How the accelerated depreciation due to the cut is apportioned between the leg and the remainder affects the marginal cost of the leg of lamb.

Apportioning the joint costs in the above three examples to the marginal cost of a product can follow many 'rules'. For example, the costs could be apportioned in the ratio of the prices of the outputs or the ratio of the sales revenues generated by the outputs. As there is an infinite number of cost apportioning rules there must be an infinite number of measures of marginal costs and therefore the marginal cost of a product is not defined.

We turn now to how the discrete nature of production and sales also renders marginal costs and revenues undefined. What is the marginal cost of producing one extra unit of a car? The answer depends, in part, on the time period involved. The marginal cost of one car per year will be different from one car per quarter, month, week, day, eight hour shift, hour, minute, or second. The time period involved will determine if the firm needs to employ more labour, capital, or even change the production process. Marginal costs depend on an arbitrary choice of time period. Similarly the marginal revenue for the sale of one more car depends on the time period being considered. When production is discrete and not continuous as in macroeconomic theory, the measures of marginal costs and revenues depends on the time period involved and as the time period is not defined uniquely then marginal costs and revenues are again undefined.

Firms therefore cannot calculate unique 'true' measures of marginal costs and revenues. Each measure depends on assumptions concerning the time period and how to allocate costs to joint products. It follows therefore that marginal costs and revenues are undefined and the assumption that marginal costs equal marginal revenues in equilibrium is perfectly incorrect.

4.2 The Business Cycle is due to Mistaken Inflation Expectations

In the Friedman and Phelps expectations augmented Phillips curve the business cycle is driven by mistakes made by labour (the asymmetric information assumption) concerning the expected rate of inflation and the subsequent correction to those mistakes (adaptive expectations). These fundamental assumptions help define the F-P theory of the Phillips curve but are unobservable and therefore the test statistic and decision rule to reject the assumptions are undefined everywhere. However, taken together these two model defining assumptions make very strong predictions about the behaviour of agents that can be examined empirically. For example, if we start in long-run equilibrium the dynamics of inflation over the business cycle begins with a loosening in monetary policy leading labour to make mistakes concerning the relationship between actual and expected inflation. Due to the asymmetric information assumption labour mistakenly believe the real wage has increased while firms (making no mistakes) correctly understand the real wage has decreased. Consequently labour supplies more labour because of their mistake and firms are happy to employ them. The business cycle draws to an end because labour corrects their mistakes by way of adaptive expectations leading the real wage to return to its long-run level along with employment, unemployment and output.

The difficulty for the F-P model is that empirically the behaviour implied by these underlying assumptions is not observed. First, while mainly unskilled labour enters employment at the start of the business cycle and leaves at the end, it is not the exact same labour. Consistency demands that if individual ‘a’ makes the initial mistake it should be individual ‘a’ who leaves once they correct their mistake. It cannot be (as appears to be the case) that when individual ‘a’ makes the mistake at the start of the business cycle individual ‘b’ leaves employment at the end when the mistake is corrected.

Second, this author is unaware of any instance either anecdotal or recorded in the literature whereby someone has explained to their employer and others that their decision to quit their employment was due to their surprise at how high prices are.[31] This is not the same as quitting because real wages are unexpectedly low due to an unexpectedly low nominal wage but the real wage is low *because prices are unexpectedly high*.[32]

Third, measuring the business cycle in terms of ten to twenty years is inconsistent with how highly sophisticated and information rich the agents are that populate the F-P theory. Such clever agents should 'wise-up' and correct their mistaken price expectations much quicker than the length of the business cycle. For the business cycle to persist for the time that is common in the data would require repeated unanticipated loosening of monetary policy through the entire business cycle. This would allow labour to mistakenly believe the real wage has systematically risen so as to continue to supply the extra labour for the length of the business cycle. But we see no such systematic increases in mean, or long-run, rates of inflation over the business cycle. On the contrary, the standard empirical Phillips curve literature proceeds on the assumption of up to a maximum of 3 breaks in the mean rate of inflation (i.e. only three changes in monetary policy) over the past five decades in the United States. Of these three breaks one or possibly two are associated with a loosening in monetary policy. Sophisticated agents are unlikely to be 'fooled' for long by such infrequent changes in policy and having corrected their mistakes bring the business cycle to an end well before periods of ten to twenty years.

One might argue the problem is the form of the asymmetric information assumption. The Lucas (1973) 'surprise' aggregate supply function makes the reverse assumption that firms and not labour make the mistakes. Unfortunately the implied behaviour in this model is equally implausible. In this model, a person enters a shop offering \$1.05 for a loaf of bread with a price tag of \$1. The manager recognises this as a 'surprise' in the sense that the manager cannot disentangle the offered higher price between an idiosyncratic increase in the demand for bread from their shop and a general increase in the price level due to a loosening in monetary policy. If it is the former the manager should increase the quantity of bread sold but if the later then

[31] There appears to be no evidence that the reverse behaviour occurs as an economy begins to expand after passing the trough of a recession. That is, there is no record of anyone ever explaining their re-found desire for employment as due to the *unexpectedly low level of prices* (i.e. the real wage is unexpectedly high).

[32] I would like to hear from any reader who can document an incident whereby a person has quit their employment and (unprompted) explains their decision with reference to their 'surprise' concerning the high level of prices.

only one loaf of bread should be sold. Given the confusion, the manager sells one loaf of bread, and on opening another loaf, hands over two more slices of bread. Consequently, the mistaken price expectations of the firm have led to an increase in the quantity of bread sold. This behaviour is as unobserved as that implied by the F-P model and suggests it is not mistaken price expectations by either labour or firms that drive the business cycle. We can therefore conclude the predictions based on the behavioural assumption that the business cycle is driven by mistaken price expectations are perfectly incorrect. And it follows therefore that the assumptions are also perfectly incorrect.

4.3 Agents are Forward Looking

In some ways the NK literature is a response to the strange and unseen behaviour implied by the price-taking F-P model outlined in the previous section. The initial richness of the NK model is that firms set prices which coincides with the observed behaviour of firms. However, the optimising forward looking behaviour by households and firms in the NK theory introduces its own set of implausible and unseen behaviour. In particular the optimisation of consumption in the NK theory requires: (i) the inter-temporal consumption elasticity of substitution to be a positive non-zero number; (ii) agents can price all future unknown products; and (iii) agents can predict all changes in future relative prices. Each of these requirements are necessary conditions for agents to optimise their consumption over time as in the NK theory and we consider each in turn.

4.3.1 The inter-temporal consumption elasticity of substitution

In the NK theory agents optimise their consumption over time and within each time period. For this to be possible it is necessary for the motivation of the consumption in the future to be transferred over time. For example, consider the infinitely lived households that optimise consumption over an infinite time horizon. Imagine the representative household sees the relative price of a particular model of a car increases ten years into the future. The implication is the household responds by buying more 'car' now in period *t*. But this implies a car now is a substitute for a car in ten years' time which it cannot be because you need a new car in ten years' time. To be a substitute the *motivation* for consuming the car in 10 years' time must be transferable to buying a car in period *t*. If this is not possible then the price elasticity of substitution between periods t and t+10 is zero.

This argument can be explained by considering oranges and lemons which may be considered substitutes in the same time period. The motivation for consuming both of these citrus fruits is based on their similar taste and the need for calories and vitamins. An increase in the relative price of lemons will lead to an increase in the demand for oranges as the motivation for consuming lemons transfers to consuming oranges. Now consider the case where the agent is mortally allergic to oranges. In this case the motivation cannot transfer to oranges and therefore oranges are not a substitute for lemons no matter how the relative price of lemons changes. In the same way, an increase in the relative price of a car in 10 years' time does not lead to an increase in the demand for cars today because the motivation for buying the car in 10 years' time is the desire for a new car and not a 10 year old car. If people do optimise their consumption decisions over time as in the NK theory then we should frequently hear people explain their consumption habits today by saying they are buying 'more' or 'less' today because the price of the substitute in the future is high or low. This inter-temporal behaviour is almost never observed in agents and when observed it is for a very special (and very small) sub-set of economic decisions.

One might ask what the sub-set is and in particular over what time horizon agents might substitute their consumption. From the discussion above it depends on how many periods the motivation can be transferred and the product can be stored in the same condition as when purchased in the future. For example, you may need a cheese sandwich for a business lunch in 10 days' time. Even if the sandwich was free today it is unlikely that you will purchase the sandwich today because it will spoil before it is needed. While a car purchased today may be approximately in the same condition as one purchased in one months' time it is not the same as a new car in ten years' time.

4.3.2 Forecasting the price of unknown products

To optimise consumption over time agents must be able to forecast the future price of unknown products and this is, at best, fanciful. Consider the car example again. The design, make and content of the car ten years in the future are unknown even to the manufacturer today. Also unknown are the specifications of all competing cars and substitute forms of travel which depend on how technology and taste will evolve over time. Assuming the price of cars is partly related to the services they provide then it is not possible for a household to identify a price of a car with unknown characteristics among a lexicography of unknown forms of travel.

4.3.3 Agents predicting structural breaks

For agents to optimise their consumption and other economic decisions over time they need to be able to predict changes in relative prices. To do this an agent must be able to predict the structural breaks in the mean price inflation for every product. Consider another proof-by-contradiction. Assume that for N periods the price inflation of good 1 is stationary with a constant mean. If agents can predict the future shift in mean in period N-k this implies inflation in the last k periods will start to adjust to the new mean and will have a different mean to the first N-k periods. This contradicts our initial assumption that inflation has a constant mean for all N periods and so we can conclude that agents cannot predict the impending shift in mean inflation.

Return now to agents optimising over the future. In the NK theory agents are forward looking in their economic decisions and to do this they need to be able to predict changes in relative prices. This requires agents to be able to predict the structural breaks in the mean rates of inflation of each and every product. However, as argued above, agents cannot logically predict structural breaks in the mean rates of inflation and therefore logically we can conclude that agents are not forward looking in the sense used in the NK theory. This does not mean agents do not want to know the future only that they are unable to predict the structural breaks so that they can undertake the forward looking optimising behaviour underpinning the NK theory. In important ways the unknown future is exactly that, unknown.[33]

The reader can test if they conform to the NK theory of optimising consumption by asking how often they give any consideration to future prices when undertaking the myriad of transactions every day. The lack of any consideration is partly because of the inability to transfer the motivation across time periods; partly due to the difficulty in predicting the characteristics of unknown future goods and services; and partly due to the inability to identify changes in relative prices when inflation is a stationary process around a shifting mean. At first the reader may think of some very special cases, possibly over very short time horizons, where predicted future

---

[33] Caballero (2010) asserts 'real people' do not have the skills and the information to undertake the optimising behaviour attributed to agents by 'core' macroeconomists. As argued above it is logical that agents cannot predict the structural breaks in mean inflation and therefore there is a logical empirical basis to Caballero's assertion.

prices are thought relevant. However, to optimise consumption over time as in the NK theory requires agents to hold predictions of *all* future prices and relative prices out to the infinite time horizon (including prices of unknown products) and not just a very small sub-set of prices. Unless the reader believes agents hold these predictions it is necessary to accept that agents do not have the necessary information to optimise consumption as in the NK theory.

In summary, it appears the three underlying assumptions of the F-P and NK models considered above are very poor approximations of actual behaviour. First, marginal costs and revenues are undefined and therefore the assumption that they are equal in equilibrium is perfectly incorrect. Second, the model defining assumptions of both the F-P and NK models also appear to be perfectly incorrect in that the implied behaviour of agents is never observed in the F-P theory and converges on never being observed in the NK theory when considered from the perspective of an infinite time horizon. Finally, the idea that agents possess the information concerning all prices of unknown future products and structural breaks in price inflation that is necessary so that agents are forward looking in the way assumed in the NK theory is logically indefensible and therefore perfectly incorrect.

5. THE 'AS IF' DEFENCE

An economic theorist may argue that 'we know firms and agents do not actually behave in the way we model them but we can model them 'as if' they behave in this way'. The clearest exposition of this argument is Friedman (1953). Friedman begins by asserting that how well a model predicts is the best test of a theory. He then proceeds to give the example of an object dropped in a vacuum will travel a distance given by Galileo's Law of Falling Bodies with the formula $s = \frac{1}{2} g t^2$ where s is the distance travelled, g is the acceleration due to gravity, and t is the time elapsed. This theory is tested repeatedly and accepted if not rejected by the data. Importantly, Friedman states that this does not mean that the model assumes a vacuum but that 'bodies that fall in the actual atmosphere behave as if they are falling in a vacuum'.[34] Friedman acknowledges that this formula is inaccurate in some circumstances such as using a feather

[34] Friedman (1953) page 18.

instead of a ball but that the model is accepted because it successfully predicts the distance travelled by a ball when dropped.

Friedman's argument is a miss-representation of Popper (1959) and confuses the test of the theory (i.e. falsification) with the objective of the theory. If the objective of the theory is to explain some aspect of the real world so that we can study a range of policies then the 'as if' argument is less satisfying. For example, a government may approach the theorist and ask how to make a ball travel at a velocity other than that predicted by the formula above. The researcher who accepts the theory should respond by saying that nothing can be done as the model predicts all objects fall as if in a vacuum. However, we know that the vacuum assumption is perfectly incorrect and a theory that rejects this assumption and models the wind resistance of the falling object can be used to provide the policy answer.

Galileo's Law of Falling Bodies continues to be used in circumstances where the vacuum assumption is perfectly incorrect. However, how good an approximation Galileo's Law is in describing how objects fall can only be understood from the perspective of a theory that incorporates the wind resistance of the object which in turn required the perfectly incorrect vacuum assumption to be rejected. In other words C4 applies even though Galileo's Law of Falling Bodies continues to be used in a restricted range of circumstances because of its computational and theoretical simplicity. If theorists had invoked the 'as if' defence and had faithfully adhered to the vacuum assumption then a fuller theory containing wind resistance would not have evolved along with all the developments that follow from understanding wind resistance.

A corresponding macroeconomics example is to model the business cycle 'as if' it is driven by mistaken price expectations as in the F-P theory. If a government asked for policy advice concerning how to offset a large and prolonged decline in output the advisor should answer 'you have to get agents to correct their mistakes by loosening monetary policy' or 'there is nothing that can be done in practice as we have to wait for the optimising agents to correct their mistaken price expectations'. However, this advice is based on a 'perfectly incorrect' assumption that the business cycle is driven by mistaken price expectations and so the policy advice is as empty of any relevance as policy advice based on the vacuum assumption in the example above.

The difficulty with the 'as if' argument is that it violates C4 above and the modelling process loses the 'real' anchor of the empirical analysis of the data.[35] It means that any notional assumption becomes valid allowing theories to diverge from describing the 'real' world to describing the imaginary world of the theorist. Furthermore, if any particular perfectly incorrect notional assumption becomes entrenched then the theory fails to evolve at all.

6. IS MACROECONOMICS A SCIENCE?

We can consider the macroeconomics discipline in terms of the four conditions set out above. C1 appears valid as only theories that provide empirically falsifiable predictions are in general considered relevant.[36] However, as demonstrated in Section 3 it appears that dominant macroeconomic theories can persist for an extended period of time even though the predictions at an aggregate level are not consistent with the data. For example, the F-P theory evolved through new classical macroeconomics, real business cycle theories to the NK theory with the discipline not rejecting the prediction in all these theories that inflation is an integrated process even though inflation is very unlikely to be 'truly' integrated as argued above. Consequently, we might conclude that the macroeconomics discipline pays little attention to C2 even when the empirical evidence is very strong and straightforward.

Over the past eighty years macroeconomics has evolved from that of Keynes (1936), through Hicks/Hansen's IS/LM, Paul Samuelson's neo-classical synthesis, Friedman's Monetarism, Lucas/Sargent/Wallace/Prescott's New Classical Macroeconomics, Kydland/Prescott's Real Business Cycle to the dynamic stochastic general equilibrium (DSGE) technique and the New Keynesian theory. The evolution occurred in steps of various sizes and is thought to reflect scientific progress.[37] Importantly the progress has mostly followed a period of intense

---

[35] When a theorist invokes the 'as if' defence they simultaneously reveal they believe the assumption in perfectly incorrect. Otherwise the theorist would argue the empirical validity of the assumption is either unknown (because we cannot measure it), or approximately correct, or actually correct. A theorist would not willingly use the 'as if' defence unless all other justifications for the assumption have been exhausted. Consequently we can think of the 'as if' defence as the last refuge of a faith based theorist.

[36] The possible exception is the DSGE technique which is often mistakenly referred to as a 'theory'. See Section 7 below.

[37] The term 'progress' is used loosely here to describe time in the history of economic thought. See Hoover (1984) for a description of the Monetarist and New Classical Macroeconomics theories, Kydland and Prescott

intellectual analysis of the underlying assumptions and methodology of the 'old' theory.[38] In this way the foundations of the 'old' theory are found inconsistent with how individuals actually or should behave and the 'new' theory corrects for this inconsistency. This suggests that the evolution of macroeconomic theories focuses heavily on the underlying assumptions and so one might argue the discipline appears to conform with C3 and C4.

However, the assumptions have evolved in ways that are not empirically valid as demonstrated in Section 4 and so this argument is flawed. More importantly a number of the model defining assumptions are not poor approximations of the economy but perfectly incorrect. In important ways the evolution of the assumptions underpinning the theories have been severed from the anchor provided by logic and the empirical analysis of the behaviour of agents and firms and have instead evolved along notional lines. Therefore a strong case can be made that the macroeconomics discipline as represented by the two dominant macroeconomic theories considered above pretty well universally and consistently violates C3 and C4.

It appears therefore that the macroeconomics discipline conforms to C1 but most likely severely violates C2, C3 and C4. If this is indeed the case then we might expect the discipline to have the following ancillary characteristics. One, theory papers will be valued more highly than applied papers because C2 and C4 are downplayed. Similarly, papers that question the assumptions underpinning the theory (i.e. conforming to C3 and C4) will be very lowly valued. Two, journals will specialise in terms of particular theories and approaches as the downplaying of C2, C3 and C4 means that the discipline resists the outright rejection of the competing theories. This allows 'like-minded' people to interact and publish 'like-minded' work within an agreed model free from the inconvenience of critical empirical analysis of the predictions and underlying model defining assumptions. For the same reasons particular universities will be over-represented in the papers in particular 'top' journals. Three, the evolution of macroeconomics will be impeded by the lack of a proper critical and empirical analysis of the

---

(1982) for the seminal Real Business Cycle work, Hoover (2001) for DSGE technique and Woodford (2003) and Gali (2008) who provide excellent expositions of the NK theory.

[38] For example, the assumptions of perfect competition, fixed prices and fixed expectations in the Keynesian theory (in contrast with Keynes' theory) has evolved over time into flexible prices, price-setting firms and rational expectations with forward looking behaviour in the NK theory which are thought to better approximate how agents behave and the economic structure of the economy.

theories in the same way as during the Scholastic Period of the Western intellectual tradition. Four, there will be a concentration of citations that are deemed 'acceptable' within the discipline that are based on C1 and not the remaining conditions. And five, advancement (including tenure) in the profession will be heavily based on publishing in 'top' journals which in turn are heavily focused on particular macroeconomic theories. This implies researchers who empirically examine the predictions and the assumptions as well as offer alternative theories will be under-represented in the top echelons of the discipline.[39]

7. CONCLUSION

The philosophy of science considers issues surrounding what constitutes a scientific theory, falsification and empirical analysis itself and therefore relevant to our understanding and interpretation of all four conditions in the framework developed above.[40] Furthermore, an important component of the philosophy of science, whether it is within the sociological, anthropological or 'economic incentives' traditions, is the issue of how theories evolve and in particular how do groups of individuals collectively agree on how, when, and if a theory is replaced by a 'new' theory. Again in terms of the framework above we can think of the philosophy of science as also helping to understand why disciplines violate conditions 2, 3, and 4 and why the degree of violation differs between disciplines.

One might explain the large faith based component of macroeconomics by it evolving from a branch of moral philosophy which may lead it to focus more on what should happen rather than what does happen. However, the pure sciences were all able to break away from the faith based roots of the moral philosophers in the 18th and 19th Centuries by rigorously incorporating conditions 2, 3 and 4 into their scientific method.

An alternative explanation is the marginal 'revolution' referred to in Section 4.1 was such a major discipline defining intellectual event for economics that profit maximisation and utility maximisation became necessary components of all acceptable economic theories.[41] However,

---

[39] Colander (1996) and Galbraith (1996) provide similar but more detailed descriptions of how the economics academic 'profession' operates. For a humorous but no less insightful description see Leijonhufvud (1973).

[40] See the references cited in footnote 13.

[41] See Hodgson (1996).

this event violated C4 as the concepts of marginal costs and revenues are undefined.[42] So that C4 could be violated on an on-going basis it is necessary for the discipline to downplay the role of empirical evidence in general and in relation to the undefined nature of marginal costs and revenues in particular. Furthermore, the marginal revolution made it 'acceptable' to indulge in notional assumptions that routinely violate C4 when creating economic theories. As a result economics failed to rigorously apply C2, C3 and C4 so that the primacy of empirical analysis over theory is achieved and allow economics to join physics, chemistry, medicine and biology as a 'pure science' in the late 19th Century.

An observer may ask where the dynamic stochastic general equilibrium (DSGE) approach fits into this analysis. This approach proceeds under the assumption that the model is 'true' and is simply calibrated so that the means and variances of 'important' variables are similar to their actual economic values. As any model with any structure can be calibrated to produce the actual values of the nominated 'important' variables the DSGE approach eschews any formal testing of the model. Consequently, DSGE is not a scientific theory in the sense of Popper and violates C1, C2, C3 and C4. Instead of being a theory it is a modelling technique in the same way as Monte Carlo simulations are a calibrated modelling technique.

Finally, in the physical, chemical, medical and biological sciences there is the term 'science fiction' to describe ideas that are in some sense scientific but include a component of fantasy. Often the fantasy includes humans with super-human characteristics like they can fly, are indestructible, read minds, see through walls and see the future. Macroeconomic agents are often endowed with super-human characteristics in that they can measure concepts that are undefined, can 'see' the future so as to predict the structure of the economy, prices and structural breaks, and are instilled with technical abilities that elude all of the most brilliantly trained economists.[43] In every sense of the word these characteristics attributed to macroeconomic agents are fantasy and should be consigned to 'economic fiction'. If not then macroeconomics

---

[42] Note the advent of marginal analysis meant that agents were no longer able to recognise the concepts used by economists to explain their own behaviour. Implicitly economists were saying 'have faith, we know what you are doing even though you as agents will not recognise it as what you are doing'.

[43] Caballero (2010) calls the attributing of 'superhuman' skills and knowledge to agents by macroeconomists the pretence-of-knowledge syndrome.

will fail to join the ‘pure sciences’ as ‘macroeconomic science’ and will remain a largely faith based discipline made up of modern day Tron theorists.

8. REFERENCES

Baumol, W.J. (1976). ‘Scale Economies, Average Cost and the Profitability of Marginal-Cost Pricing’, in R. Grieson, ed., ‘*Essays in Urban Economics and Public Finance in Honor of William S. Vickrey*, Lexington.

Baumol, W.J. (1977). ‘On the Proper Cost Tests for Natural Monopoly in Multiproduct Industry’, *American Economic Review*, December, 67, pp. 809-22.

Bird, A. (1998). *Philosophy of Science*, McGill-Queen's University Press, Montreal.

Bloor, D. (1991). *Knowledge and the Social Imagery*, 2nd edition, University of Chicago Press, Chicago.

Caballero, R.J. (2010). Macroeconomics after the Crisis: Time to Deal with the Pretense-of-Knowledge Syndrome, *Journal of Economic Perspectives*; vol. 24, no. 4, pp. 85-102.

Clarida, R., Galí, J. and M. Gertler (1999). The Science of Monetary Policy: a New Keynesian Perspective, *Journal of Economic Literature*, vol. 37, pp. 1661-1707.

Colander, D. (1996). Surviving as a Slightly out of Sync Economist, in S.G. Medema and W.J. Samuels eds. *Foundations of Economic Research: How Do Economists Do Economics?* Edward Elgar, Northampton, MA, USA.

Dasgupta, P. and P. David (1994). Towards a New Economics of Science, *Research Policy*, vol 23, pp. 487-532.

Fair, R.C. (2012). Has Macroeconomics Progressed?, *Journal of Macroeconomics*, vol. 34, pp. 2-10.

Friedman, M. (1953). The Methodology of Positive Economics. In M. Friedman (ed.) *Essays in Positive Economics*, University of Chicago Press, Chicago, pp. 3-43.

Friedman, M. (1968). The role of monetary policy, *American Economic Review*, 58, 1 (March), pp. 1-17.

Galbraith, J.K. (1996). What is To Be Done (about Economics)? in S.G. Medema and W.J. Samuels eds. *Foundations of Economic Research: How Do Economists Do Economics?* Edward Elgar, Northampton, MA, USA.

Galí, J., (2008). *Monetary Policy, Inflation and the Business Cycle: An Introduction to the New Keynesian Framework:* Princeton University Press, Princeton.

Galí, J., and M. Gertler (1999). Inflation Dynamics: A Structural Econometric Analysis, *Journal of Monetary Economics*, vol. 44, pp. 195-222.

Galí, J., Gertler M., and J.D. Lopez-Salido (2001). European Inflation Dynamics, *European Economic Review*, vol. 45, pp. 1237-1270.

Gillies, D. (1998). The Duhem Thesis and the Quine Thesis, in M. Curd and J.A. Cover, editors, *Philosophy of Science: The Central Issues*, Norton, New York.

Hands, W.D. (2001). *Reflection without rules: economic methodology and contemporary science theory*, Cambridge University Press, Cambridge.

Hodgson, G.M. (1996). Towards a Worthwhile Economics, in S.G. Medema and W.J. Samuels eds. *Foundations of Economic Research: How Do Economists Do Economics?* Edward Elgar, Northampton, MA, USA.

Hoover, K.D. (1984). Two Types of Monetarism, *Journal of Economic Literature*, vol. 22, no. 1, pp. 58-76.

Hoover, K.D. (2001) *The Methodology of Empirical Macroeconomics*, Cambridge University Press.

Howitt, P. (2012). What have Central Bankers Learned from Modern Macroeconomic Theory?, *Journal of Macroeconomics*, vol. 34, pp. 11-22.

Jevons, W.S. (1871). *The Theory of Political Economy*, Macmillan and Co, London and New York.

Kalecki, M. (1954). *Theory of Economic Dynamics*, Allen and Unwin, London.

Kalecki, M. (1971). Class Struggle and the Distribution of National Income, *Kyklos*, vol. 24, pp. 1-9.

Keynes, J.M. (1936). *The general theory of employment, interest and money*, Macmillan, London.

Kozicki, S. (2012). Macro has Progressed, *Journal of Macroeconomics*, vol. 34, pp. 23-28.

Kuhn, T.S. (1970). *The Structure of Scientific Revolutions*, 2nd edition, University of Chicago Press, Chicago.

Kydland, F.E. and E.C. Prescott (1982). Time to Build and Aggregate Fluctuations, *Econometrica*, vol. 50, pp. 1345-1370.

Lakatos, I. (1978). *The Methodology of Scientific Research Programmes*, Philosophical Papers Volume 1, J. Worrall & G. Currie, editors, Cambridge University Press, Cambridge.

Leijonhufvud, A. (1973). Live Among the Econ, *Western Economic Journal*, vol. 11, pp. 327-37.

Lucas, R.E. Jr (1973). Some International Evidence on Output- Inflation Tradeoffs, *American Economic Review*, vol. 63(3), pp. 326-34.

Mäki, U. (1994). Reorienting the assumptions issue, in *New Directions in Economic Methodology*, Roger Backhouse (ed.), London: Routledge, pp. 236-56.

Mäki, U. (2009). Realistic realism about unrealistic models, in *The Oxford Handbook of Philosophy of Economics*, Harold Kincaid and Don Ross (eds), New York: Oxford University Press, pp. 68-98.

Marshall, A. (1920). *Principles of Economics*, London: MacMillan and Co.

Marshall, A. (1927). *Industry and Trade*, London: MacMillan and Co.

Medema, S.G. and W.J. Samuels (1996). *Foundations of Economic Research: How Do Economists Do Economics?*, Edward Elgar, Northampton, MA, USA.

Menger, C. (1871). *Principles of Economics*, translated by J. Dingwall and B.F. Hoselitz, reprinted by the Ludwig von Mises Institute, 2007, Auburn, Alabama.

Panzar, J. and R.D. Willig (1977). 'Economies of Scale in Multi-Output Production', *Quarterly Journal of Economics*, August, 91, pp. 481-94.

Perron, P., (1989). The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis, *Econometrica*, vol. 57, no. 6, November, pp. 1361-1401.

Phelps, E.S. (1967). Phillips curves, expectations of inflation, and optimal unemployment over time, *Economica*, 34, 3 (August), pp. 254-81.

Popper, K.R., (1959). The Logic of Scientific Discovery, London, Hutchinson. Originally published as *Logik der Forschung* in 1934.

Russell, B. (2016). 'Modern' Phillips Curves and the Implications for the Statistical Process of Inflation, forthcoming in *Applied Economics Letters*, and formally Dundee Discussion Papers, Economic Studies, University of Dundee, September, No. 289.

Sims, C.A. (1980). Macroeconomics and Reality, *Econometrica*, vol. 48, pp. 1-48.

Sraffa, P. (1960). *Production of Commodities by Means of Production*, University Press, Cambridge.

Sugden, R. (2000). Credible worlds: the status of theoretical models in economics, *Journal of Economic Methodology*, vol. 7, pp. 1-31

Svennson, L.E.O. (2000). Open Economy Inflation Targeting, *Journal of International Economics*, vol. 50, pp. 155-83.

Taylor, J.B. (1993). Discretion versus Policy Rules in Practice, Carnegie-Rochester Conference Series on Public Policy, 39(0), pp. 195-214.

Turvey,R. (1969). Marginal Cost, *The Economic Journal*, vol. 79, no. 314, pp. 282-99.

Turvey,R. (2000). What are Marginal Costs and how to Estimate Them?, Technical Paper 13, Centre for the Study of Regulated Industries, University of Bath School of Management.

Uhlig, H. (2012). Economics and Reality, *Journal of Macroeconomics*, vol. 34, pp. 29-41.

Walras, L. (1874-77). Elements of Pure Economics, translated from *Éléments d'économie politique pure*, 1871, and annotated by W. Jaffe, 1954, London: Allen and Unwin, Homewood, Illinois.

Willig, R.D. (1979). 'Multiproduct Technology and Market Structure', *American Economic Review*, May, 69, no. 2, pp. 346-51.

Woodford, M. (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press, Princeton.

Diagram 1: The Framework for 'Testing' an Assumption