The rebuilding macroeconomic theory project part II: multiple equilibria, toy models, and policy models in a new macroeconomic paradigm

David Vines* and Samuel Wills**

Forthcoming in the Oxford Review of Economic Policy Vol. 36, No. 3

Abstract

This issue of the *Oxford Review of Economic Policy* is the second in the Rebuilding Macroeconomic Theory project. The papers in the first issue proposed specific improvements to the New-Keynesian Dynamic Stochastic General Equilibrium (NK-DSGE) framework, in response to the events of the 2008 crisis. This issue goes further. We, as editors, now argue that a new multiple-equilibrium and diverse (MEADE) paradigm is needed for macroeconomics. It will emphasize that economies can have more than one stable outcome, and study why. It will require using both toy models, which enable a quick, first-pass, intuitive understanding of a question, and policy models (aka structural economic models) which develop a detailed empirical understanding of the economy. We argue that the seminal IS/LM, Solow–Swan, Ramsey, Real Business Cycle, Taylor, and Clarida/Galí/Gertler models have all been toys, as is the benchmark NK-DSGE model. In the past the models have adapted as the questions changed, and the NK-DSGE model must now do this since it has failed to capture both the salient aspects of the lead-up to the 2008 crisis and the slow recovery afterwards. The way forward in the MEADE paradigm will be to start with simple models, ideally two-dimensional sketches, that explain mechanisms that can cause multiple equilibria. These mechanisms should then be incorporated into larger DSGE models in a new, multiple-equilibrium synthesis. All of this will need to be informed by closer fidelity to the data, drawing on lessons obtained from detailed work on policy models.

Keywords: DSGE, multiple equilibrium, toy models, structural economic models

JEL classification: A23, A31, B22, B41, E00

* Balliol College, Economics Department, and Institute for New Economic Thinking (INET) in the Oxford Martin School, Oxford University; and Centre for Economic Policy Research, e-mail: david.vines@economics.ox.ac.uk
** School of Economics, University of Sydney; Oxford Centre for the Analysis of Resource Rich Economies, Oxford University; Centre for Applied Macroeconomic Analysis, Australian National University, e-mail: samuel.wills@sydney.edu.au
I don’t believe you can dispose of the traditional theory until you have stated your view.

(Letter from Roy Harrod to John Maynard Keyes, printed in Keynes, 1973, p. 535)

I. Introduction: replacing the NK-DSGE model with a new paradigm

Why do it all again?

The first issue of the *Oxford Review of Economic Policy (OxREP)* devoted to our Rebuilding Macroeconomic Theory Project appeared in early 2018 (vol. 34 nos. 1–2). It brought together a group of leading economists, from a wide range of traditions, to discuss and debate what the profession had achieved in the 10 years since the global financial crisis (GFC), and what to do next. The outcome was broad agreement that the benchmark new-Keynesian dynamic stochastic general equilibrium (NK-DSGE) model, along the lines of Smets and Wouters (2007), did not capture enough of the past decade’s financial boom, bust, and slow recovery to justify its claim to be the profession’s core model. Together this group of people proposed a way forward which involved updating the main behavioural equations in the model,¹ and adopting a more open, pluralist approach to research.

That issue of *OxREP* was well-received.² So why devote another issue of the journal to this same subject? Our answer is simple. We believe that the existing paradigm³ in macroeconomic theory, the one identified with the NK-DSGE model, is on its way out. In our view a multiple-equilibrium and diverse paradigm—the MEADE paradigm—is about to arrive. We want to speed its journey.⁴

---

² Three of the most highly cited papers published in *OxREP* in the last 10 years were in the double-issue of the journal dedicated to this project: Blanchard (2018), Stiglitz (2018), and Vines and Wills (2018) (in order of citations). There was also a five-part discussion of the published papers by Martin Sandbu in the *Financial Times* (15 January 2018), followed up by an article by Martin Wolf (20 March 2018).

³ In his book, *The Structure of Scientific Revolutions*, Thomas Kuhn (1962) defines a scientific paradigm as a set of universally recognized scientific achievements that, for a time, provide model problems and solutions for a community of practitioners.

⁴ The label ‘MEADE’ stands for ‘Multiple Equilibrium And DiversE’. We have given the new paradigm this label in recognition of the work of Nobel Laureate James Meade, who worked with Keynes in the 1930s and 1940s and was David Vines’s mentor at the University of Cambridge in the 1970s and 1980s.
Last time we were much more conservative (see Vines and Wills (2018)). We accepted the NK-DSGE modelling paradigm, and were optimistic about how it might be developed—providing a list of frictions to be added to the model. We even believed that a perfect Platonic form might exist: the ideal NK-DSGE model.

Olivier Blanchard (2017) shared our optimism:

I feel we still do not have the right core model . . . (s)uppose that we were building a small macroeconomic model from scratch. What are, say, the three distortions we would deem essential to have in such a model, and, by implication, to have as the core of any DSGE model? What model should we teach at the start of the first year graduate course? This may a hopeless and misguided search. Maybe even the simplest characterization of fluctuations requires many more distortions. Maybe different distortions are important at different times. Maybe there is no simple model. . . . I keep faith that there is. . . . I do not have the answer, but I have a few ideas.5

We have changed our mind, basically because of the GFC and the COVID pandemic. Joe Stiglitz sets out the argument at the beginning of his article with Martin Guzman in this issue:

[t]he central problem of macroeconomics is to explain the large, deep, and often persistent downturns, accompanied by high levels of unemployment, that episodically afflict capitalist economies. Macroeconomic crises are extreme examples of economic fluctuations. But they are the most relevant. They are the events that teach the most about the stability properties of the economic system, in a way that small inventory cycles do not. And they are the events that matter the most for the lives of millions of people. (Guzman and Stiglitz, 2020)

To analyse these large downturns we need something different from the benchmark NK-DSGE model, which— as Blanchard says— was constructed to characterize fluctuations. Why? There are two inadequacies of the NK-DSGE model which together point towards this need for something new.

First, thinking about large downturns requires models with good outcomes and bad outcomes. This means models with multiple equilibria. It is not good enough to say that bad shocks cause bad outcomes; the risk of a lurch to a permanently bad outcome is always present, even with a temporary shock. Macroeconomic policy should be designed to guard against this risk. But the NK-DSGE modelling tradition does not do lurches.6 The set of changes to the ‘core model’

5 Blanchard’s ideas included the need to retain nominal rigidities, and to introduce two new distortions. The first was finite horizons (due to bounded rationality, myopia, and the inability to think far into the future, rather than because of death and the absence of operative bequest motives). The second was allowing for the import effect of agents’ own capital on their spending decisions. But he seemed uncertain as to what these added distortions would deliver. ‘How I would actually put them together in a basic model is a much harder question. . . . whether they fit together conceptually is not obvious.’ (Blanchard, 2017)

6 For a striking exception to this claim, see McKibbin and Fernando (2020), and the paper by McKibbin et al. (2020) in this issue. The outcomes depicted by McKibbin and Fernando (2020) for the global economy in the face
which we listed in Vines and Wills (2018), and the distortions for which Blanchard was searching, cannot lead to a very bad outcome within a NK-DSGE model. That is because an NK-DSGE research programme\(^7\) involves investigating shocks in a (nearly) linear model with a stable and unique equilibrium, in which the participants have rational expectations.\(^8\) We think that it is no longer adequate to restrict the macroeconomic research programme in this way. As we see in section III, many others agree.

Second, to think about large downturns requires a detailed empirical analysis of what can go wrong, and why. This requires a new respect for what Olivier Blanchard calls ‘policy models’ (Blanchard, 2018) and what Simon Wren-Lewis, in his paper in this issue of OxREP, calls ‘structural economic models’ (Wren-Lewis, 2020; see also Wren-Lewis, 2018). These are models which come from a long tradition of careful scholarly work dedicated to obtaining a detailed empirical understanding of the economy as it actually is.\(^9\) Such models involve the kinds of behaviour—including rules of thumb, mistakes, and misunderstandings—which can lead to bad outcomes. The papers by John Muellbauer (2020) and Ray Fair (2020) in this issue are powerful examples of this modelling tradition. But NK-DSGE models do not do this. These models claim that macroeconomic outcomes can be understood in a ‘microfounded’ manner: as the outcome of optimizing behaviour by representative agents who understand what they are doing, and act freely, apart from particular, clearly specified frictions. We no longer think that this is an appropriate restriction of the macroeconomic research programme; structural economic models must be constructed alongside models of the NK-DSGE kind, in which behaviour need not be fully microfounded. As outlined in section IV of this paper, many people also now agree with this.

Because of these two inadequacies we no longer think that the NK-DSGE model can be thought of as a ‘core model’. Blanchard (2018) helpfully classifies macro models into five groups: foundational, toy, core, policy, and forecasting. In the first issue in this series we described the NK-DSGE model as a core model (Vines and Wills, 2018). According to Blanchard a core model should provide ‘a generally accepted theoretical framework for the profession’. However, the inadequacies described above mean that the NK-DSGE does not satisfy this test.

Instead, we think it best to think of the NK-DSGE model as a ‘toy model’. In Blanchard’s (2018) words, toys:

---

\(^7\) The term ‘research programme’ was put forward by Imre Lakatos to describe what he calls a hard core of theoretical assumptions that cannot be abandoned or altered without abandoning the programme altogether (Lakatos, 1970). A particular research programme is what goes on within a paradigm.

\(^8\) This research strategy has been guided by the ‘correspondence principle’, set out by Paul Samuelson in his Foundations (Samuelson, 1947). As we show below, that strategy had, in fact, been articulated 10 years earlier by James Meade in his exposition of Keynes’s General Theory (Meade, 1937).

\(^9\) We thank Christopher Adam for this form of words.
[a]llow for a quick first pass at some question, present the essence of the answer from a more complicated model or from a class of models. For the researcher, [such models] may come before writing a more elaborate model, or after, once the elaborate model has been worked out.

We think this because the aim of the NK-DSGE model is not to provide ‘a detailed empirical understanding of the economy as it actually is’. Instead, its aim is to provide the simplest possible way of seeing how the fundamental pieces of the economy fit together, subject to it being ‘properly microfounded’. It is clearly a large toy, and to play with it requires a computer. But, simulating this model allows us to take a quick first pass at important questions. And its inadequacies mean that it does no more than this.

Furthermore, we don’t think that any core models actually exist at all. Foundational models set out the deep micro-foundations of economic behaviour, and forecasting models are designed to provide the best possible predictions of the future; though we don’t consider either in this issue. Policy models clearly exist, to help decision-makers reach their conclusions in the face of all the messiness of reality. As a result, they must include a great deal of detail. In contrast, toy models strip back much of this messiness to elucidate the underlying system properties: a view of ‘what is really going on’. But that’s all there is. We argue that, once the NK-DSGE model has been properly understood, it will be clear that it is a toy, and that it cannot possibly be called a ‘core model’. And there is no other core model lurking out there.

To get around the inadequacies of the NK-DSGE toy model, we believe that a new paradigm is needed for the macroeconomics discipline: the MEADE paradigm, where, as already noted, MEADE stands for ‘Multiple Equilibrium And DiversE’. This paradigm will be diverse because its research programme will contain three components: the use of toy models, an emphasis on the importance of multiple equilibria, and a greater reliance on detailed empirical work obtained from a variety of structural models. This paper sets out these three components. Here is a brief summary.

First, toy models will play an important role in the new macroeconomic paradigm. We need to understand how we got to where we are now, why our main toy is now inadequate, and what to do about this. So, in section II we briefly set out the history of macroeconomic toy models, all the way from Hicks’s IS/LM model to the current NK–DSGE synthesis. Along the way we encounter: Samuelson’s neoclassical synthesis (Samuelson, 1955); the Solow–Swan and Ramsey models of economic growth (Solow, 1956; Swan, 1956; Ramsey, 1928) and the real business cycle (RBC) model of Kydland and Prescott, 1982); and the Taylor and Clarida, Galí, and Gertler (CGG) models of stabilization policy (Taylor, 1993: Clarida et al., 1999). The common thread of this history-of-thought discussion is that inventing toys is difficult. It requires a lot of effort to represent the world simply, in order to take a first pass at understanding the pressing question of the moment. Doing this is especially difficult if the representation needs to be microfounded. We see from history that toys need to be reinvented
because existing simplifications become obsolete when the question changes. The NK-DSGE model has been an immensely valuable toy, because it encapsulates in a very simple way nearly a hundred years of research on how the pieces of the economy fit together. It also works well for routine examination of fluctuations around a steady state, like productivity or cost-push shocks. But, it is no longer an adequate toy because it assumes a unique equilibrium and contains too many of the wrong kind of shortcuts. In particular, the model has failed both to capture the asset price boom leading to up to the global financial crisis, and to understand why the subsequent recovery was so slow. So this toy also needs to be reinvented: new toys are needed.

Second, the new toys will need to allow for multiple equilibria. Section III of the paper sets out this claim. To do this we need to start by building small toy models that capture how multiple equilibria (ME) might emerge. We think that these small toys should be as simple as possible, ideally two-dimensional sketches done with pencil and paper, just like the simple toy models examined in most of section II. However, while those simple toys—all the way from the IS/LM model to the CGG model—have curves that intersect only once, ME toys will have curves which intersect multiple times. To do this the models will need to be non-linear. We set out simple sketches of three such ME models, focusing on the relevant non-linearities. These include an explanation for the asset price boom of the early 2000s (building on Krugman, 2003); for the slow recovery from the GFC (building on Carlin and Soskice, 2018); and for the difficulty some central banks have faced in reducing inflation (building on Akerlof, 2019).

We go on to argue that these small models will provide lessons that will help us build multiple equilibrium versions of the NK-DSGE model. The NK-DSGE model has taught us that the world is not a two-dimensional sketch. We don’t want to throw this knowledge away just because we have discovered multiple equilibria. We must therefore build large toys that show how everything fits together in the presence of multiple equilibria.

In the later parts of section III we consider how the current benchmark NK-DSGE model might come to incorporate more reasons for multiple equilibria, going beyond the three examples just given. In doing this we enlist the aid of three papers included in this issue of *OxREP*. Guzman and Stiglitz (2020) give a wide-ranging account of how misperceptions are inevitable in the absence of perfect markets, and of how, when mistakes are realised, adjustments are likely to be unstable. Roger Farmer (2020) presents a model of how beliefs about future output might well be self-fulfilling. And Tony Venables (2020) studies the self-perpetuating tendency of economies like the UK to separate into a high productivity core and a low productivity periphery. Jeff Sachs adds to this discussion by reminding us of how a lack of understanding of financial frictions, particularly of non-linear interactions between credit and real estate, as described by John Muellbauer, led to the boom–bust of the GFC and to the policy mistakes in dealing with that crisis (Sachs, 2020, Muellbauer 2020). Section III concludes by revisiting our findings from Part 1 of the Rebuilding Macro Theory project (Vines and Will, 2018), which identified four ways in which the NK-DSGE needed to be improved: to allow for financial
frictions, to relax the rational expectations assumption, to introduce heterogeneous agents, and to ensure that the model has better micro-foundations. We make clear that all four of these ideas point in the direction of multiple equilibria.

Third, we must look to the variety of approaches to modelling practised by those who build structural economic models (i.e. policy models). We describe this work in section IV of the paper. We start by showing how ‘hybrid’ models can combine the theoretical insights from DSGE models with a respect for empirics, using as an example the paper by Warwick McKibbin and colleagues in this issue of OxREP (McKibbin et al., 2020). We then set out the claim that a detailed empirical understanding of the economy is essential if we are to understand why multiple equilibria might arise, and how to avoid the bad outcomes which might emerge. To do this we describe the papers by Muellbauer and Fair, each of which provides an excellent example of the necessary kind of work (Muellbauer, 2020; Fair, 2020). We then discuss the paper by Wren-Lewis—the next paper in this issue. Wren-Lewis makes it absolutely clear why, when making policy, policy-makers need to rely on structural economic models and the detailed econometric work which they embody, rather than the toy NK-DSGE model. And he shows how the methodology used for building these structural models needs to differ from that used to build NK-DSGE models (Wren Lewis, 2020).

Thus, the term ‘MEADE paradigm’ is not a slogan. It is an umbrella term which denotes a more diverse, but clearly defined, research programme in which multiple equilibria play a crucial role, toys are important, and policy models (or structural economic models) are a key part of the research activity.

In advocating a move to this new paradigm we are clearly adopting a less conservative position than we did 3 years ago in Vines and Wills (2018). But we are still not revolutionaries. We see a move to such a new paradigm as another stage in the evolution of macroeconomics as a discipline.

II. Toy story

History of thought matters, despite it having all but disappeared from graduate economics programmes.

The reader in a hurry might be tempted to skip to the end of this history-of-thought section, going to Section II (vi) where we explain what is wrong with the NK-DSGE model. But doing that would be a mistake. If one is to understand what is wrong with macroeconomics it is important to understand how we got to where we now are. That story is also important as we try to work out what to do next. This is why the history-of-thought material really does matter.

So in this Section we tell the story of the development of the five major toy models in macroeconomics: the IS/LM model; the Solow–Swan growth model; the Ramsey growth
model and the RBC model; the Clarida–Gali–Gertler (CGG) model of stabilization policy; and the NK-DSGE benchmark model. We outline the challenges faced when setting up these models, and consider why they went on to fail. Roughly speaking, as times change, so must the toys.

In what follows, we identify the changes in historical circumstances that made new toys necessary: the onset of unemployment in the 1930s; the post-war economic growth of the 1950s; the onset of stagflation in the 1970s; and now, in the early 2000s, the onset of crises. But, more than this, we show that theoretical progress has been made, as new theories—and new toys—have not merely replaced, but instead encompassed, the old ones.

New toys are needed today because the global financial crisis (GFC) and the Covid crisis have revealed that economies can have multiple stable outcomes. These crises require a new framework of thought: the MEADE paradigm, with the possibility of multiple equilibria, and with new toys as part of its structure. But creating this new infrastructure will not be easy.

(i) The invention of the IS/LM model—and its weaknesses

In the 1930s, the world was in the midst of the Great Depression. The classical, partial-equilibrium toy models of the time identified high wages as the cause of the high level of unemployment, but wage cuts did not fix this problem. A new toy was needed. Keynes’ *General Theory* (1936) eventually showed that the missing link was interaction between markets, or ‘general equilibrium’. His insight was that low demand for labour stemmed from a lack of demand for goods. This discovery was incorporated into the existing body of knowledge in a complex system of equations by Meade (1937), and refined into the far-simpler IS/LM toy by Hicks (1937). The IS/LM model became widely adopted because it struck a balance between simplicity and detail, and encompassed much of what came before. However, it cast a long shadow by perpetuating Marshall’s assumption of a unique equilibrium, the assumption which we are reconsidering in the present paper.

---

10 Mark Blaug identifies two different approaches to the history of thought (Blaug, 1968, p. xi). The first, described by Blaug as ‘ultra-Marxist’, assumes that ‘the economic theory of a given period is nothing but a reflection of the prevailing historical and political circumstances’. The second, conversely, is concerned with correcting ‘mistakes in logic and [filling] gaps in [the] analysis’ and contains no connection at all with contemporary events. We deliberately combine both methods.

11 James Meade is an outstanding example of someone whose work proceeded by encompassing the work of others (see Vines, 2008).

12 Very many people have written about the Keynesian revolution; our new focus here is on what was necessary to get to the IS/LM toy, and on the inadequacies of that toy. The brief but detailed analysis which we present uses Keynes’s own books, articles, and papers (Keynes, 1923, 1924, 1925, 1930a, 1936, 1973), four other books (Kahn, 1984; Young, 1987; Temin and Vines, 2013, 2014), and four other articles (Hicks, 1937; Meade, 1937; Rappoport, 1992; Vines, 2008). The last of these relies on extended discussions which David Vines had with James Meade in the 1980s about what had happened in the 1930s.
The Great Depression and the need for a new toy model

In the 1930s, economic analysis was dominated by the ideas of Alfred Marshall.\textsuperscript{13} The eighth edition of his \textit{Principles of Economics}\textsuperscript{14} had been published in 1920 and this remained the undisputed bible of the economics profession: Keynes referred to Marshall’s ideas as ‘classical economics’. Classical economists had a straightforward explanation for high unemployment: wages were too high. But the wage cuts advocated by Arthur Pigou\textsuperscript{15} in his \textit{Theory of Unemployment} (Pigou, 1933), were clearly not working.\textsuperscript{16} A new theory was needed.

According to Keynes, it was Marshall who invented toy models: the diagrams which he put in his \textit{Principles}.\textsuperscript{17} Keynes himself considered diagrammatic toys to be important: they provide an ‘elegant apparatus . . . which all of us use as an inspirer of, and a check on, our intuitions’ (Keynes, 1924, p. 332). Quoting Marshall, he said that ‘experience seems to show that they give a firmer grasp of many important principles than can be got without them; and that there are many problems of pure theory that no one who has once learnt to use diagrams will willingly handle in any other way’ (Keynes, 1924, p. 334).

Marshall’s toys consisted of microeconomic partial-equilibrium supply-and-demand diagrams. Marshall had written the \textit{Principles}, as he explained in his Preface, to explain the nature of ‘equilibrium’ in the ‘normal conditions of modern life’. In this normal world, partial-equilibrium diagrams seemed appropriate: the real wage adjusts to remove unemployment; the price level is pinned down by the quantity of money; and the interest rate adjusts to ensure that savings equals investment, so that there can never be a lack of demand for goods. But the Great Depression of the 1930s was clearly not a ‘normal’ world, and the toy-model story—of intersecting supply and demand curves showing that wage cuts would remove unemployment—was not helpful. A new toy was needed, along with a new theory.

As we have said, Keynes had a new \textit{insight} about unemployment: that a lack of demand for \textit{goods} would reduce the demand for \textit{labor}; that unemployment is a general equilibrium problem. This was literally impossible to conceptualize using Marshall’s \textit{Principles}. Keynes’s conceptual move required him to analyse the interconnections between two markets (the goods market and the labour market), rather than analysing each market separately. His achievement was particularly remarkable given the power which Marshall’s book held over its readers. Keynes was already clear, long before he wrote the \textit{General Theory}, that a lower level of demand for goods could cause unemployment of labour. He had explained his idea in 1930 to

\begin{itemize}
  \item[\textsuperscript{13}] Marshall was Professor of Political Economy in Cambridge from 1885 to 1908.
  \item[\textsuperscript{14}] Marshall’s \textit{Principles} (1920) was not an introductory textbook in the modern sense. Each edition contained a synthesis of what Marshall considered to be the current state of the field (see Bowles and Carlin, 2020a, p. 179, fn. 2).
  \item[\textsuperscript{15}] Pigou became Professor of Political Economy in Cambridge when Marshall retired in 1908.
  \item[\textsuperscript{16}] See Harrod (1934).
  \item[\textsuperscript{17}] Keynes writes as follows, in his obituary for Marshall (Keynes, 1924): ‘Marshall’s . . . diagrammatic exercises . . . were of such a character in their grasp, comprehensiveness and scientific accuracy, and went so far beyond the ‘bright ideas’ of his predecessors that we may justly claim him as the founder of modern diagrammatic economics’ (Keynes, 1924, p. 332).
\end{itemize}
the Macmillan Committee, over 5 days of testimony, and the idea was clearly present in *The Economic Consequences of Mr Churchill*, published in 1925. The task for Keynes was to turn this insight into a theory. The task for others would be to turn that theory into a toy.

**Reaching the General Theory**

We now know that Keynes had to do four things to move from Marshall’s worldview to his new *General Theory*: allow the short-run quantity of output to vary; abandon the quantity theory of money; assume that wages are sticky; and allow households to consume less than their whole income.

First, Keynes allowed the quantity of output to vary in the short run. He had previously assumed that output was fixed, in his *Tract on Monetary Reform* (Keynes, 1923), and in his *Treatise on Money* (Keynes, 1930a), so that a change in the demand for goods would only cause a change in the price level. Both models were therefore unable to describe the fall in output and employment that had taken place, even though Keynes had described Britain’s fall in output and employment—in detail—to the Macmillan Committee. It was James Meade who resolved Keynes’s contradiction, by arguing that both prices and output would be endogenous, since, he said, the economy was located on a short-run upward-sloping supply curve. Over time, Keynes came to the view that the short-run supply curve is almost flat, so that changes in the demand for goods primarily move the *quantity* of output to re-establish equilibrium in the goods market, rather than its *price*, as supposed in the *Treatise*.

Second, Keynes had to banish the quantity theory of money. In Marshall’s system, this had the role of ensuring that the quantity of money determined the price level. In such a set-up,

---

18 The Macmillan Committee was a formal enquiry into UK policy-making, set up by the UK Prime Minister soon after the 1929 Wall Street crash, and named after its chairman, a Scottish judge (Keynes, 1931, pp. 38–311; Temin and Vines, 2013, ch. 3; and Temin and Vines, 2014, ch. 3). *The Economic Consequences of Mr Churchill* (Keynes, 1925) was published precisely to explain why unemployment would emerge if the UK returned to the gold standard at an overvalued exchange rate.

19 Meade did this at a meeting of the ‘Circus’, a group of young economists gathered in Cambridge during the academic year 1930–31 to discuss the inadequacies of Keynes’s *Treatise* (see Keynes (1973, pp. 337–43); Vines, 2008, p. 3). Austin Robinson made Meade’s point more graphically to members of the Circus: ‘Surely a decision by somebody to have their shoes shone would increase the number of shone shoes. It wouldn’t just increase the price of shoe-shines.’

20 Richard Kahn describes how Keynes and Denis Robertson, Keynes’s close colleague, had a lengthy correspondence about the quantity theory (Kahn, 1984). He shows (on pp. 58 et seq.) that Robertson never understood that one needed to analyse equilibrium in both the money market and the goods market—i.e. that one needed to use not just the quantity theory embedded in the LM curve but also the ideas embedded in the IS curve—in order to determine macroeconomic equilibrium. Years later, Kahn (1984) described this correspondence, which noticeably slowed down the appearance of the *General Theory*, as a complete waste of time. In retrospect, it is easy to see the difficulty which Robertson faced. The LM curve displays combinations of the interest rate and income at which the demand for *stock* of money equals the supply of that stock of money, whereas the IS curve shows combinations of the interest rate and income at which the *flows* of savings and investment are equal. This stock/flow idea was not properly understood until the appearance of Phillips’s hydraulic macroeconomic model in 1950; in that model savings and investment are flows of water, whereas the stock of money is located in a tank of water (see Phillips (1950) and Vines (2000)). An algebraic description of this stock/flow system was eventually supplied by James Tobin, but not until many years later (Tobin, 1969, 1982).
with prices given, any reduction in the money wage would lead directly to a reduction in the real wage and would therefore lead to an increase in the demand for labour. Hence, money wage flexibility would lead to real wage flexibility which would, in turn, prevent any incipient tendency towards unemployment from actually emerging. Consequently, there could never be a lack of demand for goods. To banish the quantity theory, Keynes had to invent the idea of liquidity preference, which implied that equilibrium in the money market was ensured by movements in the interest rate rather than movements in the price level. This analysis was to underpin the LM curve.

Third, Keynes needed to assume that wages only fell slowly in the presence of unemployment, i.e. that wages were ‘sticky’. At the Macmillan Committee, he asserted that ‘for centuries there has existed an intense resistance to any matters of reduction in the level of money incomes’ (Keynes, 1931, ch. 2). This ‘wage stickiness’ was needed to pin down the price level, and in turn the real value of a quantity of money, and so to ensure that the interest rate could be determined according to the theory of liquidity preference described above.

Finally, Keynes needed households to consume less than their entire income, with the rest saved; he introduced a consumption function with a marginal propensity to consume of less than one. This meant that a reduction in the demand for goods, and so a fall in output and income, would reduce savings. Thus, if there were an initial fall in investment, then the equilibrium between savings and investment would be re-established by a fall in the level of output and income, rather than by a reduction in the interest rate. This analysis was to lead to the IS curve.

Taken together, these four moves were necessary and sufficient for Keynes to show that a lack of demand for goods could cause unemployment of labour. For example, consider a fall in demand for goods caused by an exogenous reduction in investment. Investment would fall below savings, which the interest rate would not correct because it is determined in the money market. Instead, output would fall (since it is not fixed, as in the Treatise), lowering income, and therefore lowering savings until it again equalled investment. The size of the fall in income would depend on the size of the multiplier which, in turn, depends on the propensity to consume. And as output fell, so too would the demand for labour, which would cause

---

21 See Kahn (1984, p. 59) quoting the Preface to the French edition of the General Theory, in which Keynes says exactly what we have just said.

22 Quoted by Temin and Vines (2014, p. 25). This view was shaped by Keynes’s experience of Britain’s return to the gold standard at an overvalued rate in 1925 which he had analysed in The Economic Consequences of Mr Churchill (Keynes, 1925).

23 The consumption function was necessary in order to connect the discussions in the Circus in 1930–31 with the analysis of the multiplier that Kahn had invented, and published, at that time (Kahn, 1931). In his article Kahn described equilibrium in the goods market as being brought about when injections—in the form of investment—are equal to leakages—which he described as payment of taxes and payments for imports. Only when it was understood that the marginal propensity to consume was less than one, so that as income rises there would be leakages to savings (as well as to taxes and imports), was it possible to describe the IS curve as showing that an equilibrium is reached in the goods market when savings equals investment. Kahn (1984) describes how this idea did not emerge until 1932.
unemployment. Wages would not adjust downwards so as to remove this unemployment, because they are sticky. Keynes tells this simple, clear story in chapters 2 and 18 of the *General Theory*.

However, the analysis in the *General Theory* does not yet encompass all that came before. Having triumphantly shown, in his new theory, that movements in output can ensure that savings equals investment, Keynes was determined not to encompass Marshall’s view—that it is the interest rate which ensures that savings equals investment—but rather to declare it to be incoherent. He simply could not understand that, *if* full employment was ensured, *then* the interest rate would necessarily be at the level which equates savings and investment. This lack of understanding frustrated Roy Harrod and led to the remark by Harrod with which we began this article.24

Exasperatingly for us now, and for other readers at the time, Keynes never formalized his system in the *General Theory*. His forbearance was deliberate; he had good reasons to resist the constraints on thought that are imposed by algebraically formalizing economic analysis.25 But the model in the *General Theory* is a complex one, with four markets in play: for goods, labour, money, and bonds. We would all agree now—and one person in particular took the view then (see below)—that there are too many balls in play for a verbal narrative to be adequate. It is obvious, with so much in play, that a set of equations was needed.

*Reaching IS/LM: the simplifying tricks*

James Meade is the person who did what was necessary. He starts his paper (Meade, 1937) by making it clear that his model is a short-period one, and that his method is one of static analysis; there is no discussion of dynamic processes.

Meade’s first, and major, achievement is to show that his system passes the encompassing test. He makes it clear that, in his static model, unemployment can emerge if wages are exogenous, but that, if wages are endogenous, then there will necessarily be full employment.26 As a

---

24 On 3 August 1935, after reading the proofs of the *General Theory*, Harrod wrote to Keynes, ‘I feel that you do less than justice to Marshall and the traditional view’ (Keynes, 1973, p. 534). There follow weeks of dogged correspondence in which Harrod becomes increasingly irritated; it becomes abundantly clear to Harrod that Keynes simply does not understand the point at issue (Keynes, 1973, pp. 534–65).

25 He explained these in his obituary for Marshall which he had published 12 years earlier: ‘Unlike physics, for example, such parts of the bare bones of economic theory as are expressible in mathematical form are extremely easy compared with the economic interpretation of the complex and incompletely known facts of experience, and lead one but a very little way towards establishing useful results’ (Keynes, 1924, p. 333). He continues in a footnote: ‘[B]ut the amalgam of logic and intuition and the wide knowledge of facts, most of which are not precise, which is required for economic interpretation in its highest form is, quite truly, overwhelmingly difficult for those whose gift mainly consists in the power to imagine and pursue to their furthest points the implications and prior conditions of comparatively simple facts which are known with a high degree of precision.’

26 ‘If we suppose that the money wage-rate would fall so long as any labour remains unemployed, the system cannot be in equilibrium without full employment. In this case the money wage rate is no longer given but the equilibrium volume of employment is now a given quantity, namely equal to the given volume of labour seeking employment.’ (Meade, 1937, fn. 1)
consequence, if there is full employment, then the interest rate will have to be at the level at which savings equals investment, as in Marshall’s system.27

But Meade’s model is very complex. In his paper, ludicrously mis-named ‘A Simplified Model of Mr Keynes’ System’, he sets out a system that includes all four of the features identified above: an upward-sloping supply curve; liquidity preference; an exogenous money wage; and a consumption function (and so the multiplier). And all four of the General Theory’s markets are in play: for goods, labour, money, and bonds. Furthermore, the markets for consumption goods and investment goods have separate supply curves.28 Doing all of this required a system of eight equations. There is no intuitive analysis at all, merely a series of stability conditions and a set of comparative static results. It is clear that Meade’s interest is in how the system works as a system. In sum, Meade’s model was not simple and it was certainly not a toy.29

However, Meade’s article did provide the tools which Hicks needed to create the IS/LM toy. This is clear from the fact that Hicks uses Keynes’s notation, and a number of Meade’s equations (Young, 1987). It fell to Hicks to do the real simplifying, and to draw out the key idea. He showed, using a two-dimensional diagram, how insufficient demand and sticky wages could together create unemployment, and he achieved this by ignoring the effects of any rise in prices, which, because of diminishing returns to labour, should have been allowed to cause prices to rise when output rises, even though wages were assumed to be exogenous. This analytical trick30—namely the assumption of fixed prices—reduced the plausibility of the model, but it enabled Hicks to boil the model down to just two equations—the IS curve and the LM curve—instead of Meade’s eight equations. Together, these equations determined just two variables—output and the interest rate—as shown in a simple diagram.

27 Without the algebra it actually took Harrod six weeks of correspondence to finally persuade Keynes of this point (see Keynes, 1973, p.557).
28 This was a hangover from Keynes Treatise. In that book the markets for consumption goods and investment goods were analysed separately, for not very good reasons.
29 Meade never did toys. There were three more times in Meade’s career as a macroeconomist when he produced a complex piece of theory that someone else collapsed into a simple toy. The first two of these is the Swan diagram (Swan, 1963) which shows how internal and external balance can be achieved in an open economy using both fiscal policy and exchange rate policy. The second is the Mundell–Fleming model which compares the operation of fiscal and monetary policy under fixed and floating exchange rates. (Boughton, 2003). Both of these were derived derived directly from Meade’s The Balance of Payments (Meade, 1951) a horrendously complex book for which Meade won the Nobel Prize. The third occasion occurred many years later. In his Nobel lecture in 1977, Meade set out his ideas for a reconstruction of macroeconomic policy using a rule-based policy for the interest rate, as a way of managing aggregate demand, so as to achieve a nominal GDP target (Meade, 1978). The results were published ten years later in a book called Macroeconomic Policy: Inflation, Wealth and the Exchange Rate on which one of us worked (Weale et al., 1989). The results in that book about nominal-GDP-targeting, using an interest-rate rule, paved the way – at least in the United Kingdom - towards the use of the interest rate as the policy instrument in a flexible inflation-targeting regime. John Taylor’s proposals for a ‘Taylor rule’ only emerged a year later (see Taylor (1993, 2000) and Carlin and Soskice (2015)).
30 This trick meant that he could also helpfully abandon the disaggregation of output into a consumption-goods sector and an investment sector. There is actually also another problem with Meade’s model, concerning his treatment of investment and the bond market, which Hicks sidesteps (see Rappoport (1992)). But that need not detain us here.
When Hicks’s IS/LM toy appeared, it immediately became the way of setting out the *General Theory*, for two reasons. It struck the required balance between simplicity and detail. And it also encompassed what had come before: the very title of Hicks’s paper, ‘Mr Keynes and the Classics’, says it all.

David Vines once heard Hicks explain how he had been able to do this. The occasion was a captivating talk which Hicks gave to the Marshall Society, the undergraduate economics club at Cambridge University, one evening in the late 1970s. During the 1930s, Hicks had been working on *Value and Capital*, his book on general equilibrium theory that would be published three years after the *General Theory* in 1939 (Hicks, 1939). He had learned, he said, that one can analyse the general equilibrium of a three-good economy (in this case goods, money, and bonds) in a two-dimensional diagram. He explained that one can ignore the fourth market, for labour, when drawing the diagram, since the money wage is assumed to be fixed and so the labour market may not clear. He went on to say that one needs to put output and the interest rate on the two axes. All that remains to be done after that, he said, is to draw the market-equilibrium lines for the goods market, the money market, and the bond market. One can then ignore the bond-market line, since, using Walras’ Law, it passes through the point at which the other two lines intersect. Hicks then smiled as he pointed out that the IS curve and the LM curve are all that remains in the diagram. And so, he said, one can show that if the IS curve moves to the left, output will fall and unemployment will emerge. And if the IS curve intersects with the LM curve at a level of output consistent with full employment, then the interest rate will be at the level which ensures that savings equals investment. Job done!

**Going beyond IS/LM: Samuelson’s neoclassical synthesis**

The IS/LM model quickly became the main model used in macroeconomics. It embodied the message of chapters 1–18 of the *General Theory*, but it assumed that price level is constant as well as exogenizing the wage. It fell to Modigliani (1944) to relax the constant-price-level assumption, reintroducing the assumption of diminishing returns (which Hicks had jettisoned). Samuelson (1955) called what emerged the ‘neo-classical synthesis’. It combines all the markets analysed in the *General Theory*—those for goods, labour, money, and bonds—in a fully consistent way. Such a set-up allowed Samuelson to produce an internally consistent integration of what he called short-term outcomes (in which unemployment is possible) and what he called longer-term outcomes (in which resources are fully employed).

Formally, Samuelson’s neoclassical synthesis generalizes the IS/LM system into a set-up with three simultaneous equations: the IS curve, the LM curve, and an upward-sloping supply curve. The illustration of this model requires a rather clunky three-quadrant diagram—one which never became wildly popular—but which nevertheless coherently describes the simultaneous determination of output, the interest rate, and the price level. It also depicts how employment

---

31 Paul Krugman sets out an IS/LM-like diagram in a piece written many, many years later, in which he goes through almost exactly the same presentational steps as Hicks did in his talk (Krugman, 2000).
is determined. The illustration shows how each of these variables will respond to shocks to aggregate demand, productivity, and the price level, and to changes in fiscal or monetary policy, in both what Samuelson calls the short term (in which wages are exogenous) and in what he calls the long term (in which flexible wages ensure that resources are fully employed).

Samuelson’s neoclassical synthesis became the consensus view of macroeconomics by the mid-1950s in the United States and it remained the dominant paradigm for another 20 years (Blanchard, 2008). The ideas in this synthesis were worked on in empirically and incorporated into the early generations of econometric models, starting with the models estimated by Klein (See for example, Goldberger and Klein, 1955).

The key oversight in the IS/LM system: dynamics and multiple equilibria
For all its success in encompassing past research, the IS/LM toy was unable to analyse one of Keynes’s crucial insights in the General Theory: the possibility of self-reinforcing dynamic processes. Such processes can create instabilities, and multiple equilibria, the central topic of the present paper.

Chapters 1–18 of the General Theory focus on the short run, which is the period to which Meade and Hicks deliberately constrained themselves. In his paper, Meade (1937) assumes that an equilibrium exists, and that it is unique. (His reason was simply that the model was linear and the number of equations was equal to the number of unknowns.) He then imposed a set of stability conditions for this equilibrium, and then, subject to these holding, went on to carry out a series of comparative static exercises. He justified his procedure as follows:

It is of course possible that in the real world the system is unstable. But in what follows we shall assume that equilibrium is stable, since it is not possible to discuss the effect of given changes on the volume of employment if any small jerk to the system may start it off in one direction or the other in search of a completely new equilibrium. (Meade, 1937, p. 102)

Hicks never mentions existence, stability, and uniqueness of equilibrium in his IS/LM paper; he implicitly assumes what Meade had made explicit. Meade’s research strategy—to assume the existence, stability, and uniqueness of equilibrium—turned into the standard rules of engagement for macroeconomic theory, rules which persist to the present day.  

---

32 The diagram can be seen, for example, in Blaug’s book (Blaug, 1968, p. 636). The equations of this system can be found in Modigliani (1944, p. 46) or, in simpler form, in Blaug (1968, p. 635).
33 This set of ideas is discussed by Farmer in his paper in this issue (Farmer, 2020).
34 Meade is following the method in Marshall’s Principles in which the supply curve and the demand curve intersect only once and stability is ensured because the supply curve slopes up and the demand curve slopes down. Marshall did consider a downward-sloping supply curve, due to increasing returns, but insisted that any such supply curve should slope down less steeply than the demand curve.
35 This procedure is often attributed to Paul Samuelson who called it the ‘correspondence principle’ in his Foundations (Samuelson, 1947). But Meade had articulated this idea 10 years earlier and it really came from Marshall.
However, in chapter 19, Keynes looked at a longer-run period of time, and used theoretical analysis to examine the possibility of an unstable process. He showed that, in the presence of unemployment, a gradual adjustment of wages (an elaboration of the idea of ‘sticky’ wages) might make unemployment self-reinforcing. If, when wages and prices fell, nominal interest rates did not fall more than one for one with prices, then the real interest rate would rise. This would cause a further decline in aggregate demand, and unemployment would rise further. Over time, the economy might move further and further away from full employment in an unstable way. Keynes writes,

(i) it would be much better that wages should be rigidly fixed and deemed incapable of material changes, than that depression should be accompanied by a gradual downward tendency of money-wages, a further moderate wage reduction being expected to signalise each increase of, say, 1 percent in the amount of unemployment. For example, the effect of an expectation that wages are going to sag by, say, 2 percent in the coming year will be roughly equivalent to the effect of a rise of 2 percent in the amount of interest payable for the same period. The same observations apply mutatis mutandis to the case of a boom. (Keynes, 1936, ch. 19, p. 265)

Keynes clearly appreciated that self-reinforcing dynamics could exist, but Meade and Hicks simply excluded this possibility. It is this exclusion that we set out to challenge in the present paper.

(ii) Inventing the Solow–Swan model to explain economic growth

The IS/LM model, and its generalization in Samuelson’s neoclassical synthesis (1955), was adept at modelling short-run fluctuations, but it did not describe longer-run movements in capital, output, or employment. In fact, the model actually takes productive potential and the full-employment level of output as given, even though investment is influencing the level of the capital stock. Soon after the publication of the General Theory, Harrod (1939) and Domar (1946) had begun to study these problems. But the Harrod–Domar growth model was unsatisfactory. In 1956, Robert Solow and Trevor Swan saw, almost simultaneously, what was needed and proposed nearly identical models in response. In what follows we tell the story of why Solow managed to get to a simple toy model, which made him famous, but why Trevor Swan, by contrast, largely disappeared from view. We reveal that there were actually good analytical reasons for what happened.

36 The macroeconomic consequences of this instability process were examined formally by James Tobin, but not until much later (Tobin, 1975).

37 The material in this section will be discussed in more detail in a forthcoming book by David Vines, about the history of the ideas underpinning macroeconomic policy-making in Australia, and the international effects of these ideas (Vines, forthcoming). The book will draw on extensive archival material in the Australian National Library, including the papers of Trevor Swan.
Robert Solow considered that the Harrod–Domar model was unsatisfactory because it describes endogenous cycles of boom and bust. The model assumes a constant capital to output ratio, $v$, and a constant savings rate, $s$. As a result, output and capital will grow at a ‘warranted rate’ of growth, $g = s/v$. But the model also assumes that, to achieve full employment, output needs to grow at the ‘natural rate’, $n + x$, where $n$ and $x$ are the exogenous rates of population growth and labour-saving technical progress. Since $s$, $v$, $n$, and $x$ are all exogenous, this would only happen by chance. Solow went on to observe that the high savings rates of the 1950s would imply that capital was growing faster than the natural rate, i.e. that $s/v > (n+x)$. This would cause a shortage of labour to emerge, and would lead to a collapse in investment, since firms would know that they would be unable to find enough workers to man their growing stock of capital. Collapsing investment would, in turn, lead to a depression, until capital depreciated enough to require new investment, thus starting the process all over again. This process, Solow maintained, was obviously not what was happening in the post-war boom years.

Solow solved the problem by relaxing the assumption of a fixed capital to output ratio, and endogenizing wages. This assumption meant that a high level of investment, such as that observed in the 1950s, would make labour scarce. That would bid up wages, encouraging more capital-intensive production, i.e. increasing $v$, thereby ensuring that $g = s/v$ would fall. Eventually $g = s/v$ would converge to the natural rate of growth, $n + x$.

Solow’s genius was to show how this story could be collapsed into a two-dimensional diagram, one which has become familiar in all undergraduate textbooks. To do this he was prepared to assume constant returns to scale in the use of labour and capital, however unrealistic this might seem to us. In the version that appears in textbooks, the diagram is made possible by re-writing the production function to show output per effective worker as a function of capital per effective worker, which is only possible if there are constant returns to scale. To draw the diagram one introduces labour-saving technical progress at an exogenous rate, which causes output per worker to grow over time. Putting capital per effective worker on the horizontal axis allows us to plot a diagram showing investment per effective worker and savings per effective worker as a function of capital per effective worker. When the savings rate rises, as

---

38 The following account is based on a conversation which David Vines had with Bob Solow in Glasgow in 1991. Solow confirmed that the account is correct in an email which he sent to David Vines in the early summer of 2020. He writes, ‘The land scarce case would lead to diminishing returns to scale in capital and labour and the model would become more Ricardian,’ and then continues in a footnote, ‘[o]ne can imagine the theory [developed here] as applying as long as arable land can be hacked out of the wilderness at essentially zero cost [so that diminishing returns do not apply]’ (Solow, 1956, p. 67).

40 The original version of the model in Solow’s paper abstracts from technical progress, but Solow already hints at how to include it at the end of his paper.

41 The first curve shows the investment per effective worker that is needed to ensure that capital keeps pace with the effective labour force. After allowing for depreciation, at rate $\delta$, this has a positive gradient of $n + x + \delta$. The second curve shows savings per effective worker. This also has a positive slope, but one which gradually falls as capital per effective worker increases, because of diminishing returns to capital. Thus, subject to some obvious conditions, the two lines necessarily intersect. At this point the capital-to-effective-worker ratio is at an equilibrium value, and the Harrod–Domar problem is solved. When the ratio is above this equilibrium value, it is falling, and vice versa, so the equilibrium is stable.
in the 1950s, the savings curve moves up, and capital per effective worker gradually accumulates to a higher level. Output per person is higher in the long run, and everyone is better off. However, the growth rate of output does not permanently increase. It returns to the warranted rate, \( n + x \), the one that was identified in the Harrod–Domar model.

This toy model has proved extremely popular: Solow went on to win the 1987 Nobel Prize. But whatever happened to Trevor Swan? Many years later Solow asked, ‘Why did the version in my paper become the standard, and attract most of the attention?’ (Solow, 2007, p. 3). And many others have wondered the same thing (e.g. Dimand and Spencer, 2008). Solow went on to provide a compelling reason: Swan’s diagram was not very clear or user friendly (Solow, 2007, p. 4).\(^{42}\)

However, the reason why Swan’s diagram was more complex than Solow’s is that Swan was interested in a different question. His wish to answer this question actually prevented him from drawing Solow’s diagram, and explains why his diagram needed to be less user friendly than Solow’s. In brief, the story goes as follows. Those not interested in this story can skip to Section II (iii) below.

In the early-to-mid-1950s, Swan was not concerned with the boom–bust problem with the Harrod–Domar model that had worried Solow. He thought that these cycles could be stabilized by Keynes’s policies for managing aggregate demand, allowing economists to focus once again on the issues of long-run growth that had interested Adam Smith, John Stewart Mill, and David Ricardo (Swan, 1956, p. 334).\(^{43}\)

Swan’s major pre-occupation was with the growth prospects of the Australian economy, which had a very high rate of immigration, and so a high rate of labour-force growth. While Solow’s concern was to reconcile the growth rate of capital with that of the effective labour force, when a high savings rate meant that the former was likely to exceed the latter, Swan’s policy question was more or less the opposite. He was concerned to show that, when such a reconciliation had happened, a more rapid rate of growth in the labour force would reduce the long-run capital to labour ratio, unless Australians were prepared to save more than they were doing at the time. Moreover, Swan was also concerned with an additional question: what was the maximum immigration rate that Australia could support?

We can immediately use Solow’s diagram, as explained in footnote 41, to provide an answer to Swan’s first question. An increase in the rate of population growth, \( n \), will rotate the

---

\(^{42}\) Solow said this again in a follow-up email which he sent to David Vines later in the summer of 2020.

\(^{43}\) Swan wrote two important theoretical papers on managing aggregate demand in the years immediately preceding his growth model. These show how boom–bust cycles can be dealt with in open economies, using a combination of fiscal policy to achieve full employment, and exchange rate policy to ensure a satisfactory international position. These papers were widely circulated in Australia at the time. Although they were not published until much later (Swan, 1960, 1963), it was clear to Swan’s Australian colleagues why he was not concerned with the boom–bust problem.
investment-per-effective-worker line anticlockwise, lowering the equilibrium value for the capital-per-effective-worker ratio and thus lowering the equilibrium level of real income per worker over time. This effect can only be avoided if the savings rate, $s$, is increased at the same time, thereby raising the line showing savings per effective worker. As Bob Solow wrote in an email to David Vines in July 2020,

The point you emphasize, that faster growth of population leads to lower capital intensity, and therefore lower output per head, unless offset by a higher saving rate, follows instantly and obviously with one pencil-stroke in the standard diagram in my 1956 paper, and much less so in the presentation that appears in Trevor’s paper, though of course it is there.

Swan’s much less transparent analysis is provided in Figure 1 in his paper. The diagram which he uses can indeed be used to demonstrate that convergence to the warranted rate of growth will happen, as in the Solow paper. But the analysis is really rather clunky.44

Swan’s second question, on maximum immigration rates, was more difficult to answer, and here lies the problem. In the immediate post-war period, Australia did not produce many advanced manufacturing goods. So, a high immigration rate required more exports to pay for the additional imports required by a larger population. At the time, Australia’s exports consisted almost entirely of primary commodities; to produce more of these would require more intensive use of the limited supply of good land. As a result, the economy would run into diminishing returns.45 Therefore Swan could not make Solow’s assumption of constant returns to capital and labour, and so he could not use Solow’s diagram.

To understand the problem intuitively, consider what happens if the level of technology is constant, and the immigration rate is permanently at rate $n_1$. If capital were also to grow at rate $n_1$, then output per person would perpetually decline because of diminishing returns. The only

---

44 Swan puts growth rates, not levels, on the vertical axis, and the ratio of output to capital, $Y/K$, on the horizontal axis. The exogenous growth rate of the labour force, $n = n_1$, is a horizontal line. The growth rate of capital $g_K$ is an upward-sloping line $g_K = sY/K$. Like Solow, Swan initially ignores technical progress and assumes constant returns to scale. He also assumes that the production function is Cobb–Douglas, so the growth rate of output $g_Y$ is a weighted average of $g_K$ and $n_1$, with weights summing to unity.

The lines cross at an equilibrium point at which $g_Y = g_K = n_1$, where the output to capital ratio equals $(Y/K)^*$. When $Y/K > (Y/K)^*$ capital grows faster than labour, and thus output, so the ratio $Y/K$ falls. The opposite is also true, and so the equilibrium is stable.

Swan then adds technical progress to the analysis, in the form of increases in total factor productivity (TFP) at rate $z$, rather than labour-saving technical progress at rate $x$, as in Solow’s case. Such technical progress will shift up the $g_Y$ line. A stable equilibrium now exists where $g_Y = g_K$. This happens at a growth rate which is greater than $n_1 + z$, because technical progress increases the productivity of capital.

Swan uses the diagram to answer his first question. If $n$ increases from $n_1$ to $n_2$, the equilibrium value of $Y/K$ will rise, implying the equilibrium value of output per person will fall, unless the savings rate $s$ rises at the same time.

45 There was certainly no “‘wilderness’ available at zero marginal cost’ as Solow had assumed; in Australia much too much of the land is of low quality.
way to offset these diminishing returns is technical progress, which might hold their effects at bay.

Swan uses the diagram of his model to tackle the problem the other way round. He shows that, for any given rate of technical progress, and any given savings rate, $s$, there is an immigration rate which would leave output per head stable. He then goes on to show that a lower immigration rate than this would enable output per head to rise forever, even in the presence of diminishing returns. Swan’s diagrammatic way of showing this is both clever and necessary, given that constant returns to scale is no longer in play.\footnote{Figure 2 in Swan’s paper, building on Figure 1, does this. He relaxes the assumption that the population growth rate $n$ is exogenous. Instead, he solves for the population growth rate that equals the output growth rate, when capital grows at $g_K = sY/K$. Swan’s Figure 2 has two upward-sloping lines plotting capital growth, $g_K$, and output growth, $g_Y$. Output growth has a flatter slope than capital growth because of diminishing returns, and a vertical intercept that depends on TFP growth. If TFP growth is zero, the lines intersect at the origin. In this case $Y/K$ perpetually falls towards zero over time, a depressing outcome, that no immigrant would sign up for. However, if TFP growth is positive, then the lines intersect at a point where $Y/K$ is positive, output per head is constant, and output and (by construction) population have permanently positive growth rates; this is the immigration rate which is consistent with keeping income per head constant. Furthermore, if policy lowers the immigration rate below this level, then output per head will be able to grow in perpetuity. This diagram thus provides an answer to Swan’s second question.

Solow (2007, p. 3) notes with some surprise that Swan used a Cobb–Douglas production function, rather than Solow’s more general production function. The above account shows that Swan needed to do this to make the growth rates of TFP, capital, and labour additive, and so amenable to a simple diagram. This account also shows that Swan needed technical progress to increase TFP, rather than to save labour, in order to obtain a steady-state equilibrium with a constant level of $Y/K$.}

It is impressive that Swan was able to carry out this analysis diagrammatically. Nevertheless, compared with Solow’s diagram, Swan’s picture looks rather awkward. And nowhere else in the world did economists see the need to build a theory of economic growth that was entangled with the problem of immigration policy, in the way that seemed necessary for Australia. So, the sacrifice in clarity which was needed in order understand this policy problem did not seem necessary. That is why we all know and love Solow’s toy, and not Swan’s toy.

Getting to a simple toy is not only difficult. Sometimes it is not possible.

(iii) The failure to invent an agreed response to the stagflation of the 1970s

In the 1970s the major macroeconomic problem was stagflation and—once again—the existing models were of no use. The existing IS/LM framework, which had done so well in explaining the unemployment of the 1930s and in recommending fiscal expansion as the appropriate policy response, now failed desperately in the face of both unemployment and inflation. And growth was faltering.
The Phillips curve had already been added to Samuelson’s neoclassical synthesis (Phillips, 1958), setting Keynes’s ‘sticky wages’ in motion. Such a set-up suggested that expansionary policies might lead to inflation (Samuelson and Solow, 1960). But it became clear that a new toy was needed that went beyond the simple Phillips curve.

The answer which an influential group of economists adopted in order to do this was to theoretically microfound the supply side of the economy as the outcome of optimizing behaviour. But doing this led to the pursuit of two very different approaches that can be loosely grouped into real business cycle (RBC) theory on the one hand and new Keynesian (NK) theory on the other. These two developments showed how toys can be useful for the creative process: by getting to the heart of difficult issues quickly. But the result was two approaches which were orthogonal to each other. This led to a period of time in which macroeconomics was an incoherent discipline.

The first preliminary step in developing a microfounded supply side was to incorporate expectations, so that expected future inflation depended on immediately past inflation (Phelps, 1968; Friedman, 1968). As a result of doing this, in 1968, Milton Friedman threw down a fundamental challenge: no active macroeconomic policy would be needed in the longer run, he said, because flexible prices would ensure a return to full employment after any shock. Even if shocks did cause short-term disturbances, in the longer term the macroeconomy would be self-stabilizing providing only that the central bank ensured a steady rate of growth of the money supply. And, as both Phillips and he had argued in the 1950s, attempts to stabilize the economy might actually be destabilizing (Friedman, 1953; Phillips, 1954, 1957).

The second preliminary step was the Lucas critique (Lucas, 1980), which argued that forward-looking expectations made the Phillips curve vertical, and so made monetary policy pointless. The focus, said Lucas, should be on modelling supply. This view was based on the idea of rational expectations, which maintained that the economy consisted of optimizing agents who set prices and wages in instantly clearing markets for labour and goods, with a full understanding of the economy’s structure. As a result of this, any attempt to use policy to stimulate the economy along an upward-sloping Phillips curve—by lifting inflation above expectations—could not possibly succeed, since the resulting inflation would immediately be anticipated, leading instantly to higher short-run inflation. Moreover, the possibility that policy-makers might take such action would also be anticipated. As a result, the outcome of a regime in which policy authorities were able to respond to inflation might well be higher long-run inflation, since, whenever inflation was low, there would always be a risk that policy-makers would seek short-run gain by stimulating the economy. Nevertheless, such an active inflation-targeting regime would never be necessary because wages and prices would always instantly adjust to ensure that resources remained fully employed (Lucas and Sargent, 1978).

---

47 As noted above, Keynes had already discussed what are essentially Phillips-curve ideas in some detail in his evidence to the Macmillan Committee in 1930. See Temin and Vines (2013, ch. 3) and Temin and Vines (2014, ch. 3).
The first analytical response to these developments was RBC theory. This approach accepted the Lucas critique and abandoned stabilization policy, in order to concentrate on modelling the supply side of the economy. The Ramsey growth model, on which this set-up was based (Ramsey, 1928), begins with a Solow–Swan model into which are introduced optimizing agents who decide on the proportion of their income to save and invest, in order for the economy to grow at an optimal rate.\footnote{Ramsey’s paper predated the papers by Solow and Swan by nearly 30 years. But it is helpful to think of his growth model as being a development of their model, even although their model appeared much later.} Quite quickly the ‘freshwater’\footnote{So called because they tended to live around the freshwater Great Lakes of the US.} macro theorists who adopted this approach added autoregressive shocks to technology in the production function, which caused the stock of capital to cycle rather than converge gradually to its optimal growth path, thereby creating RBC theory. The Ramsey model can still be comprehended in a simple two-dimensional diagram, plotting the level of consumption against the level of the capital stock, to show how consumption immediately adjusts, and capital gradually accumulates, following any productivity shock.

Because of its set of strong simplifying assumptions—which now look both extreme and wrong—we can clearly describe this Ramsey–RBC set-up as a toy model. It was designed to illustrate only one aspect of the macroeconomic process. Its protagonists believed that there was nothing else going on in the macroeconomy, which was clearly not the case.

The second analytical response to these theoretical developments—the micro-founding of the supply curve and the deployment of the Lucas critique—was New Keynesian theory. This continued to assert that wages and prices remain sticky, even in a forward-looking, rational-expectations world. The necessary work was done by Fischer (1977), Taylor (1980), and Calvo (1983) in three ground-breaking papers. They allowed the Phillips curve to regain its positive slope, and restored a role for monetary policy in a world with fully rational optimizing agents. All of these papers involved overlapping contract ideas. If agents can only optimize their prices at particular times, which are not the same for everyone, then when doing so they will have to consider the delay until other agents can do the same. This constrains the amount of adjustment which any group of agents will choose to make in response to a shock, reinstating wage and price stickiness, even in an optimizing world.

This New Keynesian school of thought ultimately led to a paper by Clarida, Galí, and Gertler (1999) called ‘The Science of Monetary Policy’ (CGG) and to Galí’s textbook entitled *Monetary Policy, Inflation, and the Business Cycle* (2015). In some ways these are strange pieces of work, in that the only private-sector agents who exist are consumers; there is no investment and so no endogenous capital stock or growth. But the justification provided for this strange and minimalist assumption was exactly what one would expect to be provided in the defence of a toy model: an economy like this is all that is needed to explain ‘Taylor-rule macro’.
Taylor-rule macro is unquestionably a toy, since at the heart of this model there are only three equations. First, the Euler equation shows behaviour by optimizing consumers, standing in for the IS curve. Second, the Calvo–Phillips curve shows how there can be price stickiness in an optimizing world. And third, an equation, standing in for the LM curve, shows how the monetary authority sets the real interest rate either by following a Taylor rule, which relates the real interest rate to the inflation rate and the output gap, or by setting the real interest rate at an optimal level, given the two aspects of private-sector behaviour described in the Euler equation and the Phillips curve. Taylor-rule macro can also still be exposited in a simple two-dimensional diagram, plotting the level of inflation against the level of the output gap (see Taylor, 2000; Carlin and Soskice, 2005, 2015, ch. 3). Such a picture can be used, for example, to show how and why output is reduced when there is a shock to inflation, in order to bring inflation back on track. It can also be used to study the effects of any shock to aggregate demand.

It is clear that the macroeconomics discipline at the end of the last century was in a very odd position. Clarida, Galí, and Gertler were busy discussing the ‘science’ of stabilization policy in a world without growth. And RBC theorists were studying the process of economic growth in a world in which there was no need for any stabilization. It appeared—to use an old English expression—that ‘never the twain shall meet’. Both approaches clearly failed to respect what had come before.

(iv) Where we are now: inventing the NK-DSGE model

The NK-DSGE model emerged as an attempt to bring some order to this shambles, and to do so in the simplest possible way. Unlike the earlier chapters in this Toy Story, the NK-DSGE model was not responding to the failures of earlier toys to analyse reality. Rather, it could be described as a peace plan for twenty-first-century macroeconomists. It involves a set of constructs designed to enable the RBC and New Keynesian camps in the macroeconomics profession to talk to each other again; what Goodfriend and King (1997) called a ‘new neoclassical synthesis’.50 51 We believe that the NK-DSGE model is a toy, albeit a large one, which modern computers allow researchers to play with as easily as two-dimensional diagrams

50 The analogy with Samuelson’s neoclassical synthesis is exact. That too was a peace plan, although turned into a road map for an ongoing research programme.
51 A reader of an earlier draft of this paper has pointed out to us how little there is in the NK-DSGE toy model about financial markets and the relation of these markets to expenditure decisions. Indeed, it is possible to argue that the short-cuts taken in constructing the NK-DSGE paradigm have served to divert the macroeconomics discipline from the theoretical and empirical insights about credit markets and the financial system which have been developed by the profession ever since the 1960s. Nobel Laureate James Tobin and his colleagues were critical in developing this understanding. (See Tobin, 1969, 1982, and Buiter, 2003) The paper by John Muellbauer in this issue of *OxREP* (Muellbauer, 2020) takes forward the empirical investigation of these insights. Nevertheless these understandings have, for whatever reason, had rather little influence on the mainstream toy–model tradition within the macroeconomics profession. The history-of-thought discussion which follows reflects this fact. Muellbauer’s paper makes clear that such an approach will need to change.
were played with by earlier generations. But as we will see, it is a toy which is not able to capture the salient features of the run-up to the 2008 crisis, or the slow recovery since then.

We take the Smets–Wouters (2007) model as the benchmark version of this model. The model by Christiano et al. (2005) is another representative example. The Appendix sets out the equations of the model in detail. Here—in three brief paragraphs—we provide a description of the framework and structure of the model.

The model incorporates the four markets identified in Samuelson’s neoclassical synthesis: for goods, money, labour, and bonds. At the core of the model is a Ramsey model of optimizing agents who engage in a dynamic process of capital accumulation and growth, as in RBC theory. But there are two important frictions to this growth process. First, investment is done by firms (not consumers) and faces adjustment costs, so savings by consumers are not costlessly invested (unlike the Ramsey model) and firms must be induced to do what is necessary (by a sufficiently attractive value of Tobin’s Q). Second, prices are slow to adjust, as in the CGG model, so price signals are not enough to encourage firms to invest what is necessary to maintain full employment of resources: active policy is needed.

In the long run, the level of output is determined by the size of the labour force and the level of technology, both of which grow exogenously. Along this equilibrium growth path, capital accumulation is driven by a representative firm’s decision to invest and a representative consumer’s decision to save. Financial intermediation ensures that consumers hold the equity created by investment, at a real interest rate which, in the long run, is given by the rate of time preference. In the short run, shocks to technology, and to its expected rate of change, disturb this growth path; as do shocks to the desire to save, and to the financial intermediation process. The model can nevertheless be solved for the ‘neutral’ real interest rate: an endogenous real interest rate in which both capital and labour remain fully employed, and there is no inflationary pressure, regardless of shocks.

But in the short run, the existence of nominal rigidities, and the deviations of the real interest rate from its neutral level, create the possibility that output can be different from the level at which capital is fully employed and labour is willingly supplied. When this happens, inflation (or disinflation) is generated. This leads to a role for monetary policy, both to pin down the rate of inflation using a Taylor rule to set the real interest rate and—subject to that objective being achieved—to ensure that demand is just sufficient for resources to be fully employed.

---

52 This model was presented in our paper in the previous issue of *OxREP* on Rebuilding Macroeconomic Theory, vol. 34 nos 1–2. We present it again in detail in our Appendix because we make extensive use of it in what follows.
(v) Why the NK-DSGE toy is important, even if difficult to use

The NK-DSGE toy encompasses all of the preceding toys we have been describing in this section of the paper. In the absence of both adjustment costs to capital and sticky prices, it collapses to the Ramsey model. If we abstract from the capital accumulation process, and ignore the role of investment in aggregate demand, then it collapses to the CGG model. But if all of these important features are allowed to operate then we can get an idea of how the economy operates as a whole.

The full NK-DSGE model—in which both of the frictions are present and in which we allow for the effects of capital accumulation—seems complicated compared to the simple toys we have described up until now. It has four differential equations (for capital, for consumption, for Tobin’s Q, and for inflation): one of which is a slowly adjusting ‘state variable’ (capital), and three of which are jump variables (for consumption, inflation, and Tobin’s Q). It does not have a simple analytic solution, or a simple two-dimensional diagram, because the jump variables create an outcome which is potentially unstable in three different dimensions. Solving it requires satisfying three transversality conditions: consumption must stop growing, Tobin’s Q must return to its base value of unity, and inflation must settle at its target value. This can be achieved by computationally intensive ‘target-shooting’—experimenting with the initial values for the three jump variables until an outcome is reached in which the transversality conditions are satisfied far into the future (see Fair and Taylor, 1983). Or the model can be linearized and solved to eliminate unstable roots, which ensures that it converges to the unique, stable equilibrium where all three of the transversality conditions hold (see Blanchard and Kahn, 1980).

Nevertheless, we argue that this creature can still be played with as a toy—albeit a moderately complicated one. The reason is that modern computer technology and programmes (particularly MATLAB and Dynare) allow the model to be simulated quickly in such a way as to provide quick and intuitive ‘first passes’ at questions. In the place of the two-dimensional diagrams and shifting curves of the small toys described above, playing with the NK-DSGE model requires linearization, use of the Blanchard–Khan method, and a study of the impulse response functions which can be found by running simulations using MATLAB.

Two illustrations of how the NK-DSGE model works as a toy, analysing fluctuations around a given steady state, are given in the Appendix. These are the kinds of valuable exercise which Blanchard was describing in the quotation at the very beginning of this paper. In our discussion in the Appendix we show how a technology shock causes all of consumption, investment, capital accumulation, and aggregate supply to move simultaneously, even if a clever central bank moves the real interest rate enough to prevent prices from fluctuating. And in response to

53 Since, in the version of the model presented in the Appendix, there is no labour force growth or technical progress.
an inflation shock, we show how all of consumption, investment, capital accumulation, and aggregate supply will move endogenously in the process by which the central bank controls inflation.

These simulations encompass, and generalize, all that has been captured in the earlier toy models we have been describing up until now. No model before the NK-DSGE model was able to analyse the effects of these shocks in the way presented in these extremely important exercises in the Appendix.

It is surprising to us that there are not many good expositions available of the NK-DSGE model.\textsuperscript{54} We say that this seems strange not just because this model is the culmination of the profession’s best work in macroeconomics. It is also strange because, as we have noted, the NK-DSGE model shows how the macroeconomy should actually be thought of as operating, once all of the important features described up to now in this paper are allowed to be at work. It was in response to this failing that we set out the model in our article 3 years ago (Vines and Wills, 2018) and do so again here, in a way designed to be simple, clear, clean, and in a form that is readily intelligible to graduate students.

Even more surprisingly, there is not an exposition of the NK-DSGE model in either Michael Woodford’s book \textit{Interest and Prices: Foundations of a Theory of Monetary Policy} (Woodford, 2003), or Jordi Galí’s book \textit{Monetary Policy, Inflation, and the Business Cycle} (Galí, 2015), both of which purport to be graduate-level textbooks that explain New Keynesian macroeconomics. The lack of a clear exposition of the benchmark NK-DSGE model may be because those writing textbooks think that, although graduate students need to learn about MATLAB code and simulation methods when writing their theses, a model like the Smets–Wouters model is not central to an understanding of macroeconomic theory. As anyone who has been reading this paper will realize, we think that this view is just plain wrong.

Perhaps it is also partly because those playing with NK-DSGE models think that they are operating in ‘core model’ territory—rather than just playing with a toy in a playground—and so they think that they only need to explain what they are doing to the cognoscenti, rather than write in a manner accessible to graduate students. As readers of this paper will also realize, we think that this view shows a misunderstanding of how the macroeconomics profession has developed in the last hundred years.

\textsuperscript{54} To be more precise in our criticism, we would say that the presentation of the model in the Smets–Wouters (2007) paper is not entirely straightforward. In particular, the interconnection between the interest rate, Tobin’s Q, investment, and the capital stock is not at all clear. See equations (A5) to (A7) in our Appendix for a clear description of this interconnection. And, in Christiano \textit{et al.} (2005), the authors spend much too much time showing how the equations emerge from optimizing microfoundations. By contrast, in Appendix section (i), we simply wrote down the equations which emerge from the relevant first-order conditions and only describe their derivation intuitively. Instead we spent our time discussing the simulations results in detail.
(vi) What is wrong with the NK-DSGE toy model

Nevertheless, it is also now clear that the NK-DSGE model is no longer fit for purpose. It failed to capture the boom in asset prices leading up to the global financial crisis, and to suggest policies that would help the economy to recover afterwards. These failures are why a new macroeconomic paradigm is needed.

We now demonstrate these failures in as simple a way as possible. The reader will notice a change of gear in what follows. Entry to the macroeconomics playground in the twenty-first century requires more than a pen and paper. A computer is now an essential entry requirement. And playing the game involves looking in detail at MATLAB simulations and understanding what is going on.

A failure to understand how an asset price boom works
In the lead-up to the 2008 crisis, it was often stated that the increase in asset prices could not be a bubble. If it were a bubble, then—it was said—speculators would see that one day it would collapse. By backwards induction, they would adjust prices today, ruling out the prospect of such a bubble.

The closest we can get to describing this no-bubble argument in the NK-DSGE framework is by simulating an upward shock in consumer demand, caused by a permanent reduction in the rate of time preference. Let us consider how the model responds to such an upward shock in consumer demand, caused by a permanent 200 basis points reduction in the rate of time discounting (an increase in \( \beta \), bringing it closer to unity, caused by consumers valuing the future more). For simplicity, we use a version of the model in which the real interest rate is manipulated by a very capable central bank so as to ensure full employment of resources at all points in time, and so no inflation. Figure 1 illustrates the relevant simulation, based on the model equations and calibration described in the Appendix.

Following the shock, consumers will wish to save more because they come to value consumption in the future more highly. In addition, profit-maximizing firms will know that the real interest rate will settle at a lower level in the long run. This will cause a temporary boom in asset prices, which is depicted by movements in Tobin’s \( Q \): the ratio of an asset’s market value to its replacement value. Tobin’s \( Q \) rises because the marginal product of capital is higher than that new, lower, long-run real interest rate. As a result, firms will wish to invest more. Over time, higher investment leads to higher output, and higher consumption.

The real interest rate falls, but by less than it will fall in the long run. It will be set sufficiently above its long-run level to cause consumption to fall initially, and so make way for the increased investment which firms wish to carry out. Output can increase a little because the labour supply increases. In the medium-to-longer term, the capital stock rises by a large amount, about 10 per cent, and so output rises by nearly 6 per cent (given the Cobb–Douglas
production function and the fact that the exponent on capital is 0.35, and given the fact that the labour supply rises by about 2 per cent). The real interest rate falls by 200 basis points to the lower level consistent with the consumer’s lower rate of time preference, and Tobin’s Q goes back to one. Consumption is higher because output is higher. The real wage rises because capital has accumulated, and so workers are more productive. Labour supply is higher because of the rise in the real wage, despite consumption rising (and so workers wanting more leisure). Investment settles down at a slightly higher level consistent with the extra depreciation which happens as a result of the higher capital stock.

**Figure 1**: Simulation of a 200 basis point reduction in the time rate of discount

![Figure 1: Simulation of a 200 basis point reduction in the time rate of discount](image)

But this is clearly not what happened in the run-up to the global financial crisis. Indeed, this exercise seems extraordinarily misleading.

*A failure to understand why growth may not resume after a financial crisis*

The NK-DSGE model was also uninformative at suggesting why recovery was so slow after the GFC. In the period following the GFC it was claimed that the entire economy—capital, output, and consumption—would return to its previous growth rate. Even if interest rates could not fall below some lower bound, the depreciation of capital would eventually enable Tobin’s Q to return towards unity, and encourage investment to recover again. The reason for this is because the NK-DSGE model is linear, it returns to a Ramsey growth path along which the long-run growth rate is unchanged.

The closest we can get to describing this kind of argument in the NK-DSGE framework is by considering what might happen in the face of a permanent increase in the risk premium attached to investment—see Figure 2. We suppose that such a simulation is a way of capturing an increase in financial frictions (assumed for simplicity to be permanent). For simplicity we also again use a version of the model in which the real interest rate is manipulated by a capable
central bank to ensure full employment of resources at all points in time, and so no inflation. Figure 2 shows that, after the shock, Tobin’s Q collapses and investment falls. The central bank cuts the real interest rate falls to fully counteract the effects of the shock on aggregate demand.\textsuperscript{55} Consumption rises as a result. But gradually, as the capital stock falls, Tobin’s Q recovers and investment returns to equilibrium, at a slightly lower level because there is less capital and so less deprecation. Employment falls, because the real wage falls. But the economy returns to equilibrium, in the sense that its growth rate is unchanged.\textsuperscript{56} If the nominal interest rate is constrained by a lower bound, then this process will be dragged out. But eventually the economy will return to its unique equilibrium.

\textbf{Figure 2:} Simulation of a permanent positive shock to the risk premium in investing

![Graphs showing Y, I, K, R, Q, C, W, L over time]

But this is also clearly not what happened in the late 2000s, in the period after the global financial crisis. Indeed, this exercise also seems extraordinarily misleading.

We now turn to ask: what needs to be done, in response to the failure of the NK-DSGE toy? In the next section we turn to some new toys that might better help us understand an asset price boom and a slow recovery from crisis. Although they are simpler, they allow for multiple equilibria, and so may actually get us closer to the truth.

\textsuperscript{55} There have been useful attempts to study the effects of a zero bound, which prevents the necessary strong movements in the real interest rate and leads to a temporary unemployment of resources. But the effects of this in the NK-DSGE model can only be temporary since the Ramsey growth path is the unique long-run attractor of the model (see Guerrieri and Iacoviello, 2015).

\textsuperscript{56} In the version of the model being used here, the rate of growth of technology and the rate of growth of the population are both assumed to be zero for simplicity. But the basic point remains clear: A model with a positive rate of growth of the labour force and positive technical progress will settle down in the longer term at an equilibrium at which the capital stock is lower, because of the permanent increase in the risk premium, and, as a result, output per person is lower. But the economy will return to growing at a rate equal to the rate of labour-force growth plus the rate of technical progress.
III. Multiple equilibria

(i) Building better toys with multiple equilibria: small ME toys

We have seen that the NK-DSGE model failed to give an appropriate first-pass at understanding the lead-up to the 2008 crisis, and the slow recovery afterwards. Just as in the 1930s and the 1970s, new toys are needed. These should be able to articulate why multiple equilibria might arise, and to show what happens when they do.

We think that a return to the world of two-dimensional sketches will be helpful in pinning down just why multiple equilibria (ME) might arise. Like the toys reviewed in the first part of section II, these should only require a piece of paper and a sharp pencil, but unlike those earlier toys the curves need to be sufficiently non-linear to allow them to intersect not once, but two or three times. In this section we set out three examples. The first shows how an asset market boom can emerge when there is a feedback loop between output and Tobin’s Q (Krugman, 2003). The second describes how an economy may be slow to recover from a sufficiently large negative shock to aggregate demand, due to its effect on technical progress (Carlin and Soskice, 2018). The third describes how the presence of multiple equilibria might inhibit the process of disinflation (Akerlof, 2019). It will become clear how, in each case, something rather similar is at work.

Of course, to understand what really happens in a world with multiple equilibria is more complex, and requires toys larger than those which can be drawn with pencil and paper. We describe how these large toys might evolve at the end of this section: by combining our contemporary understanding of the world captured in the NK-DSGE paradigm with the insights from small ME toys. This playground is still an open field.

A small ME toy model of an asset price boom and slump

The first of our small-sketch toy models with multiple equilibria comes from a paper by Paul Krugman (Krugman, 2003). In this paper Krugman discusses how the level of GDP might, over a range, move up or down in an explosive manner in response to changes in Tobin’s Q, due to the operation of collateral constraints on consumers and investors. High $Q$ stimulates investment which drives up the level of GDP, $Y$, and that in turn drives up $Q$. Over a range this $Q \rightarrow Y \rightarrow Q$ process can be cumulative. This means that there may well be a stable equilibrium with high $Q$ and high $Y$, and another stable equilibrium with low $Q$ and low $Y$.

The Krugman set-up continues to assume rational expectations for the value of Tobin’s Q, but nevertheless obtains a multiple equilibrium outcome. This means that, although the model is simple, it allows for the possibility of a self-fulfilling stock market bubble. Because this a rational expectations model, this is not a bubble in the sense of inflated asset prices staying in place only as a result of expectations of rising asset prices. Due to the multiple equilibrium
structure of the model a rise in asset prices can, over a range, be entirely rational as the economy moves towards a new equilibrium.

Krugman begins by assuming a demand-side driven economy, which implicitly means nominal stickiness, with prices more rigid than that in equation (A11) in the NK-DSGE model in the Appendix. Tobin’s Q determines investment and hence output, Y, through a multiplier process.57

\[ Y = f(Q) \]  

(1)

How is Tobin’s Q determined? We know, from equations (A3) and (A5)–(A7) in the NK-DSGE model, that an increase in output will increase employment and thus profits and Q. We also know that an increase in the nominal interest rate \( R^n \) by the central bank will—given nominal rigidity—increase the real interest rate, \( R \), depress the present discounted value of profits, and so depress Q.

\[ Q = f(Y, R) \]  

(2)

We suppose that the central bank raises the real interest rate if \( Y \) is high, and reduces it if \( Y \) is low; this is a form of Taylor rule more general than that in equation (A10), since the latter does not include an output term for the sake of parsimony (as discussed in the Appendix).

\[ R = f(Y) \]  

(3)

These three equations mean that we can think about this model in \( Y, Q \) space. Equation (1) defines a goods-market equilibrium schedule, \( GG \), showing how \( Y \) depends on \( Q \). Equations (2) and (3) together define an asset-market equilibrium schedule, \( AA \), showing the response of \( Q \) to \( Y \).

Krugman suggests that the goods-market equilibrium schedule might be non-linear, like the GG curve in Figure 3. Following an old tradition of nonlinear business cycle theory (e.g. Tobin, 1955) he suggests that below some level of output, the amount of investment desired is near zero and lowering \( Q \) will not have any effect; and above some level of output an increase in \( Q \) cannot cause any further increase in investment because it would face capacity constraints, but that, within these bounds, the effect on \( Y \) of a change in \( Q \) might be very large. Our own view is that this schedule is likely to be non-linear due to the feedbacks caused by the collateral constraints on consumption described in Muellbauer’s paper in this issue (Muellbauer, 2020; see Figure 1 in that paper in particular). As Muellbauer puts it, there can be a substantial

57 In Krugman’s set-up there are liquidity-constrained consumers and asset-constrained consumers, so consumption expenditure depends not just on the real interest rate, as shown in the Euler equation (equation (A1)); this means that there is a Keynesian multiplier process and an effect of wealth effect on expenditure, both of which increase the effect of \( Q \) on \( Y \).
increase in consumption in response to a rise in asset values as the stock market and real estate rise in value. These effects will all tend to make the GG curve very flat over some range (that is, a small increase in $Q$ might cause a very large increase in $Y$).

Figure 3: A two-dimensional sketch of multiple equilibria in $Y-Q$ space

The asset-market equilibrium schedule AA can slope either up or down. The monetary authority changes interest rates in response to movements in the output gap, as shown in equation (3). If it does this sufficiently aggressively, then that will outweigh the private-sector investment response to output shown in equation (2) and the schedule will slope downwards. In this setting there is a unique equilibrium and nothing that looks like an asset price boom.

If, however, the monetary authority intervenes less strongly in response to increases in the output gap, then the AA curve can slope upwards, creating the possibility of multiple equilibria. As shown in Figure 3 there is a lower-level stable equilibrium, $Y_1$, and an upper-level stable equilibrium, $Y_2$, and a mid-level range between these two equilibria in which any process of adjustment is unstable. Starting from the lower-level equilibrium, $Y_1$, a positive shock to consumption demand could push $Y$ sufficiently far to the right that the economy experiences a financial-market boom and moves to the upper-level equilibrium rather than returning to the lower-level equilibrium. Conversely, starting from the upper-level equilibrium, $Y_2$, a negative shock to consumption demand could push $Y$ sufficiently far to the left that the economy experiences a financial-market collapse and moves to the lower-level equilibrium rather than returning to the upper-level equilibrium. Once at the high equilibrium – or at the low equilibrium - the monetary authority is limited to stabilization in response to small deviations in output around its stable equilibrium, as shown by Equation (3). This contextualizes the concentration on small deviations from a unique stable equilibrium in the canonical NK-DSGE model.
How does this view of the world differ from that described in the NK-DSGE model? Take the case of the financial market boom. In the Krugman sketch, a permanent increase in consumer demand might push the GG curve so far to the right that the only stable equilibrium is the upper one, and if the economy begins at the lower-level equilibrium, it will lurch upwards to the upper-level equilibrium. For comparison, consider an increase in consumer demand in the NK-DSGE model, for example one caused by a permanent reduction in the rate of time preference, as was shown in Figure 1. Tobin’s Q immediately jumps up, but gradually returns to its equilibrium value. Consumption actually falls initially, because the real interest rate does not fall to the new rate of time preference quickly enough to prevent this happening. Over time consumption only gradually rises to its long-run value—there is no collateral effect of higher Q inducing an immediate rise in consumption. The first response of the rigidly inflation-targeting central bank is actually to cut the real, and thus nominal, rate despite the investment boom being under way, because of the odd behaviour of consumption just described. The economy does indeed settle at a higher level of activity, because the long-run real interest rate has fallen and the capital stock has risen. But there has been no boom in Q during the time in which this happens—it initially jumps up and is then falling in all subsequent periods.

It is obvious that the NK-DSGE model’s omissions are important. Recognition that there are these omissions opens up a research agenda for those working on the kind of policy models we discuss in the next section: to try to understand whether, to use Meade’s words, ‘we shall assume that equilibrium is stable’; or, again using Meade’s words, we shall assume that ‘any small jerk to the system may start it off in one direction or the other in search of a completely new equilibrium’. We think that it is important to discuss whether the latter outcome is possible. Muellbauer (2020) suggests that it is.

A small ME toy model of a slow recovery from the Global Financial Crisis

The second of our sketches is due to Wendy Carlin and David Soskice (CS, Carlin and Soskice, 2018). In that paper the authors discuss how there may be feedback between an economy’s long-run growth rate and its rate of technical progress. High growth means high investment, which drives up the level of productivity and that in turn drives growth up further. Over a range this process can be cumulative. This means that there may well be a stable equilibrium with high rate of growth of output and high rate of technical progress, and another stable equilibrium in which both of these are low.

Our own interpretation of these ideas is represented in the two-dimensional sketch in Figure 4 (this picture is not to be found in the Carlin and Soskice paper). We plot the long-run growth rate of per capita income, g, on the horizontal axis and the growth rate of labour-saving productivity (or ‘technical progress’), m, on the vertical axis.²⁸

---

²⁸ CS discuss the possible endogeneity of total factor productivity (or TFP) rather than labour-saving productivity. We deploy labour-saving technical progress here, simply in order to make the exposition simpler.
In Figure 4 the curve showing \( m \) as a function of \( g \), labelled \( m(g) \), has an S shape and starts from the origin. It is possible that for low (but positive) growth rates, \( g \), technical progress, \( m \), may remain at zero as a minimum amount of growth of output and demand is needed to support innovation. The CS idea here is that technical progress is embodied in new investment; in the absence of the latter, little technical progress will be observed. But, as \( g \) rises further, the resulting higher level of investment might make technical progress possible. One might add that an increase in growth might undo some of the kinds of constraints on technical progress which Gordon (2015) and Summers (2014) have emphasized in their recent work on secular stagnation. Over some range this stimulus to productivity growth might be very powerful, especially since – as noted below - more investment by one firm might stimulate additional investment by other firms, suggesting a curve which becomes steeper than 45°. But eventually diminishing returns in using resources to create and develop new ideas will be encountered, as emphasized by Bloom et al. (2020), so the curve will flatten out.

**Figure 4:** A two-dimensional sketch of multiple equilibria in \( g-m \) space

In Figure 4 the second curve, showing \( g \) as a function of \( m \), and labelled \( g=m \), represents the long-run equilibrium growth path. Along this path the growth rate of *per capita* income is determined by, and equal to, the rate of labour-saving technical progress. This clearly has a 45° slope and a zero intercept. As shown, the two curves intersect three times; there is an upper-level stable equilibrium at \( g = m = g_1 \), a lower-level stable equilibrium at \( g = m = 0 \), and a mid-level range between these two equilibria around which any process of adjustment is unstable. Starting at the high-growth equilibrium at which \( g = m = m_1 \), consider a short-run negative shock to aggregate demand caused by, say, a rise in the risk premium in the economy. This
might make $g$ fall so much that the economy moves to the left in the diagram, beyond the unstable equilibrium, and converges on the $g = m = 0$ equilibrium.\footnote{Of course this diagram abstracts from many details, showing how aggregate demand is determined in the short run, features discussed by CS, many of which are made explicit in the benchmark NK-DSGE model. Nevertheless this diagram picks up a key multiple equilibrium feature of the CS model, in a way which the authors do not fully explain.} CS argue that the existence of a zero bound to the nominal interest rate makes it possible that such collapses in output and so in investment might be very large, making a move to the lower-level equilibrium all the more likely. But, once at the low-level equilibrium, there is no way back to the high-level equilibrium, since the low-level equilibrium is a stable one. Once stuck at the low equilibrium, the central bank (or fiscal policy maker if the economy is at the ZLB) is limited to stabilization in response to small deviations around the stable equilibrium as dictated by its inflation-targeting mandate. This again contextualizes the concentration on small deviations from a unique stable equilibrium in the canonical NK DSGE model.

In their paper, CS emphasize the role of beliefs in sustaining the high or low equilibrium as discussed further in Section III(ii) below. In their article, they present a multiple equilibrium model of firm behaviour with properties similar to the one in Figure 4. It focuses on the strategic complementarity between the investment decisions of firms: a shift in beliefs from optimistic to pessimistic about the growth of demand shifts the sigmoidal (S-shaped) function down and can result in the economy moving to the low equilibrium with each firm choosing low investment (and resulting in low associated productivity growth) as the tipping point at the unstable intermediate equilibrium is passed.

How does this view of investment and technical progress differ from the NK-DSGE model? In that framework there is a unique long-run equilibrium Ramsey growth path, in which capital accumulation keeps pace with population growth and technical progress, and the latter is exogenous. So, an endogenous feedback loop between investment and technical progress just cannot arise. Consider again the increase in the risk premium discussed above in Figure 2. As shown there, the long-run outcome will have a permanently lower capital stock, and so output, because the risk premium is permanently higher. But \textit{per capita} income will grow at the same rate, come what may, determined by the exogenous rate of technical progress.

Once again the NK-DSGE model’s omissions appear to be important. Recognition that they are omissions opens up another research agenda for those working on policy models: to try to understand the interconnections between the growth rate and technical progress. However interesting the discussions by Gordon and Summers have been about the slowing down of technical progress, these authors have not explored the idea that this slow-down might have itself been caused by a collapse in the growth process. And so they have not given any insight into the possibility of multiple equilibria which we have discussed here. The ideas which we have presented here thus present a research agenda for those working on policy models of the
kind which we discuss in the next section. Might the kind of multiple equilibrium presented here really be possible?

A small ME toy model of how and why disinflation may be difficult

George Akerlof seems to be on our side. In a recent issue of the Journal of Economic Perspectives (Akerlof, 2019), he contrasts non-linear, multi-equilibrium models of financial crises with the unique-equilibrium DSGE models that are currently the benchmark, and strongly criticizes the latter. Indeed, he argues that ‘It is . . . difficult, if not impossible, to produce an aesthetically pleasing model that combines the two types of equilibrium. Such a model would be the equivalent of chicken ice cream’ (ibid, p. 179). Our aim here is to show how to cook chicken ice-cream. Akerlof points out that doing this would allow us to study policy choices, for example the choice of whether to defend a currency peg or the choice of whether to embark on a process of disinflation in a high-inflation economy. We aim to show how these things might be analysed.

As a particular example, Akerlof (2019) discusses the process of disinflation in the US in the 1980s and the idea that, starting with a high rate of inflation, there may have been two equilibria. The first possibility, he says, would see high actual inflation continue along with high ‘cost of living adjustments’ (COLAs) in wage agreements, which would reinforce high inflation expectations and so perpetuate high inflation (creating what Akerlof calls ‘norms of high inflation’). The second possibility, which is what happened in the US, would see disinflation cause firms to exclude COLAs from their wage agreements, creating a ‘norm of low inflation’ (i.e. lowered inflation expectations), and an outcome of lower actual inflation.

Akerlof complains that the NK-DSGE model has nothing of interest to say about this set of ideas. And he is right; one cannot get a two-possible-inflation-rate outcome in the NK-DSGE model since it has a unique long-run equilibrium inflation rate equal to the inflation target embodied in the Taylor rule. Conventional studies of the 1970s–1980s inflation–disinflation experience attribute it to a temporary cost-push shock in the 1970s that raised the inflation rate, which was initially accommodated by policy. There then followed—it is said—an exogenous change in the policy-makers’ inflation target, leading to a gradual downward adjustment of inflation expectations, accompanied by an extended period of high real interest rates, low output, and a time-consuming process of slow disinflation.60 In that conventional analysis the new norm is imposed by the policy-maker, rather than being just one of two possible stable equilibria.61

---

60 The simulation of the effects of disinflation after a cost-push, in the Appendix to this paper, captures part, but only part, of this story. In the simulation reported on in the Appendix all expectations are rational and so there is no gradual downward adjustment of inflation expectations. Slow disinflation only emerges in the story in the Appendix because the cost-push is slow to die out.

But we think that there may be a game-theoretic account of monetary policy which leads to a multiple-equilibrium outcome of the kind Akerlof is seeking. Suppose that policy-makers actually choose whether to lower their inflation target, instead of following a Taylor rule in which the inflation target is given exogenously. The private sector could look forwards when setting wages and prices, which will make it easy for monetary policy to bring down inflation. The ease of doing this will then justify the private sector’s forward-looking beliefs that inflation will be low. The beliefs will be self-reinforcing. But the converse could also be true: the private sector could look backwards when setting wages and prices. That would make it costly to bring down inflation. The central bank would therefore only be able to bring down inflation slowly. This would justify the private sector’s backward-looking price-setting, because that sector would not believe that monetary policy would bring down inflation rapidly. Finally, it might be the case that, in such an economy, the central bank would find disinflation too costly and so would not carry it out. So, we think, two outcomes are possible. This is just like how it may—or may not—be too costly to defend a fixed exchange rate in currency crisis models, which Akerlof refers to in his paper. He presents a compelling discussion of the history of US monetary policy in exactly this way.

We now provide another simple sketch to show how this set of ideas might be modelled.

The set-up of the model is as follows. We again assume the economy is demand-driven, and that prices are sticky. The real interest rate, $R$, determines investment and hence the level of aggregate demand, represented by the output gap, $y$:

$$ y = f(R). \quad (4) $$

We assume that the short-run Phillips curve is upward-sloping, with inflation, $\pi$, depending on the output gap, $y$, and expected inflation,

$$ \pi_t = g(y) + \pi_{t+1}^e. \quad (5) $$

Let us suppose that the private sector’s inflation expectations are a weighted average of rational expectations (i.e. expectations which are forward-looking) and backward-looking expectations (see Fuhrer and Moore, 1995):

$$ \pi_{t+1}^e = (1 - \alpha) E_t[\pi_{t+1}] + \alpha \pi_{t-1}. \quad (6) $$

Finally, we suppose that the central bank raises the real interest rate, $R$, if $\pi$ is high relative to its target $\pi^T$, and reduces $R$ if $\pi$ is low relative to its target; i.e. we suppose that the central bank follows a simple Taylor-style rule:

$$ R = h(\pi - \pi^T). \quad (7) $$
This set-up has a conventional Taylor-rule-type outcome, which will be familiar to many readers. We suppose that the rate of inflation, and the expected rate of inflation, are both initially high at $\pi_1$. When the inflation target is reduced from $\pi_1$ to $\pi_2$, the central bank will raise the real interest rate—initially large but gradually declining—because expectations (Akerlof’s ‘norms’) are partly backward-looking. The negative output gap will cause inflation to initially jump down, and then gradually fall, to converging to $\pi = \pi_2$ and $y = 0$. The speed of adjustment will depend on the slope of the Philips curve, $f'$, and the extent to which expectations are formed by looking backwards, $\alpha$. Disinflation will be slower when expectations are more backward-looking ($\alpha$ is greater), because expectations adjust more slowly in this case. And, of course, the more backward-looking are expectations the larger the ‘sacrifice ratio’ – the total amount of output loss required to bring down inflation by any given amount.$^{62}$

But, we can introduce the possibility of multiple equilibria if we make just three adjustments to the model.

First, suppose that the strength of the central bank’s reaction to inflation is not fixed, as in equation (7), but instead let the central bank minimize an intertemporal loss function, whose components are a discounted weighted sum of $(\pi - \pi^T)^2$ and $y^2$. If the degree of backward-lookingness, $\alpha$, is fixed, then the solution to this optimization problem will actually take the form$^{63}$ of equation (7). Crucially, the optimal response of $R$ to $\pi$, shown by $h$, will decline as the private sector becomes more forward-looking (i.e. as $\alpha$ declines). This is because smaller values of $\alpha$ make interest rates more effective, so they need to be raised less to achieve any given speed of disinflation. Thus we write

$$ h = \varphi(\alpha) $$

where $\varphi' > 0$, showing that the more backward-looking wage-and-price-setters are, the more aggressive the optimal monetary response to inflation, $\varphi$, will need to be. We depict equation (8) as an upward-sloping line in Figure 5, with a (tiny) positive intercept on the vertical axis at $h_1$ (denoting the idea that when agents are fully forward-looking, the central bank will only need to respond to inflation by an infinitely small amount).

Second, suppose that the degree of backward-lookingness in the private sector depends on how aggressively the central bank moves the real interest rates in response to inflation,

$$ \alpha = \phi(h). $$

$^{62}$ Of course, if inflation expectations adjusted immediately ($\alpha \rightarrow 0$) there would be no need for a reduction in output, since the economy would move to $\pi_2$ immediately that the new target was announced.

$^{63}$ This is technically true if the optimization is carried out making the assumption of discretion such that the central bank re-optimizes every period. If the central bank could make credible commitments, and so was able to rely on private-sector beliefs, the solution would be more complex. But these details need not detain us here as we set up this simple sketch. See Woodford (2003).
This relationship might be highly non-linear, as depicted in Figure 5. For example, if agents observe that the central bank only needs to make very small adjustments to inflation to keep the economy under control ($h$ very small), then the private sector might well have confidence in the inflation target and so be fully forward-looking when setting its expectations about future inflation ($\alpha = 0$). Beyond some level of $h$ ($=h_2$), though, agents will observe more volatile movements in the real interest rate $R$, in response to movements in inflation, and so may begin to lose confidence that the central bank can, in fact, control inflation. In doing so they might become more backward looking, i.e. $\alpha$ might rise as $h$ rises. This might continue until agents lose all faith in the central bank’s ability to stabilize inflation, and so become fully backward-looking when setting their expectations, i.e. $\alpha$ might tend towards unity as $h$ becomes very large.

**Figure 5:** A two-dimensional sketch of multiple equilibria in $\alpha$–$h$ space

![Diagram showing multiple equilibria in $\alpha$–$h$ space](image)

The feedback between the central bank’s reaction to inflation, and the way the private sector sets its inflation expectations, is what enables there to be two stable equilibria in Figure 5. The first is a ‘good’ equilibrium at the point ($\alpha = 0$, $h = h_1$). Here the central bank only needs to move the real interest rate very slightly whenever inflation changes, which is nevertheless effective because the private sector is forward-looking and believes the central bank will achieve its target, making the job easy. The second is a ‘bad’ equilibrium, near $\alpha = 1$, where the central bank chooses to respond aggressively to changes in inflation, which is necessary because the private sector has little faith in the inflation target and so is very backward-looking, making inflation difficult to control.

Third, we can then think about whether an optimizing policy-maker in a high-inflation economy might decide whether or not to lower the inflation target, $\pi^T$, and disinflate. There are
two possible circumstances. In an economy starting at the good equilibrium, disinflation will be low-cost but generate large gains from having a low long-run rate of inflation. Disinflation may well seem attractive. But in an economy starting at the bad equilibrium it may not.

This is exactly the kind of circumstance which Akerlof has asked us to contemplate. The decision of whether to move to a low-inflation regime can then be thought of in just the same kind of way as the decision about whether to defend a fixed exchange rate, or not. 64

**Turning small toys into large toys**

However, we know that the world is not a two-dimensional sketch. The NK-DSGE model has taught us just how much stuff must be fitted together—in relation to consumption, investment, and the inflation process—to do justice to our contemporary intuitive understanding of the macroeconomic system. We don’t want to throw away all this knowledge just because we have produced two-dimensional sketches of multiple equilibria.

So, a crucial part of the research programme in the new MEADE paradigm will involve incorporating nonlinearities like those described earlier in this section into the benchmark NK-DSGE model. The aim will be to show how and why more than one equilibrium can arise when consumption, investment, and inflation are all interacting, in general equilibrium. To do this we will need to move beyond simple two-dimensional sketches.

The first step will be to develop a simple intuitive understanding of what must change in the NK-DSGE model to generate multiple equilibria, as in the simple sketches above. Small toys are useful for this. Some equations will just need a minor alteration, or different parameter values, while preserving the linear structure of the equation. But, crucially, in other equations we will need to deliberately introduce some nonlinearities.

The second step will then be for skilled simulation specialists to learn how to find rational-expectations solutions of the resulting non-linear, multiple-equilibrium versions of the NK-DSGE model. In due course, such specialists will need to show us how to routinize the calculation of such simulations and to present easy-to-understand pictures.

This sub-section of the paper is short. That is because the work to create such multiple equilibrium large toys is all yet to do. 65

---

64 John Muellbauer has pointed out to us that South Africa may well be a striking example of the circumstances in which disinflation does not seem attractive. Backward indexation of wages is standard and the South African Reserve Bank worries greatly about the ‘sacrifice ratio’. Inflation has been high and stable for decades, though it has come down a little with global disinflation.

65 Such a research strategy will be one which continues to accept the idea of model-consistent expectations as a feature of movements around whichever equilibrium – high or low – the economy happens to be moving. This requires that the private sector and economic policymakers share a true model of the economy, which is – of course – a ridiculous assumption. Nevertheless, such a research strategy can be defended as being one which is guided by Occam’s razor: “one thing at a time and the simplest one first”. The next stage will involve going beyond that, as outlined in the section which follows.
Misperceptions and multiple equilibria

The analysis in section III(i) has shown that one way to escape from the rigid determinism of the normal DSGE set-up is to introduce (realistic) non-linearities. This can give rise to multiple equilibria, creating the possibility of a good outcome and a bad outcome. In this issue, Guzman and Stiglitz want to go much further than this and consider how expectations are formed (Guzman and Stiglitz, 2020). They argue that the evolution of the economy is likely to be misperceived by agents due to the (obvious) absence of perfect futures markets. They then claim that such ‘macroeconomic inconsistencies’ are likely to cause abrupt and large adjustments when they are discovered, which can lead to an unstable adjustment process and prevent the economy returning to an ‘equilibrium’.

For example, capitalists might inevitably wake one day and realize they are investing too much, given the perception of reality by others in the economy, and so suddenly stop investing. The unstable adjustment process means that the resulting collapse might be self-perpetuating. It is possible to think of the end of the pre-GFC boom in this manner. Doing so would produce a different, but complementary, account of the 2008 crisis to that provided by Krugman and described earlier.

This is a strong argument, and we believe that the way is open for researchers to pin down the Guzman–Stiglitz mechanism in a two-dimensional sketch, like those in section III(i), before incorporating it into a full multiple-equilibrium-general-equilibrium framework. Doing this will be a challenge because the Guzman–Stiglitz paper somewhat resembles how the drafts of Keynes’s General Theory must have appeared to Harrod, Meade, and Hicks, as we describe in section II. It is a paper with a broad scope, but one in which many elements still need to be pinned down. Our sense is that a simple sketch will elucidate the behavioural-economics arguments underlying the sources of multiple equilibria to which the authors draw attention. And we sense that this will—at least in the first instance—be one without the precise restrictions of fully explicit microfoundations.66

66 Stiglitz wrote a related paper for our previous ‘Rebuilding Macroeconomic Theory’ issue (Stiglitz, 2018). Lawrence Cristiano wrote a critical response to this paper (Christiano et al., 2019), asserting that everything Stiglitz claimed was missing from the NK-DSGE tradition had already been done. We believe that these two differing views can be reconciled if we think of Stiglitz (2018) and Guzman and Stiglitz (2020) as pointing towards where multiple equilibria can be found, even though neither paper says this explicitly. So, while researchers have started incorporating many new frictions into the NK-DSGE model, as Christiano says is the case, and especially since the GFC in 2008, there has been very little work on how to incorporate multiple equilibria, which is what (we think) Guzman and Stiglitz are saying needs to be done. Perhaps such a reconciliation of views might lead to a better understanding between these two sets of authors.
Farmer takes a different approach in his paper in this issue (Farmer, 2020) by arguing that there may be a range of equilibria, i.e. that the outcome may be indeterminate. In such a set-up different beliefs can, on their own, lead to different outcomes. The overall point of Farmer’s paper is that beliefs can be important, and that optimistic beliefs about the future can be self-fulfilling. He aims to show that this can be true even in a rational-expectations growth model, not unlike the Ramsey growth model. He sets out to show that a change in beliefs by investors about the profitability of investment can actually induce capital accumulation, long-run growth, and higher profits, in a self-fulfilling manner.

What makes self-fulfilling investment beliefs possible is the way Farmer treats the labour market. He draws a distinction between three different ways of proceeding. First, there is ‘normal’ neoclassical economics, in which labour-market equilibrium is determined by the interaction of a demand schedule and a supply schedule. Then there is ‘classical search theory’, in which equilibrium is determined by the relative bargaining power of workers and firms. And then there is what he calls ‘Keynesian search theory’, in which labour market outcomes can be determined by the demand for labour and thus by entrepreneurial ‘animal spirits’. In such a set-up, Farmer argues, beliefs about the future value of the stock market can be self-fulfilling. Farmer aims to ensure that these self-fulfilling beliefs can be possible by making two main changes to a typical NK-DSGE model as described in our Appendix. The first is to introduce a new exogenous variable describing beliefs about the value of the stock market, which is defined as the discounted present value of the returns to capital, $Z_t$. Starting from an initial equilibrium at time $t = 1$, a permanent positive shock to these expectations at time $t = 2$ can, he says, move the economy to a new equilibrium from time $t = 3$, with more capital and higher employment, output, and consumption. Profits are also higher by the amount that beliefs expected them to be. See Figure 7 of Farmer (2020). Getting to that new equilibrium requires a surge of investment in time $t = 2$. When $t = 3$ arrives, everything in the economy will be that much larger: there will be a higher capital stock, and higher employment in line with this, and the marginal product of capital will have returned to a position in which it equals consumers’ time rate of discount, and the real wage will have returned to its initial position.

The surge of belief-driven investment in Farmer’s model would be impossible in a standard NK-DSGE set-up like the one in our Appendix. The reason for that lies in the labour market. In the standard model, equilibrium in the labour market is achieved when the disutility of labour (equation A2) equals the marginal product of labour (equation A4), which gives rise to a unique real wage. Starting at equilibrium, if the representative firm wants to invest more then it will need to pay its workers a higher wage, to induce them to work harder and produce the extra investment goods which are demanded. However, if workers produce more, then their marginal product will fall (because of the model’s assumption of diminishing marginal returns). So, the firm would need to be paying a wage that is higher than labour’s marginal product. This isn’t

---

67 Broadly, an indeterminate equilibrium is one where, for the same values of the exogenous variables, there may be not just one, or two, or three, but infinitely many equilibria.
in the firm’s interest, so it wouldn’t happen. Regardless of investors’ optimism, output is pinned down by the equilibrium in the labour market.

So, as we have already said, Farmer makes it possible for investment beliefs to be self-fulfilling by making a second change to the standard NK-DSGE model, namely a change to the way in which the labour market operates: he models the labour market using Keynesian Search Theory. For the purposes of his experiment this means that he assumes workers are price-takers: they will do the work assigned to them for the price that they are offered. Therefore, firms don’t have to raise the real wage to encourage workers to build the additional capital at time $t = 2$. Indeed, because of diminishing returns, all workers, including the extra ones, will get a lower real wage at time $t = 2$. In Farmer’s model, if the lords of industry decide that output will be higher, then by Jove output will be higher.

While Farmer’s model departs from some of the core assumptions of the current NK-DSGE paradigm, it is not implausible. One can imagine that workers could temporarily be induced to work harder for the same pay by an inspiring leader talking of sunlit uplands at time $t = 3$. Or, indeed, by a scoundrel who exploits the lack of bargaining power of workers who are disparate, replaceable, and need to keep a roof over their heads and a meal on the table. But it should be noted that this was what was assumed by economists who believed in the Phillips curve in the 1960s, before the expectations-augmented Phillips curve put paid to this set of ideas. Who knows? Maybe those earlier economists were right after all.

This is not to be dismissive. The central idea in this paper—that beliefs might be self-fulfilling—is an important one. However, it would also be interesting to know how the ideas in this paper would work if we introduced the typical convexity into the labour-market bargaining process. Would this idea survive if workers were not prepared to work harder for less pay at time $t = 2$, and work harder indefinitely at the same real wage as the one which ruled initially? It would be useful to consider a middle ground between the perfectly optimizing, leisure-loving household of the NK-DSGE model, and the perfectly pliant, price-taking worker of the Farmer model. It is an open research question as to whether optimistic beliefs of investors could be self-fulfilling in those circumstances.

It would also be valuable to consider alternative mechanisms which might make self-fulfilling investment expectations possible. One example might be a game-theoretic approach that takes account of the agglomeration externalities of investment, of the kind discussed by Tony Venables in his paper in this issue (Venables, 2020), which we discuss below. In such a set-up, increased investment by one firm lowers the costs of production of other firms. Therefore, if they all invest together, this would offer an alternative way for firms to escape the constraint of diminishing returns to labour. Firms might be able to reach a high-output equilibrium by colluding (or being centrally coordinated) to boost investment, at the same time as their beliefs improved, even if they would remain in their initial position if their beliefs stayed unchanged.

---

68 The surge in investment happens only in time $t = 2$.
69 See, for example, Blanchard and Galí (2009).
The policy implications of these ideas are crucial. When beliefs can determine outcomes, then market economies do not have a unique full-employment position. Policy is not therefore about ensuring as rapid as possible a return to a given equilibrium, as in the NK-DSGE model in our Appendix. Instead, beliefs and policy will work together to shape what the location of that equilibrium will be.

(iv) New economic geography and multiple equilibria

And now for something completely different, which, nevertheless, provides another way of thinking about the possibility of multiple equilibria in macroeconomic analysis.

In this issue of *OxREP*, Venables draws on the lessons of New Economic Geography to study the uneven distribution of economic activity and income across space (Venables, 2020). In this setting, some sectors benefit from agglomeration economies, and so tend to cluster together in certain cities. This leads to persistent inequalities between cities. And in this setting—and here is the link with the key multiple equilibrium ideas in this paper—a temporary negative shock can cause an economy to shift to a new permanent equilibrium, in which there are fewer high-productivity cities.

The model is set up as follows. There are two types of firms, and two types of cities, which specialize in producing either tradable or non-tradable goods. Production of tradable goods by a firm is subject to agglomeration effects which can lead to high productivity, if the firm is located in a city with enough other tradable firms. Non-tradable goods production is not subject to agglomeration effects, so firms producing non-tradable goods necessarily suffer from low productivity. It is the agglomeration effects which cause cities to specialize in only one type of good. In cities producing tradable goods, firms will be more productive and pay higher wages, making it too expensive for any firms to produce non-tradable goods in such a city. So, non-tradables will be produced in other cities, where wages will be lower. It will not be profitable for a firm to produce tradable goods in any of these other cities: there will not be any productivity-boosting agglomeration effects available there, since no other firms will be producing tradable goods there.

The equilibrium proportion of tradable cities in this economy, $x$, can be described using an elegant two-dimensional sketch (see Figure 1 in Venables, 2020). Labour in tradable cities is paid a higher wage, but incurs a higher cost of living. The more tradable cities there are the lower the supply, and the higher the relative price, of non-tradable goods. This in turn lifts the wage in non-tradable cities, attracting labour away from tradable cities, reducing agglomeration effects there, and so reducing the wage to be obtained from working there. Thus, as $x$ rises, the wage gap between the two types of cities falls. The upper limit on $x$ occurs when so few non-tradable goods are produced that their price rises enough to equate the wage in non-tradable cities with that in tradable cities. Conversely, the lower limit on $x$ is when the relative supply
of non-tradable goods is so large, and price so low, that the wage in non-tradable cities falls enough for a single firm to profitably produce tradable goods there, even without agglomeration benefits. Within this range, any outcome for $x$ can be an equilibrium. Consider an increase in import competition, or a loss of export opportunities (like a no-deal Brexit). This will irreversibly destroy the ability of some cities to go on producing tradable goods. Those cities will switch to producing non-tradables, in which labour-productivity is much lower. The important point is that, if the import competition goes away—or opportunity to compete in export markets is fully regained—then it will not be possible for these cities to resume tradable-good production. No individual firm will find it profitable to switch to tradable production in such a city, because there are no other firms already producing tradables there, and so no agglomeration opportunities will be available. The economy will have therefore switched to a new, stable, lower-productivity equilibrium.

This is a very serious case of hysteresis, resulting from the non-linearity associated with the agglomeration economies which are supposed to be at work in tradable good production.

The next steps in this research will be to incorporate these ideas into a full general-equilibrium model of the NK-DSGE kind. The model would need to have two sectors, each with a capital stock and an investment function. There would need to be international trade and international movements of capital, as in the model in the paper by McKibbin et al. (2020) which is discussed in the next section. The outcome will be a general equilibrium growth model with sectoral and geographic characteristics, and multiple stable outcomes. We encourage researchers to pursue this goal.

(v) Other ways of generating multiple equilibria

Of course, very many additions and alterations have been made to the NK-DSGE model. As Blanchard (2018) says, the model is—quite properly—used as a platform into which new elements can be added.\(^{v}\) In our paper for the first OxREP issue on ‘Rebuilding Macroeconomic Theory’ (Vines and Wills, 2018), we identified four main ways in which new elements need to be added: by incorporating financial frictions, by relaxing rational expectations, by allowing for heterogeneous agents, and by improving the behavioural micro-foundations underpinning the model.

We now realize that each of these additions might provide an insight into how multiple equilibria can arise. This possibility of multiple equilibria is not yet widely discussed, other than by the authors whose papers we discuss in this paper. For example, Blanchard (2018) does not mention this idea, even though his paper is called ‘On the Future of Macroeconomic

\(^{v}\) ‘The specific role of DSGEs in the panoply of general equilibrium models is to provide a basic macroeconomic Meccano set. By this I mean a formal, analytical platform for discussion and integration of new elements.’ (Blanchard, 2018, p. 49).
Nor do Christiano et al. (2019), when commenting on the work being done by Joseph Stiglitz. In what follows we discuss how each of the potential additions to the NK-DSGE model that we identified in the previous issue of OxREP might actually point towards kinds of multiple equilibria that we have been discussing in the present paper.

We finish this section by discussing, very briefly, one other source of multiple equilibria: climate change and climate change policies.

**Financial frictions**

Financial frictions are obviously important: the GFC in 2008 was a crisis of financial intermediation. But the financial sector plays no causal role in the benchmark NK-DSGE toy model: it simply intermediates transactions in the real economy. There is no consideration of the frictions and non-linearities that exist in the real world because of incomplete markets, liquidity constraints, and feedback loops. The events of 2008 and the effects of extraordinary monetary easing during the 2020 Covid crisis all show this treatment to be inadequate.

The sketch model due to Krugman presented above is an initial response to this challenge. Jeffrey Sachs, in his paper in this issue of *OxREP* (Sachs, 2020), takes the ideas further by describing the 2008 crisis as having been brought on by a credit cycle in the housing market. Low interest rates encouraged over-investment, relative to what would be sustainable at the long-term neutral rate, leading to a crisis, and then to underinvestment in the future. The possibility that there will be such a boom–bust credit cycle is, he says, brought about by feedback between asset prices, collateral, and consumption. This is similar to the claims made by Muellbauer (2020) in his paper, discussed in section IV. Sachs offers Austrian Business Cycle Theory as a way of thinking about this boom–bust process. He argues that it is important to be able to defend an economy susceptible to boom and bust against a range of shocks: not just against aggregate demand shocks, but also against financial shocks and supply-side shocks. Financial policy needs, he says, to be designed with this possibility in mind.

The next step in properly incorporating financial frictions will be to include feedback loops into the benchmark NK-DGSE model, in the manner outlined at the end of section III(i). An indication of why this is so important is provided by the BGG model (Bernanke, Gertler, and Gilchrist, 1999)—in many ways the classic exercise in adding financial frictions to the NK-DSGE set-up. The BGG model has the usual unique-equilibrium feature of NK-DSGE models. Because of this, the financial frictions studied in that paper turn out to have very little effect on

---

71 This is not surprising, since he was explicit in limiting himself to thinking about fluctuations, rather than thinking about the potential for crises, which is our concern here.

72 See our earlier footnote describing the criticism by Christiano et al. (2019) of work by Stiglitz.

73 The widespread reliance in the models on simple probability distributions—namely the normal and log-normal distributions—assumes away the fat tails and tail correlations that appear to have first-order effects in many real-world phenomena. One example of this is the fact that geometric Brownian motion, and log-normal distributions, are, to this day, used in so much of the academic literature on finance, despite the lack of empirical fit of these distributions to historical data and their inability to capture the ‘volatility smile’ that is observed in reality.
the simulation outcomes of the model. Such an approach to the study of financial crises is clearly not good enough.

This exercise—adding a credit cycle to the NK-DSGE model in such a way as to lead to multiple equilibria—will need to be done in a parsimonious way, so that the relevant mechanisms can be understood. But this work will also need to be sufficiently detailed to incorporate convincing microfoundations, and to allow for the necessary feedbacks between asset prices, collateral, credit conditions, and consumption. There is an obvious tension here between parsimony and comprehensiveness.

And this theoretical research will need to be supported empirically. We discuss Muellbauer’s suggestions on how to embed such work in a structural economic model in section IV of this paper.

**Relaxing the rational expectations assumption**

How can we escape from the rigid constraints imposed by the treatment of expectations in the NK-DSGE model? In that model there is a unique equilibrium, and, as a result, a unique set of expectations by consumers and firms that are consistent with that equilibrium.

The three sketch models described in section III(i) create the possibility that there might be more than one rational-expectations equilibrium, and these equilibria might be ‘self-fulfilling prophecies’. George Soros (2008) terms this concept ‘reflexivity’. Furthermore, as we have seen in section III(ii), different agents may also have inconsistent expectations of the future (Guzman and Stiglitz, 2020). If that is so, then obviously only some expectations can be self-fulfilling.

More generally, the self-fulfilling indeterminacy examined by Farmer, discussed in section III(iii), might also be important (Farmer, 2020). Simply imagining a better outcome may actually be enough to bring it about (Tuckett, 2012). That possibility can give a role for beliefs, or ‘animal spirits’, that better captures the kind of leap of faith that ultimately underpins any real entrepreneur’s decision to invest, in the face of high opacity and uncertainty. This idea has underpinned Shiller’s (1978) work on bubbles. Much earlier, Keynes’ *General Theory* (1936) clearly foresaw the importance of multiple equilibria when describing markets as a beauty contest with the winner determined by collective opinion; indeed, such ideas seem to underpin Keynes’s conception that the economy as a whole might have many possible equilibrium unemployment rates, with better animal spirits stimulating investment and reducing unemployment. As described in section II of this paper, Keynes’s observation was obscured by the rational-expectations revolution of the 1970s. It was eventually obliterated by the unique-equilibrium NK-DSGE model of the early 2000s. These instincts are worth keeping alive.

The implications for policy are—of course—crucial. As we have already noted in our discussion of the Farmer paper, if self-fulfilling prophesies exist, then the role of policy may
not be to simply hasten the inevitable return of the economy to a given equilibrium after a shock. Policy might actually determine the equilibrium that the economy returns to. Some policy-makers already understand this. An example is Mario Draghi’s ‘whatever it takes’ speech in 2012 which helped return Eurozone debt markets to a ‘good’ equilibrium, or at least to avoid a ‘very bad’ equilibrium.

**Heterogeneous agents**

Incorporating a variety of agents into economic models allows us to study distributional effects that are important in their own right. But doing so can also have macroeconomic implications and give rise to something completely different from traditional concepts of equilibria. Muellbauer (2020) argues that consumption should be modelled heterogeneously. He points to microeconomic evidence that the marginal propensity to consume varies with the household balance sheet: being highest for the asset poor, and lowest for the asset rich. As Muellbauer shows, balance sheet effects like the redistribution of wealth can play an important role in the response to shocks, and are capable of making these responses destabilizing.

The representative agent set-up of the NK-DSGE model means that there is a tight relationship between individual behaviour and aggregate behaviour. This ignores the important effects on aggregate behaviour that can come from large individual agents and firms. For example, power laws are observed in the distribution of many aspects of human endeavour, from the distributions of stock returns, wealth, and income to the sizes of firms and cities (Farmer and Geanakoplos, 2008; Gabaix, 2011, 2016). As a result, the decisions of wealthy individuals and large firms can have macroeconomically relevant effects; the central-limit theorem doesn’t hold. Many economic variables are not normally distributed, and this matters for aggregate behaviour, including for the way in which destabilizing responses to shocks can develop.

Finally, a new generation of macroeconomic models under the umbrella of ‘complexity economics’ do not just lead to alternative equilibria, but can instead create something fundamentally different. Complexity economics combines the behaviour of many heterogeneous agents using computer simulations. Creating a realistic representation of a national or global economy in this way is difficult, given the sheer number of agents and scale of their interactions. However, doing so can give rise to emergent behaviour that is qualitatively different from the behaviour of each individual, including limit cycles and chaotic attractors. (See Beinhocker, 2007, especially chs 5, 8, and 9; Farmer and Geanakoplos, 2009; Baptista et al., 2017; and Haldane and Turrell, 2018.) Understanding how this happens may become extraordinarily important as we seek to understand, and manage, the changes in the structure of the economy that will emerge after the Covid crisis.
Better micro-foundations

Finally, it is easy to say that the NK-DSGE models should have better microfoundations, but it is difficult to know what this actually means.

One research programme will go on seeking a deeper understanding of how consumers and firms behave in the model. Three of the key equations in the NK-DSGE model—the Euler equation for consumption, the Tobin’s Q investment function, and the Calvo–Phillips curve—are all the outcomes of optimizing behaviour in the face of what appear to be relevant constraints. But as Olivier Blanchard reminded us all 3 years ago, ‘what feels like micro-foundations to one economist feels like total ad-hocery to another’ (Blanchard, 2018, p. 52). Thus, in reality, the intertemporal budget constraint is clearly not the only constraint that binds consumers, as supposed by the Euler equation. Nor are adjustment costs of capital the only thing that prevents firms rapidly changing investment in response to shocks. And the absence of a Calvo fairy is clearly not all that prevents complete price flexibility. But the appropriate analytical response to each of these three criticisms is not yet clear.74 Nor do we understand the effect such responses will have on the multiple-equilibrium question which we have been examining in this paper.

There is also a much deeper research programme under way, not just about the microfoundations of DSGE models, but also about the microfoundations of economic analysis as a whole. In this paper we have not yet mentioned motives other than self-interest, such as the norms of fairness and reciprocity emphasized by Adam Smith in his Theory of Moral Sentiments (Smith, 1790); nor have we discussed how strategic interaction by consumers stemming from these motives might influence their behaviour. And we have not considered how firms’ internal organizational processes affect their behaviour, nor how the strategic interactions between firms affect the economy as a whole. Nor have we examined the power of firms over workers, particularly in a world of incomplete contracts, and the resulting inequality in the economy. Nevertheless, the teaching of elementary economics is in the process of being revolutionized75 by focusing on these features, in a way which has not been done up until now. This is well described by Bowles and Carlin (2020a); see especially Table 9 in their paper. It is apparent that, in reality, these features greatly influence the nature of the economy as a whole (Bowles and Carlin, 2020b). It is also clear that they will influence macroeconomic behaviour, the likelihood of multiple equilibria, and the stability of adjustment processes in response to shocks. But we have, as yet, rather little understanding of how to model these influences.

74 This analytical work will be informed by the remarkable progress which empirical micro research has made since the GFC on many questions of macro-relevance. John Muellbauer discusses a number of these in his piece in this OxREP issue (Muellbauer, 2020), for example in relation to the marginal propensity to consume, and in connection with the role of collateral constraints in linking house price changes to aggregate demand.

75 The CORE Project—where CORE stands for Curriculum Open Access Resources in Economics—was launched at Her Majesty’s Treasury in London in November 2013. Since 2015 there has been an introductory economics textbook freely available on the web which is now being used in many universities throughout the world.
Incorporating climate-change features

There is now a large literature on tipping points and non-linear feedback loops in the global climate system. Many climate scientists fear there is a high probability that the world is hurtling towards a catastrophic tipping point in the global climate. Amplifying feedback loops threaten mass species extinction, major sea-level rises and other disasters.\(^7\)

It has been common until now to treat these climate-induced shocks, and the policy responses which they induce, as orthogonal to the macro-modelling issues discussed in this paper. The UK’s Stern report considered the macroeconomic cost – in terms of investment as a proportion of GDP - of the investments in renewable energy which are required to address the climate challenge. (See Stern, 2013) And the paper by Warwick McKibbin and colleagues included in this issue of OxREP (McKibbin, et.al. 2020) is a valuable piece of work designed to study the implications of the necessary climate-change policy, including the imposition of carbon-taxes, for the conduct of monetary policy, and – conversely - the implications of monetary policy for the design of such taxes. But up until now climate policy has been thought of as a microeconomic subject: the Stern Report was a microeconomic exercise in resource allocation – how much to spend on renewable energy and what kind of taxes and quantitative controls are necessary to bring about the necessary reduction in emissions. It is true that some of the discussions about the Stern Report were macroeconomic in the sense of using an intertemporal Ramsey-type model to discuss welfare maximisation over time. But none of the discussions of the Stern Report were about the interconnections, and feedbacks, between climate policy and the process of macroeconomic growth and the necessary policies for macroeconomic stabilisation.

Nevertheless the paper by Cameron Hepburn and his colleagues on policies for a green recovery from the Covid crisis (Hepburn et.al. 2020) shows that this is beginning to change. Lurking within that paper is the optimistic idea that a green recovery from the Covid crisis might be self-generating, because the future rate of return on renewables is very high, with the result that the private sector might lead the way. It is true that such a process would be assisted by microeconomic policies (i) if there are the right relative prices – carbon tax etc – and regulations – requirements for electric cars etc., and/or (ii) if there is the right infrastructure – e.g. power supply infrastructure etc. Nevertheless, the optimistic story goes, the huge opportunities provided by enormous technical change in the provision of renewables might generate a huge self-fulfilling global boom, independent of whether there is good

\(^7\) Here are five examples of these concerns.
(i) Melting of polar ice cap will reduce reflection of sun’s rays, increasing global warming.
(ii) Melting of permafrost in the Arctic tundra will release methane, trapped under the ice, in the short run about 30 times more potent greenhouse gas than CO\(_2\).
(iii) Large amounts of carbon are buried in the tundra soil. Warming will release much of this carbon, over a longer time scale.
(iv) When rain forests are stressed by drought, a major risk, the carbon cycle reverses: they release CO\(_2\) instead of absorbing it. The dramatic rise in forest burning, triggered by human activity and by the greater frequency of electric storms, is an aspect of this
(v) The oceans are major absorbers of CO\(_2\). Their warming reduces this stabilising capacity.
policy or not, a boom not unlike that experienced during the golden age after the Second World War.

But lurking in that same paper is the rather different, and more pessimistic possibility that, although fixing the climate is essential, the benefits from doing this cannot be easily captured privately and so the private rate of return on renewables is not really very high, and so the possibility of a green recovery from Covid, led by the private sector, is not really there. In particular, this more pessimistic view suggests that a private-sector led recovery will not happen unless there big price signals provided by policy, through e.g. a carbon tax and unless there are very strong regulations; and also that (ii) a private-sector led recovery will not unless there is the right infrastructure in place. This more pessimistic view runs something like this. Although there are opportunities provided by technical climate in the provision of renewables, these will not, on their own be will enough to generate anything like the postwar golden age. And they will certainly not led to a spontaneous post-covid recovery. Large fiscal expenditures will be necessary.

This same paper contains one other idea. This is that, even if good microeconomic policies are put in place, none of the move to renewables will come about unless there is confidence that there really will be a global recovery from the huge global macroeconomic shock caused by Covid.

There are many analytical and empirical issues here. We do not know the answers to these questions. But they do suggest that, for any given set of policies, there might be more than one outcome possible. That is, they show how important our multiple equilibrium ideas will be in understanding both how to think about climate change policy and how to think about post-Covid recovery policy.

IV. Building policy models

We now turn to what Olivier Blanchard (2018) calls policy models and what Wren-Lewis (2020, this issue), calls structural economic models (SEMs). The name given by Wren-Lewis describes what these models do: ‘careful scholarly work that is dedicated to obtaining a detailed empirical understanding of the economy as it actually is’. By contrast, Blanchard’s name describes what the models are used for. The fact that there are two different names for the same sort of model is central to the arguments of our paper.

Wren-Lewis provides a compelling account of why SEMs, rather than DSGE models, are the models that are used for policy-making. The papers in this issue by Muellbauer and Fair provide valuable evidence in support of his argument (Muellbauer, 2020; Fair, 2020). All three

---

77 This was the description which we used in section I.
authors show how important it is to have this ‘detailed empirical understanding of the economy as it actually is’ if one is to make sensible policy.

It is not just, says Wren Lewis, that the benchmark NK-DSGE model contains too many short-cuts, as we have emphasized up to now. It is also that the theoretical microfoundations underpinning the NK-DSGE model actually inhibit the study of detailed features of the economy which policy-makers need. This is a strong claim, one which, as we shall see, is actively supported by all of Wren-Lewis, Fair, and Muellbauer. All three authors show, constructively, how the empirical analysis involved in building SEMs makes it possible to overcome this constraint. This is important.

More than this, we would go one step further. We argue that the work involved in building SEMs will enable us to understand how the economy is vulnerable to multiple equilibria (ME), and so to see why bad outcomes might arise, and then to discover what kinds of policy regimes might be needed to prevent this happening. In particular, the detailed, careful empirical study of the macroeconomy that is undertaken when SEMs are built may well point researchers in the direction of new small ME toys. Furthermore, it will suggest how the ideas embodied in these small ME toys might be incorporated into multiple-equilibrium version(s) of the NK-DSGE model. And, to turn a full circle, these last two activities will point the way towards what next to do when building SEMs.

But first we discuss hybrid models. As Wren-Lewis says, ‘some SEMs . . . are built by starting with a DSGE model, and develop by allowing the data to suggest where the DSGE needs some elaboration’. The result is ‘hybrid’ models—models which have the features of both a micro-founded NK-DSGE toy model and of a structural economic model.

(i) Hybrid models—the G-Cubed example

A leading example of a hybrid model is Warwick McKibbin’s G-Cubed model, which is described in detail in the paper in this issue by McKibbin et al. (2020). The model is very deliberately built within a new-Keynesian DSGE framework but, in the interests of empirical realism, that framework is expanded within the model to include many frictions, many countries, and many sectors. Doing this has allowed the G-Cubed model to shed light on a number of important global policy issues. The paper uses a particular version of the G-Cubed model to analyse why climate-change issues might require a change in conventional monetary policy frameworks. In the present paper we do not discuss those policy issues. We instead describe what the G-Cubed hybrid model looks like, show how its structure actually assists in

78 See for example McKibbin and Fernando (2020) for a study of the effects of the Covid crisis and McKibbin and Vines (2020) for analysis of a proposal for enhanced fiscal support of many emerging-market economies in the face of this crisis.
the study of climate-policy issues, and examine what the model says about the multiple-equilibrium issue on which we have been focusing in this paper.

As already noted, this model is designed within the constraint of the framework of the benchmark NK-DSGE model. This means that, at the core of the model, there is a Ramsey model of profit-maximizing firms engaged in a dynamic process of capital accumulation and growth, and alongside these firms there are consumers who allocate wellbeing optimally across time, by following an Euler equation. But in the short term there are frictions imposed on this growth process. These include not only the two frictions present in the NK-DSGE model (namely capital adjustment costs and nominal rigidities), but also three other sets of frictions, all of which appear empirically necessary.

First, some consumers are liquidity constrained. In the NK-DSGE model, consumers respond to both the contemporaneous real interest rate and to forward-looking expectations of future consumption, which are rational in the context of the model, by following an Euler equation. In the G-Cubed model only 30 per cent of consumers do this. The remaining 70 per cent are assumed to have limited access to financial markets and so are supposed to follow a simple rule of thumb which leads them to consume their entire income in each period.

Second, some firms take a backward-looking approach to investment. In the NK-DSGE model, investment is driven by profit-maximizing firms that decide how much to invest in the face of quadratic adjustment costs to capital. As a result, investment in each sector of the economy responds gradually to the value of Tobin’s Q, leading to a gradual adjustment of the capital stock over time. That is true here, too. But here there is a separate one of these investment functions for each of the 20 different sectors identified in the model, because relative investment opportunities in each sector will depend not just on the overall level of aggregate demand but also on the relative price of that sector’s output. And an important additional friction is imposed: only 30 per cent of firms respond to a forward-looking Q which evolves in a model-consistent manner; the remaining 70 per cent of firms respond to a backward-looking Q.

Third, some wages are set in a backward-looking manner. In the G-Cubed model, prices in each sector are set by profit-maximizing firms in the face of downward-sloping demand curves, so that the real wage in each sector depends both on the marginal product of labour in that sector and on the conditions of aggregate demand. Nominal wages are, in turn, driven by a Calvo–Rotemberg-style Philips curve. But there is an additional labour-market friction: some workers are assumed to be backward-looking. This means that nominal wages are even

---

79 The downward slope of the demand curve for each firm’s goods comes not only for the Dixit–Stiglitz reasons present in the NK-DSGE model but also from the substitution in consumption and production carried out by consumers and firms in the face of changes in the relative prices of the 20 different goods and services which, as described below, are separately identified in the model.

80 The way in which this feature is implemented is allowed to differ from country to country because labour contracts differ across countries.
more sticky than in the Calvo–Rotemberg set-up. Any excess supply of labour enters the pool of unemployed workers. In the long run the nominal wage will adjust to remove unemployment from the labour market. But, in the short run, unemployment can arise both because of supply-side shocks and because of changes to aggregate demand in the economy as a whole.

Taken together, these extra frictions mean that, in the short term, the behaviour of the G-Cubed model departs from the real business-cycle (RBC) model even more than does the benchmark NK-DSGE model. But, just like the NK-DSGE model, the G-Cubed model reverts to behaving like an RBC model in the longer term. This is because each of the extra frictions has been deliberately designed not to influence the long run outcomes of the model, and so to disappear over time, just like the capital adjustment costs and nominal rigidities do in the benchmark NK-DSGE model.

This ‘hybrid’ model displays two further crucial differences from the benchmark NK-DSGE model.

First, since it is a global model it must be disaggregated into countries and regions. In each country or region there are markets for goods and services, equities, bonds, household capital, and foreign exchange. These regional markets all interact through international trade and capital flows. Trade flows are determined by import and export functions for each country, which map on to the exports and imports of other countries, all of which respond to real exchange rates. The model supposes perfect international mobility of capital between countries, and the exchange rate is determined, à la Dornbusch, by the uncovered interest parity (UIP) condition.

Second, as already noted, the model disaggregates the economy into 20 different sectors. In each country (or region), there are multiple firms, as well as household and government sectors, all of which interact in markets for goods, services, and primary inputs. Production is carried out by an explicit set of heterogeneous firms, one representative firm in each sector, but therefore many representative firms in the economy as a whole. The prices of energy (of various types) and mining are set differently to those of manufactured goods and services, allowing for the flexible prices of those goods. And the model also captures inter-industry linkages, since some sectors produce inputs into other sectors, as happens in computable general equilibrium (CGE) models.

These final two features enable the model to yield important insights into climate-change policies. Disaggregating output into many sectors makes it possible to show that climate policies can affect the relative prices of goods and services in important ways. And the model is sufficiently detailed to show how a carbon tax in one country might lower investment there, depreciate its real exchange rate, lead to a current account surplus, and thereby cause capital to be exported to other countries where investment has become relatively more attractive. The

---

81 In McKibbin et al. (2020) there are 10 countries (or regions), which are listed in the paper.
82 The 20 different sectors are listed in the paper by McKibbin et al. (2020).
model might also be used to examine the first-mover advantage that a country might gain by being an early-adopter of climate-change policies, in such a way as to actually stimulate investment there.

It should be apparent that this model has moved a long way from its DSGE origins towards becoming a structural economic model. It acknowledges that most consumers are liquidity constrained and that some wage-setters and investors are backward-looking, and it covers multiple countries and multiple sectors. It is also big. There are hundreds of equations. The model is solved under rational expectations, and includes 229 forward-looking jump variables. The model is non-linear and so must be solved using a version of the Fair–Taylor target-shooting algorithm. With so many jump variables this is a non-trivial task.

Nevertheless, the model always returns to a unique, long-run equilibrium Ramsey growth path; it doesn’t do multiple equilibria. The model, therefore, cannot countenance the possibility that a temporary shock could cause a permanently good or bad outcome. Given any set of exogenous conditions, only one equilibrium outcome is possible. This means that, if someone wants to use the G-Cubed model to study the possibility of very bad outcomes, he or she can only do this by inputting exogenous shocks that move the unique equilibrium. For example, in McKibbin and Fernando’s (2020) study of the Covid crisis, it was shocks to equity risk premia that were needed to move the equilibrium to the very bad position which was displayed. The interesting thing about this study is that these plausibly sized shocks moved the equilibrium by a very large amount, and so produced very bad outcomes, even though the exogenous shocks to risk premia were ones which might—at first sight—have seemed rather small. That all makes sense because, in the model, shocks to equity risk premia have a very major effect on the capital accumulation process within a country, and so on capital flows between countries, as capital moves to where it is better rewarded. They thus have very large effects on global outcomes. The G-Cubed model is one of the best policy models we have available. But, we are tempted to ask, should not these risk premia that are applied to the model become things which are endogenously explained? Doing this might enable the model to display multiple equilibria. In particular, it might allow permanently bad outcomes to emerge endogenously from temporary shocks to risk premia. This is the kind of thing that is hinted at in the paper by Guzman and Stiglitz (2020).

83 Nevertheless there are still very evident shortcuts. They mean, for example, that the 70% of liquidity constrained households cannot own a house except by inheritance, or have a mortgage, or hold any buffer stock savings. So it is only the 30% who drive the link between financial asset stocks and real flows of expenditure in the economy as far as households are concerned. As a result, time-varying gaps will build up between the economic outcomes for the 30% and the 70%. It is hard to see how this time-varying heterogeneity is handled.

84 There are 10 Euler equations (one in each of 10 regions), 10 forward-looking wage rates, one for each region, 200 Tobin’s Q variables (in each of 20 sectors in each of 10 regions), and 9 exchange rates relative to the US dollar.

85 Do the maths. A permanent increase in the risk premium attached to capital investment which permanently increases the cost of capital by 20 per cent—say an increase in the required rate of return on capital from 10 per cent to 12 per cent—will reduce the long-run capital stock by 20 per cent (since the production function is Cobb–Douglas). That will have an extremely large effect on economic outcomes.
(ii) Going beyond a hybrid model—creating a full structural economic model

Not all SEM models are hybrids—most are deliberately not located within a NK-DSGE framework. It is a pleasure to include in this issue of OxREP a paper by Ray Fair (2020) describing a model of this kind which he has built and maintained for many years. What is particularly attractive about the SEM which Fair describes is that it is independent of any policy institution and so is an exemplar of academic engagement with this class of model. Fair’s model fits directly into the framework laid out by Wren-Lewis; in section II of his paper he argues that the principle aim of his SEM is to describe the dynamic behaviour of the economy, and that fidelity to the data is central to that purpose (‘the methodology is empirically data driven and the data rule’).

Section III of Fair’s paper offers a very specific list of features that he thinks an SEM should include, things that are not found in the benchmark NK-DSGE model. These include: adding financial and housing wealth to the consumption function, modelling prices and wages carefully, allowing for different fiscal multipliers for taxes and government purchases, and recognizing the short-run endogeneity of labour and capital productivity as firms retain workers and capital in the face of negative demand shocks.

The Appendix to Fair’s paper then goes on to outline what he thinks a full SEM should contain. It is instructive to compare how a full SEM differs from that of the current benchmark NK-DSGE model that we have included in our Appendix. Accommodating the features described by Fair into a microfounded NK-DSGE model would clearly be a big task, something which nobody has yet done. If journals require microfoundations, then these important features must—for the time being—be thrown away. You can have one or the other, but not both. This clearly supports the Wren-Lewis argument on why SEMs must be used for policy purposes, rather than a new-Keynesian DSGE model.

Section IV of Fair’s paper further outlines how and why a good SEM might be useful: by providing a revealing description, and explanation, of historical events. One is tempted to infer—although Fair does not say this explicitly—that he believes that this historical tracking can only be done with a model that includes all of the features he has identified in section III. Or to put the point another way, one is tempted to think that Fair does not believe that such a tracking exercise would be possible with the benchmark NK-DSGE model, or indeed with a fully micro-founded extension of that model.

Nevertheless, Fair’s paper raises two questions.

First, is fidelity to the data enough? Fair defends SEMs because they fit the data. But it is important to recall the criticisms which emerged of models like Fair’s model, way back in the 1970s and 1980s, as a result of the rational expectations revolution and the Lucas critique (see section II, above). Fair responds to this criticism by arguing that the data reject rational
expectations in favour of the adaptive expectations used in his model, in which case the Lucas critique does not hold. We think his response is only partially successful.\footnote{Fair responds to the Lucas critique in footnote 1. He writes: ‘a SEM model that is based on the assumption of adaptive expectations and thus not subject to the Lucas critique can be used for both forecasting and policy analysis.’ But adaptive expectations were precisely Lucas’s target when he launched his critique. The defence lurking in Fair’s paper, but not explicitly stated, is the one we describe above, that the data actually reject a rational-expectations specification in favour of the adaptive expectations specification which he adopts. There has been much discussion of this question in the literature, which we think remains unresolved. We deliberately do not discuss it here.}

As a consequence, we wonder whether there are contemporary areas of enquiry to which Fair’s kind of SEM cannot contribute. For example, there has been a recent debate about ‘is the Phillips Curve dead?’—that is, whether the historically observed relationship between inflation and the output gap still holds. This is highly relevant for policy-makers. Resolution of this debate rests on how the Phillips Curve is identified. Even if there is a strong Phillips Curve relationship, activist central banks will be acting to make it disappear, which runs right into the Lucas critique. Properly identifying the slope of the Phillips Curve in the face of this problem requires us both to isolate ‘policy shocks’ in the Taylor rule (or policy rule) and to have a theory of expectations more subtle than merely adaptive expectations. Doing this may require much more attachment to theoretical priors than Fair’s type of SEM can accommodate. Might this be a place where DSGE models could usefully be deployed (see Tenreyro, 2018)\

Fair’s answer to this first question has profound—and troubling—implications for what Fair says about the conduct of monetary policy. In Fair’s model, inflation does not follow an accelerationist-expectations-augmented Phillips Curve. Instead, the level of unemployment influences the level of wages and prices, rather than their change, as in the conventional set-up. But, that conventional set-up underpins all of Taylor-rule macroeconomics, including our discussion above of whether the Phillips Curve is dead. We have corresponded with Ray Fair about this aspect of his model. His response, when pressed, is to say firmly that the data do not support an accelerationist structure. ‘There is strong evidence,’ he says in section III(iii) of his paper, ‘against the NAIRU specification. This would appear to be a first-order problem for new Keynesian Phillips curves: they have the wrong dynamics.’ Such a view raises serious concerns about Taylor rule macro and the policy process to which it leads. Fair recognizes this. But he does not offer a clear alternative.

Fair’s approach is very different from that of a typical NK-DSGE advocate. Such a person would insist, as a theoretically motivated prior, that wages and prices be estimated within an accelerationist structure. If necessary, ‘fiddles’ would then be found to ensure that such a structure was not rejected by the data.\footnote{Fiddles of the kind brutally described by Romer (2016).} It is not at all clear that Fair is adopting the right answer to this question. But it is also not clear that fiddles provide the right answer.
The second question is this. What should we do when the builders of SEM disagree with each other? Fair insists that financial and housing wealth should be included in the consumption function, creating something more complex than the Euler equation. But we will see below that Muellbauer (2020) argues for something even more complex. This includes income expectations and a credit channel, using the wealth-to-income ratio rather than net worth, and disaggregating both wealth and consumption as part of the explanation. How much detail is necessary?

In our view this is exactly the type of discussion that the profession needs to have. How detailed does an SEM need to be in order to do what is required of it? And in our view, what is required of it includes revealing the risks of multiple equilibria and bad outcomes.

We will see by the end of this section that this discussion will need to be much more subtle and complex than the discussions which normally take place around the estimation of an NK-DSGE model.

(iii) Creating a consumption function for use in a structural economic model

Muellbauer’s paper (2020) occupies a central place in our Rebuilding Macroeconomic Theory project because it provides a careful and detailed account of the kind of consumption function needed in an SEM. This is something which he calls a ‘credit-augmented consumption function’ (see equation (1) in the paper). Such a consumption function is a generalization of the ‘life-cycle-permanent-income-hypothesis’ (LCPIH) consumption function. By contrast, the Euler equation commonly deployed in NK-DSGE models is a special case of the LCPIH consumption function.88

The LCPIH consumption function involves rationally optimizing, forward-looking individuals choosing their consumption, based on the permanent income they expect over their lifetime. Appendix 2 of Muellbauer’s paper provides a helpful narrative of the four steps it has taken for his consumption function to emerge from the LCPIH model: (i) it adds a credit channel to allow credit conditions to alter consumption; (ii) it separates liquid and illiquid assets in household balance sheets, and shows that the marginal propensity to consume (MPC) is much higher out of income which takes a liquid form; (iii) it applies a much higher discount rate to future income than used in the LCPIH model and the Euler equation; and (iv) it allows consumption to be influenced in the short term by income insecurity (proxied by the unemployment rate), and by the cash-flows of indebted households (changes in which are captured by changes in the interest rate).

88 The version of the LCPIH model which leads to the Euler equation in the NK-DSGE model assumes an infinite time horizon (and thus bequests). And the implementation of that equation in the NK-DSGE model assumes rational expectations.
These four extensions really matter when fitting consumption data, as Muellbauer demonstrates in sections V and VI of his paper. It is often argued that there is a trade-off at work here: that better explanation of data requires a move away from theoretical coherence (Pagan, 2003). But, in Appendix 1, Muellbauer argues that his better fit of the data has not required a sacrifice of theoretical coherence, but has, instead, shifted the research frontier outwards. In this Appendix, Muellbauer compares his approach with the simpler, more conventional, implementation of the LCPIH model, stemming from Ando and Modigliani (1963). In that approach consumption is related to just two variables: non-property income and net worth. Muellbauer documents the tricks which enable this conventional approach to use much less data to describe consumption than his approach. He compares these tricks to the epicycles of Ptolemy that were thought by sixteenth-century observers to provide an adequate explanation of the small amount of data on planetary movements that was available at the time. It does seem that these tricks are implausible.

These four extensions also really matter for policy analysis. They are essential for an understanding of how both fiscal and monetary policy might stabilize the economy. The size of the Keynesian fiscal multiplier relies on the MPC, and so getting it right is crucial. Muellbauer shows that the MPC is much larger than the extremely low value implied by the NK-DSGE model, and therefore that fiscal policy is much more powerful than is thought conventionally. Another conventional idea is that monetary policy relies on the channel of intertemporal substitution evident in the Euler equation, a set-up in which low real interest rates bring consumption forward from the future. Muellbauer shows that this does not happen because, by and large, households are credit constrained, unlike in the Euler-equation analysis. He shows that, instead, monetary policy operates through a credit channel. Taken together, these findings show that both fiscal policy and monetary policy can have strong effects, unlike in the NK-DSGE model, and that these effects arise for reasons not captured by that model.

These four extensions are also needed to understand how shocks propagate, and cumulate, through the economy. As we have made clear, Muellbauer allows for feedback between credit conditions and asset prices in his model of consumption. This can cause shocks to have a cumulative effect on consumption, which can lead to instability and the possibility that the economy lurches from good equilibria to bad. In contrast, the NK-DSGE model obscures this possibility because it treats credit flows and asset prices as frictionless side-shows; or, as Muellbauer says, ‘memo items’ which proxy expectations of future growth but play no role in system dynamics or in long-run behaviour.

---

89 As discussed above, this is what Ray Fair does.
90 Ad hoc extensions to the conventional approach try to address this inadequacy by assuming that a fraction of households simply spend current income, rather than being guided by the LCPIH. As we described above, this is what McKibbin does.
91 The Euler equation implies that consumption growth is driven by expectations about future income.
There are three complementary ways in which Muellbauer argues that shocks can propagate, and cumulate, through the economy.

First, there is a non-linear interaction between credit conditions and the marginal effect of housing collateral on consumption. Loose credit conditions both raise the price of houses (which form a large part of household wealth) and increase the responsiveness of consumption to those house-price increases. Tight conditions do the opposite, as happened during the global financial crisis in 2008, when there was a dramatic collapse in the consumption-to-income ratio. Muellbauer provides compelling evidence in support of this story and also a link to a paper (Berger et al., 2018) which provides some relevant theory.

Second, as average property prices fall, the tail-risk of mortgage defaults rises very non-linearly. Defaults, in turn, reduce the capital and liquidity of banks, which affects their ability to extend credit, raises their borrowing costs, and increases the risks of contagion between them. This further reduces asset prices. Consumption then falls because credit is less available. As Muellbauer notes, ‘models with these kinds of non-linearities will replicate [the low growth of] volatility . . . in “normal” times but can warn of potential crisis risks and handle crises without having to adopt a different modelling framework.’

Third, ‘user costs’ also affect house prices non-linearly. Muellbauer argues that any expected appreciation of house prices reduces the user costs of housing. Expected appreciation depends on past appreciation minus the interest rate, and so when house prices are rising rapidly the user cost of housing can become very small. Conversely, in a downturn, these user costs can jump very sharply, especially if the risk premium jumps too. As the expected return from a housing investment is amplified by leverage, a tightening of credit conditions will raise user costs even further. This is part of the reason why house prices, and consumption, can overshoot when credit is easy, and collapse when it is tightened.

It is difficult to simulate all these effects in a full macroeconomic policy model. Muellbauer provides intuitive descriptions of what can happen. Work now needs to be done to strip the effects down to a simple two-dimensional multiple-equilibrium (ME) toy model (or models), as outlined in section III. Once this is done, there will be additional work to do bringing the insights from Muellbauer’s work, and these small ME toys, creating larger ME versions of the NK-DSGE model. We think that such efforts will be very valuable.

Taking a step back, the Muellbauer paper tells us that the behaviour of the financial system is intrinsically related to the instability of modern economies, and he offers some hints about how macroeconomic theory ought to respond. Since the global financial crisis, work on interconnectedness and leverage within the financial system has shown us that contagion between

---

92 ‘User costs’ are the costs of owning housing, defined by Muellbauer as ‘an adjusted real interest rate given by the tax-adjusted mortgage interest rate minus the expected rate of house price inflation, plus the depreciation rate for housing, the tax rate on property and a risk premium’.
banks can lead to out-of-control macroeconomic outcomes (May et al., 2008; Haldane, 2009; Gai et al., 2011). A large literature outlines how financial stability can be improved by regulating banks, and using the central bank’s ‘lender of last resort’ function when this is insufficient. This financial regulation must not ignore the non-linearities, and the potential for amplification, that are discussed in the Muellbauer paper—non-linearities that arise from interactions between the financial economy and the real economy through the behaviour of consumption in response to the prices of real estate. Since 2008 counter-cyclical macroprudential regulation has started moving in the right direction. Whether it has done so enough is a topic for discussion that is way beyond the scope of this paper.

(iv) Some necessary remarks on methodology

In his paper, Muellbauer (2020) summarizes his view of NK-DSGE models by saying that they are insufficiently dynamic, stochastic, general equilibrium, new, or Keynesian. They are therefore, he says, not suited for policy purposes. Instead, both he and Fair argue that policy should be informed by structural economic models (SEMs), which—presumably—try to pay better attention to all of the above list of words. But how should these SEMs actually be constructed and estimated?

Wren-Lewis (2020) argues that, in doing their estimation work, SEM modellers will need to move away from the ‘puzzle’ methodology that is common in the estimation of NK-DSGE models, towards improvements targeted more closely to individual equations. If, for example, Muellbauer (2020) is correct and the benchmark consumption Euler equation does not fit the data, then one would think to start by looking at consumption. This is just like the mechanic in a garage who, when the engine will not start, first looks at the spark plugs. But, as Wren-Lewis says, this typically does not happen. Instead, DSGE modellers tend to follow a ‘puzzle methodology’, by looking at how their model matches the overall properties of the economy, and labelling discrepancies as ‘puzzles’. These puzzles are often not resolved by partial-equilibrium estimation, but instead by adding generic properties like inertia in GDP. A common approach is to add wealth to the consumption function as a proxy for the entire housing sector, because wealth is already in the model and doing so is straightforward. The top journals tend to encourage this, favouring exploration of full model properties over detailed estimation of individual equations.

We agree with both Wren-Lewis and with Blanchard (2018) that good journals need to adapt what they publish to the work which is clearly needed: they need to encourage more detailed, albeit partial, estimation of particular equations. In some ways, the status quo makes institutional sense: documenting that an empirical fact exists, without fully understanding why, may be more important for a policy-maker than an academic. This explains why SEMs tend to live in central banks, and why journals may overlook partly worked-out big ideas in favour of
the more clearly defined task of resolving a puzzle. But science can only progress by respecting the data.

And there is more that needs to be said about the actual methods of estimation. The starting point is that SEMs should be estimated in single equations or small blocks. Of course, in doing this, those who estimate the models will attempt to allow for simultaneity, through use of instrumental-variables techniques and other methods. Such an approach to simultaneity is very different from the way in which it is normally dealt with when NK-DSGE models are estimated. This is normally done by estimating the full model, all at once, in a set-up in which the full structure of the new-Keynesian DSGE model is imposed on the estimation process. Even for small models, like that in our Appendix, this is a very large task since there are very many parameters to be estimated. The task is often made manageable by imposing strong priors on some of the parameters, where these priors come from theory and from initial off-model knowledge. Romer (2016) has complained bitterly that this can lead to very odd results. As he shows, it is not just the identification restrictions, coming from the imposed structure of the model, which create difficulties. It is also the fact that imposing tight Bayesian priors in some equations can cause very strange outcomes for parameter estimates in other equations with weaker priors. Single-equation, or small-block, estimation avoids this possibility.

We think that there is a much deeper problem than this. As we have noted, Muellbauer (2020) says that the NK-DSGE model is like Ptolemy’s pre-Copernican framework for looking at the universe: it may be logically coherent on its own terms, but fundamentally flawed because it is so contradicted by the evidence. He takes particular aim at the treatment of consumption, arguing that the Euler equation, with fully rational expectations, is an analytical straight-jacket, which needs to be shaken off if we are to be able to estimate a sensible consumption function and thereby describe reality. Muellbauer’s complaint can be levelled at the way in which the NK-DSGE structure impedes the carrying out of important empirical work on other important parts of the macroeconomy, including—among other things—the decisions of firms to invest and choices by firms and worker of what wages and prices to set. There too the constraint needs to be shaken off.

Only when we escape from the constraint that the NK-DSGE model imposes will we be able to carry out the ‘careful scholarly work that is dedicated to obtaining a detailed empirical understanding of the economy as it actually is’. Only when we have done this will we be able to properly investigate the possibility of unstable cumulative processes that lead to multiple equilibria and to bad outcomes. And only then will we be able to design good macroeconomic policy.
V. Conclusion

We conclude the paper with some meta-level comments about the research programme of the new MEADE paradigm.

We argue that the NK-DSGE model needs to be displaced from its pre-eminent position in teaching, publishing, and informing policy in central bank research departments. It has fatal flaws tied to the unique rational expectations equilibrium path and it takes too many inappropriate short cuts.

The three kinds of research activities which we have identified—a more liberal use of toy models, understanding multiple equilibria, and building policy models through detailed empirical work on particular aspects of the economy—together make for an exciting research programme. They will accommodate many types of scholarship. And they will provide the possibility of doing something rather more interesting than just adding yet more frictions to the benchmark NK-DSGE model.

There will need to be much better communication between the different kinds of modellers and their models. New empirical relationships may first appear in structural economic models before leading to new suggestions about simple toys. Integrating these simple toys into the full NK-DSGE model may, in turn, suggest new empirical relationships.

Such a displacement of the NK-DSGE model will require a change in the macroeconomic research culture. In Vines and Wills (2018) we pointed to the need for tolerance and acceptance of different models for different purposes, and for a culture which allocates publications, positions, and funding accordingly. This will be a culture that is less insular: it will break down the ‘classical dichotomy’ which exists in current research practice: one which separates the study of short-term disturbances from the study of long-term growth. It will be a culture that more flexibly incorporates data and theory. And it would be a culture that recognizes the particular need for models to assist with policy-making.

This new culture will require a significant reform of the current publication process, which has many failings. It currently pushes researchers, particularly young ones, towards an acceptance of received wisdom. This process is, in theory, double-blind. But it is, in practice, single-blind, due to the existence of online working papers. And it often exposes authors to reviewers who have an interest in defending models in which they have invested decades of intellectual capital, something which, of course, incentivizes an inherent conservatism (Akerlof, 2018). Furthermore, there is a tendency of economics journals to publish work by authors who are connected to the editors. (Laband and Piette, 1994; Colussi, 2018; Heckman and Moktan, 2018) These problems are exacerbated by a slow refereeing process, which is neither adequately recognized nor sufficiently remunerated, dis incentivizing the effort involved. There is a rigid hierarchy of journals that is disproportionately weighted towards the top five (Heckman and
Moktan, 2018), a process which has become more exclusive over time (Card and DellaVigna, 2013). And there are also problems with the oligopolistic structure of the academic publishing industry, which does not pay authors for their contributions, and collects rents by controlling access to publicly funded research. Two helpful responses may be to place more weight on working papers, as is done in other scientific communities like physics and computer science, and to use votes to measure impact, as occurs in online social networks like Reddit.

Nevertheless, there are real grounds for optimism, which one can see by looking beyond our own backyard. The sense of possible excitement is well conveyed in an interview with Bob May, one of the greatest Australian scientists of the twentieth century, who became the UK’s Chief Scientist and President of the Royal Society, and who died just a few months ago. May was outstanding in mathematics (doing fundamental early work on chaos theory), in physics, in engineering, and then in zoology. And at the end of his career he also did important work in macroeconomics, using his epidemiological knowledge to study contagion in financial markets in work with Andy Haldane at the Bank of England.

May was interviewed in London in 2011 by the science journalist Robin Williams. At the beginning of the interview Williams asked, ‘What exactly are you? Is it Professor of Everything?’ May replied revealingly:

I am a scientist with a short attention span. There are different kinds of people in science, not just theorists and experimentalists, [but also] people who pick on one problem and devote their life to it, [and] people who accidentally stumble across something. I like to get in early, when you can do nice simple things that are important, and then as the field grows and it becomes more a matter of important and systematic elaboration, I find that less congenial. It is not that I go round deliberately thinking of what’s a different thing to do, it is just that my scientific career—from the fact that it even exists onward—has been a sequence of accidents.

‘Yes, I see’ said Williams. ‘So what is the essential ingredient that has been secret of it all, is it maths?’

I wouldn’t say that it’s mathematics, to put a label on it. Mathematics is ultimately no more than thinking very clearly about something. And I like puzzles, but in the sense that I like thinking about complicated things, asking what are the potential simplicities hidden in them, and expressing that tentative thought in mathematical terms, and seeing where it leads me, in testable ways. (Williams, 2011)

Like May said, there are different kinds of people in economics, too. Some will write papers like Guzman and Stiglitz’s contribution to this issue, which describes far-reaching ideas while leaving room for others to pin down the mechanics. Others, like Paul Krugman, will come to specialize in doing exactly the opposite, by collapsing the world into simple intuitive sketches,
as Hicks did with IS/LM. The particular kind of genius which those people display—armed with a piece of paper and a sharp pencil—involves working out what to include and what to leave out.\textsuperscript{93} Then there will be the technically skilled simulation specialists who will find the multiple-equilibrium solutions to non-linear NK-DSGE models and present them in easy-to-understand pictures. We will all come to regard these rather large models as the twenty-first-century version of what toys now look like, and we will all learn to play with them. There will be others, like Muellbauer and Fair, who do the really hard work of building structural economic models to empirically describe the economy as it really is, and who help us to understand where bad equilibria might really lie. And there will be people like Warwick Mckibbin who build hybrid models in which they impose a theoretical structure around policy models—but not so much as to kill their usefulness.

The meta-level claim of this introductory essay is that all of these rather different kinds of work will belong within the research programme of the multiple-equilibrium and diverse (MEADE) macroeconomic paradigm that is now emerging.

We hope that this paper of ours might encourage a bit more mutual respect. Those who build NK-DSGE models have come to think of themselves as being at the centre of things; as being the king of the castle. It will be helpful if these researchers came to recognize that other people are doing equally valuable things. Indeed, they might even become more humble. As John Maynard Keynes said in 1930, ‘[T]he economic problem . . . should be a matter for specialists—like dentistry. If economists could manage to get themselves thought of as humble, competent people on a level with dentists, that would be splendid’ (Keynes, 1930b, p. 373). The outcome might be greater parity of esteem among macroeconomists, something which Simon Wren-Lewis calls for, loud and clear, in the next paper in this volume.

Appendix: The new-Keynesian benchmark DSGE model

(i) The real components of the model

We first set out the components of the real model that underlies the new-Keynesian benchmark model. For brevity and simplicity, we do not set out the utility-maximization problem of the representative consumer or the profit-maximization problem of the representative firm; we simply set out the first-order conditions, and equilibrium conditions, derived from these two sets of optimization decisions.\textsuperscript{94} And, also for simplicity, we assume a constant labour force and an unchanging level of technology; the model can be generalized in a straightforward way to incorporate labour force growth and exogenous technical progress.

\textsuperscript{93} The rather marvellous thing about Joe Stiglitz’s career is that he has also done this kind of thing. The economics of information is littered with Stiglitz’s sketches, as are a number of other parts of the subject.

\textsuperscript{94} For relevant details see Woodford (2003, ch. 5, section 3) and Schmitt-Grohé and Uribe (2006).
Notation is as follows: C, I, Y, L, K, w, R, and Q represent (respectively) consumption, investment, output, labour supply, the capital stock, the real wage, the (gross) real interest rate, and Tobin’s Q. The model consists of the following eight equations.

\[
\begin{align*}
\frac{1}{c_t} &= \beta R_t E_t \left( \frac{1}{c_{t+1}} \right) \\
w_t &= \chi c_t L_t^\rho \\
Y_t &= A_t K_t^{\alpha} L_t^{1-\alpha} \\
w_t &= (1 - \alpha) \frac{Y_t}{L_t} \\
R_t &= E_t \left[ \alpha \frac{K_{t+1}^{1-\delta} + (Q_{t+1} - 1) + \frac{1}{2\xi}(Q_{t+1} - 1)^2}{Q_t} \right] \\
I_t &= K_{t+1} - (1 - \delta) K_t + \frac{\xi}{2} \frac{(K_{t+1} - K_t)^2}{K_t} \\
Q_t &= 1 - \xi + \xi \frac{K_{t+1}}{K_t} \\
Y_t &= C_t + I_t
\end{align*}
\]

Equation (A1) is the intertemporal Euler equation for the representative consumer; equation (A2) is the intra-temporal labour-supply equation of the representative consumer, which equates the (real) wage to the marginal disutility of labour.\(^{95}\) Equation (A3) shows aggregate supply\(^{96}\) and equation (A4) shows that the representative firm employs labour up to the point where the marginal product of labour is equal to the wage. Equation (A5) shows that the representative firm carries out investment up to the point at which the marginal product of capital is equal to the real interest rate plus an allowance for the depreciation of the capital stock, minus any anticipated capital gains on the capital stock, plus an allowance for the marginal costs of capital adjustment.\(^{97}\) Equation (A6) shows that capital accumulation is equal to investment minus depreciation minus the resources wasted whenever the capital stock is adjusted; adjustment costs are convex and quadratic—the bigger is \(\xi\) the greater are these costs. Equation (A7) determines Tobin’s Q; the equation shows that the larger are adjustment costs of capital (i.e. the larger is \(\xi\)) the further will Q deviate from unity when the capital stock is away from its equilibrium. Equations (A5)–(A7), when taken together, mean that, when the model is simulated, the larger is \(\xi\), the more gradual will be any adjustment of the capital stock to its desired level. Equation (A8) shows that aggregate demand is always equal to aggregate supply; the version of the model without nominal rigidities solves for the real interest rate \(R\) which brings this about—effectively by making saving equal to investment. The model is log-linearized around the non-stochastic steady state and solved using the Blanchard–Kahn method; \(K\) is a predetermined variable and \(C\) and \(Q\) are jump variables.

\(^{95}\) The consumer’s utility function is logarithmic in consumption and decreasing and convex in labour.

\(^{96}\) The production function of the representative firm is Cobb–Douglas.

\(^{97}\) Because of the interaction of equations (A5)–(A7), this term has an effect which is increasing in \(\xi\) when the model is simulated, even although \(\xi\) appears in the denominator of equation (A5).
Notice from equations (A5)–(A7) that, whenever there is a shock to the system, the real interest rate will move away from the marginal product of capital plus an allowance for depreciation, given that $\xi > 0$, in order to maintain an equilibrium in which investment goes on being equal to savings. This is precisely because there are costs in adjusting the capital stock.

In this model, although all markets clear, there are general-equilibrium interactions between markets. Keynes found that such interactions happened between markets when he postulated sticky wages—implying that the labour market might not clear—and so effectively turned Marshallian analysis into IS-LM. Here the general-equilibrium story is an additional one: the microfoundations of this model mean that interactions between markets happen, even if all markets clear. For example, consider the fall in productivity analysed in the first of the simulations presented below. This initially leads to lower consumption, even though the real interest rate falls, and will lead, at any given real wage, to a reduction in the demand for labour. But leisure will also fall because consumption has fallen and so there will be an increase in labour supply. The initial outcome in the labour market, and in particular the initial movement in the real wage, will depend on the relative strengths of these movements in labour demand and labour supply, both of which are caused by developments in the product market.

Notice that, in the (implausible) limiting case, in which $\xi \to 0$, the model converges to a Ramsey model with a real-business-cycle structure, in which whatever is saved is costlessly invested. In such a set-up, equation (A7) shows that Tobin’s Q always remains equal to unity, equations (A6) and (A8) together show that additions to the capital stock are exactly equal to what is saved minus depreciation, and equation (A5) shows that the real interest rate is always just equal to the marginal product of capital plus an allowance for depreciation. But with adjustment costs to capital, i.e. with $\xi > 0$, then, after any shock, Tobin’s Q will move away from unity, some of what is invested will be wasted, and the real interest rate must diverge from the marginal productivity of capital by whatever amount is necessary to make savings equal to investment. This chain of reasoning shows why it is essential to include capital adjustment costs in the model to get away from the Ramsey model structure.

The parameters used in the simulations reported on in the paper are displayed in Table A1; these parameters correspond to the idea of a quarterly model.

**Table A1: Parameter values**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\delta$</th>
<th>$\chi$</th>
<th>$\phi$</th>
<th>$\xi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numerical value</td>
<td>0.35</td>
<td>0.99</td>
<td>0.025</td>
<td>0.3</td>
<td>0.33</td>
<td>10</td>
</tr>
</tbody>
</table>

(ii) **Generalizing the model to one which includes nominal rigidities**

We now add nominal rigidities to the model, and a Taylor rule, and so arrive at the full new-Keynesian benchmark model. We can use this to study the behaviour of inflation.
We proceed by introducing differentiated goods and nominal rigidities in the manner of Calvo (1983). Specifically, we assume a continuum of goods producing firms, \( i \), where \( i \) lies in the range \([0,1]\); these firms operate under monopolistic competition. These goods are aggregated into a single ‘final’ good by means of a Dixit–Stiglitz aggregator in which the elasticity of substitution between the varieties of goods is equal to \( \varepsilon \). Any firm re-optimizes its price in a period with fixed probability \((1-\kappa)\). (This is the Calvo fairy at work.) With probability \( \kappa \) prices are not re-optimized and are assumed to rise at the average rate of inflation. Finally, we assume—as is standard in the literature—that there is a subsidy to the firms (paid by a lump-sum tax) so that in the steady state firms produce the same quantities as in the flex-price economy, even though each firm faces a downward-sloping demand curve and so can restrict its output. This means that the steady-state values for \( C, I, Y, L, K, w, R, \) and \( Q \) are the same in the full new-Keynesian benchmark model as they were in the real model presented above.

The three new variables in this full model are \( \pi \), the rate of inflation, \( R^n \), the nominal (gross) interest rate, and \( mc \), which represents real marginal cost.

Of course, the model no longer solves for a real interest, \( R \), that would make savings equal to investment in a way which would ensure the absence of any inflationary pressure. Instead, it allows aggregate demand to move relative to aggregate supply, in a way which may create inflationary pressure. It then allows the central bank to determine the real interest rate by setting the nominal interest rate, in response to developments in the inflation rate, according to a Taylor rule. Equation (A9) expresses the real interest rate in terms of the nominal interest rate; equation (A10) shows the way in which the central bank sets the nominal interest rate according to a Taylor rule, one in which—for simplicity—there is no output term.\(^{98}\)

\[
R_t = E_t \left( \frac{R^n_t}{\pi_{t+1}} \right) \quad \text{(A9)}
\]

\[
R^n_t = \beta^{-1} \left( \frac{\pi_t}{\pi} \right) \phi_{\pi} \quad \text{(A10)}
\]

The term \( \beta^{-1} \) at the front of the Taylor rule determines the constant term when the rule is written in linear form; it shows that the real interest rate will converge to its equilibrium value \( \beta^{-1} \) when any disturbances to inflation have disappeared. As equation (A1) shows, such a restriction on the parameters is necessary in order for it to be possible for consumption to be in equilibrium.

The first eight equations of this new system are identical to those in the real system presented above, except for two crucial alterations. We replace equation (A4) with the following equation:

\[
w_t = mc_t (1 - \alpha) \frac{Y_t}{L_t} \quad \text{(A4a)}
\]

\(^{98}\) In a normal Taylor rule there is a term in the output gap as well as the inflation rate. But since, with an endogenous capital stock the output gap varies endogenously, not just due to changes in demand but also due to changes in supply, this term is complex to calculate. And a feedback on inflation alone is all that is necessary to create a working toy model.
This equation solves for levels of real marginal cost, \( mc \), different from unity, thereby enabling aggregate demand to move in a way which is not constrained by aggregate supply. We know, from equation (A2), that the economy always lies on the labour supply curve. We also know, from equation (A8), that output is demand determined. Consider the effects of an increase in demand caused by, say, a demand shock. This will cause output to rise, which from the production function—equation (A3)—will lead to higher employment. The real wage must rise as the economy moves up the labour supply curve (equation A2), \textit{ceteris paribus}. But, as more labour is employed, the marginal product of labour must fall (because of diminishing returns). Thus, because the wage is rising (as the economy moves up the labour supply curve), but the marginal product of labour is falling (because of diminishing returns), real marginal cost (which equals the wage divided by the marginal product of labour) must be rising. That is what equation (A4a) shows.

We replace equation (5) with the following equation

\[
R_t = E_t \left[ mc_{t+1} \alpha_k \left( Y_{t+1} - \delta + (Q_{t+1} - 1) + \frac{1}{2} (Q_{t+1} - 1)^2 \right)^\kappa \right].
\]  

(A5a)

The reasoning underlying this equation is similar to that applying to equation (A4a), but now the direction of causation is different. Consider again the effects of an increase in demand, caused by a demand shock. This will cause output to rise and so lead to higher employment and a higher wage. But higher output will also raise the marginal product of capital and so encourage investment. Nevertheless the extent to which investment rises will be constrained by the wedge between the wage and the marginal product of labour, which is what is captured by the real marginal cost variable, \( mc \). Equation (A5a) shows the effect of this wedge on \( Q \) and thus, through equations (A6) and (A7), on investment.

Finally, equation (A11) shows the new-Keynesian Phillips curve, in which inflation in any period depends on the (discounted) expected future rate of inflation plus a term which depends on the logarithm of the real marginal cost of labour to firms\(^99\)

\[
\pi_t = \beta E_t \pi_{t+1} + \frac{(1-\beta)(1-\kappa)}{\kappa} \ln(mc_t) + \mu. \tag{A11}
\]

This equation effectively shows, for example, that whenever aggregate demand rises above aggregate supply, and so \( mc \) rises above unity, any firm visited by the Calvo fairy will raise its price level. It will do this to an extent which depends on how many other firms are being visited at the same time — i.e. on the size of \( \kappa \) — which shows the proportion of firms not being visited in any period of time, and also on how high \( \pi \) is expected to be in the next period.

\(^{99}\) The logarithm of the real marginal cost of labour shows the proportional deviation of real marginal cost from the value of unity which it would have at the intersection of the labour supply and labour demand curves, at which point the wage equals the marginal product of labour and there would be no inflationary pressure. This argument is complicated, in an inessential manner, by the term \( \beta \) in front of the term \( E_t \pi_{t+1} \) on the right-hand side of this equation. For an explanation, which relates to how expected future events are discounted, see Galí (2015).
The model is solved using the Blanchard–Kahn method; the inflation rate, \( \pi \), is an additional jump variable.\(^{100}\)

We use the following calibration for the two new parameters that we introduce; \( \phi_\pi = 1.5, \kappa = 0.75 \). The first of these corresponds to the parameter used by Taylor when he first introduced his rule; this value ensures that whenever the inflation rate changes the real interest rate moves in the same direction. A value of \( \kappa \) of 0.75 ensures that the Calvo fairy arrives at each firm, on average, once a year. We do not need to calibrate \( \varepsilon \). This is because, although the size of \( \varepsilon \) influences the size of mark-up of prices over marginal cost, this mark-up is constant and so disappears when the system is log-linearized.

The version of the model in which price-flexibility ensures that resources remain fully employed may be obtained by letting \( \phi_\pi \to \infty \). When this happens, \( \pi \to 0 \), \( Rn \to R \), and \( mc \to 1 \), and the model converges to the real model described above.

(iii) Simulating the new-Keynesian DSGE model

We now describe how the new-Keynesian DSGE model can be simulated by describing a productivity shock and an inflation shock. Doing this shows the way in which the model encompasses both the RBC model and the CGG model. Note that exactly these simulations were included in Vines and Wills (2018).

A productivity shock

Simulating a negative productivity shock shows how the NK-DSGE model encapsulates the insights of the Ramsey–RBC model. The simulation in Figure A.1 demonstrates this, using a permanent shock of 10 per cent to total factor productivity.

First, a reminder of what would happen in the Ramsey model is helpful. In that model, after such a productivity shock, consumers decide to save less, so as to reduce the capital stock to its new longer-run equilibrium. By contrast, in the SW model it is firms which decide on investment and they do so in response to changes in the value of Tobin’s \( Q \). These depend both on what has happened to the productivity of capital—it has fallen—and to the real interest rate.

For simplicity, in presenting this simulation in Figure A1, we override the Taylor rule and instead imagine that a perfectly capable central bank sets the interest rate necessary to keep aggregate demand equal to aggregate supply. This is, we suppose that \( \phi_\pi \to \infty \), so that the model corresponds to that described in Appendix section (i). When the shock happens, it causes the value of Tobin’s \( Q \) to fall, and so profit-maximizing firms will want to invest less; the

\(^{100}\) Note that, as \( \kappa \) tends towards unity, the setting of prices tends towards full flexibility. If—in the case of a value of \( \kappa \) close to unity—demand pressures were to make real marginal cost differ from unity, then that would cause a very large disturbance to inflation. In an economy with a nominal anchor provided by a Taylor rule, such as Equation (A10), the central bank would raise the real interest rate, and that would, of course, moderate the demand pressures. In such an economy, as \( \kappa \) tended towards unity, the economy would tend towards one in which there was perfect price flexibility, output always equalled its natural level, and inflation moved instantaneously to a level at which the real interest was such as to bring about exactly that level of output.
length and intensity of the period of disinvestment will depend on the costs of capital adjustment. As a result of this disinvestment, the level of aggregate supply will fall more in the long run than in the short run. The sustainable level of consumption will fall and, in the interests of consumption smoothing, consumers will cut their level of consumption immediately. Because the long-run fall in output will be larger than the short-run fall, the fall in consumption will be large, even if the real interest rate does not change. It is clear that, even if investment did not fall at all, the fall in aggregate demand might be larger than the fall in aggregate supply. If adjustment costs in investment are sufficiently small—as assumed here—then it is clear that the fall in demand will be larger than the fall in supply, and this is what is depicted here.\footnote{There has been a debate in the literature about whether this assumption is appropriate.} This means that the real rate of interest will need to fall to ensure that savings remains equal to investment; and that aggregate demand remains exactly equal to the lower level of aggregate supply.

**Figure A1**: A permanent negative shock of 10 per cent to TFP

![Graph showing the effects of a negative shock to TFP on Y, I, K, R, Q, C, W, and L](image)


In the long run, the real interest rate will return to its equilibrium level which is equal to the rate at which consumers discount the future; in the short run, when it is less than this, the difference will be exactly equal to the marginal costs of adjusting the capital stock. In the long run investment will be a little lower, because the capital stock is lower and so depreciation is lower; consumption will be lower in line with the reduction in supply.
An inflation shock

Simulating an inflation (or cost-push) shock shows how the new-Keynesian DSGE benchmark model successfully encompasses the insights of the CGG model. The simulations in Figure A2 are based on the full model set out above in which the real interest rate is set according to the Taylor rule as shown. They show how the model responds to a 1 per cent increase in the cost-push-shock term, $\mu$, which we assume follows an AR(1) process with persistence 0.8, as is common in the literature. This prevents the results shown in the pictures merely having a spike, as the result of a spike in inflation, which would happen if there were no such persistence.

After the shock, all firms are assumed to raise their prices as a result of the shock. Higher inflation induces an immediate rise in the nominal (and real) interest rate, because the policymaker is following a Taylor rule. Consumption falls, and so does investment, the latter because Tobin’s Q falls. Output, employment, and the real wage fall, because workers are always on their labour supply curve. Inflation comes down only gradually, partly because the cost-push shock is autoregressive, and also because, in the light of this, each firm will moderate its price cuts because they know that some other firms are not yet reducing their own prices because of the autoregressive nature of the inflation shock.

**Figure A2**: A 1 per cent cost-push shock with a persistence of 0.8

![Graph showing the response of various economic variables to a 1 per cent cost-push shock](image)

*Source: Vines and Wills, 2018.*

After the shock, all firms are assumed to raise their prices as a result of the shock. Higher inflation induces an immediate rise in the nominal (and real) interest rate, because the policymaker is following a Taylor rule. Consumption falls, and so does investment, the latter because Tobin’s Q falls. Output, employment, and the real wage fall, because workers are always on their labour supply curve. Inflation comes down only gradually, partly because the cost-push shock is autoregressive, and also because, in the light of this, each firm will moderate its price cuts because they know that some other firms are not yet reducing their own prices because of the autoregressive nature of the inflation shock.
cuts because they know that some other firms are not yet reducing their own prices because of the autoregressive nature of the inflation shock.

A comparison with the CGG model is helpful. This model shows that the capital stock is depressed as a result of the disinvestment which happens during the adjustment process, but is gradually rebuilt to its initial level. This feature—that a fall in investment and so in capital will be part of any disinflation process—is something which Galí (2015) completely obscures in his textbook, by presenting a version of the new-Keynesian model in which there is no study of endogenous investment.

Luk and Vines (2015) show that the method of controlling inflation by monetary policy, which is examined here, rather than partly controlling inflation by fiscal policy, is costly precisely because it causes disinvestment during the adjustment process, something which needs to be reversed in a way which is costly. But they also find that the welfare effects of doing this are small. This is somewhat surprising, until it is recognized that there is no hysteresis in the model.

References


