Institute for
New Economic Thinking
AT THE OXFORD MARTIN SCHOOL

THE FUTURE OF MACROECONOMICS

John Muellbauer

Revised 10 November 2018

INET Oxford Working Paper No. 2018-10

INET Oxford Programme: Economic Modelling

Paper prepared for the ECB colloquium held in honour of Vítor Constâncio: 'The Future of Central Banking'

The Future of Macroeconomics

John Muellbauer

Nuffield College, and Institute for New Economic Thinking at the Oxford Martin School, University of Oxford, U.K.*
(john.muellbauer@nuffield.ox.ac.uk)

Paper prepared for the ECB colloquium held in honour of Vítor Constâncio: 'The Future of Central Banking'

Wednesday, 16 May 2018

Revised 10 November 2018

Keywords: DSGE, central banks, macroeconomic policy models, finance and the real economy, financial crisis, consumption, credit constraints, household portfolios, Phillips curve, forecasting inflation.

JEL classification: E17, E21, E31, E32, E37, E44, E51, E52, E58, G01.

*This paper draws on research over years with co-authors Janine Aron, the late Gavin Cameron, Valerie Chauvin, John Duca, Felix Geiger, David Hendry, Keiko Murata, Anthony Murphy, Manuel Rupprecht and David Williams and I am grateful to them all. Section 4 borrows from current work with Janine Aron and I am immensely grateful to her. Recent discussions with John Duca have also been very helpful as have comments from Philipp Hartmann and David Hendry. The research in Section 3 was partly the product of a project pursued at the ECB during my tenure of a Wim Duisenberg Visiting Fellowship. Longer-term research support from the Open Society Foundation via INET at the Oxford Martin School is gratefully acknowledged.

Abstract

New Keynesian DSGE models, fashionable with central banks, have too long ignored the insights of the Information Economics Revolution to which Joseph Stiglitz made seminal contributions. Further, these models have failed to assimilate important research insights from encompassing alternative theories, model selection and the implications of structural breaks, to which econometric literature David Hendry is a key contributor. This paper summarises the critique. After responding to a recent counter-attack on critics of New Keynesian DSGE models, it shows how evidence-based research can improve quantitative policy models allowing central banks to have a better understanding of financial stability, and to improve inflation models and forecasts. In particular, household sector equation systems for consumption, balances sheets and asset prices are able to explain the amplification mechanisms that put some countries at greater risk of financial instability. And new evidence from forecasting US core inflation suggests that with appropriate controls, a stable relationship between unemployment and inflation is revealed, though not between the policy rate and inflation. This empirical evidence is consistent with arguments advanced in recent years by the ECB's retiring Vice President Vítor Constâncio.

1. Introduction

Given the wider theme of the colloquium, I concentrate on quantitative models used or potentially used at central banks to guide policy. My thinking on these issues since the 1970s was particularly influenced by the work of Joseph Stiglitz on 'information economics' and of David Hendry on how better to learn from data and avoid the key pitfalls when forecasting. These are my two most significant intellectual influences in quantitative macroeconomics. In his Nobel lecture, Stiglitz (2001) wrote: "In the 70s, economists became increasingly critical of traditional Keynesian ideas, partly because of their assumed lack of micro-foundations. The attempt made to construct a new macroeconomics based on *traditional* microeconomics, with its assumptions of well-functioning markets, was doomed to failure. Recessions and depressions, accompanied by massive unemployment, were symptomatic of … market failures….If individuals could easily smooth their consumption by borrowing at safe rates of interest, then the relatively slight loss of lifetime income caused by an interruption of work of six months or a year would hardly be a problem; but the unemployed do not have access to capital markets, at least not at reasonable terms, and thus unemployment is a cause of enormous stress. If markets were perfect, individuals could buy private insurance against these risks; yet it is obvious that they cannot. Thus, one of the main developments to follow from this line of research into the consequences of information imperfections for the functioning of markets is the construction of *macroeconomic* models that help explain why the economy amplifies shocks and makes them persistent, and why there may be, even in competitive equilibrium, unemployment and credit rationing."

In the same lecture, he suggests: "Information economics has alerted us to the fact that history matters; there are important hysteresis effects. Random events – the black plague – have consequences that are irreversible. Dynamics may be better described by evolutionary processes and models, than by equilibrium processes. And while it may be difficult to describe fully these evolutionary processes, this much is already clear: there is no reason to believe that they are, in any general sense, 'optimal'."

On learning from data, as his research over many decades has demonstrated, Hendry (2017) argues: "There is a false belief that data-based model selection is a subterfuge of scoundrels—rather than the key to understanding the complexities of macro-economies". On forecasting, he wrote in 2011: "Economic forecasting is inevitably carried out using (unknowingly) mis-specified models, in the face of unanticipated location shifts….much of the

existing empirical evidence on forecast performance can be accounted for by a theory which allows for both features, using parameter estimates based on mis-measured data."

There are very important insights embodied in these quotations, though largely ignored among macroeconomists and at central banks in the quantitative models and research methodologies until recently fashionable. The 2018 (Spring) issue of the *Oxford Review of Economic Policy*, marking 10 years since the global financial crisis, is devoted to the controversies around different methodological approaches in macroeconomics. It includes critiques from Olivier Blanchard, Joseph Stiglitz, Simon Wren-Lewis and David Hendry and myself of the New Keynesian, representative agent, dynamic stochastic general equilibrium (DSGE) models used in many central banks in recent years. Defences in the issue include those from Ricardo Reis and Jesper Linde.

The June issue of the *Journal of Economic Perspectives,* also marking this 10-year anniversary, is another significant mile-stone. Gertler and Gilchrist trace the history of the crisis, and note the importance of the financial accelerator operating through the household sector and interacting with spill-over effects within the banking system. Mian and Sufi review the extensive microeconomic evidence for the role of credit shifts in the US sub-prime crisis and the constraining effect of high household debt levels. They summarise this as the "credit-driven household demand channel". Kaplan and Violante spell out implications of heterogeneous agent models incorporating idiosyncratic, uninsurable income risk, credit and liquidity constraints, and discuss the limitations of existing models and unresolved research questions, for example on asset pricing and labour market income risk. They acknowledge that current versions of the heterogeneous agent New Keynesian model still "miss the potentially large wealth effects on consumption for wealthy households that can arise from changes in asset prices".

In the same issue, Cristiano et al. revise and tone down their widely circulated defence of the DSGE approach, Christiano et al. (2017). In what might be described as a sleight of hand, Christiano et al. (2017), label all quantitative models aiming to capture general equilibrium phenomena, as 'DSGE', and call those who work outside this framework, 'dilettantes'. This ignores the fact that increasing numbers of central banks use large 'semi-structural' econometric models of the macro-economy that do *not* assume specific micro-optimising behaviour and do *not* adopt the assumption of rational expectations common to all agents, both generally adopted in DSGE models. It also ignores the 'suite of models' approach, including forecasting models for specific sectors, widely used at central banks. Finally, they mistakenly conflate general equilibrium theory models, designed to give often useful insights

into a particular set of mechanisms, with the quantitative models that encompass the main variables of interest to central banks. In Section 2 of this paper, I tackle head-on the issues they raise.

In the remainder of the paper, I address two central concerns of central bankers: financial stability, and the understanding and forecasting of inflation and of business cycles as represented by fluctuations in GDP. I argue that the *conventional practices* in how the macro-profession is allowed to learn from data have led to failures by central bankers in making adequate progress on such concerns. The profession has been forced into a schizophrenic condition. On one hand, the combination of the Lucas critique and the demand for tractable micro-foundations for full general equilibrium forces heavy *restrictions* on models. These are imposed by the omission of major channels and imposing Bayesian priors in model estimation, and at the extreme, resorting to calibration. On the other hand, Sims 'incredible restrictions' critique of large econometric policy models in his 1980 paper on *Macroeconomics and Reality*, proposed to let data speak with minimal *a priori* restrictions, by estimation of loosely-parameterised VAR models. However, the curse of dimensionality then leads to making restrictive assumptions by limiting both the set of relevant variables and the lag-length, and once again to Bayesian restrictions to reduce the impact of that curse. *Both* methodological options create mis-specified models through omissions and heavily limit what can be learned from data.

These issues are critically addressed in Sections 3 to 5. Section 3 discusses financial stability, and how a broader perspective on empirical evidence for the personal sector and its linkages with providers of credit can provide clearer insights into economic issues, particularly with regard to the financial accelerator. For instance, such a perspective can explain how differences in institutions between countries are related to differing risks of instability, rather than assuming away these key institutional differences as is often done.

In Section 4, the empirical evidence on modelling and forecasting US inflation (currently pre-occupying markets) is re-examined from this broader empirical perspective. In Aron and Muellbauer (2018), substantial improvements in forecasting performance are found using indicators of institutional changes in market power (union density and industry concentration), a broad set of relevant drivers of core inflation including the levels of relative prices, and 'Parsimonious Longer Lags' (PLL). Simple restrictions on the lag structure save parameters and allow larger sets of candidate regressors and longer lags to be considered. We find strong evidence that the inflation process involves longer lags than conventionally thought relevant, supporting similar results by Gordon (2009, 2013). We also reconsider the highly

debated question of whether there is a relationship, let alone a stable relationship, between core inflation and the level of US unemployment. Farmer and Nicolò (2018) have recently proposed a Keynesian model without a Phillips curve, by which they mean the New Keynesian Phillips Curve (NKPC). We agree that the NKPC is unstable, but conclude that at least since the late 1970s there has been a relatively stable relation between US core inflation and the level of unemployment, *provided the right controls are included for other drivers of inflation*.

Section 5 gives an example of how a more flexible approach to learning from data can also substantially improve forecasts of household income and of GDP. The findings hint that path dependence is a major issue for growth and the business cycle.

Empirical evidence on the personal sector and its linkages with credit supply and these forecasting results throw light on economic processes, which should enhance the formulation of the new generation of large macro-econometric models. In particular, section 3 points to the need to relax the net worth constraint commonly imposed to model consumption, to control for shifts in credit conditions, and use better models of house prices and residential investment. The Federal Reserve's FRB-US model could be greatly improved in these directions. Section 4 suggests that the price and wage sector of such models could be improved from the insights of more comprehensive inflation models. Section 5 suggests a more flexible approach is needed to explain the evolution of capacity and that growth and the business cycle might be more intimately connected than often assumed. This also affects how one can model households' evolving perceptions of 'permanent income' in a world of evolutionary and sometimes abrupt change.

2. DSGE models and 'dilettantism'

The New Keynesian DSGE models that have dominated the macroeconomic profession and central bank thinking for the last two decades are based on the principle of formal derivation from micro-foundations.[1] They assume an optimising behaviour of consumers and firms consistent with rational or 'model-consistent' expectations of future conditions. To produce a tractable model, it is further assumed that the behaviour of firms corresponds to that of a 'representative' firm and of consumers to a 'representative' consumer. In turn, this entails ignoring the necessarily heterogeneous credit or liquidity constraints that in practice bind both on firms and consumers. Another assumption to obtain tractable solutions is that of a stable

[1] An exposition of "the celebration of the 'Science of Monetary Policy'" is given in Clarida et al. (1999).

long-run equilibrium trend path for the economy. If the economy is never far from such a path, the role of uncertainty would necessarily be limited. Hence, popular pre-global financial crisis versions of the model could exclude banking and finance, taking as given that finance and asset prices were a mere by-product of the real economy. A competitive economy is assumed but with a number of distortions, including nominal rigidities (sluggish price adjustment) and monopolistic competition. This distinguishes New Keynesian DSGE models from the preceding general equilibrium Real Business Cycle models (RBC). The range of stochastic shocks that could disturb the economy is also expanded relative to the productivity or taste shocks of the RBC model. Some DSGE models calibrate values of the parameters; but where the parameters are estimated, Bayesian system-wide estimation is used, imposing substantial prior restrictions to achieve parameter values that are deemed 'reasonable'.

Taking a plain-spoken and critical view of the New Keynesian Dynamic Stochastic General Equilibrium models, one finds they are *not stochastic enough* – as they trivialise the role of uncertainty and heterogeneity, are *not dynamic enough* – as they miss key lags in relationships, and are *not really general equilibrium* – as they ignore important feed-back loops, seen for example in the global financial crisis. They are *scarcely new*, being based on ideas made redundant by the asymmetric information revolution of the 1970s and 80s, and *hardly Keynesian*, as they miss co-ordination failures in labour and financial markets. As argued in Hendry and Muellbauer (2018), and by Stiglitz (2018), the specific flaws[2] of the NK-DSGE model can be summarised under six headings:

1. The micro-foundations are built on sand, ignoring the information economics revolution and assuming complete markets

Stiglitz's prescient 2001 Nobel Prize lecture and Stiglitz (2018) spell out the importance of incomplete markets[3] and asymmetric information in real-world economics. Of particular interest are the implications of credit and liquidity constraints for household behaviour. Research by Deaton (1991) and Carroll, in a series of papers beginning with Carroll (1992), shows that given uninsurable individual income risk and the liquidity constraints that result from asymmetric information, households will engage in buffer-stock behaviour to ameliorate income risk, and they discount expected future income at higher rates than assumed by the textbook model. This behaviour has profound implications for the effectiveness of monetary

[2] Muellbauer (2010) puts forward a broadly similar critique.
[3] Buiter (2009), in his trenchant post-crisis critique of standard macroeconomics, highlights the complete market assumption as the central failing of New Keynesian DSGE models.

and fiscal policy. In contrast, in NK-DSGE models, households discount temporary fluctuations in income to maintain spending in the face of shocks, thus providing a stabilising anchor to the economy.

2. The micro-foundations are built on sand, assuming representative agents

In the real macro-economy, heterogeneity rules. It is well-known that the conditions under which 'average behaviour' of households is the same as that of an individually-optimising household are highly restrictive. There are many examples of the general point that even the *functional form* which holds at the micro-level can be radically different from that holding at the aggregate level, when aggregating across heterogeneous agents.[4] The combination of asymmetric information and household inequality mean that income processes, liquidity constraints faced by households, and asset ownership are all highly heterogeneous. It is nevertheless facile to suggest that aggregate data is therefore uninformative: but aggregate models need to be richer and encompass the extensive as well as intensive margins of behaviour. To illustrate, the aggregate mortgage debt to GDP ratio depends *both* on the average debt of households which hold mortgages (the intensive margin) and the fraction of those with a mortgage (the extensive margin). There is no reason why a function designed to explain the former should also explain the latter. But an aggregate model for mortgage debt should capture both features of the data. Aggregate models could use micro-data or assumptions about the forms of the micro distribution of data to address the implications of stochastic aggregation.[5]

Another example illustrates the informational content in aggregate data. The incidence of unemployment is highly heterogeneous and so is price setting. For example, pricing practices for rents, medical and insurance services, clothing and electronic products are all likely to differ. Despite this heterogeneity, as we show in Section 4, there is a surprisingly stable relationship between aggregate unemployment and aggregate core inflation, provided the right controls are included in the inflation equation.

3. The rational expectations assumption is incompatible with a world of structural breaks and radical uncertainty.

[4] Hendry and Muellbauer (2018, p.295) give two such examples: Houthakker (1956) and Bertola and Caballero (1990).

[5] One example is Aron and Muellbauer (2016) who estimate a distributional form for negative equity from a micro-snapshot to generate time-series estimates of the aggregate proportion of mortgages with negative equity based on aggregate debt/equity ratios. This proves highly significant in explaining the proportion of defaults and foreclosures in the UK.

Structural changes, frequent and widespread in every economy, are a key source of forecast error. It is therefore wildly implausible to endow agents with rational expectations that foresee such breaks. To quote: "*The mathematical basis of DSGEs fails when events suddenly shift the underlying distributions of relevant variables. The 'law of iterated expectations' becomes invalid because an essential, but usually unstated, assumption in its derivation is that the distributions involved stay the same over time. Economic analyses with conditional expectations ('rational expectations') and inter-temporal derivations then also fail, so DSGEs become unreliable when they are most needed*."[6] Heterogeneity at the micro-level suggests that for model-consistent expectations, individuals would need not only to forecast their own *individual* heterogeneous circumstances but also *aggregate* circumstances. The latter, and potentially both, are subject to structural breaks in, for example, technology, credit conditions, monetary and fiscal policy rules, globalization, and trade union power. As Caballero (2010) pointed out, the 'pretence of knowledge' was extreme in the NK-DSGE approach, both on the side of the modeller and on the side of the agents populating the model economy.

Forecasting failures are well-known in macro-economics (Loungani, 2001; Ahir and Loungani, 2014), so that even *pre*-crisis, model-consistent expectations are highly unrealistic. Loungani (2001) concludes: "the record of failure to predict recessions is virtually unblemished". This does not preclude economic agents from forecasting. But it does mean those forecasts are best represented by simpler, limited information forecasting models, even if agents often get those forecasts wrong.

4. Omitting shifting credit constraints, household balance sheets, and asset prices, and hence ignoring the financial accelerator

The omission of money, credit, banks and asset prices from the NK-DSGE model led Charles Goodhart to comment: "It excludes everything I am interested in", see Buiter (2009). The asymmetric information revolution of the 1970s provided micro-foundations for the application of credit constraints by the banking system (Stiglitz, 2018). The use of collateral became increasingly widespread from the 1960s to reduce the asymmetric information problem. This led to a huge rise in real-estate backed lending, transforming the banking sector in most OECD countries (Jordà et al., 2016). For many countries, shifting constraints as applied to collateral became among the most important structural changes in the economy. An unintended side effect was the *shifting of risk*: the micro-risk of asymmetric information between lender and

[6] As reported by Hendry and Muellbauer (2018, p.299-301), summarising the conclusions of Hendry and Mizon (2014).

borrower, e.g. adverse selection and moral hazard, was shifted to the macro-risk of real estate price collapses. With hindsight, many US bankers say that their biggest mistake during the mortgage credit boom was not in mis-managing micro-risk but in failing to appreciate that average house prices faced serious down-side risks. Section 3 returns to these themes of financial stability and house price overvaluation.

5. The failure to be structural in the Cowles Commission sense

Haavelmo's (1944) classic Cowles Commission article on the probability approach in econometrics is the first systematic definition of 'structural'. Haavelmo contrasts the potential lack of autonomy of empirical regularities in an econometric model with the laws of thermodynamics, friction, and so forth, which are autonomous or 'structural' because they 'describe the functioning of some parts of the mechanism *irrespective* of what happens in some *other* parts'. Tracing how the New Classical Revolution in macroeconomics gained dominance and the reason large econometric policy models of the 1970s were displaced as policy models by DSGE models, the Lucas (1976) critique proved to be the key (see Wren-Lewis (2018) and Hoover (1994)). Since the parameters of such large policy models were supposedly not 'structural', they necessarily changed whenever policy changed. In the reign of the DSGE models, however, there was a subtle shift in the *meaning* of 'structural': the definition came to imply 'micro-founded in individual optimising behaviour'. Hendry and Muellbauer (2018, p.292-4) observe that with shaky micro-foundations in a world of structural breaks, DSGE models typically fail to maintain parameter stability. Therefore, the DSGE models also fail to be *structural* in the more fundamental sense of the Cowles Commission.

6. The lack of flexibility of even the newest representative agent DSGEs

The consumption Euler equation is the key mechanism for the operation of model consistent-expectations. This makes it the main *straitjacket* of the representative agent DSGE approach. A much more modular approach, as for example adopted in the non-model consistent version of FRB-US, allows heterogeneity in expectations between households and firms. Within DSGE models resting on an aggregate or sub-aggregate Euler equation[7], this kind of modularity is hard to achieve.

[7] More recent models sometimes assume two representative households, one is an intertemporal optimiser, and the other is myopic and just spends current income.

The attack on critics of New Keynesian DSGE

A recent attack on critics of DSGE models by Christiano et al. (2017) begins with the hard-line label: "people who don't like DSGE models are dilettantes".[8] To define the term, "the dilettante would be content to point out the existence of competing forces" faced with a range of policy questions, such as whether exchange rate depreciation will stimulate an economy. In what looks like a deliberate attempt to obfuscate in their Section 2.1 they suggest: "as a practical matter, people often use the term DSGE models to refer to quantitative models of growth or business cycle fluctuations". This completely misses the point made by Blanchard (2018)[9] that DSGE models (with their heavy restrictions) should not crowd out research into alternative quantitative models that allow data to speak more freely. Many central banks now use or are developing estimated econometric general equilibrium models that aim to capture the key dynamics and feedback mechanisms operating in real economies. To label such models as 'DSGE' can only mislead the uninitiated.

Christiano et al. begin by reviewing DSGE models with their origins and extensions, which attempts to bridge the chasm between their economic assumptions and the real world. Reviewing RBC models and their failings, it is acknowledged that micro-data cast doubt on key assumptions, including perfect credit and insurance markets. New Keynesian DSGE models are briefly introduced, which added nominal frictions in labour and good markets to the RBC model to be able to answer questions such as what are the consequences for output of a monetary policy shock. The Christiano et al. (2005) NK-DSGE model is explained, which has much in common with that of Smets and Wouters (2003, 2007). Christiano et al. acknowledge the representative agent and complete asset market assumptions; they admit that Calvo-style pricing (Calvo, 1983) contradicts aspects of micro-data and makes sense only in moderate inflation environments. In discussing the consumption Euler equation and the need to assume habits to obtain data-coherent responses to interest rate changes, they admit that it has been known for decades that the Euler equation is rejected by macro-data[10], even with habit formation. They defend its use because the representative agent NK-DSGE model gives "roughly the right answer" to the question of how consumption responds to a cut in interest rates: "a useful reduced-form way of capturing the implications of …more realistic, micro-

[8] The sentence is absent in a more recent version of the paper.
[9] This article was preceded by widely-read blogs, Blanchard (2016, 2017).
[10] At the 2017 Oxford-NYRB conference, Christiano referred to the Euler equation as "the most rejected equation in economics".

founded models". However, if reduced-form methods *are permitted* by the hard-line DSGE school, there are surely better ways of learning from the data, see Section 3 below.

They then attempt to convince that financial frictions were adequately dealt with by some of these models. They castigate Stiglitz (2018) amongst others who "…asserted that pre-crisis DSGE models did not allow for financial frictions, liquidity-constrained consumers, or a housing sector". They cite selected counter-examples, including models with two types of consumers: those both unconstrained by liquidity and rational and forward-looking, and those who are credit–constrained and just spend income. But micro-theory does not say that credit-rationed consumers only spend income; this assertion is not micro-founded, and the example is inadequate to try to justify financial frictions in these models. This kind of ad hocery has long been anathema to the developers of the micro-founded buffer stock model, Deaton and Carroll.

Christiano et al. refer also to DSGE models with firm-based credit market frictions introduced by Bernanke et al. (1999) and incorporated in Christiano et al. (2003). However, for reasons of tractability, these are 'Mickey Mouse' *single*-period financial frictions. They cannot adequately capture reality when debt contracts are typically multi-period and bankruptcy and defaults have major scarring effects on economies. More damagingly, these DSGE models are almost always linearised around steady states, which assume for this Utopian world that major financial crises *cannot occur*.

Finally, they appeal to the introduction of a housing market and a financial friction into a DSGE model by Iacoviello (2005), later calibrated to US data in Iacoviello and Neri (2010). There is no denying that this widely-cited work is an technical *tour de force*. However, these extended DSGE models actually illustrate the weaknesses of representative agent, efficient market rational expectations models. They assume two representative households, patient and impatient, and present in a fixed proportion. Patient households apply a loan-to-value constraint when offering mortgage loans to the impatient households, which is a kind of financial friction. However, given their assumption of infinitely-lived or dynastic households, saving for a down-payment, one of the most important saving motives in industrial countries, is omitted from the model. Their closed economy model lacks banks and housing foreclosures, and assumes a frictionless and efficient housing market; the transmission and amplification of monetary or other shocks via housing is therefore *extremely* limited. Their model implies that aggregate home equity withdrawal (i.e. the excess of households' mortgage borrowing over acquisitions of housing), is always negative. However in practice, US home equity withdrawal was *strongly*

positive for much of 2001 to 2006, and in the peak quarters, it was of the order of 10 per cent of that quarter's aggregate household income.

This important fact and the realised foreclosures were not in the set of salient data chosen by Iacoviello and Neri for their model calibration. For their calibrated model, they compare the correlation between consumption growth and house price growth, with and without the financial friction. Without the friction, the correlation is 0.099, the result of the common influence of the shocks on house prices and consumption. With the friction, the correlation rises to 0.123. The difference is tiny and the implication is that the much-lauded financial friction hardly matters. Even more troubling is the assumption in these models that there are random changes in peoples' tastes which made them strongly prefer housing compared to their taste in the past in order to explain the house price boom of the 2000s. Since there are no credit shocks in the model, these *preference shocks* are the main variables that can explain movements in house prices. Unsurprisingly the house price dynamics in these models poorly capture the persistence and volatility in the *actual* data. To treat preference shocks as a euphemism for credit market shocks, including the financial innovations and credit crunches induced by bad loans, distracts seriously from understanding the real world.

After the partisan review, Christiano et al. discuss the estimation and evaluation of DSGE models. One strategy is to minimise the distance between a subset of model-implied second moments and their analogues in the data, or minimising the distance between model and data impulses to economic shocks, using partially identified VAR models. The choice of which moments to match is subjective, and analogous to the problem of omission of salient data in the Iacoviello-Neri model just discussed. A second problem is that many VAR models are grossly mis-specified by omitting longer lags[11], creating the delusion that estimated DSGE parameters are data-coherent, when they also fail to capture longer lags. The Bayesian methods they recommend for DSGE models therefore suffer from similar problems to the Bayesian VARs. Both, as noted earlier, are far from immune to structural breaks in the data and the Hendry-Mizon critique (Hendry and Mizon, 2014).

Stiglitz is right to criticise the use of the Hodrick-Prescott filter to de-trend the data which was common before the financial crisis. The crisis made its weaknesses obvious, though authors such as Canova had earlier pointed to the problem of sweeping "medium-term business cycles into the trend". Christiano et al. cite one example of a pre-crisis paper that did not fall into this trap. Stiglitz, citing Korinek (2017), is surely right to point to the subjectivity involved

[11] I argue this point in more detail in Sections 4 and 5 below.

in choosing which moments to match, when that is the method used for estimation. Stiglitz argues, citing Korinek that "for a given set of moments, there is no well-defined statistic to measure goodness of fit of a DSGE model". Of course, classical maximum likelihood would provide such a measure but, quoting Christiano et al. themselves: "the data used for estimation is relatively uninformative about the value of some of the parameters in DSGE models". Bayesian methods used to deal with this rely on subjectivity. The need for tractability means the choice of variables to be modelled is typically quite narrow and also subjective, usually excluding credit aggregates, asset prices and balance sheets. Stiglitz's last point is that "DSGE models frequently impose a number of restrictions that are in direct conflict with micro evidence". The defence that all models have such inconsistencies hardly washes when those inconsistencies are often gross, in particular contradicting evidence supporting the insights of the information economics revolution.

Christiano et al. ask why DSGE models failed to predict the financial crisis, though almost all quantitative models failed. They admit that pre-crisis models did not sufficiently emphasise financial frictions, noting that those included in DSGE models, e.g. those based on Bernanke et al. (1999) had only modest quantitative effects. Some of the reasons for this were discussed above and overlap with the same issues faced by Iacoviello (2005). They consider the post-crisis developments in the DSGE literature, starting with the introduction of more serious financial frictions. One example, allowing a roll-over crisis in the shadow-banking sector, is Gertler et al. (2016), though this is more of a calibrated theory model designed to understand a mechanism than a general policy model. Such models have an important function. A second example is Christiano et al. (2014) which introduces a time-varying variance of technology shocks amplified by a Bernanke-Gertler-Gilchrist-style financial friction and estimated on data including equity prices and interest rate spreads. They conclude that their second-moment shock substantially reduces the implausible pre-crisis attribution of most of business cycle variation in GDP to technology shocks. However, it is doubtful if any plausible model of the US economy can exclude the housing market and the structural changes in credit markets, discussed in section 3 below. Fortunately, there is now a spate of research, much of it framed in terms of heterogeneous agents, taking housing and the mortgage market more seriously, including Hedlund et al. (2016), of which more below.

They turn to the zero lower bound (ZLB) and other non-linearities. They note that DSGE models with enough frictions can generate strong fiscal policy effectiveness when the ZLB holds and discuss solution methods that do not rely on the log-linear approximations criticised by Stiglitz and many others. They complain that very recent DSGE literature deals

with such criticisms and Stiglitz needs to catch up on his reading. They also criticise his comment "the inability of the DSGE model to…provide policy guidance on how to deal with the consequences [of the crisis], precipitated current dissatisfaction with the model". However, while the DSGE literature on the ZLB gives the right advice on fiscal policy effectiveness, it remains silent on bank rescues, easing constraints in the mortgage market and helping borrowers with mortgage payment difficulties. In practice, rescues of financial institutions and QE or 'credit easing', particularly purchases of agency bonds secured on housing collateral, proved important. The Federal Home Administration system, in providing publicly-backed mortgage credit to partially compensate for the almost complete collapse of the private mortgage-backed securities market, played a pivotal role in stabilising the US economy. These facts are not mentioned by Christiano et al. (2018).

New developments of heterogeneous agent models at last begin to incorporate key insights from the asymmetric information revolution to which Stiglitz was the key contributor. It is clear that Stiglitz's critique was focused on representative agent NK-DSGE models, and he surely must welcome this new generation of heterogeneous agent models. Nevertheless, these models are still at the stage of being calibrated theory models designed to explain key mechanisms. For example, the important Heterogeneous Agent New Keynesian (HANK) paper by Kaplan et al. (2018) gives major insights into monetary transmission without relying on implausible substitution effects of the representative agent NK-DSGE model, see the discussion in Kaplan and Violante (2018). However, housing is treated like an illiquid financial asset subject to major trading costs and credit constraints and with an exogenous price. This therefore omits the important monetary transmission channel that exists through the US housing market. Hedlund et al. (2016) are making progress designing a model with more realistic housing market features, including matching frictions and sticky prices.

Christiano et al. discuss how DSGE models are used at policy institutions, but fail to mention that the non-DSGE FRB-US model continues to be heavily used (despite its defects discussed in the next section), and that many including the Bank of Canada, ECB, Dutch National Bank, and the Australian RBA[12] have developed or are developing non-DSGE policy models.

[12] The paper (Cusbert and Kendall, 2018) introducing the new model says: "A weakness of DSGE models is that they often do not fit the data as well as other models, and the causal mechanisms do not always correspond to how economists and policymakers think the economy really works. In order to more easily manage these models, they typically focus on only a few key variables, which can limit the range of situations where they are useful. The key strength of full-system econometric models like MARTIN is that they are flexible enough to incorporate the causal mechanisms that policymakers believe are important and fit the observable relationships in the data reasonably well. They can also be applied very broadly to model a wide range of variables. This flexibility reflects that the

In their conclusion, Christiano et al. mention "exciting work on deviations from conventional rational expectations (including) social learning, adaptive learning and relaxing the assumptions of common knowledge" as part of the organic development of the DSGE enterprise. One can only celebrate such developments. Without doubt, much useful and creative work is being done with structural models to better understand particular mechanisms. Whether embedding them in a necessarily over-simplified full general equilibrium setting is *always* helpful is questionable.[13] There are, of course, fine examples where it *is* necessary, for example in Brunnermeier and Sannikov (2014). In their general equilibrium model, which includes a banking sector, low fundamental risk leads to higher equilibrium leverage: low risk environments are conducive to a greater build-up of systemic risk, arguably relevant to the Great Moderation period from the mid-1980s to 2006. There have been other major advances in post-crisis general equilibrium macro theory but translation into usable policy models for central banks has not yet occurred.

Carlin and Soskice (2018) set out a stylised macroeconomic model with two well-defined solutions, one Keynesian and one Wicksellian, and a graphical tool-kit which aids comprehension. It includes some of the essential ingredients with which a large econometric policy model should be consistent and therefore capable of providing an explanation of the slowness of the recovery since the global financial crisis. Their Keynesian unemployment equilibrium is underpinned by five assumptions: a zero bound to interest rates; the absence of disinflation in the presence of high unemployment; strategic complementarities among investors capable of giving rise to multiple equilibria; the assumption that technical progress is embodied in investment so that a low-investment outcome will give rise to a low rate of technical progress and finally, agents who discount expected future income at higher rates than assumed by the textbook model, which as noted above has profound implications for the effectiveness of monetary and fiscal policy. In chapter 6 of Carlin and Soskice (2015), they introduce a leveraged banking sector and hence the possibility of a destabilizing financial cycle, see section 3 below, that can take the economy away from the Wicksellian equilibrium.

Finally, given heterogeneity, the DSGE programme is far from offering the only way forward. Haldane and Turrell (2018) make a strong case for agent-based modelling, which is

model is not derived from a single theoretical framework, which can make causal mechanisms less clear than in DSGE models."

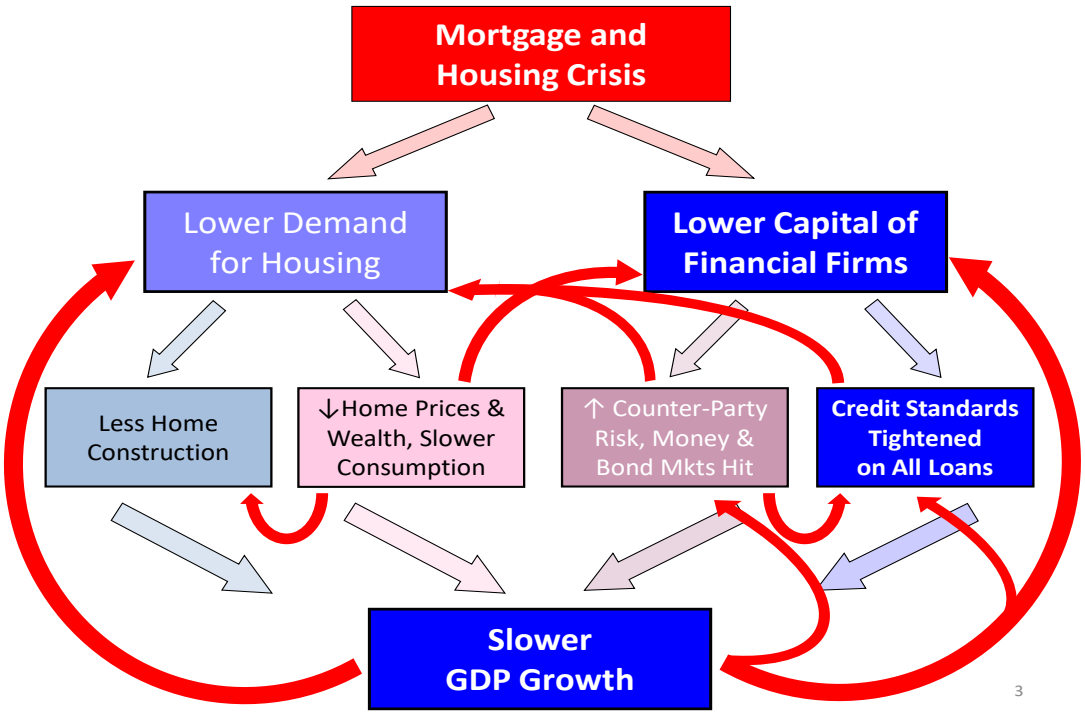
[13] For example, the structural heterogeneous agent model of leverage and mortgage foreclosures of Corbae and Quintin (2015) assumes exogenous house prices. Its admirably crafted balance between realism concerning the banking sector and heterogeneity of borrowers and tractability could not have been achieved had it also been required to supply a full general equilibrium treatment.

firmly founded on micro-data. Economic agents in these models are assumed to use heuristic behavioural rules given the limited information they face, rather than assuming that they can successfully and continuously accomplish the heroic task of solving complex intertemporal optimising problems.

3. Financial stability: lessons from modelling consumption, debt and house prices

The US sub-prime crisis began with a serious problem of *over-valuation* of asset prices, especially of housing. The different types of over-valuation are discussed below. The consequence of overvaluation, eventually is falling house prices. Falling house prices reduce residential investment and lower consumer spending in the US, where housing collateral is an important driver of consumption. These are the two left channels shown in Figure 1, illustrating the feedback loops in the US sub-prime crisis. Falling house prices increased bad loans and lowered the capital of financial firms. This raised risk spreads in credit markets and impaired the ability of banks to extend credit, shown in the two right channels of Figure 1. This fed back on residential investment and household spending, as shown in the figure, increasing unemployment and reducing GDP, and this, in turn, fed back further to reduce the demand for housing, shown in the extreme left arrow in the figure, and the capital of financial firms, shown in the extreme right arrow.

Such feedback loops involve non-linearities as well as amplification. For example, a fall in house prices that drives up the incidence of negative equity, can, via bad loans, cause a *sharper* contraction in credit availability than the expansion of credit availability caused by an equivalent rise in house prices.



Source: John Duca, from Duca and Muellbauer, 2013

Figure 1: The financial accelerator in the US sub-prime crisis

One can classify the causes of overvaluation of asset prices into three broad groups: exogenous macroeconomic shocks, fragile fundamentals and endogenous dynamic processes, see Muellbauer (2012) for discussion.

Exogenous negative macroeconomic shocks to economic fundamentals are one reason why, with hindsight, house prices can be seen to have been overvalued. Such shocks can include a deterioration in the terms of trade, a rise in oil prices for net oil importers (e.g. in the 1970s), a collapse of export markets (e.g. Finland just after the collapse of the Soviet Union), a rise in global interest rates for a small open economy,[14] external credit supply shocks for small open economies, political risk, and natural disasters. Such shocks are arguably close to unforeseeable.

A dislocation of economic fundamentals, which historical experience *should* have flagged as increasingly fragile, is a second category promoting overvaluation of prices. Examples are: duration mis-match in credit supply (e.g. mortgage funding in Ireland and the UK disrupted by the money-market sudden stop in August 2007); currency mis-match of

[14] Examples include the Reagan fiscal shock in the early 1980s and the German unification shock of the early 1990s). Such shocks have particularly strong effects for such economies where floating rate debt dominates.

debt;[15] and unsustainably weak financial regulation (e.g., the misuse of securitisation and fraud in the UK and the US before the global financial crisis).

The third group concerns endogenous, dynamic economic feedbacks, which in some contexts, as in the US sub-prime crisis, can amplify the impact of external shocks. These feedbacks can vary both across countries and over time. The speeds of the different feedbacks can be very important: for example, if an amplifying feedback is large and immediate, but a negative feedback slow but persistent, financial instability is more likely.

The role of endogenous feedbacks

Consider first the negative endogenous feedbacks that can be generated during a house price boom. One instance of such a feedback on prices stems from a build-up in debt levels as the *quality* of loans to households and property developers deteriorates, especially in liberal-credit environments. High debt levels limit spending and access to further credit. A second endogenous, negative and persistent feedback arises if there are large, credit-funded expansions in housing stocks, which then weigh down on house prices. Examples include Ireland, Spain and parts of the US in 2000-2006, where oversupply remained a problem for several years. A third negative feedback via aggregate demand occurs in economies in which high down-payment ratios are required of mortgage borrowers: then saving rates of would-be home-owners increase when house prices rise relative to income. A fourth negative feedback occurs in economies where property taxes are closely linked with very recent market values, dampening aggregate demand and sapping returns from housing investment. Such negative feedbacks would be stabilising if they operated quickly enough or were not overwhelmed by the amplifying feedbacks, that *can* boost the upswing in that phase of the financial cycle, discussed next.

Turning to the amplifying endogenous feedbacks, consider first the role of extrapolative expectations. The role of user cost, which measures the interest cost of borrowing relative to expected appreciation, in the demand for durable goods such as housing has long been studied. There is now much evidence that market participants often tend to extrapolate past rates of appreciation in forming expectations of appreciation to come. Then a series of positive shocks, for example in access to credit or in falling interest rates, increasing the rate of appreciation of house prices, later generates a fall in user cost, increasing the demand for

[15] For example, in the Baltic republics and Hungary in the global financial crisis, and the Asian financial crisis of the late 1990s.

housing and feeding back onto house prices. This was an important ingredient in the over-valuation in US house prices from the mid-2000s according to evidence from Duca et al. (2011, 2016).[16] Since lower down-payment ratios enhance the returns from a housing investment financed by a mortgage, this kind of over-shooting in house price dynamics is likely to be greater in countries where low down-payment ratios are prevalent and time varying as loan conditions are eased.[17] Such amplifying feedbacks also exacerbate down-turns of the cycle after prices have started falling.

Though Figure 1 shows consequences of negative shocks on an overvalued housing market, it also reflects similar, but opposite signed, feedbacks operating in the *upswing* of a housing market.[18] A second amplifying feedback potentially comes from the mechanisms illustrated on the right hand side of Figure 1. Higher expected appreciation can make lenders more keen to lend as borrowers would have more equity against their mortgage, making the loan safer from the lender's point of view. Moreover, in a rising market, lending is more profitable and previous bad loans shrink, enhancing the capital of financial firms. As Geanakoplos (2010) has argued, an endogenous leverage cycle can simultaneously drive growth in debt and house prices.

A third amplifying feedback, illustrated in the extreme left channel of Figure 1 comes with an increase in residential investment, which boosts employment and household income, and therefore aggregate demand, including demand for housing. In countries where planning constraints are severe or the planning process slow, this short-term feedback is likely to be smaller, though the impact of demand shocks on house prices in the presence of inelastic supply is greater.

A fourth, but also *far from universal*, amplifying feedback comes from the consumption channel feeding into aggregate demand, illustrated in the second from left channel in Figure 1. This tends to be greater where down-payment constraints are loose, i.e. household leverage higher, access to home equity loans is easy and rates of owner-occupation are high, as in the US. Research on consumption functions to check for balance sheet effects, including from housing, is helpful in establishing in which countries amplifying feedback loops are more likely, see Hendry and Muellbauer (2018) for a discussion. Thus, there *can* be pronounced

[16] Regular quarterly surveys of the housing-price expectations of potential housing-market participants should help assess overshooting linked to extrapolative expectations, and Case and Shiller's expectations surveys led to their real-time judgment of overvaluation in the US in the mid-2000s.
[17] Empirical evidence for this point comes from Muellbauer and Murphy (1996) and Chauvin and Muellbauer (2018).
[18] Evidence in Duca et al (2016) suggests an asymmetry, with stronger effects of falling house prices on the contraction of credit supply, than of rising prices on its expansion.

overshooting of house prices induced by a series of strong positive shocks, amplified by the four mechanisms just discussed.

As explained above, a house price boom *can* also generate some negative feedbacks which would be stabilising if they acted quickly enough. If they do not, their very perisistence can then create a double whammy of a crisis, when combined with the quickly acting feedbacks that amplify house prices falls in the downturn.

The role of leverage at the level of households in these feedback mechanisms was noted. The more a financial system permits or fosters high leverage at financial firms, the more likely is it that it will also be high at the level of households or non-financial firms. For financial firms, high leverage increases risks, particularly those arising from sizable overvaluation of property prices, given the important role of real estate collateral for lending. A factor that increased leverage at financial firms was the shift in governance within large investment banks, mainly in the 1980s, from partnerships, where managers were owners who retained substantial 'skin in the game,' to public corporations where managers had incentives to design asymmetric contracts for their private gain. Duca et al. (2016) attribute the rise in loan-to-value ratios for first-time buyers in the US in the 2000s to leverage-increasing interventions: the 2000 Financial Futures Modernisation Act[19], lower bank capital requirements on 'investment grade' nonprime mortgage backed securities and the 2004 Securities and Exchange Commission decision to loosen leverage restrictions on investment and other banks. Tendencies for excess debt leverage can also be exacerbated by tax systems which incentivise borrowing (e.g. through tax relief on mortgage payments available for owner-occupiers in the U.S. and the Netherlands, though not in Canada or Australia) and legal frameworks that protect borrowers with limited or no-recourse loan contracts (as is still the case in many U.S. states, but rare elsewhere).

Macro-evidence has accumulated for the role of leverage and of real estate connected financial instability (Cerutti et al. (2017) and Mian et al. (2017)). Mian and Sufi (2014) have provided extensive microeconomic evidence for the role of credit shifts in the US sub-prime crisis and the constraining effect of high household debt levels. Turner (2015) analyses the role of debt internationally with more general mechanisms, as well as in explaining the poor recovery from the global financial crisis. Jordà et al. (2016) have drawn attention to the increasing role of real estate collateral in bank lending in most advanced countries and in financial crises. The IMF's October 2017 Financial Stability Report provides further evidence,

[19] This gave derivatives first priority in claims on a company's assets ahead of other claimants, thus encouraging the use of derivatives to back the expansion of mortgage-backed securities.

highlighting the critical role of mortgage debt and non-linearity, finding more pronounced effects at high debt ratios, and larger effects in countries with open capital accounts, fixed exchange rate regimes, less transparent credit registries (information), and less strict financial supervision. The IMF also found that easy monetary policy during a credit boom likely exacerbated the subsequent down-turn when booms turn into busts.

Implications for econometric policy models

For policy models to be useful in allowing for the mechanisms discussed above, they need well-specified household sector equations, including for consumption, mortgage debt and house prices, a well-specified residential investment equation and a linkage between the financial sector and credit availability for the household and investment sectors. As noted in Section 2, New Keynesian DSGE models, omitted debt and household balance sheets, including housing, together with shifts in credit availability, crucial for understanding consumption and macroeconomic fluctuations. The US Federal Reserve did not abandon its large non-DSGE econometric policy model FRB-US. However, it too was defective in that it also relied on the representative agent permanent income hypothesis for the major part of consumption[20], which ignored shifts in credit constraints and mistakenly lumped all elements of household balance sheets, debt, liquid assets, illiquid financial assets (including pension assets), and housing wealth into a single net worth measure of wealth.

This is wrong for the following reasons. First, housing is a consumption good as well as an asset, so consumption responds differently to a rise in housing wealth than to an increase in financial wealth, Aron et al. (2012). Second, different assets have different degrees of 'spendability'. It is indisputable that cash is more spendable than pension or stock market wealth, the latter subject to asset price uncertainty and access restrictions or trading costs. This suggests estimating separate marginal propensities to spend out of liquid and illiquid financial assets. Third, the marginal effect of debt on spending is unlikely just to be minus that of either illiquid financial or housing wealth. The reason is that debt is not subject to price uncertainty and it has long-term servicing and default risk implications, with typically highly adverse consequences, disproportionately affecting the most leveraged households.

There is now strong micro evidence that the effect of housing wealth on consumption, where it exists, is much more of a collateral effect than a wealth effect, see Browning et al.

[20] It allowed a fraction of households just to spend income.

(2013), Mian et al. (2013), Windsor et al. (2015), Mian and Sufi (2016) and Burrows (2018). As mortgage credit constraints vary over time, this contradicts the time-invariant housing wealth effect embodied in FRB-US.

Of structural changes, the evolution and revolution of credit market architecture is often the single most important. In the US, credit card ownership and instalment credit spread between the 1960s and the 2000s. The government-sponsored enterprises—Fannie Mae and Freddie Mac—were recast after 1968 to underwrite mortgages. Interest rate ceilings were lifted in the early 1980s. Falling IT costs transformed payment and credit screening systems in the 1980s and 1990s. As the discussion of factors permitting increased leverage in the US showed, there were major shifts in credit availability in the late 1990s and early 2000s. These shifts occurred in the political context of pressure to extend credit to the poor.

Given the lessons of the information revolution and the work on liquidity constraints of Deaton and Carroll, it is clear that the text-book micro-foundations of the standard life-cycle/permanent income model do not stand up. Using the log-linear approximation as in Muellbauer and Lattimore (1995), the text-book model takes the form

$$\ln(c_t/y_t)= \alpha_0+ \ln (y_t^p/y_t)+ \gamma A_{t-1}/y_t \qquad (1)$$

where c is consumption, y is non-property income, y^p is permanent non-property income using a discount rate equal to the real rate of interest, and A is net worth. The marginal propensity to spend out of net worth is γ. If one is unsure about the theoretical foundations, the 'encompassing principle', see Hendry and Muellbauer (2018), p.313, suggests relaxing and testing the restrictions implied by a model. Thus, the asset to income ratio can be split into the main components, e.g. liquid assets, debt, illiquid financial wealth and housing wealth; the coefficient on the log ratio of permanent to current income can be freely estimated instead of being imposed at one, and a higher average discount rate checked, as implied by Deaton and Carroll's work on buffer-stock saving; the intercept can be allowed to vary with access to credit since this would affect the saving for a down-payment motive, the size of which might also depend on house prices relative to income; and the marginal propensity of housing wealth can be allowed to shift with access to borrowing on home equity. In a series of papers, my co-authors and I have found support for these generalisations and major differences between economies in the connection between house prices or housing wealth and consumption. For Japan, Germany, France and even Canada, there appears to be a negative effect of higher house prices/income on consumption, though there is also a small housing wealth effect in France. In all these countries apart from Japan, credit liberalisation for households has increased the

consumption to income ratio, though Germany had only a quite modest degree of liberalisation. In contrast, for the US, UK, Australia and South Africa, the marginal propensity to spend out of housing is positive and varies strongly with access to credit. In all cases, the marginal propensity to spend out of liquid assets is higher than out of illiquid assets, and debt has a far more powerful negative effect on consumption than implied by the net worth restriction.

These papers are examples of the looser, more relevant, application of theory. In contrast to the FRB-US consumption function which incorporates no shifts in credit constraints and aggregates the household balance sheet into a single net worth concept, contradicted by micro evidence, it no longer corresponds to a representative agent optimizing model. The FRB-US model followed Muellbauer and Lattimore (1995) in assuming two types of agents, one following a life-cycle model on the lines of equation (1), albeit with a higher risk-adjusted discount rate to compute permanent income, and the other simply spending income. However, it disregarded our evidence that the marginal propensity to consume (*mpc*) is higher for liquid assets and that the *mpc* for debt is large and negative, and our theoretical explanation of why the housing wealth effect is different from a financial wealth effect (Muellbauer and Lattimore, 1995, p.268-271). It took no account of our arguments that "Financial liberalization, by making asset backed credit more easily available, made these illiquid assets more spendable" (p.281) and "….improved access to… home loans, and reduced down payment to house price ratios" (p.289).

The claimed micro-foundations of the FRB-US consumption function do not save it from parameter instability: the estimated speed of adjustment for data up to 2009 of 0.19 falls to 0.10 for recent data. This is clear evidence against treating the FRB-US consumption function as a 'structural' equation in the classical sense of invariant to shifts in the economic environment.

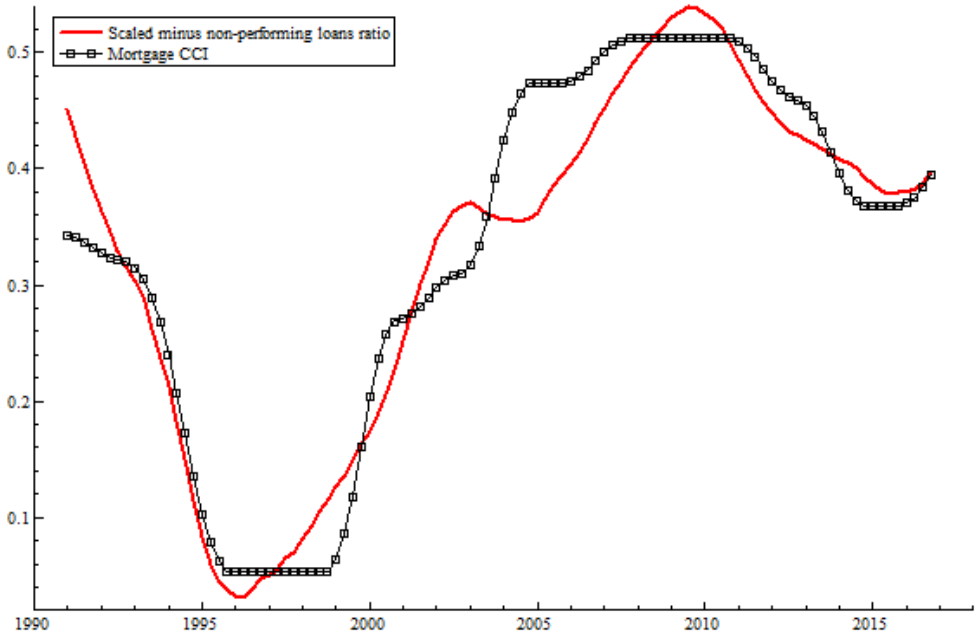
As a result of its omissions, the FRB-US model failed to give proper warning of risks faced by the US economy after 2007. At the Jackson Hole conference in 2007, Mishkin (2008) reported the results of FRB-US simulations of a 20 per cent decline in real house prices spread over 2007–8. The standard version of the model simulated GDP lower than the baseline by 0.25 per cent in early 2009 and consumption lower by only 0.6 per cent in late 2009 and 2010. The simulations suggested a rapid recovery of residential investment given the lowering of the policy rate in response to the slowing economy. FRB-US failed to include a plausible model of house prices and so also missed the feedback from the credit crunch back on to house prices modelled in Duca et al. (2011, 2016). Consistent with this time series evidence, Favara and

Imbs (2015) and Anundsen and Heeboll (2016) provide strong micro-evidence for the causal link between credit supply and house prices in the US.

The LIVES approach to modelling the household sector

A number of our papers use the 'latent interactive variable equation system (LIVES)' set out in Duca and Muellbauer (2013), with the fullest application in the six-equation system in Geiger et al. (2016) and Chauvin and Muellbauer (2018). Consumption, consumer credit, housing loans, house prices, liquid assets and permanent income are jointly modelled with two latent variables, representing shifts in consumer credit and mortgage credit availability, common to a number of the equations. This model takes to the macro data what Mian and Sufi (2018) call the 'credit-driven household demand channel' and quantifies the role of household balance sheets in the financial accelerator, emphasised by Gertler and Gilchrist (2018). Including a measure for permanent income, a measure of households' income growth expectations, is important, providing protection against the Lucas critique. Section 5 below will say more about the practicalities of modelling it.

Mortgage credit conditions help drive house prices, housing loans and consumption in France, though demography is also important. Without controlling for mortgage credit conditions, it is impossible to obtain coherent estimates of a house price equation for France. The latent variables can be estimated using state space methods or spline functions. Because they represent any joint drivers of the three variables not otherwise controlled for, there can be doubts about whether they reflect specifically shifts in credit conditions. However, the mortgage credit conditions index for France in Chauvin and Muellbauer (2018) turns up strongly from 1984, when widely documented financial liberalisation began in France. And after 1990, it is strongly negatively correlated with the ratio of non-performing loans, particularly when credit availability contracts, as the following Figure demonstrates.



Source : Chauvin and Muellbauer (2018)

Figure 2: Scaled negative non-performing loan ratio (8-quarter moving average, lagged 2 quarters) and estimated mortgage credit conditions index.

The six-equation model is highly relevant for thinking about potential risks for financial stability in France from the housing-credit nexus. The consumption estimates for France suggest that in the house price boom between 1996 and 2008, the positive effect on the ratio of consumption to income of higher housing wealth relative to income, a small but positive housing wealth effect, and looser mortgage credit conditions, was largely offset by the negative effect on consumption of higher house prices and higher debt relative to income. France, like Germany where the negative effect of higher house prices to income is even larger, see Geiger et al. (2016), is therefore very different from the Anglo-Saxon economies where home equity loans produced large collateral effects of housing wealth on consumption. As a result, despite higher house prices, France did not experience an Anglo-Saxon-style consumption boom in which the financial accelerator via home equity loans proved powerful and destabilising. Moreover, the induced increase in household debt will weigh negatively on future consumption.

Extrapolative expectations of capital gains, which enter user cost, a driver of demand for housing and hence of house prices, are potentially a powerful endogenous source of house price over-valuations. They were an important factor in the US boom of the 2000s, see Duca et al. (2011, 2016), and probably contributed to excess credit growth. The scale of extrapolative

expectations in France was moderate even at the height of the French boom, as shown in the estimated user cost contribution in Chauvin and Muellbauer (2018). Since higher leverage amplifies returns from house price appreciation, the moderate contribution of such expectations is probably the result of the far lower level of leverage permitted to households by French financial regulators.

Our six-equation model does not endogenise credit conditions, but Figure 2 suggests there would strong potential in endogenising the NPL ratio of the banking system, data permitting, to quantify the link between the household and banking sectors. Moreover, the substantial lag between the NPL ratio and the mortgage credit conditions index, implies that in real time, early warnings would be flagged up well before credit conditions turned down, with negative consequences for house prices and consumption. A top-down macro approach needs to be integrated[21] with micro evidence on potential household vulnerabilities and individual bank stress tests data to better tune macro-prudential policies, see Constâncio (2017a, 2018a). Improving the quality of the top-down approach, taking proper account of institutional differences between countries as seen in France compared to Anglo-Saxon countries, would make an important contribution to this endeavour.

It is sometimes argued that the global financial crisis was such a rare event that there is little to be gained in more normal times for building mechanisms into models that trace how such a crisis might affect the household sector. However, not only can such risks not be precluded, but better models of the household and housing sectors throw important light on monetary policy transmission in more normal business cycle fluctuations and on potential obstacles to a strong recovery from high levels of debt and changing demography. They also illuminate potential risks for and via the household sector from other sources, such as a rise in global interest rates and/or a substantial fall in equity prices, see Constâncio (2018a).

4. Lessons from forecasting US core inflation

Forecasting is one area of applied economics where practitioners can escape the constraints of conventional practice. The true test of a useful model is its out-of-sample forecast performance, regardless of convention - at least in the absence of major structural breaks, when all forecasting models break down (Clements and Hendry, 2011a). Central banks are perennially interested in

[21] Constâncio (2017b) says: "Stress tests of the banking and financial system must not be limited to microprudential supervision but need to be embedded in a macro-financial environment and take a macroprudential dimension."

forecasting inflation, but especially in 2018, when the US economy may be at a turning point. Aron and Muellbauer (2013) is a forecasting paper that challenged the conventional wisdom, and demonstrated substantial out-performance in forecasting US inflation, including against the unobserved components-stochastic volatility model of Stock and Watson (2007). Forecasting models were estimated for the 12-month-ahead US rate of inflation, measured by the chain-weighted consumer expenditure deflator, for 1974–98 and subsequent pseudo out of-sample forecasting performance examined. Alternative forecasting approaches for different information sets were compared with several benchmark univariate autoregressive models. Three key ingredients to the out-performance were: including equilibrium correction component terms in relative prices[22]; introducing nonlinearities to proxy pre-1983 state-dependence in the inflation process and using a 'parsimonious longer lags' parameterization to permit long lags without running into the curse of dimensionality. It was established that applying the standard Bayesian information criterion, commonly used in unrestricted VARs to select lag length, results in throwing away highly relevant longer information. This was a remarkably robust finding, true for each of seven information sets considered. In common with much forecasting literature, it was also concluded that forecast pooling or averaging improves forecast performance.

A paper currently in progress uses similar methods to develop forecasting models for core inflation[23], defined by the Federal Reserve's favourite measure, the PCE deflator, excluding food and energy (Aron and Muellbauer, 2018). These new models for 12-month-ahead core inflation demonstrate similar outperformance against benchmark univariate autoregressive models as those in the 2013 paper. In the earlier paper, three relative price terms were highly important over the 12-month forecasting horizon: two measures of domestic costs relative to the PCE deflator, namely unit labour costs and house prices (a key drivers of rents), and a measure of foreign prices and the exchange rate embedded in the real exchange rate, also a relative price.[24] The hypothesis is that, in a long-run equilibrium, the PCE deflator is a function of unit labour costs, house prices (as a proxy for rental costs) and foreign prices, both

[22] In Sargan (1964), on wages and prices in the UK, the equilibrium correction mechanism that underpins much macroeconomic empirical research was introduced. Interestingly, its first application was also to modelling inflation. Hendry (2001) develops this equilibrium correction approach to modelling annual data on the UK GDP deflator back to 1875.

[23] See Yellen (2017) for an excellent and comprehensive discussion of controversies and uncertainties around the drivers of core inflation.

[24] This has close parallels with Hendry (2001) for whom the long-run solution for the UK price level depends on unit labour costs, foreign prices and commodity prices. Given the UK's small open economy, the weight on foreign prices is larger than for the US. Allowing for structural breaks, e.g. because of wars, this long-run solution is remarkably stable for long periods.

of imported raw materials and of final consumption goods. Starting from an equilibrium position, if one of these changes, the equilibrium is disturbed and gradual adjustment to a new equilibrium occurs. Since this takes time, these 'equilibrium correction terms' account for a good deal of observed inflation persistence. Lags in the inflation process are likely to be long and complex for at least three reasons. One is that inflation expectations in the form of private sector forecasts of inflation based on past data are likely to be an element in price setting. A second is that house prices feed gradually into rents, given the preponderance of 6 and 12 month rental contracts. A third is that in a multi-sector economy, similar equilibrium adjustment processes will be occurring within and between sectors, for example, related to the input-output structure of the economy as argued by Huang and Liu (2000).

The need for long and complex lags in forecasting inflation is even more apparent with the US core price index, at forecasting horizons of 3, 6, 12 and 24 months. Conventional wisdom among central bankers suggests that there is little information in economic data relevant for forecasting 24 months ahead, beyond the recent inflation history. By contrast, Aron and Muellbauer (2018) show that because real world lags are rather longer than most economists have assumed[25], there is relevant information for forecasting US core inflation two years ahead.

A relationship between unit labour costs and unemployment allows the derivation of a strong empirical relationship, given two other controls, between the US core price index and unemployment, relative import prices (alternatively, the real exchange rate) and house prices.[26] The two other controls are first, a measure of pricing power[27] based on the Herfindahl-Hirschman index of concentration of US public companies, and second, the rate of unionisation (as also found relevant in our 2013 paper). There are short-run drivers including the changes in (logs) of import prices, hourly earnings, the nominal exchange rate, trade-weighted foreign prices, oil prices and non-core prices. A parallel model without union density, but with a strong equilibrium correction term to unit labour costs, also gives a reasonable

[25] Gordon (2009, 2013) and in earlier work, and Hendry (2001) are exceptions.
[26] Castle and Hendry (2009) examine very long-run determinant of UK wages, finding a strong but non-linear relationship with the unemployment rate.
[27] Grullon et al. (2017) track increasing concentration in US industries from 1972 to 2014. They show a strong association between profit margins and the degree of concentration. We are grateful to them for making available their measure of concentration based on the 3-digit NAICS industrial classification. Dickson (2006) separates the efficiency and market power effects of higher concentration for a panel of 253 US manufacturing industries covering the years 1963 to 1992. Specifically, for given average cost, higher concentration leads to higher prices, but when average cost are relegated to the error term, higher concentration leads to lower prices. Salinger (1990) provides earlier evidence of a link between price increases and firm concentration in US manufacturing.

account of the data.[28] This points to a strong relationship between unit labour costs and the unemployment rate, controlling for union density.

A challenge faced in the new paper was dealing with the structural break of the global financial crisis. Clements and Hendry (1998, 2011a) have repeatedly argued that the major cause of forecast failure are structural breaks resulting in a mean shift in I(1) variables. The global financial crisis was a huge shock that also caused major shifts in the structure of the US economy, including lower productivity growth. Real GDP in the US and other OECD countries, even in 2018, was everywhere below pre-crisis trends. A post-crisis forecasting model needs to incorporate dummies to allow for the possibility of a shift in the long-run relationship between the relative prices, as well as the temporary disruptions such as those due to inventory liquidations. Post-crisis, forecasting can be resumed, and the comparative pseudo-out-of-sample performance of alternative models evaluated using data up to early 2018.

A key aspect of improving forecasting performance is down to the use of longer lags, and specifically, the 'Parsimonious Longer Lags' (PLL) referred to in various contexts in this paper. As noted in the introduction, while Vector Auto Regression *(*VAR) models aim to preserve generality by not imposing an *a priori* structure on models, (Sims, 1980), they suffer from the 'curse of dimensionality', as increases in lag lengths or in the number of variables included rapidly raise the number of parameters to be estimated. In practice, the gain in generality comes at the cost of restricting the number of variables and lag lengths.

An improved trade-off between these objectives in the context of a VAR can be obtained through the use of PLL which imposes credible restrictions. Since a 'multi-step' model[29] for forecasting inflation, here 12-months-ahead, can be regarded as *one* equation in a reduced form of a VAR, applying the PLL restrictions achieves such an improved trade-off in the forecasting context. The PLL used in the 2013 and 2018 versions of the US forecasting equations takes the following form: for variables in differences, free coefficients are allowed on the first 3 monthly changes; for lags at three to six months, these are restricted to the 3 month change or Δ_3, and to Δ_6 if six to 12 months, and to Δ_{12} if 12 months or longer.[30] The intuition behind these simple and easy to implement restrictions is that precision in the exact

[28] In this version of the model, the effect of the level of the unemployment rate is less strong but the previous 6-month change in the unemployment rate has a strong negative effect.
[29] Multi-step models for inflation forecasting were popularised by Stock and Watson (1999, 2003).
[30] Gordon (2013), and in his earlier work, uses a similar idea to PLL in a quarterly context. In his inflation model, moving averages of lagged inflation at lags of 1, 5, 9, and even longer, summarise the longer run information in the data. His model does not include levels variables such as relative prices but does include a stochastic trend: the deviation of the unemployment rate from this stochastic trend is central to explaining inflation.

timing of impulse responses is harder and harder to achieve, the longer the lag. However, rather than ignoring longer lags, it is better to include them even if the timing is not exactly optimal. Compared to unrestricted lags on changes in variables up to 23 months, 24 parameters are thus replaced by six parameters. Several different information sets of increasing complexity are modelled, and in the most general set of these, the levels of the relative prices (in logs), the unemployment rate, the concentration index and trade union density also appear.

The forecasting performance of the models for core US inflation are contrasted by comparing three different methodologies in terms of reported root mean square forecast errors. Each methodology is applied to the same range of information sets, including at its simplest, a univariate set in core inflation. The first methodology uses an unrestricted standard AR(k+1) or VAR(k+1) specification for inflation rates, with the lag length selected by the Bayesian information criterion.[31] This model is contrasted with a restricted model applying the parsimonious and longer lag (PLL) structure, allowing lags up to 24 months. The third methodology averages the forecasts from the second method with those from a PLL version of the univariate model. Dummies are included to address the global financial crisis for all the models.[32] The models forecast recursively, adding one observation at a time, and re-estimating the model each time from 1982 to generate the 12-month-ahead forecast from 2012:1[33] to 2017:3, with data up to March 2018. This gives 63 observations to compute the root mean squared forecast error (RMSE) for each methodology and each of the range of information sets.

The environment of low and fairly stable inflation during this period makes it particularly challenging to beat the benchmark model. Nevertheless, the finding is that the outperformance of US core inflation models with PLL restrictions applied to the widest of the information sets is broadly similar to the 2013 paper. As in that paper, model averaging improves forecast performance. The RSMEs of the best models using PLL are around 30-35 percent lower than the RMSE of the univariate BIC-selected benchmark inflation model. In almost all information sets there is a considerable loss of forecasting performance from using the BIC criterion applied to conventionally unrestricted lags as compared with the longer lags

[31] It is standard in the VAR and the multi-step forecasting literature based on VARs to use the Akaike or Schwarz information criterion (AIC or BIC) to choose the maximum lag length of the model.

[32] Given a maximum lag of 24 months for the economic variables, impulse dummies for the global financial crisis period from 2008:9 to 2010:9 are included in all models. To forecast 12 months ahead, it is necessary also to include leads up to 12 months of the impulse dummies; together this is equivalent to omitting 36 months around the crisis period. In one information set, we also include a 2008:9 step dummy to allow for a possible step-change in the mean of the process (i.e. a permanent change, say in productivity).

[33] For example, the 12-month ahead forecast for 2012:1 uses parameters estimated up to 2011:1.

allowed with the PLL restriction. Clearly, the application of standard methods throws away a great deal of useful information for forecasting.

This has implications for the forecasting literature. Much effort has gone into dynamic factor models, which combine the information from many *variables* to improve forecasting performance, see Stock and Watson (2011). But little effort has been expended to assess the considerable information content of longer *lags*, which is regrettable in view of the potential forecasting gains. The combination of PLL with dynamic factors has considerable potential, in our view.

Having established the forecasting power of extended information sets with both long-run variables and lags up to two years, a parsimonious form of the model to aid the economic interpretation was sought using model selection following Doornik (2009).[34] The fit of the resulting model is shown in Figure 3.

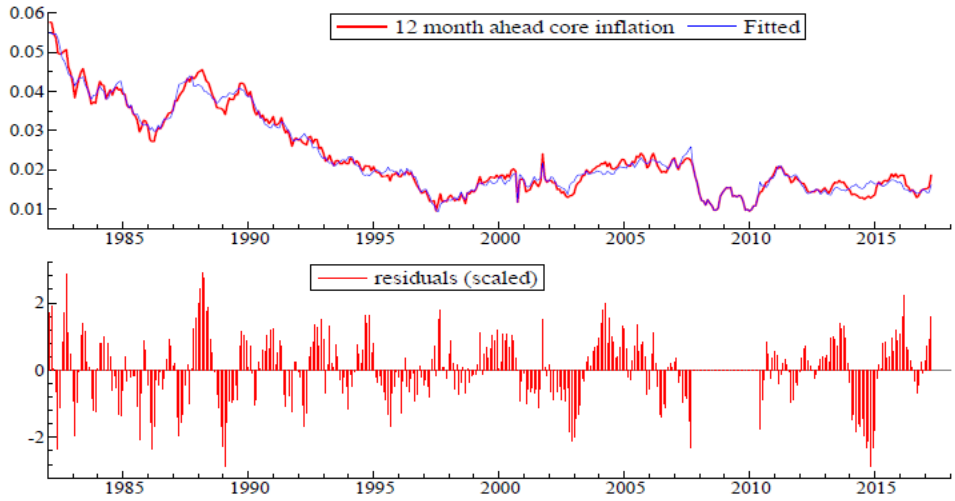


Figure 3: Fitted vs actual values of the 12-month-ahead model of core inflation, and scaled residuals.

[34] The Doornik-Hendry Autometrics software, Doornik (2009), is an objective and easily reproducible tool for parsimonious model selection, not affected by the subjective choices of the modeller. Any other investigator with the same data and the same specification of the 'general unrestricted model' (GUM), will then make the same model selection, given the chosen settings in Autometrics. This software examines a full set of general to simple reduction paths to select a parsimonious form of the GUM to satisfy a set of test criteria. In our context, the overlapping nature of the dependent variable means that residuals will be auto-correlated and so the corresponding tests, including portmanteau tests, are switched off. Heteroscedasticity could be endemic: the corresponding tests are therefore switched off, but we use heteroscedasticity and autocorrelation corrected (HAC) t-ratios and F-tests of stability for model selection. This results in a considerable reduction in the number of parameters, and further reductions are possible, for example by combining two 3-month change variables at t and t-3 into a 6-month change at t.

There is satisfactory stability of the long-run parameters over different samples, given the range of uncertainty provided by robust t-ratios. Figure 4 shows recursive estimates from 1994 to 2017 of the long-term elements. Reading from the top, left to right, the estimates are for the constant term, log (import prices/core index), the unemployment rate, log (house prices/core index), the concentration ratio and union density. The unemployment effect has a robust t-ratio of around -11 at the end of the sample. This constitutes impressive evidence that a version of the Phillips curve is alive and well.[35] When the data are taken back to 1977, although short-run dynamics are somewhat different, the coefficient on the unemployment rate is almost identical.

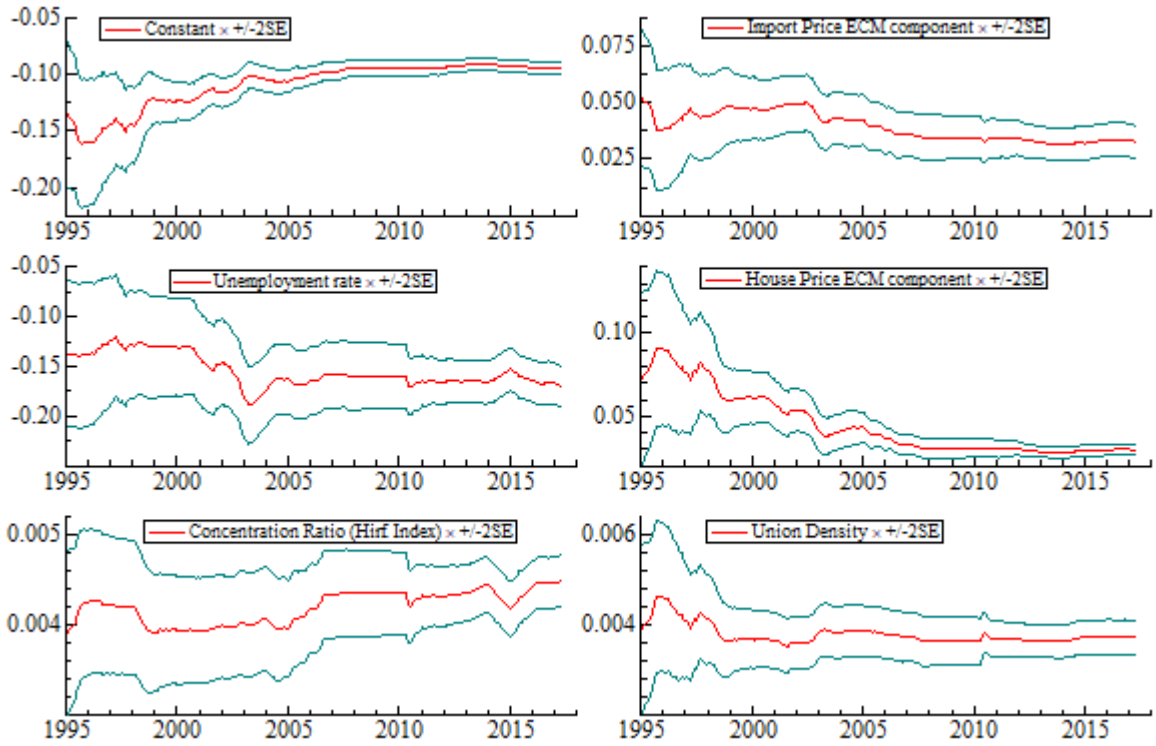


Figure 4: Recursive parameter estimates of the long-run parameters in the core inflation forecasting model

[35] This was also a conclusion reached by Gordon (2013) in a model including a stochastic trend, which might well be reflecting longer run forces such as the decline in trade union power.

Our evidence is that the marginal effect of the unemployment rate on inflation has NOT 'flattened' (fallen): rather, the intercept has shifted. However, it is likely that the marginal effect of the Federal Reserve's policy rate on inflation HAS flattened, and shifted[36]. Our research has established a link from monetary policy to inflation not only via labour market slack but via the exchange rate and house prices. But the latter links are complex and time-dependent and involve other major drivers such as credit liberalisation. For example, section 3 suggested the effect of user cost on house prices depends on time-varying levels of household leverage. Moreover, Duca et al (2011, 2016) estimate an important non-linearity in the effect of user costs so that at low levels, after significant capital gains, the marginal effect of the mortgage rate on house prices is greater, while the marginal effect is smaller after significant capital losses, for example in the 2007-2012 period. Furthermore, the linkage between the Federal Funds rate and the mortgage rate is also state dependent, with risk spreads dominating the effect of the policy rate in the 2007-2012 period. The US mortgage market, dominated by fixed rate mortgages, offers a cheap refinance option for those with adequate net equity. But in the house price collapse, after net equity evaporated, many were no longer able to refinance, which slowed the transmission of the policy rate to the effective mortgage rate available to households. These are three reasons why the monetary transmission to house prices was necessarily weaker and slower in the 2007-2012 period. This made it all the more important that credit easing and the use of FHA-underwritten mortgages to replace new private label mortgage backed securities, which had virtually ceased to be available, came to the rescue of the system. Since our research on consumption also demonstrates an important time-varying effect of housing collateral on consumption, and hence on aggregate demand and labour market slack, these arguments suggest that conventional monetary policy was particularly ineffective in the 2007-2012 period in the US.[37]

In the forecasting model above, part of inflation dynamics come from the adjustment of consumer prices, when the long-run drivers shift. The long-run equilibrium solution for the log core consumption deflator is a linear combination of the log import price index and the log house price index, with the unemployment rate and union density mainly proxying log unit labour costs and the concentration index best interpreted as part of the mark-up between

[36] I am grateful to Ben Friedman for raising the contrasting questions.

[37] As noted in section 2, this diagnosis and the policy solutions could not have been suggested by the New Keynesian DSGE models. In contrast, the household sector models discussed in section 3, incorporating time-varying credit conditions, can explain such time variations in monetary policy effectiveness. Embedding them in larger models to more fully capture the feedback loops is necessary.

consumer prices and costs of production.[38] Figure 5 plots the log consumption deflator minus the weighted average of log import prices and log house prices against the other longer run components of the solution: the unemployment rate, union density and firm concentration.[39] If union density and firm concentration stabilise, some upward pressure on the *relative* consumption deflator should be emerging from the lower unemployment rate. Onto the consumption deflator itself, as of 2018, there should be a feed-through still to come from higher house prices, and for import prices, a battle between higher tariffs increasing prices and an appreciating exchange rate lowering prices.

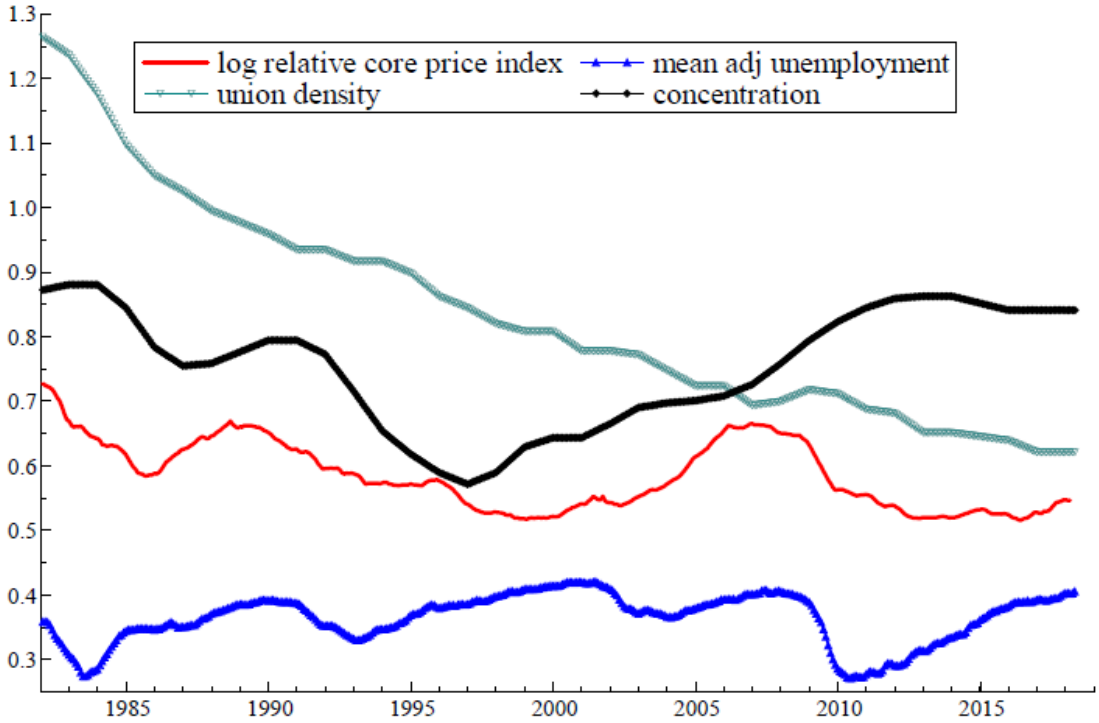


Figure 5: long-run contributions to the log relative core price index of the unemployment rate, union density and firm concentration.

The three longer run drivers shown in Figure 5 do *not* add up to the relative consumption deflator because of lagged adjustment and because of other complex, lagged inflation

[38] This does not exclude the possibility that, in this reduced form forecasting model, the unemployment rate might also affect the mark-up and that the concentration ratio might also proxy monopsony power of firms in the labour market, so affecting unit labour costs.

[39] For example the concentration effect is the Herfindahl concentration index times the end-sample value of its coefficient plotted in Figure 4, divided by the sum of the end-sample coefficients on log import prices and log house prices.

dynamics, the reasons for which were discussed above. One of the intriguing findings on these dynamics is the *negative* effect of core inflation in the current and previous year on the next year's core inflation, given equilibrium correction and conventionally signed effects of recent growth of earnings, import prices, other foreign prices and the exchange rate. One interpretation of this finding is as part of an anti-inflation feedback rule of the Federal Reserve, tightening policy in response to recently higher core inflation. But it is also possible that businesses, where prices have increased more than in potentially competitive sectors, find themselves losing market share and so are forced to price more moderately. Note that this negative feedback is additional to that already represented through the equilibrium correction terms based on relative prices. These finding puts an entirely different perspective on the apparent persistence of inflation implied by naïve univariate forecasting models.

Farmer and Nicolo (2018) propose a Keynesian model *without* the Phillips curve. They argue: "Central bankers use the concept of a time-varying natural rate of unemployment before deciding when and if to raise the nominal interest rate. The difficulty of estimating the natural rate arises, in practice, because the economy displays no tendency to return to its natural rate. That fact has led to much recent skepticism about the usefulness of the Phillips curve in policy analysis." While they are surely right to want to abolish the *conventional* New Keynesian Phillips Curve, there is nevertheless strong evidence, see above, that unemployment has an important effect on the core consumption deflator and hence on core inflation, *provided the right controls, including long-run variables and at long enough lags, are included*.[40] Thus, a broader and more economically satisfactory concept of the Phillips Curve could be resurrected which we could christen 'the Sargan-Phillips Curve', as Sargan (1964) was the first to emphasise the role of equilibrium correction in the inflation context. In other words, he was the first to formalise the idea that part of the inflation process is dynamic adjustment of relative prices as explained above. Further powerful long-run evidence for this view for the UK is provided by Hendry (2001).

The results of Aron and Muellbauer (2013, 2018) for models of US PCE and the core consumption deflator, and in Hendry (2001) for the UK GDP deflator, throw serious doubt on the usefulness of the NAIRU concept. The empirical results strongly contradict the accelerationist view of inflation. As noted above, there is a tendency for high annual rates of

[40] The models in Aron and Muellbauer (2013, 2018) with the widest information sets also include the short-term dynamics of hourly earnings in the US. As these are also affected by unemployment, the *net* effect is even larger than indicated above.

inflation in the US to be followed by *lower* inflation a year later, given the other controls. One could define a kind of 'natural rate of unemployment' as the rate consistent with say 2 percent inflation. However, as our results interpret a substantial part of inflation as being the process of adjustment of the price level towards a long-run level dependent on several other factors as well as the unemployment rate, it is wrong to focus narrowly on the unemployment rate. Of the long-run relative price variables, import prices or the real exchange rate and the level of house prices are also potentially responsive to monetary policy. Moreover, the level of concentration among public companies, and union density, are potentially affected by regulatory policies. For policy analysis, therefore, it is far better to embed a comprehensive inflation model, embodying these more general drivers of inflation and their longer lags, into an econometric model in which alternative policy options can be dynamically simulated. Constâncio (2018), reviewing some of the empirical evidence against the New Keynesian Phillips Curve, also expresses severe scepticism about the usefulness of the NAIRU and natural rate concepts.

5. Forecasting per capita household income and GDP for France

A further example of how Parsimonious Longer Lags can improve forecasts is provided by models of real GDP and of household income. Aron and Muellbauer (2002) used the PLL tool to develop a forecasting model for real GDP in South Africa, where multiple regime shifts also had to be taken into account.[41] For the consumption function, as noted in section 3, it is important to incorporate permanent income, capturing income growth expectations. This permits the separation of effects on consumption of current income from longer-term expectations. Since the stock market turns out to have predictive power for permanent income, the pure wealth effect can be separated from the income expectations effect. The NK-DSGE model assumed, by the omission of equity prices and balance sheets, that equity prices are a side-show, merely reflecting income growth expectations. Instead, we find significant illiquid financial wealth effects in every country studied even after controlling for permanent income.

Chauvin and Muellbauer (2018) develop a model for forecasting permanent real per capita household income assuming a 5% per quarter discount rate over a 10-year horizon. The FRB-US model assumes a similarly high discount rate, which is broadly consistent with the insights of the buffer-stock models of household saving, averaging over households. In a quarterly

[41] In turn, this built on insights developed for forecasting annual real GDP for the US in Muellbauer (1996).

consumption function, the weight for France is around 0.55 for log permanent income and 0.45 for log current income, given controls including asset to income ratios and credit conditions. This is consistent with a spate of micro-evidence on the effects, for example, of temporary tax cuts or income tax credits, though of course, there is a great deal of heterogeneity at the micro level.

The forecasting model includes a rich set of determinants, using 4-quarter moving averages to parsimoniously capture longer lags. The economic variables with a negative effect on the log permanent real per capita income include changes in nominal, and levels of real interest rates, current log real per capita income because of trend reversion, the unemployment rate (because it weakens the power of workers in wage negotiations and hence their share of national income), log real oil prices and the log real exchange rate, indicating worsening competitiveness. The real oil price, the real interest rate and the real exchange rate have particularly persistent effects on the ratio of permanent to current income, possibly reflecting their impact on investment and innovation, but perhaps also on the wage bargaining power of workers. The long lags may be partly the result of the high degree of state intervention in France, e.g. in social benefits, to stabilise household incomes. This is likely to slow the speed at which shocks are transmitted to income.

Variables with a positive effect include a survey measure of short-term household expectations, recent changes in log real per capita income, possibly indicating some growth momentum, the log stock market index in real terms (it indicates expectations of productivity growth and is one of the drivers of capital investment which expands future capacity), and finally the ratio of the working age population divided by the total population. Its future decline is likely to limit per capita income growth.

One cannot literally interpret such a complex model in terms of households having information on all the variables included. However, the media environment is one in which economic prospects and many expert views, for example, from the OECD, the IMF and financial commentators, are widely discussed. The negative effects on income prospects of high real oil prices, high real interest rates and an uncompetitive exchange rate are widely understood by experts. Arguably such a model represents a maximally 'rational' interpretation of household income expectations, given a world subject to evolution and structural breaks.

Linear trends are included, including one in the late 1970s and one beginning with the financial crisis. However, households cannot possibly have foreseen the financial crisis. Hence for a model like this, which builds in the ex-post effects of the crisis, to have a chance of representing their expectations, it is necessary to remove, before the crisis occurred, the effect

of the downshift in the crisis of the linear trend from the generated permanent income measure used to model consumption and house prices. We then assume a gradual, two-year process of learning, in which households absorb information about the downshift in trend growth.

In a parallel model to forecast French per capita real GDP one and two years ahead, similar variables are relevant, with the exception of the unemployment rate. The lags are long and application of the Bayesian information criterion to the standard unrestricted lag structure as used in VARs, results in a drastic deterioration in fit and forecasting performance. However, lags are typically not quite as long as for equivalent one and two-year ahead household income equations, and real GDP per capita is considerably more volatile and harder to forecast. This is consistent with the stabilising role for household income of the state in France discussed above.

6. Conclusions

This paper has provided argument and macroeconomic evidence from diverse research in favour of the relevance of the 'information economics revolution' to which Joseph Stiglitz made such important contributions. A strong implication is that credit constraints are a ubiquitous feature of the macro economy. It then follows, that given the post-war transformations of banking and credit institutions, *shifts* in access to credit need to be taken into account in building macroeconomic models. This is an insight that was widely ignored not only in the New Keynesian DSGE literature but is still neglected in the large non-DSGE macro-econometric models used at many central banks.

Financial stability continues to be a major issue for central banks and better macro-models are needed in order to understand the feedback mechanisms that can amplify shocks and lead to financial instability. Systems of equations that my co-authors and I have been developing for the household sector, including consumption, portfolio composition and house prices, do take shifting credit constraints into account. Often, it is impossible to obtain coherent models of consumption, house prices and mortgage debt without such shifts. These models highlight institutional differences between countries and over time, helping to understand why some are much more prone to amplifying feedbacks that reduce financial stability. Macro models incorporating these new models of the household sector would be a useful complement to stress-testing and micro-evidence on household vulnerabilities to enhance the application, where necessary of macro-prudential tools.

The information economics revolution supports Stiglitz's crucial insight that path dependence is a key to understanding the evolution of the macro-economy.[42] The evolutionary nature of economic processes and the highly limited ability of agents to foresee the future suggests the macro-economic profession needs to take account of the information embedded in far longer lags than is conventional.

The fashions in macroeconomics, with DSGE models on the one hand, and Bayesian-VAR models on the other, have led to economists paying *far too little attention* to the long-run. As a result, they have not been able to learn fully the lessons buried in macroeconomic data. This refers both to mis-specification through the omission of variables that reflect long-run relationships and the omission of sufficiently long lags. One example concerns the difficult issues of identifying empirically the slowly evolving influence of demography on income growth, household saving behaviour and potentially on future real interest rates. The household systems approach discussed above, applied to long samples of data and more than one country, promises to help resolve these difficulties.

Many economists understand the concept of cointegration of non-stationary data and the importance of the long-run relationships. This is not a new point. But in the context of inflation modelling and forecasting, it is too often neglected, as the vast majority of models are structured in the differenced mode, without long-run variables such as relative prices, despite the classic paper by Sargan (1964).

The simple 'Parsimonious Longer Lag' tool, combined with the model selection and forecasting insights that David Hendry has developed over many years, promise significant future improvements in learning from data when modelling and forecasting the macro-economy. Evidence has been provided that the application of these tools greatly improves the understanding and forecasting of US inflation and of fluctuations in household income and GDP in France.

We find that while the link from unemployment to inflation has not become flatter, that between the US policy rate and US inflation *has* become flatter, with conventional policy particularly ineffective in the 2009 – 2012 period. These findings are consistent with a more general point made by Constâncio (2017c). He asks whether the balance sheet measures adopted by major central banks during the financial crisis should remain permanent parts of the central bank toolkit. He says: "there are good arguments to preserve the instrument in the policy toolkit. They stem from structural changes that have occurred in financial markets." To

[42] Durlauf (1991) has highlighted some of the specific mechanisms by which path dependence influences GDP.

those changes mentioned by Constâncio, one should add the massive shift by banks to real estate collateral backed lending, tracked by Jorda et al (2016), and the rise in leverage discussed in section 3, which new financial regulations have far from fully reversed.

References

Ahir, H. and Loungani, P. (2014), ‘‘There will be growth in the spring’: How well do economists predict turning points?', https://voxeu.org/article/predicting-economic-turning-points

Anundsen, A. K., and C. Heeboll. (2016), ‘Supply restrictions, subprime lending and regional US house prices’, Journal of Housing Economics, 31 (1): 54-72.

Aron, J. and J. Muellbauer (2002) ‘Interest Rate Effects on Output: Evidence from a GDP Forecasting Model for South Africa’, IMF Staff Papers, 49 (S1), 185–213.

Aron, J., and Muellbauer, J. (2013), ‘New Methods for Forecasting Inflation, Applied to the USA’, Oxford Bulletin of Economics and Statistics, 75, 637–61.

Aron, J., and J. Muellbauer, J. (2016), ‘Modelling and Forecasting Mortgage Delinquency and Foreclosure in the UK’, Journal of Urban Economics, 94 (C): 32–53.

Aron, J., and Muellbauer, J. (2018), ‘The reincarnation of the Phillips Curve, with new forecasting evidence for US core inflation’, MS in preparation.

Aron, J., Duca, J., Muellbauer, J., Murata, K., and Murphy, A. (2012), ‘Credit, housing collateral, and consumption: Evidence from Japan, the UK, and the US’, Review of Income and Wealth, 58(3), 397–423.

Bernanke, B., Gertler, M., and Gilchrist, S. (1999), ‘The Financial Accelerator in a Quantitative Business Cycle Framework’, in J. B. Taylor and M. Woodford (eds), Handbook of Macroeconomics: Volume 1, Amsterdam, Elsevier, 1341–93.

Bertola, G., and Caballero, R. J. (1990), ‘Kinked Adjustment Costs and Aggregate Dynamics’, NBER Macroeconomics Annual, 5, 237–96.

Blanchard, O. J. (2016), ‘Do DSGE Models Have a Future?’, at https://piie.com/publications/policy-briefs/do-dsge-models-have-future .

Blanchard, O. J. (2017), ‘Distortions in Macroeconomics’, at https://piie.com/commentary/speeches-papers/distortions-macroeconomics .

Blanchard, O. (2018), ‘On the Future of Macroeconomic Models’, Oxford Review of Economic Policy, 34(1–2), 43–54.

Browning, M., Gørtz, M. and Leth-Petersen, S. (2013), ‘Housing Wealth and Consumption: A Micro Panel Study.’ Economic Journal, 123, 401–428.

Brunnermeier, M. K., and Y. Sannikov. (2014), ‘A Macroeconomic Model with a Financial Sector’, American Economic Review, 104(2), 379-421.

Buiter, W. (2009), ‘The Unfortunate Uselessness of Most ‘State of the Art’ Academic Monetary Economics’, VOXEU, March, http://voxeu.org/article/macroeconomics-crisis-irrelevance .

Burrows, V. (2018), ‘The Impact of House Prices on Consumption in the UK: a New Perspective’, Economica, 85 (337), 92-123.

Caballero, R. J. (2010), ‘Macroeconomics after the Crisis: Time to Deal with the Pretense-of-Knowledge Syndrome,’ Journal of Economic Perspectives, American Economic Association, 24(4), 85-102, Fall.

Calvo, G. A. (1983), ‘Staggered Pricing in a Utility Maximizing Framework’, Journal of Monetary Economics, 12, 383-398.

Canova, F. (1994), ‘Detrending and Turning-Points’, European Economic Review, 38 (3/4), 614–623.

Canova, F. (1998), ‘Detrending and Business Cycle Facts’, Journal of Monetary Economics, 41(3), 475–512.

Carlin, W., and Soskice, D. W. (2015), Macroeconomics: Institutions, Instability, and the Financial System, Oxford, Oxford University Press.

Carlin, W., and Soskice, D. W. (2018), ‘Stagnant Productivity and Low Unemployment: Stuck in a Keynesian Equilibrium’, Oxford Review of Economic Policy, **34**(1–2), 169–94.

Carroll, C. D., (1992), ‘The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence’, Brookings Papers on Economic Activity, 23, 61–156.

Case, K. E., and R. J. Shiller. (2003), ‘Is There a Bubble in the Housing Market?’ Brookings Papers on Economic Activity 2: 299–342.

Castle, J. L. & Hendry, D. F., 2009. ‘The long-run determinants of UK wages, 1860-2004,’ Journal of Macroeconomics, 31(1), 5-28.

Castle, J. L., Fawcett, N. W. P., and Hendry, D. F. (2011), ‘Forecasting breaks and during breaks’, in M.

M. P. Clements and D. F. Hendry (eds.), Oxford Handbook of Economic Forecasting, Oxford: Oxford University Press, 5–353.

Cerutti, E., J. Dagher, and Dell’Ariccia, G. (2017), ‘Housing Finance and Real Estate Booms: A Cross-Country Perspective’, Journal of Housing Finance, 38, 1-13.

Christiano, L. R., M. S. Eichenbaum and M. Trabandt (2017), 'On DSGE Models', Nov. 9, http://faculty.wcas.northwestern.edu/~yona/research/DSGE.pdf
Christiano, L. R., M. S. Eichenbaum and M. Trabandt (2018), 'On DSGE Models', Journal of Economic Perspectives, 32(3), 113–140.
Christiano, L. R. Motto, and Rostagno, M. (2014), 'The Great 'Risk Shocks', American Economic Review, 104 (1), 27–65.
Chauvin, V. and J. Muellbauer (2018) 'Consumption, household portfolios and the housing market in France', Economie et Statistique/Economics and Statistics, 500-501-502, 151-178.
Christiano, L., Eichenbaum, M., and Evans, C. (2005), 'Nominal rigidities and the dynamics effects of a shock to monetary policy', Journal of Political Economy, 113, 1–45.
Christiano, L. R., Motto, R., and Rostagno, M. (2003), 'The Great Depression and the Friedman–Schwartz Hypothesis', Journal of Money, Credit and Banking, 35(6), 1119–97.
Clarida, R., Galí, J., and Gertler, M. (1999), 'The Science of Monetary Policy: A New Keynesian Perspective', Journal of Economic Literature, 37, 1661–707.
Clements, M. P., and Hendry, D. F. (1996), 'Intercept Corrections and Structural Change', Journal of Applied Econometrics, 11(5), 475–94.
Clements, M. P., and Hendry, D. F. (1999), 'Forecasting Non-stationary Economic Time Series', Cambridge, MA, MIT Press.
Constâncio, V. (2017a), 'Macroprudential stress tests: a new analytical tool', VoxEu, February 2017.
Constâncio, V. (2017b), 'Macroprudential policy in a changing financial system', https://www.ecb.europa.eu/press/key/date/2017/html/ecb.sp170511.en.html.
Constâncio, V. (2017c), 'The future of monetary policy frameworks', lecture at the Instituto Superior de Economia e Gestão, Lisbon, 25 May 2017. https://www.ecb.europa.eu/press/key/date/2017/html/ecb.sp170525_lecture.en.pdf
Constâncio, V. (2018a), 'Financial stability risks and macroprudential policy in the euro area', https://www.ecb.europa.eu/press/key/date/2018/html/ecb.sp180314_3.en.html
Constâncio, V. (2018b), 'Past and future of the ECB monetary policy', Speech at the Conference on 'Central Banks in Historical Perspective: What Changed After the Financial Crisis?', at the Central Bank of Malta, Valletta, 4 May 2018.
Corbae, D. and E. Quintin (2015), 'Leverage and the Foreclosure Crisis, Journal of Political Economy', 123, 1-65.
Cusbert, T. and E. Kendall (2018), 'Meet MARTIN, the RBA's New Macroeconomic Model', Australian Reserve Bank Bulletin, March, 31-44.
Deaton, A. S. (1991), 'Saving and Liquidity Constraints', Econometrica, 59, 1221–48.
Dickson, V. (2006), 'Price-cost margins, prices and concentration in US manufacturing: a panel study', Applied Economics Letters, 12 (2), 79-83.
Doornik, J.A. (2009), 'Autometrics', in J.L. Castle and N. Shephard (eds.), The Methodology and Practice of Econometrics, Oxford: Oxford University Press.
Duca, J., Muellbauer, J., and A. Murphy. (2011), 'House Prices and Credit Constraints: Making Sense of the US Experience', The Economic Journal, 121(552), 533-551.
Duca, J., and J. Muellbauer. 2013. 'Tobin LIVES: Integrating Evolving Credit Market Architecture into Flow of Funds Based Macro-Models.' in B. Winkler, A. van Riet and P. Bull (eds), A flow-of-funds perspective on the financial crisis Volume 2, London: Palgrave-Macmillan. Also ECB Working Paper 1581, August.
Duca, J. V., Muellbauer, J., and Murphy, A. (2016), 'How Mortgage Finance Reform Could Affect Housing', American Economic Review, 106(5), 620–4.
Durlauf, S.N. (1991) 'Path Dependence in Aggregate Output' NBER Working Paper No. 3718, published in Industrial and Corporate Change (1994), 3(1) 149–171.
Farmer, R.E.A. and Nicolò, G. (2018), 'Keynesian Economics without the Phillips Curve', Journal of Economic Dynamics and Control, 89, 137-150.
Favara, G., and Imbs, J. (2015), 'Credit Supply and the Price of Housing', American Economic Review, 105(3), 958–92.
Fisher, I. (1933), 'The Debt-Deflation Theory of Great Depressions', Econometrica, 1, 337–57.
Geanakoplos, J. (2010), 'Solving the Present Crisis and Managing the Leverage Cycle', New York Federal Reserve Economic Policy Review, 16 (1), 101–31.
Gertler, M. and Gilchrist, S. 2018. "What Happened: Financial Factors in the Great Recession." Journal of Economic Perspectives, 32 (3): 3-30.

Gertler, M., N. Kiyotaki and Prestipino, A. (2016), 'Wholesale banking and bank runs in macroeconomic modelling of financial crises,' Handbook of Macroeconomics, 2, 1345–1425.
Geiger, F., Muellbauer, J., and Rupprecht, M. (2016), 'The Housing Market, Household Portfolios and the German Consumer', Working Paper Series No. 1904, European Central Bank.
Gordon, R. (2009), 'The history of the Phillips Curve: consensus and bifurcation', Economica 78(309):10-50.
Gordon, R. J. (2013). 'The Phillips Curve is alive and well: Inflation and the NAIRU during the slow recovery', NBER Working Paper 19390.
Grullon, G., Y. Larkin and R. Michaely (2017) 'Are U.S. Industries Becoming More Concentrated?' Technical Report. https://ssrn.com/abstract=2612047 or http://dx.doi.org/10.2139/ssrn.2612047
Haavelmo, T. (1944), 'The Probability Approach in Econometrics', Econometrica, 12, 1–115.
Haldane, A. G., and Turrell, A. E. (2018), 'An Interdisciplinary Model for Macroeconomics', Oxford Review of Economic Policy, 34(1–2), 219–51.
Hedlund, A., Karahan, F., Mitman, K., and Ozkan, S. (2016), 'Monetary Policy, Heterogeneity, and the Housing Channel', Working Paper, University of Missouri.
Hendry, D. F. (2001). Modelling UK inflation, 1875–1991. Journal of Applied Econometrics, 16, 255–275.
Hendry, D. (2011), 'Forecasting from Mis-specified Models in the Presence of Unanticipated Location Shifts', in M. P. Clements and D. F. Hendry (eds.), Oxford Handbook of Economic Forecasting, Oxford: Oxford University Press, 271-314.
Hendry, D. (2017), 'Deciding between Alternative Approaches in Macroeconomics', International Journal of Forecasting, 34, 119-135.
Hendry, D. F., and Mizon, G. E. (2014), 'Unpredictability in Economic Analysis, Econometric Modelling and Forecasting', Journal of Econometrics, 182, 186–95. Also see http://www.voxeu.org/article/why-standard-macro-models-fail-crises.
Hendry, D. F., and Muellbauer, J. N. J. (2018), 'The Future of Macroeconomics: Macro Theory and Models at the Bank of England', Oxford Review of Economic Policy, 34(1–2), 287–328.
Hoover, K. D. (1994), 'Econometrics as Observation: The Lucas Critique and the Nature of Econometric Inference', Journal of Economic Methodology, 1, 65–80.
Houthakker, H. S. (1956), 'The Pareto Distribution and the Cobb–Douglas Production Function in Activity Analysis', Review of Economic Studies, 23, 27–31.
Huang, Kevin X.D. and Liu, Zheng, "Input-Output Structure and the General Equilibrium Dynamics of Inflation and Output" (2000). Economic Research Institute Study Papers. Paper 186.
Iacoviello, M. (2005), 'House Prices, Borrowing Constraints, and Monetary Policy in the Business Cycle', American Economic Review, 95, 739–64.
Iacoviello, M., and Neri, S. (2010), 'Housing Market Spillovers: Evidence from an Estimated DSGE Model', American Economic Journal: Macroeconomics, 2, 125–64.
Jordà, O., Schularick, M., and Taylor, A. M. (2016), 'The Great Mortgaging: Housing Finance, Crises and Business Cycles', Economic Policy, 31(85), 107–52.
Lindé, J. (2018), 'DSGE Models: Still Useful in Policy Analysis?', Oxford Review of Economic Policy, 34(1–2), 269–86.
Jordà, Ò, Schularick, M. and Taylor, A. M. (2016), 'The great mortgaging: housing finance, crises and business cycles', Economic Policy, 31 (85), 107–152.
Kaplan, G., Moll, B., and Violante, G. L. (2016), 'Monetary Policy According to HANK', American Economic Review, 108(3): 697-743.
Kaplan, G. and L. Violante, G.L. 2018. "Microeconomic Heterogeneity and Macroeconomic Shocks." Journal of Economic Perspectives, 32 (3): 167-94.
Korinek, A. (2015), 'Thoughts on DSGE Macroeconomics: Matching the Moment, But Missing the Point?' Working Paper, Johns Hopkins University, August.
Loungani, P. (2001), 'How accurate are private sector forecasts?', Cross-country evidence from consensus forecasts of output growth', International Journal of Forecasting, 17(3), 419-432.
Lucas, R. E. (1976), 'Econometric Policy Evaluation: A Critique', in K. Brunner and A. Meltzer (eds), The Phillips Curve and Labor Markets, Carnegie-Rochester Conferences on Public Policy, Amsterdam, North-Holland, 19–46.
Mian, A., and Sufi, A. (2014), House of Debt, Chicago, IL, University of Chicago Press.
Mian, A., and Sufi, A. (2016), 'Who Bears the Cost of Recessions? The Role of House Prices and Household Debt', NBER Working Paper 2225.
Mian, A., Rao, K., and Sufi, A. (2013), 'Household balance sheets, consumption, and the economic slump', The Quarterly Journal of Economics, 128(4), 1687–1726.
Mian, A., A. Sufi and Verner, E. (2017), 'Household Debt and Business Cycles Worldwide', The Quarterly Journal of Economics, 132 (4), 1755-1817.

Mian, A. and Sufi, A. 2018. "Finance and Business Cycles: The Credit-Driven Household Demand Channel." Journal of Economic Perspectives, 32 (3): 31-58.
Muellbauer, J. (1996), 'Income Persistence and Macropolicy Feedbacks in the U.S.', Oxford Bulletin of Economics and Statistics, 58, 703–33.
Muellbauer, J. and A. Murphy (1997), 'Booms and Busts in the UK Housing Market,' Economic Journal 107: 1701-27.
Muellbauer, J. (2010), 'Household Decisions, Credit Markets and the Macroeconomy, Implications for the Design of Central Bank Models', Bank for International Settlements Discussion Paper 306. http//ideas.repec.org/p/bis/biswps/306.html
Muellbauer, J. (2012), 'When is a Housing Market Overheated Enough to Threaten Stability?,' RBA Annual Conference Volume, in: A. Heath, F. Packer and C. Windsor (eds.), Property Markets and Financial Stability, Reserve Bank of Australia. http://www.rba.gov.au/publication s/confs/2012/index.html
Muellbauer, J., and Lattimore, R. (1995), 'The consumption function: a theoretical and empirical overview', in M. H. Pesaran and M. Wickens (eds.), Handbook of Applied Econometrics: Macroeconomics, Blackwell, Oxford, 221-311.
Reis, R. (2018), 'Is Something Really Wrong with Macroeconomics?', Oxford Review of Economic Policy, 34 (1–2), 132–55.
Salinger, M. (1990), 'The Concentration-Margins Relationship Reconsidered', Brookings Papers in Economic Activity, 287-335.
Sargan, J. D. (1964), 'Wages and Prices in the United Kingdom: A Study of Econometric Methodology', In Hart, P. E., G. Mills and J. K. Whitaker (eds.), Econometric Analysis for National Economic Planning, 25—63. Butterworth Co., London.
Sims, C. A. (1980), 'Macroeconomics and reality', Econometrica, 48 (January), 1–48.
Smets, F. and Wouters, R. (2003), 'An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area', Journal of the European Economic Association,1(5), 1123–1175.
Smets, F. and Wouters, R. (2007), 'Shocks and frictions in U.S. business cycles: a Bayesian DSGE approach', American Economic Review, 97 (3), 586–606.
Stiglitz, J. E. (2001), 'Information and the change in the paradigm in economics', Nobel lecture. https://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/2001/stiglitz-lecture.html
Stiglitz, J. E. (2018), 'Where Modern Macroeconomics Went Wrong', Oxford Review of Economic Policy, 34(1–2), 70–106.
Stock, J. H. and Watson, M. W. (2011), 'Dynamic Factor Models', in M. P. Clements and D. F. Hendry (eds.), Oxford Handbook of Economic Forecasting, Oxford: Oxford University Press, 35-60.
Stock, J. H. and Watson, M. W. (1999), 'Forecasting Inflation', Journal of Monetary Economics, 44, 293–335.
Stock, J. H. and Watson, M. W. (2003), 'Forecasting Output and Inflation: The Role of Asset Prices', Journal of Economic Literature, 41, 788-829.
Stock, J. H. and Watson, M. W. (2007), 'Why Has U.S. Inflation Become Harder to Forecast?', Journal of Money, Credit, and Banking, 39, 3-34.
Turner, A. (2015), Between Debt and the Devil: Money, Credit, and Fixing Global Finance, Princeton, NJ: Princeton University Press.
Windsor, C., Jääskelä, J.P. and Finlay, R. (2015),'Housing Wealth Effects: Evidence from an Australian Panel', Economica, 82, 552-577.
Wren-Lewis, S. (2018), 'Ending the Microfoundations Hegemony', Oxford Review of Economic Policy, 34 (1–2), 55–69.
Yellen, Janet L. (2017), 'Inflation, Uncertainty, and Monetary Policy', at 'Prospects for Growth: Reassessing the Fundamentals' 59th Annual Meeting of the National Association for Business Economics Cleveland, Ohio, September 26, 2017.